Ruprecht-Karls-Universität Heidelberg
Institut für Mathematik
Professor Dr. Jan JOHANNES

UNIVERSITÄT HEIDELBERG
ZUKUNFT SEIT 1386

STATISTICS OF INVERSE PROBLEMS

Preliminary version: July 6, 2023

If you find **errors in the outline**, please send a short note
by email to johannes@math.uni-heidelberg.de

MATHEMATIKON, Im Neuenheimer Feld 205, 69120 Heidelberg
phone: +49 6221 54.14.190 – Fax: +49 6221 54.14.101
email: johannes@math.uni-heidelberg.de
webpage: sip.math.uni-heidelberg.de

Table of contents

Chapter 1

Statistical inverse problems

The observable signal $g = \mathrm{T}\theta$ corrupted with an additive noise is first formalised in this chapter and secondly the noisy observation of the operator.

Overview

§01 Noisy image and known operator

Let $(\mathbb{H}, \langle\cdot,\cdot\rangle_{\mathbb{H}})$ and $(\mathbb{G}, \langle\cdot,\cdot\rangle_{\mathbb{G}})$ be separable real Hilbert spaces and let $\mathrm{T} : \mathbb{H} \to \mathbb{G}$ be a known linear, bounded operator, $\mathrm{T} \in \mathbb{L}(\mathbb{H},\mathbb{G})$ in short. We are interested in the reconstruction of $\theta \in \mathbb{H}$ from a noisy version of $g = \mathrm{T}\theta$, which we formalise first in this section by introducing stochastic processes.

§01|00.01 **Notation**. For $x, y \in \mathbb{R}$ we agree on the following notations $\lfloor x \rfloor := \max\{k \in \mathbb{Z}\colon k \in (-\infty, x]\}$ (integer part), $x \vee y = \max(x, y)$ (maximum), $x \wedge y = \min(x, y)$ (minimum), $\{x\}_+ = \max(x, 0)$ (positive part), $\{x\}_- = \max(-x, 0)$ (negative part) and $|x| = \{x\}_+ + \{x\}_-$ (modulus).

(a) For $c \in \mathbb{R}$ and $\mathbb{A} \subseteq \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\} = [-\infty, \infty]$ we set $\mathbb{A}_{\geqslant c} := \mathbb{A} \cap [c, \infty]$, $\mathbb{A}_{\leqslant c} := \mathbb{A} \cap [-\infty, c]$, $\mathbb{A}_{>c} := \mathbb{A} \cap (c, \infty]$, $\mathbb{A}_{<c} := \mathbb{A} \cap [\infty, c)$, $\mathbb{A}_{\setminus c} := \mathbb{A} \setminus \{c\}$, and $\overline{\mathbb{A}} := \mathbb{A} \cup \{\pm\infty\}$.

(b) For $b \in \overline{\mathbb{R}}$ and $a \in \overline{\mathbb{R}}_{<b}$ we write $[\![a,b]\!] := [a,b] \cap \overline{\mathbb{Z}}$, $[\![a,b)\!) := [a,b) \cap \overline{\mathbb{Z}}$, $(\!(a,b]\!] := (a,b] \cap \overline{\mathbb{Z}}$, and $(\!(a,b)\!) := (a,b) \cap \overline{\mathbb{Z}}$. Moreover, let $[\![n]\!] := [\![1,n]\!]$ and $[\![n)\!) := [\![1,n)\!)$ for $n \in \mathbb{N} = \mathbb{Z}_{>0}$.

(c) For a *σ-algebra* $\mathscr{A}$ we denote by $\mathscr{A}_{\mathbb{A}} := \mathscr{A} \cap \mathbb{A}$ the trace of $\mathscr{A}$ over a *set* $\mathbb{A}$ which is for $\mathbb{A} \in \mathscr{A}$ a σ-algebra too. For $c \in \mathbb{R}$ we set $\mathscr{A}_{\geqslant c} := \mathscr{A} \cap [c,\infty]$, $\mathscr{A}_{>c} := \mathscr{A} \cap (c,\infty]$, $\mathscr{A}_{\leqslant c} := \mathscr{A} \cap [\infty, c]$, and $\mathscr{A}_{<c} := \mathscr{A} \cap [\infty, c)$. We denote by $\overline{\mathscr{B}} := \mathscr{B}_{\overline{\mathbb{R}}}$ the Borel-σ-algebra over the compactified real line $\overline{\mathbb{R}}$, where the sets $\{-\infty\}$, $\{\infty\}$ and $\mathbb{R}$ are in $\overline{\mathbb{R}}$ closed and open, respectively, and hence Borel-measurable. Note that $\mathscr{B} := \mathscr{B}_{\mathbb{R}}$ is the Borel-σ-algebra over $\mathbb{R}$.

(d) Given two measurable space $(\Omega, \mathscr{A})$ and $(\Omega_2, \mathscr{A}_2)$ we denote by $\mathcal{M}(\mathscr{A}, \mathscr{A}_2)$ the set of all $\mathscr{A}$-$\mathscr{A}_2$ measurable functions mapping Ω into Ω_2. We call $f \in \mathcal{M}(\mathscr{A}) := \mathcal{M}(\mathscr{A}, \mathscr{B})$ and $f \in \overline{\mathcal{M}}(\mathscr{A}) := \mathcal{M}(\mathscr{A}, \overline{\mathscr{B}})$ *real* and *numerical*, respectively. Similarly, $f \in \mathcal{M}_{\geqslant 0}(\mathscr{A}) := \mathcal{M}(\mathscr{A}, \mathscr{B}_{\geqslant 0})$ (or $f \in \overline{\mathcal{M}}_{\geqslant 0}(\mathscr{A}) := \mathcal{M}(\mathscr{A}, \overline{\mathscr{B}}_{\geqslant 0})$) and $f \in \mathcal{M}_{>0}(\mathscr{A}) := \mathcal{M}(\mathscr{A}, \mathscr{B}_{>0})$ (or $f \in \overline{\mathcal{M}}_{>0}(\mathscr{A}) := \mathcal{M}(\mathscr{A}, \overline{\mathscr{B}}_{>0})$) is called *positiv* and *strictly positive*. If $\mathscr{A} = \mathscr{B}$ then we write $\mathcal{M}_{\geqslant 0} := \mathcal{M}_{\geqslant 0}(\mathscr{B})$, $\mathcal{M}_{>0} := \mathcal{M}_{>0}(\mathscr{B})$, $\overline{\mathcal{M}}_{\geqslant 0} := \overline{\mathcal{M}}_{\geqslant 0}(\mathscr{B})$, and $\overline{\mathcal{M}}_{>0} := \overline{\mathcal{M}}_{>0}(\mathscr{B})$ for short. □

§01|01 Stochastic process

§01|01.01 **Notation**. Here and subsequently, a non-empty and generally non-finite subset $\mathcal{J}$ of $\mathbb{N}$, $\mathbb{Z}$ or $\mathbb{R}$ and a subset $\mathcal{U}$ of $\mathcal{J}$ denote an index set. We consider the product spaces $\mathbb{R}^{\mathcal{J}} = \times_{j \in \mathcal{J}} \mathbb{R}$ and $\mathbb{R}^{\mathcal{U}} = \times_{u \in \mathcal{U}} \mathbb{R}$, where we identify the family $y_{\bullet} = (y_j)_{j \in \mathcal{J}} \in \mathbb{R}^{\mathcal{J}}$ and the map $y_{\bullet} : \mathcal{J} \to \mathbb{R}$ with $j \mapsto y_j$. Eventually, we define arithmetic operations on elements of $\mathbb{R}^{\mathcal{J}}$ coordinate-wise, for example meaning $a_{\bullet} b_{\bullet} = (a_j b_j)_{j \in \mathcal{J}}$ and $r a_{\bullet} = (r a_j)_{j \in \mathcal{J}}$ for $a_{\bullet}, b_{\bullet} \in \mathbb{R}^{\mathcal{J}}$ and $r \in \mathbb{R}$. Let us further introduce $0_{\bullet} := (0)_{j \in \mathcal{J}}$ and $1_{\bullet} := (1)_{j \in \mathcal{J}}$. The map $\Pi_{\mathcal{U}} : \mathbb{R}^{\mathcal{J}} \to \mathbb{R}^{\mathcal{U}}$ given by $y_{\bullet} = (y_j)_{j \in \mathcal{J}} \mapsto (y_j)_{j \in \mathcal{U}} =: \Pi_{\mathcal{U}} y_{\bullet}$ is called *canonical projection*. In particular, for each $j \in \mathcal{J}$ the *coordinate map* $\Pi_j := \Pi_{\{j\}} : \mathbb{R}^{\mathcal{J}} \to \mathbb{R}$ is given by $y_{\bullet} = (y_{j'})_{j' \in \mathcal{J}} \mapsto y_j =: \Pi_j y_{\bullet}$. Moreover, $\mathbb{R}^{\mathcal{J}}$ is equipped with the product Borel-σ-algebra $\mathscr{B}^{\otimes \mathcal{J}} := \bigotimes_{j \in \mathcal{J}} \mathscr{B}$. Recall that $\mathscr{B}^{\otimes \mathcal{J}}$ equals the smallest σ-algebra on $\mathbb{R}^{\mathcal{J}}$ such that all coordinate maps Π_j, $j \in \mathcal{J}$ are measurable. i.e., $\mathscr{B}^{\otimes \mathcal{J}} = \sigma(\Pi_j, j \in \mathcal{J})$. Moreover, let $(\mathcal{J}, \mathscr{J}, \nu)$ be a measure space with σ-algebra $\mathscr{J}$ over $\mathcal{J}$ containing all elemenatry events $\{j\}$, $j \in \mathcal{J}$, and σ-finite measure $\nu \in \mathscr{M}_{\sigma}(\mathscr{J})$. We denote by $\mathcal{L}_2(\nu) := \mathcal{L}_2(\mathscr{J}, \nu) := \mathcal{L}_2(\mathcal{J}, \mathscr{J}, \nu) \subseteq \overline{\mathcal{M}}(\mathscr{J})$ the usual set of square integrable numerical functions defined on $(\mathcal{J}, \mathscr{J}, \nu)$. Define the set of equivalence classes $\mathbb{J} := \mathbb{L}_2(\nu) := \mathbb{L}_2(\mathcal{J}, \mathscr{J}, \nu)$, which forms a Hilbert space endowed with usual inner product $\langle \cdot, \cdot \rangle_{\mathbb{J}} := \langle \cdot, \cdot \rangle_{\mathbb{L}_2(\nu)}$ and induced norm $\|\cdot\|_{\mathbb{J}} := \|\cdot\|_{\mathbb{L}_2(\nu)}$. □

§01|01.02 **Comment**. Given a measurable space $(\Omega, \mathscr{A}, \mu)$, $s \in \overline{\mathbb{R}}_{>0}$ and the usual space $\mathcal{L}_s(\Omega, \mathscr{A}, \mu)$ of $\mathcal{L}_s(\mu)$-integrable functions introduce for each $h \in \overline{\mathcal{M}}(\mathscr{A})$, the μ-equivalence class $\{h\}_{\mu} := \{h_{\circ} \in \overline{\mathcal{M}}(\mathscr{A}) \colon h = h_{\circ}\ \mu\text{-a.e.}\}$. Define the set of equivalence classes $\mathbb{L}_s(\mu) := \mathbb{L}_s(\mathscr{A}, \mu) := \mathbb{L}_s(\Omega, \mathscr{A}, \mu)$ $:= \{\{h\}_{\mu} \colon h \in \mathcal{L}_s(\mathscr{A}, \mu)\}$ and $\|\{h\}_{\mu}\|_{\mathbb{L}_s(\mu)} := \|h\|_{\mathcal{L}_s(\mu)}$ for $\{h\}_{\mu} \in \mathbb{L}_s(\mu)$. For $s \in \overline{\mathbb{R}}_{\geqslant 1}$, $(\mathbb{L}_s(\mu), \|\cdot\|_{\mathbb{L}_s(\mu)})$ is a complete normed vector space, i.e. a Banach space. Formally, we denote by $\{\bullet\}_{\mu} : \mathcal{L}_s(\mu) \to \mathbb{L}_s(\mu)$ the natural injection $h \mapsto \{h\}_{\mu}$. In case $s = 2$ the norm $\|\{h\}_{\mu}\|_{\mathbb{L}_2(\mu)} := \|h\|_{\mathcal{L}_2(\mu)} = (\mu(|h|^2))^{1/2}$ is induced by the inner product $(\{h\}_{\mu}, \{h_{\circ}\}_{\mu}) \mapsto \langle \{h\}_{\mu}, \{h_{\circ}\}_{\mu} \rangle_{\mathbb{L}_2(\mu)} := \mu(h h_{\circ})$, and hence $(\mathbb{L}_2(\mu), \langle \cdot, \cdot \rangle_{\mathbb{L}_2(\mu)})$ is a Hilbert space. As usual we identify the equivalence class $\{h\}_{\mu}$ with its representative h, and write $h \in \mathbb{L}_2(\mu)$ for short. If $\lambda = \mu$ is the Lebesgue-measure then we write shortly $(\mathbb{L}_2, \langle \cdot, \cdot \rangle_{\mathbb{L}_2})$ and $\{\bullet\} : \mathcal{L}_2 \to \mathbb{L}_2$. □

§01|01.03 **Stochastic process**. Let $(Y_j)_{j \in \mathcal{J}}$ be a family of real-valued random variables on a common probability space $(\Omega, \mathscr{A}, \mathbb{P})$, that is, $Y_j \in \mathscr{A}$ for each $j \in \mathcal{J}$. Consider the $\mathbb{R}^{\mathcal{J}}$-valued random variable $Y_{\bullet} := (Y_j)_{j\in\mathcal{J}} \in \mathcal{M}(\mathscr{A}, \mathscr{B}^{\otimes\mathcal{J}})$, i.e. $Y_{\bullet} : \Omega \to \mathbb{R}^{\mathcal{J}}$ is a $\mathscr{A}$-$\mathscr{B}^{\otimes\mathcal{J}}$-measurable map given by $\omega \mapsto (Y_j(\omega))_{j\in\mathcal{J}} =: Y_{\bullet}(\omega)$. $Y_{\bullet}$ is called a *stochastic process*. Its *distribution* $\mathbb{P}^{Y_{\bullet}} := \mathbb{P} \circ Y_{\bullet}^{-1}$ is the image probability measure of $\mathbb{P}$ under the map $Y_{\bullet}$, i.e. $Y_{\bullet} \sim \mathbb{P}^{Y_{\bullet}}$ or $\mathbb{P}^{Y_{\bullet}} \in \mathscr{W}(\mathscr{B}^{\otimes\mathcal{J}})$ for short. Further, denote by $\mathbb{P}^{Y_{\mathcal{U}}} = \mathbb{P} \circ Y_{\mathcal{U}}^{-1} = \mathbb{P}^{Y_{\bullet}} \circ \Pi_{\mathcal{U}}^{-1}$ the distribution of the stochastic process $Y_{\mathcal{U}} := \Pi_{\mathcal{U}} Y_{\bullet} = (Y_u)_{u\in\mathcal{U}}$ on $\mathcal{U} \subseteq \mathcal{J}$. The family $(\mathbb{P}^{Y_{\mathcal{U}}})_{\mathcal{U}\subseteq\mathcal{J} \text{ finite}}$ is called *family of finite-dimensional distributions* of $Y_{\bullet}$ or $\mathbb{P}^{Y_{\bullet}}$. In particular, $\mathbb{P}^{Y_j} = \mathbb{P}^{\Pi_j Y_{\bullet}} = \mathbb{P}^{Y_{\bullet}} \circ \Pi_j^{-1} \in \mathscr{W}(\mathscr{B})$ denotes the distribution of $Y_j = \Pi_j Y_{\bullet}$. Furthermore, for $j, j_o \in \mathbb{H}$ we write $\mathbb{P}(Y_j) = \mathbb{P}^{Y_{\bullet}}(\Pi_j)$ and $\mathbb{C}\mathrm{ov}(Y_j, Y_{j_o}) := \mathbb{P}(Y_j Y_{j_o}) - \mathbb{P}(Y_j)\mathbb{P}(Y_{j_o}) = \mathbb{P}^{Y_{\bullet}}(\Pi_j \Pi_{j_o}) - \mathbb{P}^{Y_{\bullet}}(\Pi_j)\mathbb{P}^{Y_{\bullet}}(\Pi_{j_o})$, if it exists, for the expectation of Y_j and the covariance of Y_j and Y_{j_o} with respect to $\mathbb{P}$. □

§01|01.04 **Assumption**. The stochastic process $Y_{\bullet} = (Y_j)_{j\in\mathcal{J}}$ on a measurable space $(\Omega, \mathscr{A})$ as a function $\Omega \times \mathcal{J} \to \mathbb{R}$ with $(\omega, j) \mapsto Y_j(\omega)$ is $\mathscr{A} \otimes \mathscr{J}$-$\mathscr{B}$-measurable, $Y_{\bullet} \in \mathcal{M}(\mathscr{A} \otimes \mathscr{J})$ for short. □

§01|01.05 **Definition**. Let $Y_{\bullet} = (Y_j)_{j\in\mathcal{J}} \sim \mathbb{P}^{Y_{\bullet}}$ be a stochastic process satisfying Assumption §01|01.04. If $\mathbb{P}(|Y_j|) \in \mathbb{R}_{\geqslant 0}$, i.e. $Y_j \in \mathcal{L}_1(\mathbb{P})$ or $\Pi_j \in \mathcal{L}_1(\mathbb{P}^{Y_{\bullet}})$ in equal, for each $j \in \mathcal{J}$, then $\mathrm{m}_{\bullet} := (\mathrm{m}_j := \mathbb{P}(Y_j))_{j\in\mathcal{J}} \in \mathbb{R}^{\mathcal{J}}$ is called *mean function* of $Y_{\bullet}$ where $\mathrm{m}_{\bullet} \in \mathcal{M}(\mathscr{J})$ due to Assumption §01|01.04. If in addition $\nu(\mathrm{m}_{\bullet}^2) \in \mathbb{R}_{\geqslant 0}$, hence $\mathrm{m}_{\bullet} \in \mathbb{J}$, then $\mathrm{m}_{\bullet}$ is called *($\mathbb{J}$-)mean*. If $\mathbb{P}(|Y_j|^2) < \infty$, i.e., $Y_j \in \mathcal{L}_2(\mathbb{P})$ or $\Pi_j \in \mathcal{L}_2(\mathbb{P}^{Y_{\bullet}})$ in equal, for each $j \in \mathcal{J}$, then $\mathrm{cov}_{\bullet,\bullet} = (\mathrm{cov}_{j,j_o} := \mathbb{C}\mathrm{ov}(Y_j, Y_{j_o}))_{j,j_o\in\mathcal{J}} \in \mathbb{R}^{\mathcal{J}^2}$ is called *covariance function* of $Y_{\bullet}$, where $\mathrm{cov}_{\bullet,\bullet} \in \mathcal{M}(\mathscr{J}^2)$ due to Assumption §01|01.04. A linear and bounded (continuous) operator from $\mathbb{J}$ into itself, $\Gamma \in \mathbb{L}(\mathbb{J})$ for short, satisfying $\langle \Gamma x_{\bullet}, y_{\bullet} \rangle_{\mathbb{J}} = \int_{\mathcal{J}} \int_{\mathcal{J}} y_j \,\mathrm{cov}_{j,j_o}\, x_{j_o} \nu(dj)\nu(dj_o)$ for all $y_{\bullet}, x_{\bullet} \in \mathbb{J} = \mathbb{L}_2(\nu)$ is called *covariance operator*of $Y_{\bullet}$ or $\mathbb{P}^{Y_{\bullet}}$. If $Y_{\bullet}$ admits a mean function $\mathrm{m}_{\bullet} \in \mathcal{M}(\mathscr{J})$ (respectively mean $\mathrm{m}_{\bullet} \in \mathbb{J}$) and a covariance function $\mathrm{cov}_{\bullet,\bullet} \in \mathcal{M}(\mathscr{J}^2)$ (respectively covariance operator $\Gamma \in \mathbb{L}(\mathbb{J})$) then we write shortly $Y_{\bullet} \sim \mathrm{P}_{(\mathrm{m}_{\bullet}, \mathrm{cov}_{\bullet,\bullet})}$ (respectively $Y_{\bullet} \sim \mathrm{P}_{(\mathrm{m}_{\bullet}, \Gamma)}$). □

§01|01.06 **Notation**. For notional convenience we eventually identify Y_j and Π_j, i.e. $Y_j \sim \mathbb{P}$ for short. We denote by $\mathscr{W}(\mathscr{B})$ the set of all probability measures on $(\mathbb{R}, \mathscr{B})$, by $\mathscr{W}_2(\mathscr{B}) \subseteq \mathscr{W}(\mathscr{B})$ the subset of all probability measures with finite second moment, by $\mathrm{P}_{(\mu,\sigma^2)} \in \mathscr{W}_2(\mathscr{B})$ a probability measure with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 \in \mathbb{R}_{\geqslant 0}$, and by $\mathrm{P}_{\{0\}\times\mathbb{R}_{\geqslant 0}} = \{\mathrm{P}_{(0,\sigma^2)} \in \mathscr{W}_2(\mathscr{B}) \colon \sigma \in \mathbb{R}_{\geqslant 0}\}$ the subset of all probability distributions with finite second moment and mean zero. For $\mathrm{P}_{(\mu_j,\sigma_j^2)} \in \mathscr{W}_2(\mathscr{B})$, $j \in \mathbb{N}$, we denote by $\otimes_{j\in\mathbb{N}} \mathrm{P}_{(\mu_j,\sigma_j^2)}$ the associated product measure on $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}^{\otimes\mathbb{N}})$. □

§01|01.07 **Remark**. A covariance operator $\Gamma \in \mathbb{L}(\mathbb{J})$ associated with a stochastic process $Y_{\bullet} \sim \mathbb{P}^{Y_{\bullet}}$ is self-adjoint and non-negative definite, $\Gamma \in \mathbb{L}^{\geqslant}(\mathbb{J})$ for short. If

$$\sup\left\{\mathbb{P}(|\nu(y_{\bullet}Y_{\bullet})|^2) \colon y_{\bullet} \in \mathbb{J} = \mathbb{L}_2(\nu), \|y_{\bullet}\|_{\mathbb{J}} \leqslant 1\right\} \in \mathbb{R}_{\geqslant 0},$$

which holds for example if $\mathbb{P}(\|Y_{\bullet}\|_{\mathbb{J}}^2) \in \mathbb{R}_{\geqslant 0}$ or in equal $\|Y_{\bullet}\|_{\mathbb{J}} \in \mathcal{L}_2(\mathbb{P})$ (implying $Y_{\bullet} \in \mathbb{J}$ $\mathbb{P}$-a.s.), then there exists a covariance operator $\Gamma \in \mathbb{L}^{\geqslant}(\mathbb{J})$ satisfying $\langle \Gamma x_{\bullet}, y_{\bullet} \rangle_{\mathbb{J}} = \mathbb{C}\mathrm{ov}(\nu(x_{\bullet}Y_{\bullet}), \nu(y_{\bullet}Y_{\bullet}))$ for all $x_{\bullet}, y_{\bullet} \in \mathbb{J}$. Observe that $\|Y_{\bullet}\|_{\mathbb{J}}^2 = \sup\left\{|\nu(y_{\bullet}Y_{\bullet})|^2 \colon y_{\bullet} \in \mathbb{J}, \|y_{\bullet}\|_{\mathbb{J}} \leqslant 1\right\}$. Note that $\|Y_{\bullet}\|_{\mathbb{J}} \in \mathcal{L}_2(\mathbb{P})$ is a sufficent condition for the existence of a covariance operator, but it is not a necessary condition (see Lemma §01|01.18). □

§01|01.08 **Lemma**. *Let $Y_{\bullet} = (Y_j)_{j\in\mathcal{J}} \sim \mathbb{P}^{Y_{\bullet}}$ be a stochastic process satisfying Assumption §01|01.04 and $Y_j \in \mathcal{L}_2(\mathbb{P})$ for each $j \in \mathcal{J}$, and let $\mathrm{v} \in \mathbb{R}_{\geqslant 1}$.*

(i) *If for all $h_{\bullet} \in \mathbb{J}$*

$$\mathbb{P}\left(|\nu(h_{\bullet}Y_{\bullet})|^2\right) \leqslant \mathrm{v}\|h_{\bullet}\|_{\mathbb{J}}^2 \tag{01.01}$$

then $Y_{\bullet}$ admits a covariance operator $\Gamma \in \mathbb{L}^{\geqslant}(\mathbb{J})$ satisfying $\|\Gamma\|_{\mathbb{L}(\mathbb{J})} \leqslant \mathrm{v}$.

(ii) *If for all $h_{\bullet} \in \mathbb{J}$ in addition to* (01.01) *we have also*

$$\mathbb{P}\big(|\nu(h_{\bullet}Y_{\bullet})|^2\big) - \big|\mathbb{P}\big(\nu(h_{\bullet}Y_{\bullet})\big)\big|^2 \geqslant \mathrm{v}^{-1}\|h_{\bullet}\|_{\mathbb{J}}^2 \tag{01.02}$$

then $\Gamma \in \mathbb{L}^{\geqslant}(\mathbb{J})$ is invertible with inverse $\Gamma^{-1} \in \mathbb{L}(\mathbb{J})$ where $\|\Gamma^{-1}\|_{\mathbb{L}(\mathbb{J})} \leqslant \mathrm{v}$.
Consequently, if (01.01) *and* (01.01) *are satisfied for all $h_{\bullet} \in \mathbb{J}$ then we have*

$$\mathrm{v}^{-1}\|h_{\bullet}\|_{\mathbb{J}}^2 \leqslant \|h_{\bullet}\|_{\Gamma}^2 = \langle \Gamma h_{\bullet}, h_{\bullet}\rangle_{\mathbb{J}} \leqslant \mathrm{v}\|h_{\bullet}\|_{\ell_2}^2 \quad \forall h_{\bullet} \in \mathbb{J}. \tag{01.03}$$

§01|01.09 **Proof** of Lemma §01|01.08. Given in the lecture. □

§01|01.10 **Empirical mean function**. Assume a probability space $(\mathcal{Z}, \mathscr{Z}, \mathbb{P})$ and a stochastic process $\psi_{\bullet} = (\psi_j)_{j\in\mathcal{J}} \in \mathcal{M}(\mathscr{Z}\otimes\mathscr{J})$, i.e. $\mathcal{Z}\times\mathcal{J} \ni (z,j) \mapsto \psi_j(z) \in \mathbb{R}$ is $\mathscr{Z}\otimes\mathscr{J}$-$\mathscr{B}$-measurable, satisfying in addition $\psi_j \in \mathcal{L}_1(\mathbb{P}) := \mathcal{L}_1(\mathcal{Z}, \mathscr{Z}, \mathbb{P})$ for each $j \in \mathcal{J}$. Consider the product probability space $(\mathcal{Z}^n, \mathscr{Z}^{\otimes n}, \mathbb{P}^{\otimes n})$ and $Y_{\bullet} = (Y_j)_{j\in\mathcal{J}}$ with $Y_j := \widehat{\mathbb{P}}_n(\psi_j) \in \mathscr{Z}^{\otimes n}$ where $z = (z_i)_{i\in[\![n]\!]} \mapsto Y_j(z) = (\widehat{\mathbb{P}}_n(\psi_j))(z) = \frac{1}{n}\sum_{i\in[\![n]\!]}\psi_j(z_i)$ for each $j \in \mathcal{J}$. By construction $\mathrm{m}_{\bullet} = (\mathrm{m}_j = \mathbb{P}(\psi_j))_{j\in\mathcal{J}} = \mathbb{P}(\psi_{\bullet}) \in \mathcal{M}(\mathscr{J})$ is the mean function of $Y_{\bullet}$. The statistic $\dot{\varepsilon}_j := n^{1/2}(\widehat{\mathbb{P}}_n(\psi_j) - \mathbb{P}(\psi_j)) \in \mathcal{M}(\mathscr{Z}^{\otimes n})$ is centred, i.e. $\dot{\varepsilon}_j \in \mathcal{L}_1(\mathbb{P}^{\otimes n})$ with $\mathbb{P}^{\otimes n}(\dot{\varepsilon}_j) = 0$, and we have

$$\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathcal{J}} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P})(\psi_{\bullet}) = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_{\bullet}) - \mathbb{P}(\psi_{\bullet})) \in \mathcal{M}(\mathscr{Z}^{\otimes n}\otimes\mathscr{J}).$$

exploiting $\psi_{\bullet} \in \mathcal{M}(\mathscr{Z}\otimes\mathscr{J})$. Since $Y_j = \mathrm{m}_j + n^{-1/2}\dot{\varepsilon}_j$ for each $j \in \mathcal{J}$ by construction we write shortly $Y_{\bullet} = \mathrm{m}_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ and call $Y_{\bullet}$ *empirical mean function*. If for each $j \in \mathcal{J}$ in addition we assume $\psi_j \in \mathcal{L}_2(\mathbb{P})$ then we obtain $Y_j = \widehat{\mathbb{P}}_n(\psi_j) \in \mathcal{L}_2(\mathbb{P}^{\otimes n})$ and, hence $\dot{\varepsilon}_j \in \mathcal{L}_2(\mathbb{P}^{\otimes n})$ by construction. By exploiting $\psi_{\bullet} \in \mathcal{M}(\mathscr{Z}\otimes\mathscr{J})$ the *covariance function* $\mathrm{cov}_{\bullet,\bullet} \in \mathcal{M}(\mathscr{J}^2)$ of $\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathcal{J}}$ is given for each $j, j_{\circ} \in \mathcal{J}$ by

$$\mathrm{cov}_{j,j_{\circ}} = \mathbb{C}\mathrm{ov}(\dot{\varepsilon}_j, \dot{\varepsilon}_{j_{\circ}}) = \mathbb{P}(\psi_j\psi_{j_{\circ}}) - \mathbb{P}(\psi_j)\mathbb{P}(\psi_{j_{\circ}}) = n\,\mathbb{C}\mathrm{ov}(Y_j, Y_{j_{\circ}}).$$

Consequently, we have $\dot{\varepsilon}_{\bullet} \sim \mathrm{P}_{(0,\mathrm{cov}_{\bullet,\bullet})}$ and $Y_{\bullet} = \mathrm{m}_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} \sim \mathrm{P}_{(\mathrm{m}_{\bullet}, n^{-1}\mathrm{cov}_{\bullet,\bullet})}$. There exists a covariance operator $\Gamma \in \mathbb{L}^{\geqslant}(\mathbb{J})$, if in addition $\sup\big\{\mathbb{P}\big(\big|\nu(a_{\bullet}\psi_{\bullet})\big|^2\big) : a_{\bullet} \in \mathbb{J} = \mathbb{L}_2(\nu), \|a_{\bullet}\|_{\mathbb{J}} \leqslant 1\big\} \in \mathbb{R}_{\geqslant 0}$, which holds whenever $\|\psi_{\bullet}\|_{\mathbb{J}} \in \mathcal{L}_2(\mathbb{P})$ or in equal $\mathbb{P}(\|\psi_{\bullet}\|_{\mathbb{J}}^2) \in \mathbb{R}_{\geqslant 0}$. Observe that $\|\psi_{\bullet}\|_{\mathbb{J}}^2 = \sup\big\{\big|\nu(a_{\bullet}\psi_{\bullet})\big|^2 : a_{\bullet} \in \mathbb{J}, \|a_{\bullet}\|_{\mathbb{J}} \leqslant 1\big\}$. Note that $\|\psi_{\bullet}\|_{\mathbb{J}} \in \mathcal{L}_2(\mathbb{P})$ is a sufficent condition for the existence of a covariance operator, but it is not necessary. □

§01|01.11 **White noise process**. A stochastic process $\dot{\mathrm{W}}_{\bullet} = (\dot{\mathrm{W}}_j)_{j\in\mathcal{J}}$ is called *white noise process*, if $(\dot{\mathrm{W}}_j)_{j\in\mathcal{J}}$ is a family of independent and identically $\mathrm{P}_{(0,1)}$-distributed real random variables, where each $\dot{\mathrm{W}}_j$ has zero mean and variance one, $\dot{\mathrm{W}}_j \sim \mathrm{P}_{(0,1)}$ and $\dot{\mathrm{W}}_{\bullet} \sim \mathrm{P}_{(0,1)}^{\otimes\mathcal{J}}$ in short. □

§01|01.12 **Notation**. In other words, the distribution $\mathbb{P}^{\dot{\mathrm{W}}_{\bullet}}$ of a white noise process $\dot{\mathrm{W}}_{\bullet} = (\dot{\mathrm{W}}_j)_{j\in\mathcal{J}} \sim \mathbb{P}^{\dot{\mathrm{W}}_{\bullet}}$ equals the product of its marginal $\mathrm{P}_{(0,1)}$-distributions, i.e. $\mathbb{P}^{\dot{\mathrm{W}}_{\bullet}} = \otimes_{j\in\mathcal{J}}\mathbb{P}^{\dot{\mathrm{W}}_j} = \otimes_{j\in\mathcal{J}}\mathrm{P}_{(0,1)} = \mathrm{P}_{(0,1)}^{\otimes\mathcal{J}}$. □

§01|01.13 **Remark**. The centred stochastic process $\dot{\varepsilon}_{\bullet} := (\dot{\varepsilon}_j)_{j\in\mathcal{J}}$ of error terms considered in an Empirical mean function §01|01.10 is in general not a white noise process. □

§01|01.14 **Notation**. We denote by $\ell_2 := \mathbb{L}_2(\nu_{\mathbb{N}}) = \mathbb{L}_2(\mathbb{N}, 2^{\mathbb{N}}, \nu_{\mathbb{N}})$ the space of all square-summable real-valued sequences endowed with counting measure $\nu_{\mathbb{N}} := \sum_{j\in\mathbb{N}}\delta_{\{j\}}$ over the index set $\mathbb{N}$.

§01|01.15 **Property**. *Let $\dot{\mathrm{W}}_{\bullet} := (\dot{\mathrm{W}}_j)_{j\in\mathbb{N}} \sim \mathrm{P}_{(0,1)}^{\otimes\mathbb{N}}$ be a white noise process. By assumption $\dot{\mathrm{W}}_{\bullet}$ admits $0_{\bullet} := (0)_{j\in\mathbb{N}}$ as ℓ_2-mean and $\Gamma = \mathrm{id}_{\ell_2} \in \mathbb{L}^{\geqslant}(\ell_2)$ as covariance operator, i.e. $\dot{\mathrm{W}}_{\bullet} \sim \mathrm{P}_{(0_{\bullet},\mathrm{id}_{\ell_2})}$, since $\langle a_{\bullet}, b_{\bullet}\rangle_{\ell_2} = \sum_{j\in\mathbb{N}} a_j b_j = \sum_{j\in\mathbb{N}} a_j \sum_{j_{\circ}\in\mathbb{N}} \mathrm{cov}_{j,j_{\circ}} b_{j_{\circ}} = \langle \Gamma a_{\bullet}, b_{\bullet}\rangle_{\ell_2}$ for all $a_{\bullet}, b_{\bullet} \in \ell_2$.* □

§01|01.16 **Gaussian process**. A stochastic process $Y_{\bullet} = (Y_j)_{j\in\mathcal{J}} \sim \mathrm{P}_{(\mathrm{m}_{\bullet},\mathrm{cov}_{\bullet\bullet})}$ satisfying Assumption §01|01.04 with mean function $\mathrm{m}_{\bullet} \in \mathcal{M}(\mathscr{J})$ and covariance function $\mathrm{cov}_{\bullet\bullet} \in \mathcal{M}(\mathscr{J}^2)$ is called a *Gaussian process*, if the family of finite-dimensional distributions $(\mathbb{P}^{Y_{\mathcal{U}}})_{\mathcal{U}\subseteq\mathcal{J}\text{ finite}}$ consists of normal distributions, that is, $Y_{\mathcal{U}} = (Y_u)_{u\in\mathcal{U}}$ is normally distributed with mean vector $(\mathrm{m}_u)_{u\in\mathcal{U}}$ and covariance matrix $(\mathrm{cov}_{u,u'})_{u,u'\in\mathcal{U}}$. We write shortly $Y_{\bullet} \sim \mathrm{N}_{(\mathrm{m}_{\bullet},\mathrm{cov}_{\bullet\bullet})}$ or $Y_{\bullet} \sim \mathrm{N}_{(\mathrm{m}_{\bullet},\Gamma)}$, if in addition there exist a covariance operator $\Gamma \in \mathbb{L}^{\geq}(\mathbb{J})$ associated with $Y_{\bullet}$. The Gaussian process $\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}_{(0_{\bullet},\mathrm{id}_{\mathbb{J}})}$ with $\mathbb{J}$-mean zero and covariance operator $\mathrm{id}_{\mathbb{J}}$ is called *iso-Gaussian process* or *Gaussian white noise process*, which equals $\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}$ in the particular case $\mathbb{J} = \mathbb{L}_2(\nu_{\mathbb{N}}) = \ell_2$. □

§01|01.17 **Definition** Random function. Let $(\mathbb{H}, \langle\cdot,\cdot\rangle_{\mathbb{H}})$ be an Hilbert space equipped with its Borel-σ-algebra $\mathscr{B}_{\mathbb{H}}$, which is induced by its topology. A random variable $Y \in \mathcal{M}(\mathscr{A},\mathscr{B}_{\mathbb{H}})$, i.e. an $\mathscr{A}$-$\mathscr{B}_{\mathbb{H}}$-measurable map $Y : (\Omega,\mathscr{A}) \to (\mathbb{H},\mathscr{B}_{\mathbb{H}})$, is called an $\mathbb{H}$-valued random variable or a *random function* in $\mathbb{H}$. □

§01|01.18 **Lemma**. *Consider $(\ell_2, \langle\cdot,\cdot\rangle_{\ell_2})$. There does not exist a non-zero random function $Y_{\bullet} = (Y_j)_{j\in\mathbb{N}}$ in ℓ_2 which is a Gaussian white noise process.*

§01|01.19 **Proof** of Lemma §01|01.18. Given in the lecture. □

§01|02 Noisy image

§01|02.01 **Assumption**. The Hilbert space $\mathbb{J} = \mathbb{L}_2(\mathcal{J},\mathscr{J},\nu)$ with σ-finite measure $\nu \in \mathscr{M}_{\sigma}(\mathscr{J})$, σ-algebra $\mathscr{J}$ over $\mathcal{J}$ containing all elementary events $\{j\}$, $j \in \mathcal{J}$, and the surjective partial isometry $\mathrm{V} \in \mathbb{L}(\mathbb{G},\mathbb{J})$, i.e. $\mathrm{V}\mathrm{V}^{\star} = \mathrm{id}_{\mathbb{J}}$, are *fixed* and presumed to be *known in advance*. □

§01|02.02 **Notation**. Come back to the reconstruction of $\theta \in \mathbb{H}$ from a noisy version of $g = \mathrm{T}\theta \in \mathbb{G}$. Under Assumption §01|02.01 setting $\mathrm{A} := \mathrm{V}\mathrm{T} \in \mathbb{L}(\mathbb{H},\mathbb{J})$ and $g_{\bullet} = (g_j)_{j\in\mathcal{J}} := \mathrm{V}g \in \mathbb{J}$ we write $g_{\bullet} = \mathrm{A}\theta$. Keep in mind, that we identify the equivalence class and its representative $g_{\bullet}$. □

§01|02.03 **Noisy image**. Let $\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathcal{J}}$ be a stochastic process satisfying Assumption §01|01.04 with mean zero and let $n \in \mathbb{N}$ be a sample size. The stochastic process $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ with $\mathbb{J}$-mean $g_{\bullet}$ is called a *noisy version* of the image $g_{\bullet} = \mathrm{V}g \in \mathbb{J}$, or *noisy image* for short. We denote by $\mathbb{P}_g^n$ the distribution of $\widehat{g}_{\bullet}$. If $\dot{\varepsilon}_{\bullet}$ admits (possibly depending on g) a covariance function, say $\mathrm{cov}_{\bullet\bullet}^g \in \mathcal{M}(\mathscr{J}^2)$, or a covariance operator, say $\Gamma_g \in \mathbb{L}^{\geq}(\mathbb{J})$, then we eventually write $\dot{\varepsilon}_{\bullet} \sim \mathrm{P}_{(0_{\bullet},\mathrm{cov}_{\bullet\bullet}^g)}$ and $\widehat{g}_{\bullet} \sim \mathrm{P}_{(g_{\bullet},n^{-1}\mathrm{cov}_{\bullet\bullet}^g)}$ or $\dot{\varepsilon}_{\bullet} \sim \mathrm{P}_{(0_{\bullet},\Gamma_g)}$ and $\widehat{g}_{\bullet} \sim \mathrm{P}_{(g_{\bullet},n^{-1}\Gamma_g)}$ for short. □

§01|02.04 **Empirical mean model**. For each $g \in \mathbb{G}$ let $\mathbb{P}_g \in \mathscr{W}(\mathscr{Z})$ be a probability measure on a measurable space $(\mathcal{Z},\mathscr{Z})$. Similar to an Empirical mean function §01|01.10 consider a stochastic process $\psi_{\bullet} = (\psi_j)_{j\in\mathcal{J}} \in \mathcal{M}(\mathscr{Z}\otimes\mathscr{J})$ which in addition for all $g \in \mathbb{G}$ satisfies $\psi_j \in \mathcal{L}_1(\mathbb{P}_g)$ for each $j \in \mathcal{J}$ and $\mathbb{P}_g(\psi_{\bullet}) = (g_j = \mathbb{P}_g(\psi_j))_{j\in\mathcal{J}} = g_{\bullet} = \mathrm{V}g$. Considering a statistical product experiment $(\mathcal{Z}^n,\mathscr{Z}^{\otimes n},\mathbb{P}_{\mathbb{G}}^{\otimes} = (\mathbb{P}_g^{\otimes})_{g\in\mathbb{G}})$ as in an Empirical mean function §01|01.10 we define $\widehat{g}_{\bullet} = (\widehat{g}_j := \widehat{\mathbb{P}}_n(\psi_j))_{j\in\mathcal{J}} = \widehat{\mathbb{P}}_n(\psi_{\bullet}) \in \mathcal{M}(\mathscr{Z}^{\otimes n}\otimes\mathscr{J})$. For $g \in \mathbb{G}$ assuming a $\mathbb{P}_g^{\otimes n}$-sample the $\mathbb{J}$-mean of $\widehat{g}_{\bullet}$ is by construction $\mathbb{P}_g(\psi_{\bullet}) = g_{\bullet} = \mathrm{V}g \in \mathbb{J}$. Moreover, the stochastic process

$$\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathcal{J}} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_g)(\psi_{\bullet}) = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_{\bullet}) - \mathbb{P}_g(\psi_{\bullet})) \in \mathcal{M}(\mathscr{Z}^{\otimes n}\otimes\mathscr{J}).$$

is centred, i.e. $\dot{\varepsilon}_j \in \mathbb{L}_1(\mathbb{P}_g^{\otimes n}) = \mathbb{L}_1(\mathcal{Z}^n,\mathscr{Z}^{\otimes n},\mathbb{P}_g^{\otimes n})$ with $\mathbb{P}_g^{\otimes n}(\dot{\varepsilon}_j) = 0$ for each $j \in \mathcal{J}$, and exploiting $\psi_{\bullet} \in \mathcal{M}(\mathscr{Z}\otimes\mathscr{J})$ it satisfies Assumption §01|01.04. Since $\widehat{g}_j = g_j + n^{-1/2}\dot{\varepsilon}_j$ for each $j \in \mathcal{J}$ the stochastic process $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ is a noisy version of the image $g_{\bullet} = \mathrm{V}g \in \mathbb{J}$. □

§01|02.05 **Sequence model**. Consider $\mathbb{J} = \ell_2 = \mathbb{L}_2(\nu_{\mathbb{N}})$ as in §01|01.14. Let $\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathbb{N}}$ be a real-valued stochastic process (satisfying always Assumption §01|01.04) with mean zero and let $n \in \mathbb{N}$ be a sample size. The observable noisy version $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} \sim \mathbb{P}^n_{g_{\bullet}}$ with ℓ_2-mean $g_{\bullet} \in \ell_2$ as in §01|01.14 takes the form of a *sequence model*

$$\widehat{g}_j = g_j + n^{-1/2}\dot{\varepsilon}_j, \quad j \in \mathbb{N}. \tag{01.04}$$

If $\dot{\varepsilon}_{\bullet}$ admits a covariance function (possibly depending on $g_{\bullet}$), say $\mathrm{cov}^{g_{\bullet}}_{\bullet\bullet} \in \mathbb{R}^{\mathbb{N}^2}$, then we eventually write $\widehat{g}_{\bullet} \sim \mathrm{P}_{(g_{\bullet}, n^{-1}\mathrm{cov}^{g_{\bullet}}_{\bullet\bullet})}$ for short. If in addition $\dot{\varepsilon}_{\bullet}$ admits a covariance operator $\Gamma_{g_{\bullet}} \in \mathbb{L}^{\geq}(\ell_2)$ (an infinite matrix) then we write $\widehat{g}_{\bullet} \sim \mathrm{P}_{(g_{\bullet}, n^{-1}\Gamma_{g_{\bullet}})}$. □

§01|02.06 **Gaussian sequence model.** Let $\dot{\mathrm{B}}_{\bullet} := (\dot{\mathrm{B}}_j)_{j\in\mathbb{N}} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$ be a Gaussian white noise process. The observable noisy version $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\mathrm{B}}_{\bullet}$ with ℓ_2-mean $g_{\bullet} \in \ell_2$ takes the form of a *Gaussian sequence model*

$$\widehat{g}_j = g_j + n^{-1/2}\dot{\mathrm{B}}_j, \; j \in \mathbb{N} \quad \text{with} \quad (\dot{\mathrm{B}}_j)_{j\in\mathbb{N}} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)} \tag{01.05}$$

and we denote by $\mathrm{N}^n_{g_{\bullet}}$ the distribution of the stochastic process $\widehat{g}_{\bullet}$. □

§01|02|01 Examples of empirical mean models

§01|02.07 **Notation.** Consider over $\mathcal{D} \in \mathscr{B}$ the measure space $(\mathcal{D}, \mathscr{B}_{\mathcal{D}}, \lambda_{\mathcal{D}})$ where $\lambda_{\mathcal{D}} \in \mathscr{M}_{\sigma}(\mathscr{B}_{\mathcal{D}})$ denotes the restriction of the Lebesgue measure $\lambda \in \mathscr{M}_{\sigma}(\mathscr{B})$ to the Borel-σ-algebra $\mathscr{B}_{\mathcal{D}} = \mathscr{B} \cap \mathcal{D}$, and the Hilbert space $\mathbb{L}_2(\lambda_{\mathcal{D}}) := \mathbb{L}_2(\mathcal{D}, \mathscr{B}_{\mathcal{D}}, \lambda_{\mathcal{D}}) =: \mathbb{G}$. Let $(\mathrm{v}_j)_{j\in\mathbb{N}}$ be an *orthonormal system* in $\mathbb{L}_2(\lambda_{\mathcal{D}})$. The linear operator $\mathrm{V} : \mathbb{L}_2(\lambda_{\mathcal{D}}) \to \ell_2$ with $g \mapsto \mathrm{V}g := g_{\bullet} = (g_j := \langle g, \mathrm{v}_j\rangle_{\mathbb{L}_2(\lambda_{\mathcal{D}})})_{j\in\mathbb{N}}$ is a surjective partial isometry $\mathrm{V} \in \mathbb{L}(\mathbb{L}_2(\lambda_{\mathcal{D}}), \ell_2)$. Its adjoint operator $\mathrm{V}^{\star} \in \mathbb{L}(\ell_2, \mathbb{L}_2(\lambda_{\mathcal{D}}))$ satisfies $\mathrm{V}^{\star}a_{\bullet} = \sum_{j\in\mathbb{N}} a_j \mathrm{v}_j = \nu_{\mathbb{N}}(a_{\bullet}\mathrm{v}_{\bullet}) \in \mathbb{L}_2(\lambda_{\mathcal{D}})$ for all $a_{\bullet} \in \ell_2$ (the limit is taken in ℓ_2). We call $g_{\bullet} = (g_j)_{j\in\mathbb{N}}$ *(generalised) Fourier coefficients* and V *(generalised) Fourier series transform*. □

§01|02.08 **Density estimation on $\mathcal{D}$.** Let $\mathbb{D}_2$ be a set of square-integrable Lebesgue densities on $(\mathcal{D}, \mathscr{B}_{\mathcal{D}})$, and hence $\mathbb{D}_2 \subseteq \mathbb{L}_2(\lambda_{\mathcal{D}}) =: \mathbb{G}$. We denote for each density $g \in \mathbb{D}_2$ by $\mathbb{P}_g := g\lambda_{\mathcal{D}} \in \mathscr{W}(\mathscr{B}_{\mathcal{D}})$ the associated probability measure. Assuming an iid. sample $(X_i)_{i\in[\![n]\!]}$ of size $n \in \mathbb{N}$ we consider the statistical product experiment $\big(\mathcal{D}^n, \mathscr{B}^{\otimes n}_{\mathcal{D}}, \mathbb{P}^{\otimes n}_{\mathbb{D}_2} := (\mathbb{P}^{\otimes n}_g)_{g\in\mathbb{D}_2}\big)$. Let $\mathrm{V} \in \mathbb{L}(\mathbb{L}_2(\lambda_{\mathcal{D}}), \ell_2)$ be a generalised Fourier series transform (see Notation §01|02.07) which is fixed and known in advanced. Evidently, for each density $g \in \mathbb{D}_2 \subseteq \mathbb{L}_2(\lambda_{\mathcal{D}})$ the generalised Fourier coefficients $g_{\bullet} = (g_j)_{j\in\mathbb{N}} = \mathrm{V}g$ for each $j \in \mathbb{N}$ satisfy

$$g_j = \langle g, \mathrm{v}_j\rangle_{\mathbb{L}_2(\lambda_{\mathcal{D}})} = \lambda_{\mathcal{D}}(g\mathrm{v}_j) = g\lambda_{\mathcal{D}}(\mathrm{v}_j) = \mathbb{P}_g(\mathrm{v}_j),$$

i.e. $\mathrm{v}_j \in \mathbb{L}_1(\mathcal{D}, \mathscr{B}_{\mathcal{D}}, \mathbb{P}_g) =: \mathbb{L}_1(\mathbb{P}_g)$. Moreover, the stochastic process $\mathrm{v}_{\bullet} = (\mathrm{v}_j)_{j\in\mathbb{N}}$ on $(\mathcal{D}, \mathscr{B}_{\mathcal{D}}, \mathbb{P}_g)$ is $\mathscr{B}_{\mathcal{D}} \otimes 2^{\mathbb{N}}$-$\mathscr{B}$-measurable, i.e $\mathrm{v}_{\bullet} \in \mathcal{M}(\mathscr{B}_{\mathcal{D}} \otimes 2^{\mathbb{N}})$. Similar to an Empirical mean model §01|02.04 we define $\widehat{g}_{\bullet} = (\widehat{g}_j := \widehat{\mathbb{P}}_n(\mathrm{v}_j))_{j\in\mathbb{N}} = \widehat{\mathbb{P}}_n(\mathrm{v}_{\bullet}) \in \mathcal{M}(\mathscr{B}^{\otimes n}_{\mathcal{D}} \otimes 2^{\mathbb{N}})$ where for each $j \in \mathbb{N}$

$$x = (x_i)_{i\in[\![n]\!]} \mapsto \widehat{g}_j(x) = (\widehat{\mathbb{P}}_n(\mathrm{v}_j))(x) = n^{-1}\sum_{i\in[\![n]\!]} \mathrm{v}_j(x_i).$$

By construction $g_{\bullet} = (g_j = \mathbb{P}_g(\mathrm{v}_j))_{j\in\mathbb{N}} = \mathbb{P}_g(\mathrm{v}_{\bullet}) \in \mathcal{M}(2^{\mathbb{N}})$ is the ℓ_2-mean of $\widehat{g}_{\bullet}$. For each $j \in \mathbb{N}$ the statistic $\dot{\varepsilon}_j := n^{1/2}(\widehat{\mathbb{P}}_n(\mathrm{v}_j) - \mathbb{P}_g(\mathrm{v}_j)) \in \mathcal{M}(\mathscr{B}^{\otimes n}_{\mathcal{D}})$ is centred, i.e. $\dot{\varepsilon}_j \in \mathbb{L}_1(\mathcal{D}^n, \mathscr{B}^{\otimes n}_{\mathcal{D}}, \mathbb{P}^{\otimes n}_g) =: \mathbb{L}_1(\mathbb{P}^{\otimes n}_g)$ with $\mathbb{P}^{\otimes n}_g(\dot{\varepsilon}_j) = 0$, and exploiting $\mathrm{v}_{\bullet} \in \mathcal{M}(\mathscr{B}_{\mathcal{D}} \otimes 2^{\mathbb{N}})$ the stochastic process

$$\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathbb{N}} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_g)(\mathrm{v}_{\bullet}) = n^{1/2}(\widehat{\mathbb{P}}_n(\mathrm{v}_{\bullet}) - \mathbb{P}_g(\mathrm{v}_{\bullet})) \in \mathcal{M}(\mathscr{B}^{\otimes n}_{\mathcal{D}} \otimes 2^{\mathbb{N}})$$

satisfies Assumption §01|01.04. Since $\widehat{g}_j = g_j + n^{-1/2}\dot{\varepsilon}_j$ for each $j \in \mathbb{N}$ by construction $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ is a noisy version of $g_{\bullet}$. □

§01|02.09 **Regression with uniform design.** Consider the measure space $([0,1], \mathscr{B}_{[0,1]}, \lambda_{[0,1]})$ where $\lambda_{[0,1]}$ denotes the restriction of the Lebesgue measure to the Borel-σ-algebra $\mathscr{B}_{[0,1]}$ over $[0,1]$, and the Hilbert space $\mathbb{L}_2(\lambda_{[0,1]}) := \mathbb{L}_2([0,1], \mathscr{B}_{[0,1]}, \lambda_{[0,1]})$ of square Lebesgue-integrable functions. Let (X, Y) be a $[0,1] \times \mathbb{R}$-valued random vector. We denote by $\mathbb{P}^X \in \mathscr{W}(\mathscr{B}_{[0,1]})$ the marginal distribution of X, by $\mathbb{P}^{Y|X}$ a regular conditional distribution of Y given X, and by $\mathbb{P}^{X,Y} = \mathbb{P}^X \odot \mathbb{P}^{Y|X} \in \mathscr{W}(\mathscr{B}_{[0,1]} \otimes \mathscr{B})$ the joint distribution of (X, Y). We tactically identify X and Y with the coordinate map $\Pi_{[0,1]}$ and $\Pi_{\mathbb{R}}$, respectively, and thus (X, Y) with the identity $\mathrm{id}_{[0,1]\times\mathbb{R}}$ such that $\mathbb{P} = \mathbb{P}^{X,Y} \in \mathscr{W}(\mathscr{B}_{[0,1]} \otimes \mathscr{B})$. If in addition $Y \in \mathbb{L}_1(\mathbb{P}) = \mathbb{L}_1([0,1] \times \mathbb{R}, \mathscr{B}_{[0,1]} \otimes \mathscr{B}, \mathbb{P})$ then $\mathbb{P}^{Y|X}(\mathrm{id}_{\mathbb{R}}) = \mathbb{P}\big(Y \big| X\big) =: g \in \mathcal{M}(\mathscr{B}_{[0,1]})$ is unique up to $\mathbb{P}^X$-a.s. equality. Moreover, we have $g \in \mathbb{L}_1(\mathbb{P}^X) = \mathbb{L}_1([0,1], \mathscr{B}_{[0,1]}, \mathbb{P}^X)$ and the error term $\xi := Y - g(X)$ satisfies $\xi \in \mathbb{L}_1(\mathbb{P})$ with $\mathbb{P}(\xi) = 0$. Let us denote in this situation by $\mathbb{P}_g^{Y|X}$ and $\mathbb{P}_g := \mathbb{P}^X \odot \mathbb{P}_g^{Y|X} \in \mathscr{W}(\mathscr{B}_{[0,1]} \otimes \mathscr{B})$, respectively, a regular conditional distribution of Y given X and the joint distribution of (X, Y). Keep however in mind, that even if $g \in \mathbb{L}_1(\mathbb{P}^X)$ is fixed the conditional distribution $\mathbb{P}_g^{Y|X}$ is still not fully specified. We assume in what follows that the regressor X is uniformly distributed on the interval $[0,1]$, i.e. $X \sim \mathrm{U}_{[0,1]} = \lambda_{[0,1]} = \mathbb{P}^X$ and that $g \in \mathbb{F}_2 \subseteq \mathbb{L}_2(\mathbb{P}^X) = \mathbb{L}_2(\lambda_{[0,1]}) =: \mathbb{G}$ identifying again equivalence classes and their representatives. Denote by $\mathrm{U}_g := \mathrm{U}_{[0,1]} \odot \mathbb{P}_g^{Y|X}$ the joint distribution of (X, Y) without fully specifying the conditional distribution $\mathbb{P}_g^{Y|X}$. For $g, h \in \mathbb{L}_2(\mathbb{P}^X) \subseteq \mathbb{L}_1(\mathbb{P}^X)$ we have $gh \in \mathbb{L}_1(\mathbb{P}^X)$ and thus $\mathbb{P}^X(gh) \in \mathbb{R}$. Keep in mind that X and Y equals the coordinate map $\Pi_{[0,1]}$ and $\Pi_{\mathbb{R}}$, respectively. Consequently, if $Y \in \mathbb{L}_2(\mathrm{U}_g)$ and $h \in \mathbb{L}_2(\mathbb{P}^X) = \mathbb{L}_2(\lambda_{[0,1]})$, hence $h(X) \in \mathbb{L}_2(\mathrm{U}_g)$, then we obtain $Yh(X) \in \mathbb{L}_1(\mathrm{U}_g)$ and

$$\mathrm{U}_g(Yh(X)) = \mathbb{P}^X(\mathbb{P}_g^{Y|X}(Y)h) = \mathbb{P}^X(gh) = \lambda_{[0,1]}(gh) = \langle g, h\rangle_{\mathbb{L}_2(\lambda_{[0,1]})} \in \mathbb{R}$$

identifying again equivalence classes and their representatives. We consider the statistical product experiment $\big(([0,1] \times \mathbb{R})^n, (\mathscr{B}_{[0,1]} \otimes \mathscr{B})^{\otimes n}, \mathrm{U}_{\mathbb{F}_2}^{\otimes n} := (\mathrm{U}_g^{\otimes n})_{g \in \mathbb{F}_2}\big)$ of size $n \in \mathbb{N}$ and for $g \in \mathbb{F}_2$ we denote by $((X_i, Y_i))_{i \in [\![n]\!]} \sim \mathrm{U}_g^{\otimes n}$ an iid. sample of $(X, Y) \sim \mathrm{U}_g = \mathrm{U}_{[0,1]} \odot \mathbb{P}_g^{Y|X}$. Let $\mathrm{V} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1]}), \ell_2)$ be a generalised Fourier series transform as in Notation §01|02.07 which is fixed and known in advanced. Evidently, for each $g \in \mathbb{F}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1]})$ the generalised Fourier coefficients $g_\bullet = (g_j)_{j \in \mathbb{N}} = \mathrm{V}g$ for each $j \in \mathbb{N}$ satisfy

$$g_j = \langle g, \mathrm{v}_j\rangle_{\mathbb{G}} = \lambda_{[0,1]}(g\mathrm{v}_j) = \mathrm{U}_g(Y\mathrm{v}_j(X)).$$

Therefore the stochastic process $\psi_\bullet = (\psi_j(X, Y) := Y\mathrm{v}_j(X))_{j \in \mathbb{N}} \in \mathcal{M}((\mathscr{B}_{[0,1]} \otimes \mathscr{B}) \otimes 2^{\mathbb{N}})$ fulfils Assumption §01|01.04 and $g_\bullet = \mathrm{U}_g(\psi_\bullet)$. Similar to an Empirical mean model §01|02.04 we define $\widehat{g}_\bullet = (\widehat{g}_j := \widehat{\mathbb{P}}_n(\psi_j))_{j \in \mathbb{N}} = \widehat{\mathbb{P}}_n(\psi_\bullet) \in \mathcal{M}((\mathscr{B}_{[0,1]} \otimes \mathscr{B})^{\otimes n} \otimes 2^{\mathbb{N}})$. By construction $g_\bullet = \mathrm{U}_g(\psi_\bullet) \in \mathcal{M}(2^{\mathbb{N}})$ is the ℓ_2-mean of $\widehat{g}_\bullet$. For each $j \in \mathbb{N}$ the statistic $\dot{\varepsilon}_j := n^{1/2}(\widehat{\mathbb{P}}_n(\psi_j) - \mathrm{U}_g(\psi_j)) \in \mathcal{M}((\mathscr{B}_{[0,1]} \otimes \mathscr{B})^{\otimes n})$ is centred, i.e. $\dot{\varepsilon}_j \in \mathbb{L}_1(\mathrm{U}_g^{\otimes n})$ with $\mathrm{U}_g^{\otimes n}(\dot{\varepsilon}_j) = 0$, and exploiting $\psi_\bullet \in \mathcal{M}((\mathscr{B}_{[0,1]} \otimes \mathscr{B}) \otimes 2^{\mathbb{N}})$ the stochastic process

$$\dot{\varepsilon}_\bullet = (\dot{\varepsilon}_j)_{j \in \mathbb{N}} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathrm{U}_g)(\psi_\bullet) = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_\bullet) - \mathrm{U}_g(\psi_\bullet)) \in \mathcal{M}((\mathscr{B}_{[0,1]} \otimes \mathscr{B})^{\otimes n} \otimes 2^{\mathbb{N}})$$

satisfies Assumption §01|01.04. Since $\widehat{g}_j = g_j + n^{-1/2}\dot{\varepsilon}_j$ for each $j \in \mathbb{N}$ by construction $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ is a noisy version of $g_\bullet$. □

§01|02|02 Extension to complex-valued models

§01|02.10 **Notation** Reminder. Given a non-empty and generally non-finite subset $\mathcal{J}$ of $\mathbb{N}$, $\mathbb{Z}$ or $\mathbb{R}$ and a subset $\mathcal{U}$ of $\mathcal{J}$ as an index set consider the complex product spaces $\mathbb{C}^{\mathcal{J}} = \times_{j \in \mathcal{J}} \mathbb{C}$ and $\mathbb{C}^{\mathcal{U}} = \times_{u \in \mathcal{U}} \mathbb{C}$, where we identify the family $y_\bullet = (y_j)_{j \in \mathcal{J}} \in \mathbb{C}^{\mathcal{J}}$ and the map $y_\bullet : \mathcal{J} \to \mathbb{C}$ with $j \mapsto y_j$. Eventually, we define arithmetic operations on elements of $\mathbb{C}^{\mathcal{J}}$ coordinate-wise, for example meaning $a_\bullet b_\bullet = (a_j b_j)_{j \in \mathcal{J}}$ and $ra_\bullet = (ra_j)_{j \in \mathcal{J}}$ for $a_\bullet, b_\bullet \in \mathbb{C}^{\mathcal{J}}$ and $r \in \mathbb{C}$. Let us further

introduce $\mathbb{0}_\bullet := (0)_{j\in\mathcal{J}}$, $\mathbb{1}_\bullet := (1)_{j\in\mathcal{J}}$, and the imaginary unit $\imath$. The map $\Pi_{\mathcal{U}} : \mathbb{C}^{\mathcal{J}} \to \mathbb{C}^{\mathcal{U}}$ given by $y_\bullet = (y_j)_{j\in\mathcal{J}} \mapsto (y_j)_{j\in\mathcal{U}} =: \Pi_{\mathcal{U}} y_\bullet$ is called *canonical projection*. In particular, for each $j \in \mathcal{J}$ the *coordinate map* $\Pi_j := \Pi_{\{j\}} : \mathbb{C}^{\mathcal{J}} \to \mathbb{C}$ is given by $y_\bullet = (y_{j'})_{j'\in\mathcal{J}} \mapsto y_j =: \Pi_j y_\bullet$. Let $\mathscr{B}$ denote the Borel-σ-algebra over $\mathbb{C}$ (with a slight abuse of notation). Moreover, $\mathbb{C}^{\mathcal{J}}$ is equipped with the product Borel-σ-algebra $\mathscr{B}^{\otimes\mathcal{J}} := \bigotimes_{j\in\mathcal{J}} \mathscr{B}$. Recall that $\mathscr{B}^{\otimes\mathcal{J}}$ equals the smallest σ-algebra on $\mathbb{C}^{\mathcal{J}}$ such that all coordinate maps Π_j, $j \in \mathcal{J}$ are measurable. i.e., $\mathscr{B}^{\otimes\mathcal{J}} = \sigma(\Pi_j, j \in \mathcal{J})$. Moreover, let $(\mathcal{J}, \mathscr{J}, \nu)$ be a measure space with σ-finite measure $\nu \in \mathscr{M}_\sigma(\mathscr{J})$. We write for each $\mathscr{J}$-$\mathscr{B}$-measurable $h : \mathcal{J} \to \mathbb{C}$ shortly $h \in \mathcal{M}(\mathscr{J})$ with a slight abuse of notation. For $s \in \overline{\mathbb{R}}_{\geqslant 1} = [1, \infty]$ we introduce the usual space $\mathcal{L}_s(\nu) := \mathcal{L}_s(\mathcal{J}, \mathscr{J}, \nu)$ of $\mathcal{L}_s(\nu)$-integrable complex-valued functions. Define further the set of equivalence classes $\mathbb{L}_s(\nu) := \mathbb{L}_s(\mathcal{J}, \mathscr{J}, \nu) := \big\{\{h\}_\nu : h \in \mathcal{L}_s(\nu)\big\}$ (see Comment §01|01.02). In case $s = 2$ the norm $\|\{h\}_\mu\|_{\mathbb{L}_2(\nu)} := \|h\|_{\mathcal{L}_2(\nu)} = (\nu(|h|^2))^{1/2}$ is induced by the inner product $(\{h\}_\nu, \{h_o\}_\nu) \mapsto \langle\{h\}_\nu, \{h_o\}_\nu\rangle_{\mathbb{L}_2(\nu)} := \nu(h\overline{h_o})$ (denoting by $\overline{z}$ the complex conjugate of $z \in \mathbb{C}$), and hence $(\mathbb{L}_2(\nu), \langle\cdot,\cdot\rangle_{\mathbb{L}_2(\nu)})$ is a complex Hilbert space. As usual we identify the equivalence class $\{h\}_\nu$ with its representative h, and write $h \in \mathbb{L}_2(\nu)$ for short. If $\lambda = \nu$ is the Lebesgue-measure then we write also shortly $(\mathbb{L}_s, \|\cdot\|_{\mathbb{L}_s})$. and $(\mathbb{L}_2, \langle\cdot,\cdot\rangle_{\mathbb{L}_2})$.
Let $(Y_j)_{j\in\mathcal{J}}$ be a family of complex-valued random variables on a common probability space $(\Omega, \mathscr{A}, \mathbb{P})$, that is, $Y_j \in \mathcal{M}(\mathscr{A})$ for each $j \in \mathcal{J}$. Consider the $\mathbb{C}^{\mathcal{J}}$-valued random variable $Y_\bullet := (Y_j)_{j\in\mathcal{J}} \in \mathcal{M}(\mathscr{A}, \mathscr{B}^{\otimes\mathcal{J}})$ where $Y_\bullet : \Omega \to \mathbb{C}^{\mathcal{J}}$ is a $\mathscr{A}$-$\mathscr{B}^{\otimes\mathcal{J}}$-measurable map given by $\omega \mapsto (Y_j(\omega))_{j\in\mathcal{J}} =: Y_\bullet(\omega)$. $Y_\bullet$ is called a (complex-valued) *stochastic process*. Its *distribution* $\mathbb{P}^{Y_\bullet} := \mathbb{P} \circ Y_\bullet^{-1}$ is the image probability measure of $\mathbb{P}$ under the map $Y_\bullet$, i.e. $Y_\bullet \sim \mathbb{P}^{Y_\bullet}$ or $\mathbb{P}^{Y_\bullet} \in \mathscr{W}(\mathscr{B}^{\otimes\mathcal{J}})$ for short. Further, denote by $\mathbb{P}^{Y_{\mathcal{U}}} = \mathbb{P} \circ Y_{\mathcal{U}}^{-1} = \mathbb{P}^{Y_\bullet} \circ \Pi_{\mathcal{U}}^{-1}$ the distribution of the stochastic process $Y_{\mathcal{U}} := \Pi_{\mathcal{U}} Y_\bullet = (Y_u)_{u\in\mathcal{U}}$ on $\mathcal{U} \subseteq \mathcal{J}$. The family $(\mathbb{P}^{Y_{\mathcal{U}}})_{\mathcal{U}\subseteq\mathcal{J} \text{ finite}}$ is called *family of finite-dimensional distributions* of $Y_\bullet$ or $\mathbb{P}^{Y_\bullet}$. In particular, $\mathbb{P}^{Y_j} = \mathbb{P}^{\Pi_j Y_\bullet} = \mathbb{P}^{Y_\bullet} \circ \Pi_j^{-1} \in \mathscr{W}(\mathscr{B})$ denotes the distribution of $Y_j = \Pi_j Y_\bullet$. Furthermore, for $j, j_o \in \mathcal{J}$ we write $\mathbb{P}(Y_j) = \mathbb{P}^{Y_\bullet}(\Pi_j)$ and $\mathbb{C}\mathrm{ov}(Y_j, Y_{j_o}) := \mathbb{P}(Y_j\overline{Y_{j_o}}) - \mathbb{P}(Y_j)\mathbb{P}(\overline{Y_{j_o}})$, if it exists, for the expectation of Y_j and the covariance of Y_j and Y_{j_o} with respect to $\mathbb{P}$. □

§01|02.11 **Assumption**. The complex-valued stochastic process $Y_\bullet = (Y_j)_{j\in\mathcal{J}}$ on a common measurable space $(\Omega, \mathscr{A})$ as a function $\Omega \times \mathcal{J} \to \mathbb{C}$ with $(\omega, j) \mapsto Y_j(\omega)$ is $\mathscr{A} \otimes \mathscr{J}$-$\mathscr{B}$-measurable, $Y_\bullet \in \mathcal{M}(\mathscr{A} \otimes \mathscr{J})$ for short. □

§01|02.12 **Notation**. Consider the *complex* Hilbert spaces $\mathbb{L}_2(\lambda_{[0,1)}) := \mathbb{L}_2([0,1), \mathscr{B}_{[0,1)}, \lambda_{[0,1)})$ and $\mathbb{J} := \ell_2(\mathbb{Z}) = \mathbb{L}_2(\nu_{\mathbb{Z}}) = \mathbb{L}_2(\mathbb{Z}, 2^{\mathbb{Z}}, \nu_{\mathbb{Z}})$ where the latter is the space of all square-summable *complex-valued* sequences endowed with counting measure $\nu_{\mathbb{Z}} := \sum_{j\in\mathbb{Z}} \delta_{\{j\}}$ over the index set $\mathbb{Z}$. For each $j \in \mathbb{Z}$ introduce the exponential $\mathrm{e}_j \in \mathcal{M}(\mathscr{B}_{[0,1)})$ with $\mathrm{e}_j(x) := \exp(-\imath 2\pi x j)$ for $x \in [0,1)$ forming together the exponential basis $(\mathrm{e}_j)_{j\in\mathbb{Z}}$ in $\mathbb{L}_2(\lambda_{[0,1)})$. Moreover, the complex-valued stochastic process $\mathrm{e}_\bullet = (\mathrm{e}_j)_{j\in\mathbb{Z}}$ on $([0,1), \mathscr{B}_{[0,1)})$ is $\mathscr{B}_{[0,1)} \otimes 2^{\mathbb{Z}}$-$\mathscr{B}$-measurable, i.e. satisfies Assumption §01|02.11. The linear operator $\mathrm{F} : \mathbb{L}_2(\lambda_{[0,1)}) \to \ell_2(\mathbb{Z})$ with $g \mapsto \mathrm{F}g := g_\bullet = (g_j := \langle g, \mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})})_{j\in\mathbb{Z}} = \lambda_{[0,1)}(g\overline{\mathrm{e}_\bullet})$ is a bijective isometry (unitary) $\mathrm{F} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}), \ell_2(\mathbb{Z}))$. Its adjoint operator $\mathrm{F}^\star \in \mathbb{L}(\ell_2(\mathbb{Z}), \mathbb{L}_2(\lambda_{[0,1)}))$ satisfies

$$\nu_{\mathbb{Z}}(\lambda_{[0,1)}(g\overline{\mathrm{e}_\bullet})\overline{a_\bullet}) = \nu_{\mathbb{Z}}((\mathrm{F}g)\overline{a_\bullet}) = \langle \mathrm{F}g, a_\bullet\rangle_{\ell_2} = \langle g, \mathrm{F}^\star a_\bullet\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} = \lambda_{[0,1)}(g\overline{\mathrm{F}^\star a_\bullet}) = \lambda_{[0,1)}(g\nu_{\mathbb{Z}}(\overline{a_\bullet \mathrm{e}_\bullet}))$$

and hence $\mathrm{F}^\star a_\bullet = \sum_{j\in\mathbb{Z}} a_j \mathrm{e}_j = \nu_{\mathbb{Z}}(a_\bullet \mathrm{e}_\bullet) \in \mathbb{L}_2(\lambda_{[0,1)})$ for all $a_\bullet \in \ell_2(\mathbb{Z})$ (the limit is taken in $\mathbb{L}_2(\lambda_{[0,1)})$). We call $g_\bullet = (g_j)_{j\in\mathbb{Z}}$ *Fourier coefficients* and F *Fourier-series transform*. □

§01|02.13 **Density estimation on** $[0,1)$. Let $\mathbb{D}_2$ be a set of square-integrable Lebesgue densities on $([0,1), \mathscr{B}_{[0,1]})$, and hence $\mathbb{D}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)})$ (by the usual embedding of real-valued functions) as in Notation §01|02.10. We denote for each Lebesgue density g on $([0,1), \mathscr{B}_{[0,1]})$ by $\mathbb{P}_g := g\lambda_{[0,1)} \in \mathscr{W}(\mathscr{B}_{[0,1)})$ the associated probability measure. We consider the statistical product experiment $\big([0,1)^n, \mathscr{B}_{[0,1)}^{\otimes n}, \mathbb{P}_{\mathbb{D}_2}^{\otimes n} :=$

$(\mathbb{P}_g^{\otimes n})_{g\in\mathbb{D}_2})$. Let $\mathrm{F} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}), \ell_2(\mathbb{Z}))$ be the Fourier-series transform (see Notation §01|02.12). Evidently, for each $g \in \mathbb{D}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)}) \subseteq \mathbb{L}_1(\lambda_{[0,1)})$ its Fourier-series $g_{\bullet} = (g_j)_{j\in\mathbb{Z}} = \mathrm{F}g$ for each $j \in \mathbb{Z}$ satisfy

$$g_j = \langle g, \mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} = \lambda_{[0,1)}(g\bar{\mathrm{e}}_j) = \mathbb{P}_g(\bar{\mathrm{e}}_j).$$

The complex-valued stochastic process $\bar{\mathrm{e}}_{\bullet} = (\bar{\mathrm{e}}_j)_{j\in\mathbb{Z}}$ on $([0,1), \mathscr{B}_{[0,1)})$ is $(\mathscr{B}_{[0,1]}\otimes 2^{\mathbb{Z}})$-$\mathscr{B}$-measurable, i.e. $\bar{\mathrm{e}}_{\bullet} \in \mathcal{M}(\mathscr{B}_{[0,1)} \otimes 2^{\mathbb{Z}})$ for short. We define $\widehat{g}_{\bullet} = (\widehat{g}_j := \widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_j))_{j\in\mathbb{Z}} = \widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_{\bullet}) \in \mathcal{M}(\mathscr{B}_{[0,1)}^{\otimes n} \otimes 2^{\mathbb{Z}})$ similar to an Empirical mean model §01|02.04 where for each $j \in \mathbb{Z}$

$$x = (x_i)_{i\in[\![n]\!]} \mapsto \widehat{g}_j(x) = (\widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_j))(x) = n^{-1}\sum_{i\in[\![n]\!]} \bar{\mathrm{e}}_j(x_i) = n^{-1}\sum_{i\in[\![n]\!]} \exp(\imath 2\pi j x_i).$$

By construction $g_{\bullet} = (g_j = \mathbb{P}_g(\bar{\mathrm{e}}_j))_{j\in\mathbb{Z}} = \mathbb{P}_g(\bar{\mathrm{e}}_{\bullet}) \in \mathcal{M}(2^{\mathbb{Z}})$ is the $\ell_2(\mathbb{Z})$-mean of $\widehat{g}_{\bullet}$. For each $j \in \mathbb{Z}$ the statistic $\dot{\varepsilon}_j := n^{1/2}(\widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_j) - \mathbb{P}_g(\bar{\mathrm{e}}_j)) \in \mathcal{M}(\mathscr{B}_{[0,1)}^{\otimes n})$ is centred, i.e. $\dot{\varepsilon}_j \in \mathbb{L}_1(\mathbb{P}_g^{\otimes n})$ with $\mathbb{P}_g^{\otimes n}(\dot{\varepsilon}_j) = 0$, and exploiting $\bar{\mathrm{e}}_{\bullet} = (\bar{\mathrm{e}}_j)_{j\in\mathbb{Z}} \in \mathcal{M}(\mathscr{B}_{[0,1)} \otimes 2^{\mathbb{Z}})$ the complex valued stochastic process

$$\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathbb{Z}} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_g)(\bar{\mathrm{e}}_{\bullet}) = n^{1/2}(\widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_{\bullet}) - \mathbb{P}_g(\bar{\mathrm{e}}_{\bullet})) \in \mathcal{M}(\mathscr{B}_{[0,1)}^{\otimes n} \otimes 2^{\mathbb{Z}})$$

satisfies Assumption §01|02.11. Since $\widehat{g}_j = g_j + n^{-1/2}\dot{\varepsilon}_j$ for each $j \in \mathbb{Z}$ by construction $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ is a noisy version of $g_{\bullet}$. □

§01|02.14 **Regression with uniform design.** Consider the measure space $([0,1), \mathscr{B}_{[0,1)}, \lambda_{[0,1)})$ and the *complex* Hilbert space $\mathbb{L}_2(\lambda_{[0,1)})$ as in Notation §01|02.10. Let $\mathbb{F}_2$ be a set of square-integrable real-valued regression function on $([0,1], \mathscr{B}_{[0,1]})$, and hence $\mathbb{F}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)}) =: \mathbb{G}$ (by the usual embedding of real-valued functions). We consider as in Regression with uniform design §01|02.09 the statistical product experiment $(([0,1) \times \mathbb{R})^n, (\mathscr{B}_{[0,1)} \otimes \mathscr{B})^{\otimes n}, \mathrm{U}_{\mathbb{F}_2}^{\otimes n} := (\mathrm{U}_g^{\otimes n})_{g\in\mathbb{F}_2})$ of size $n \in \mathbb{N}$ and for $g \in \mathbb{F}_2$ we denote by $((X_i, Y_i))_{i\in[\![n]\!]} \sim \mathrm{U}_g^{\otimes n}$ an iid. sample of $(X, Y) \sim \mathrm{U}_g = \mathrm{U}_{[0,1)} \odot \mathbb{P}_g^{Y|X}$. Let $\mathrm{F} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}), \ell_2(\mathbb{Z}))$ be the Fourier-series transform (see Notation §01|02.12). For each $g \in \mathbb{F}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)})$ the Fourier coefficients $g_{\bullet} = (g_j)_{j\in\mathbb{Z}} = \mathrm{F}g$ for each $j \in \mathbb{Z}$ satisfy

$$g_j = \langle g, \mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} = \lambda_{[0,1)}(g\bar{\mathrm{e}}_j) = \mathrm{U}_g(Y\bar{\mathrm{e}}_j(X)).$$

The complex-valued stochastic process $\psi_{\bullet} = (\psi_j(X, Y) := Y\bar{\mathrm{e}}_j(X))_{j\in\mathbb{Z}} \in \mathcal{M}((\mathscr{B}_{[0,1)} \otimes \mathscr{B}) \otimes 2^{\mathbb{Z}})$ fulfils Assumption §01|02.11 and $g_{\bullet} = \mathrm{U}_g(\psi_{\bullet})$. Similar to an Empirical mean model §01|02.04 we define $\widehat{g}_{\bullet} = (\widehat{g}_j := \widehat{\mathbb{P}}_n(\psi_j))_{j\in\mathbb{Z}} = \widehat{\mathbb{P}}_n(\psi_{\bullet}) \in \mathcal{M}((\mathscr{B}_{[0,1)} \otimes \mathscr{B})^{\otimes n} \otimes 2^{\mathbb{Z}})$ where for each $j \in \mathbb{Z}$

$$\widehat{g}_j = \widehat{\mathbb{P}}_n(\psi_j) = n^{-1}\sum_{i\in[\![n]\!]} Y_i\bar{\mathrm{e}}_j(X_i) = n^{-1}\sum_{i\in[\![n]\!]} Y_i\exp(\imath 2\pi j X_i).$$

By construction $g_{\bullet} = \mathrm{U}_g(\psi_{\bullet}) \in \mathcal{M}(2^{\mathbb{Z}})$ is the $\ell_2(\mathbb{Z})$-mean of $\widehat{g}_{\bullet}$. For each $j \in \mathbb{Z}$ the statistic $\dot{\varepsilon}_j := n^{1/2}(\widehat{\mathbb{P}}_n(\psi_j) - \mathrm{U}_g(\psi_j)) \in \mathcal{M}((\mathscr{B}_{[0,1)} \otimes \mathscr{B})^{\otimes n})$ is centred, i.e. $\dot{\varepsilon}_j \in \mathbb{L}_1(\mathrm{U}_g^{\otimes n})$ with $\mathrm{U}_g^{\otimes n}(\dot{\varepsilon}_j) = 0$, and exploiting $\psi_{\bullet} \in \mathcal{M}((\mathscr{B}_{[0,1)} \otimes \mathscr{B}) \otimes 2^{\mathbb{Z}})$ the complex-valued stochastic process

$$\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathbb{Z}} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathrm{U}_g)(\psi_{\bullet}) = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_{\bullet}) - \mathrm{U}_g(\psi_{\bullet})) \in \mathcal{M}((\mathscr{B}_{[0,1)} \otimes \mathscr{B})^{\otimes n} \otimes 2^{\mathbb{Z}})$$

satisfies Assumption §01|02.11. Since $\widehat{g}_j = g_j + n^{-1/2}\dot{\varepsilon}_j$ for each $j \in \mathbb{Z}$ by construction $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ is a noisy version of $g_{\bullet}$. □

§01|02.15 **Notation.** Consider the complex Hilbert space $\mathbb{L}_2 := \mathbb{L}_2(\lambda) = \mathbb{L}_2(\mathbb{R}, \mathscr{B}, \lambda)$ as in Notation §01|02.10. Let $\mathrm{F} \in$ denote the *Fourier-Plancherel transform* satisfying

$$g_j := (\mathrm{F}g)_j = \int_{\mathbb{R}} g(x)\exp(\imath 2\pi x j)\lambda(dx), \quad j \in \mathbb{R}, \quad \forall g \in \mathbb{L}_1 \cap \mathbb{L}_2.$$

Introducing $\mathrm{e}_\bullet = (\mathrm{e}_j)_{j\in\mathbb{R}} \in \mathcal{M}(\mathscr{B}^2)$ given by $\mathrm{e}_j(x) := \exp(-\imath 2\pi xj)$ for $x, j \in \mathbb{R}$ we evidently have $\mathrm{e}_j, \mathrm{e}_\bullet(x) \in \mathcal{M}(\mathscr{B})$ and (keep for each $j \in \mathbb{R}$ in mind that $\bar{\mathrm{e}}_j \in \mathbb{L}_\infty$ but $\bar{\mathrm{e}}_j \notin \mathbb{L}_2$)

$$g_\bullet = (g_j)_{j\in\mathbb{R}} = \mathrm{F}g = (\mathrm{F}g)_\bullet = ((\mathrm{F}g)_j = \lambda(g\bar{\mathrm{e}}_j))_{j\in\mathbb{R}} = \lambda(g\bar{\mathrm{e}}_\bullet), \quad \forall g \in \mathbb{L}_1 \cap \mathbb{L}_2.$$

Moreover, F is unitary with adjoint $\mathrm{F}^\star \in \mathbb{L}(\mathbb{L}_2)$ satisfying

$$\lambda(\lambda(g\bar{\mathrm{e}}_\bullet)\overline{h_\bullet}) = \lambda((\mathrm{F}g)_\bullet\overline{h_\bullet}) = \langle \mathrm{F}g, h_\bullet\rangle_{\mathbb{L}_2} = \langle g, \mathrm{F}^\star h_\bullet\rangle_{\mathbb{L}_2} = \lambda(g\overline{\mathrm{F}^\star h_\bullet}) = \lambda(g\lambda(\overline{h_\bullet}\bar{\mathrm{e}}_\bullet))$$

and hence $(\mathrm{F}^\star h_\bullet)(x) = \lambda(h_\bullet \mathrm{e}_\bullet(x))$, $x \in \mathbb{R}$, for all $h_\bullet \in \mathbb{L}_1 \cap \mathbb{L}_2$. For $g \in \mathbb{L}_1 \cap \mathbb{L}_2$ we write $g_\bullet := (g_j := \lambda(g\bar{\mathrm{e}}_j))_{j\in\mathbb{R}} = \lambda(g\bar{\mathrm{e}}_\bullet) = \mathrm{F}g$ such that $g = \mathrm{F}^\star g_\bullet$ (with a slight abuse of notation). We note that the complex-valued stochastic process $\mathrm{e}_\bullet = (\mathrm{e}_j)_{j\in\mathbb{R}}$ on $(\mathbb{R}, \mathscr{B})$ is $\mathscr{B}^2$-$\mathscr{B}$-measurable, i.e. $\mathrm{e}_\bullet = (\mathrm{e}_j)_{j\in\mathbb{R}} \in \mathcal{M}(\mathscr{B}^2)$ for short, and thus satisfies Assumption §01|02.11. □

§01|02.16 **Density estimation on $\mathbb{R}$.** Let $\mathbb{D}_2$ be a set of square-integrable Lebesgue densities on $(\mathbb{R}, \mathscr{B})$, and hence $\mathbb{D}_2 \subseteq \mathbb{L}_2 =: \mathbb{G}$ (by the usual embedding of real-valued functions). We denote for each density $g \in \mathbb{D}_2$ by $\mathbb{P}_g := g\lambda \in \mathscr{W}(\mathscr{B})$ the associated probability measure. We consider the statistical product experiment $\big(\mathbb{R}^n, \mathscr{B}^{\otimes n}, \mathbb{P}^{\otimes n}_{\mathbb{D}_2} := (\mathbb{P}^{\otimes n}_g)_{g\in\mathbb{D}_2}\big)$. Let $\mathrm{F} \in \mathbb{L}(\mathbb{L}_2)$ be the Fourier-Plancherel transform (see Notation §01|02.15). Evidently, for each $g \in \mathbb{D}_2 \subseteq \mathbb{L}_2$ and hence $g \in \mathbb{L}_1 \cap \mathbb{L}_2$ its Fourier-Plancherel transform $g_\bullet = (g_j)_{j\in\mathbb{R}} = \mathrm{F}g$ for each $j \in \mathbb{R}$ satisfies

$$g_j = \lambda(g\bar{\mathrm{e}}_j) = \mathbb{P}_g(\bar{\mathrm{e}}_j).$$

The stochastic process $(\bar{\mathrm{e}}_j)_{j\in\mathbb{R}}$ on $(\mathbb{R}, \mathscr{B})$ is $\mathscr{B}^2$-$\mathscr{B}$-measurable, i.e. $\bar{\mathrm{e}}_\bullet = (\bar{\mathrm{e}}_j)_{j\in\mathbb{R}} \in \mathcal{M}(\mathscr{B}^2)$ for short. Similar to an Empirical mean model §01|02.04 we define $\widehat{g}_\bullet = (\widehat{g}_j := \widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_j))_{j\in\mathbb{R}} = \widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_\bullet) \in \mathcal{M}(\mathscr{B}^{\otimes n} \otimes \mathscr{B})$ where for each $j \in \mathbb{R}$

$$x = (x_i)_{i\in[\![n]\!]} \mapsto \widehat{g}_j(x) = (\widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_j))(x) = n^{-1}\sum_{i\in[\![n]\!]} \bar{\mathrm{e}}_j(x_i) = n^{-1}\sum_{i\in[\![n]\!]} \exp(\imath 2\pi j x_i).$$

By construction $g_\bullet = (g_j = \mathbb{P}_g(\bar{\mathrm{e}}_j))_{j\in\mathbb{R}} = \mathbb{P}_g(\bar{\mathrm{e}}_\bullet) \in \mathcal{M}(\mathscr{B})$ is the $\mathbb{L}_2$-mean of $\widehat{g}_\bullet$. For each $j \in \mathbb{R}$ the statistic $\dot{\varepsilon}_j := n^{1/2}(\widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_j) - \mathbb{P}_g(\bar{\mathrm{e}}_j)) \in \mathcal{M}(\mathscr{B}^{\otimes n})$ is centred, i.e. $\dot{\varepsilon}_j \in \mathbb{L}_1(\mathbb{P}^{\otimes n}_g)$ with $\mathbb{P}^{\otimes n}_g(\dot{\varepsilon}_j) = 0$. Since $\bar{\mathrm{e}}_\bullet = (\bar{\mathrm{e}}_j)_{j\in\mathbb{R}} \in \mathcal{M}(\mathscr{B}^2)$ the stochastic process

$$\dot{\varepsilon}_\bullet = (\dot{\varepsilon}_j)_{j\in\mathbb{R}} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_g)(\bar{\mathrm{e}}_\bullet) = n^{1/2}(\widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_\bullet) - \mathbb{P}_g(\bar{\mathrm{e}}_\bullet)) \in \mathcal{M}(\mathscr{B}^{\otimes n} \otimes \mathscr{B})$$

satisfies Assumption §01|02.11 and, by construction $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ is a noisy version of $g_\bullet$. □

§01|02.17 **Notation.** Consider on the measurable space $(\mathbb{R}_{>0}, \mathscr{B}_{>0})$ the restriction $\lambda_{>0} \in \mathscr{M}_\sigma(\mathscr{B}_{>0})$ of the Lebesgue-measure λ on $\mathbb{R}_{>0}$, and for $\mathrm{c} \in \mathbb{R}$ the σ-finite measure $\mathrm{x}^{\mathrm{c}}\lambda_{>0} \in \mathscr{M}_\sigma(\mathscr{B}_{>0})$ with Lebesgue-density $\mathrm{x}^{\mathrm{c}} \in \mathcal{M}(\mathscr{B}_{>0})$ given by $x \mapsto \mathrm{x}^{\mathrm{c}}(x) := x^{\mathrm{c}}$. For $s \in \mathbb{R}_{\geqslant 1}$ introduce the *complex* vector space $\mathbb{L}_s(\mathrm{x}^{\mathrm{c}}) := \mathbb{L}_s(\mathrm{x}^{\mathrm{c}}\lambda_{>0}) := \mathbb{L}_s(\mathbb{R}_{>0}, \mathscr{B}_{>0}, \mathrm{x}^{\mathrm{c}}\lambda_{>0})$ of all *complex-valued $\mathbb{L}_s(\mathrm{x}^{\mathrm{c}}\lambda_{>0})$-integrable functions*. Given the *complex* Hilbert space $\mathbb{L}_2 := \mathbb{L}_2(\lambda) := \mathbb{L}_2(\mathbb{R}, \mathscr{B}, \lambda)$ of all *complex-valued square-Lebesgue-integrable functions* let $\mathrm{M}_{\mathrm{c}} \in \mathbb{L}(\mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1}), \mathbb{L}_2)$ denote the *Mellin transform* satisfying

$$\begin{aligned} g_j := (\mathrm{M}_{\mathrm{c}}g)_j &= \int_{\mathbb{R}_{>0}} x^{\mathrm{c}-1+\imath 2\pi j} g(x)\lambda_{>0}(dx) = \int_{\mathbb{R}_{>0}} x^{\imath 2\pi j} g(x)(\mathrm{x}^{\mathrm{c}-1}\lambda_{>0})(dx) \\ &= \int_{\mathbb{R}_{>0}} x^{-\mathrm{c}+\imath 2\pi j} g(x)(\mathrm{x}^{2\mathrm{c}-1}\lambda_{>0})(dx), \quad j \in \mathbb{R}, \quad \forall g \in \mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1}) \cap \mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1}). \end{aligned}$$

Introducing $\mathrm{x}_\bullet = (\mathrm{x}_j)_{j\in\mathbb{R}} \in \mathcal{M}(\mathscr{B}_{>0}\otimes\mathscr{B})$ given by $\mathrm{x}_j(x) := x^{-\imath 2\pi j}$ for $x \in \mathbb{R}_{>0}$, $j \in \mathbb{R}$ we evidently have $\mathrm{x}_\bullet(x) \in \mathscr{B}$, $\bar{\mathrm{e}}_j \circ \log = \mathrm{x}^{\imath 2\pi j} = \overline{\mathrm{x}}_j \in \mathscr{B}_{>0}$, and

$$g_\bullet = (g_j)_{j\in\mathbb{R}} = \mathrm{M}_\mathrm{c} g = (\mathrm{M}_\mathrm{c} g)_\bullet = ((\mathrm{M}_\mathrm{c} g)_j = \mathrm{x}^{\mathrm{c}-1}\lambda_{>0}(\overline{\mathrm{x}}_j g))_{j\in\mathbb{R}}$$
$$= (\mathrm{x}^{2\mathrm{c}-1}\lambda_{>0}(\mathrm{x}^{-\mathrm{c}}\overline{\mathrm{x}}_j g))_{j\in\mathbb{R}}, \quad \forall g \in \mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1}) \cap \mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1}).$$

Moreover, M_c is unitary with adjoint $\mathrm{M}_\mathrm{c}^\star \in \mathbb{L}(\mathbb{L}_2, \mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1}))$ satisfying

$$\lambda((\mathrm{x}^{2\mathrm{c}-1}\lambda_{>0})(\mathrm{x}^{-\mathrm{c}}\overline{\mathrm{x}}_\bullet g)\overline{h_\bullet}) = \lambda((\mathrm{M}_\mathrm{c} g)_\bullet \overline{h_\bullet}) = \langle \mathrm{M}_\mathrm{c} g, h_\bullet\rangle_{\mathbb{L}_2} = \langle g, \mathrm{M}_\mathrm{c}^\star h_\bullet\rangle_{\mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1})}$$
$$= (\mathrm{x}^{2\mathrm{c}-1}\lambda_{>0})(g\overline{\mathrm{M}_\mathrm{c}^\star h_\bullet}) = (\mathrm{x}^{2\mathrm{c}-1}\lambda_{>0})(g\lambda(\overline{\mathrm{x}}_\bullet \overline{h_\bullet})\mathrm{x}^{-\mathrm{c}}), \quad \forall g_\bullet \in \mathbb{L}_1 \cap \mathbb{L}_2$$

and hence $(\mathrm{M}_\mathrm{c}^\star h_\bullet)(x) = \lambda(\mathrm{x}_\bullet(x) h_\bullet) x^{-\mathrm{c}} = (\lambda(\mathrm{x}_\bullet h_\bullet)\mathrm{x}^{-\mathrm{c}})(x)$, $x \in \mathbb{R}_{>0}$ for all $h_\bullet \in \mathbb{L}_1 \cap \mathbb{L}_2$. For $g \in \mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1}) \cap \mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1})$ we write $g_\bullet := (g_j := \mathrm{x}^{2\mathrm{c}-1}\lambda_{>0}(\mathrm{x}^{-\mathrm{c}}\overline{\mathrm{x}}_j g) = \lambda_{>0}(\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_j g))_{j\in\mathbb{R}} = \mathrm{M}_\mathrm{c} g$ such that $g = \mathrm{M}_\mathrm{c}^\star g_\bullet$ (with a slight abuse of notation). We note that for each $\mathrm{c} \in \mathbb{R}$ the complex-valued stochastic process $\mathrm{x}^\mathrm{c}\overline{\mathrm{x}}_\bullet = (\mathrm{x}^\mathrm{c}\overline{\mathrm{x}}_j)_{j\in\mathbb{R}}$ on $(\mathbb{R}_{>0}, \mathscr{B}_{>0})$ is $\mathscr{B}_{>0}\otimes\mathscr{B}$-$\mathscr{B}$-measurable, i.e. $\mathrm{x}^\mathrm{c}\overline{\mathrm{x}}_\bullet = (\mathrm{x}^\mathrm{c}\overline{\mathrm{x}}_j)_{j\in\mathbb{R}} \in \mathcal{M}(\mathscr{B}_{>0}\otimes\mathscr{B})$ for short, and thus satisfies Assumption §01|02.11. □

§01|02.18 **Density estimation on** $\mathbb{R}_{>0}$. Let $\mathbb{D}_2 \subseteq \mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1}) \cap \mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1})$ with $\mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1}) =: \mathbb{G}$ (by the usual embedding of real-valued functions) be a set of densities on $(\mathbb{R}_{>0}, \mathscr{B}_{>0})$ for some $\mathrm{c} \in \mathbb{R}$ fixed and presumed to be *known in advance*. We denote for each density $g \in \mathbb{D}_2$ by $\mathbb{P}_g := g\lambda_{>0} \in \mathscr{W}(\mathscr{B}_{>0})$ the associated probability measure. We consider the statistical product experiment $(\mathbb{R}_{>0}^n, \mathscr{B}_{>0}^{\otimes n}, \mathbb{P}_{\mathbb{D}_2}^{\otimes n} := (\mathbb{P}_g^{\otimes n})_{g\in\mathbb{D}_2})$. Let $\mathrm{M}_\mathrm{c} \in \mathbb{L}(\mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1}), \mathbb{L}_2)$ be the Mellin transform (see Notation §01|02.17). Evidently, for each $g \in \mathbb{D}_2 \subseteq \mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1}) \cap \mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1})$ its Mellin transform $g_\bullet = (g_j)_{j\in\mathbb{R}} = \mathrm{M}_\mathrm{c} g$ for each $j \in \mathbb{R}$ satisfies

$$g_j = \lambda_{>0}(\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_j g) = \mathbb{P}_g(\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_j).$$

The complex-valued stochastic process $\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_\bullet = (\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_j)_{j\in\mathbb{R}}$ on $(\mathbb{R}_{>0}, \mathscr{B}_{>0})$ is $(\mathscr{B}_{>0}\otimes\mathscr{B})$-$\mathscr{B}$-measurable, i.e. $\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_\bullet = (\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_j)_{j\in\mathbb{R}} \in \mathcal{M}(\mathscr{B}_{>0}\otimes\mathscr{B})$ for short. Similar to an Empirical mean model §01|02.04 we define $\widehat{g}_\bullet = (\widehat{g}_j := \widehat{\mathbb{P}}_n(\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_j))_{j\in\mathbb{R}} = \widehat{\mathbb{P}}_n(\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_\bullet) \in \mathcal{M}(\mathscr{B}_{>0}^{\otimes n}\otimes\mathscr{B})$ where for each $j \in \mathbb{R}$

$$x = (x_i)_{i\in[\![n]\!]} \mapsto \widehat{g}_j(x) = (\widehat{\mathbb{P}}_n(\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_j))(x) = n^{-1}\sum_{i\in[\![n]\!]} \mathrm{x}^{\mathrm{c}-1}(x_i)\overline{\mathrm{x}}_j(x_i) = n^{-1}\sum_{i\in[\![n]\!]} x_i^{\mathrm{c}-1+\imath 2\pi j}.$$

By construction $g_\bullet = (g_j = \mathbb{P}_g(\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_j))_{j\in\mathbb{R}} \in \mathcal{M}(\mathscr{B})$ is the $\mathbb{L}_2$-mean of $\widehat{g}_\bullet$. For each $j \in \mathbb{R}$ the statistic $\dot{\varepsilon}_j := n^{1/2}(\widehat{\mathbb{P}}_n(\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_j) - \mathbb{P}_g(\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_j)) \in \mathcal{M}(\mathscr{B}_{>0}^{\otimes n})$ is centred, i.e. $\dot{\varepsilon}_j \in \mathbb{L}_1(\mathbb{P}_g^{\otimes n})$ with $\mathbb{P}_g^{\otimes n}(\dot{\varepsilon}_j) = 0$. By exploiting $\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_\bullet = (\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_j)_{j\in\mathbb{R}} \in \mathcal{M}(\mathscr{B}_{>0}\otimes\mathscr{B})$ the stochastic process

$$\dot{\varepsilon}_\bullet = (\dot{\varepsilon}_j)_{j\in\mathbb{R}} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_g)(\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_\bullet) = n^{1/2}(\widehat{\mathbb{P}}_n(\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_\bullet) - \mathbb{P}_g(\mathrm{x}^{\mathrm{c}-1}\overline{\mathrm{x}}_\bullet)) \in \mathcal{M}(\mathscr{B}_{>0}^{\otimes n}\otimes\mathscr{B})$$

satisfies Assumption §01|02.11 and, by construction $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ is a noisy version of $g_\bullet$. □

§01|03 Statistical direct problem

§01|03.01 **Assumption**. The Hilbert space $\mathbb{J} = \mathbb{L}_2(\mathcal{J}, \mathscr{J}, \nu)$ with σ-finite measure $\nu \in \mathscr{M}_\sigma(\mathscr{J})$ and the surjective partial isometries $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \mathbb{J})$ and $\mathrm{U} := \mathrm{A} = \mathrm{VT} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$, i.e. $\mathrm{VV}^\star = \mathrm{id}_\mathbb{J} = \mathrm{UU}^\star$, are fixed and presumed to be *known in advance*. □

§01|03.02 **Notation**. Under Assumption §01|03.01 we consider the reconstruction of $\theta_\bullet = \mathrm{U}\theta \in \mathbb{J}$ (or in equal $\theta = \mathrm{U}^\star\theta_\bullet \in \mathbb{H}$) from a noisy version of $g_\bullet = \mathrm{V}g = \mathrm{A}\theta = \mathrm{U}\theta = \theta_\bullet \in \mathbb{J}$. Keep in mind, that we identify the equivalence class and its representative $\theta_\bullet$. □

§01|03.03 **Statistical direct problem**. Consider as in Definition §01|02.03 a stochastic process $\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathcal{J}}$ satisfying Assumption §01|01.04 with mean zero and a sample size $n \in \mathbb{N}$. Under Assumption §01|03.01 the observable noisy image has $\mathbb{J}$-mean $\theta_{\bullet} = \mathrm{U}\theta \in \mathbb{J}$, takes the form $\widehat{\theta}_{\bullet} = \theta_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ and is called a *noisy version* of the parameter $\theta \in \mathbb{H}$, or *noisy parameter* for short. We denote by $\mathbb{P}^n_{\theta}$ the distribution of $\widehat{\theta}_{\bullet}$. If $\dot{\varepsilon}_{\bullet}$ admits (possibly depending on θ) a covariance function, say $\mathrm{cov}^{\theta}_{\bullet\bullet} \in \mathcal{M}(\mathscr{J}^2)$, or a covariance operator, say $\Gamma_{\theta} \in \mathbb{L}^{\geqslant}(\mathbb{J})$, then we eventually write $\dot{\varepsilon}_{\bullet} \sim \mathrm{P}_{(0_{\bullet},\mathrm{cov}^{\theta}_{\bullet\bullet})}$ and $\widehat{\theta}_{\bullet} \sim \mathrm{P}_{(\theta_{\bullet},n^{-1}\mathrm{cov}^{\theta}_{\bullet\bullet})}$ or $\dot{\varepsilon}_{\bullet} \sim \mathrm{P}_{(0_{\bullet},\Gamma_{\theta})}$ and $\widehat{\theta}_{\bullet} \sim \mathrm{P}_{(\theta_{\bullet},n^{-1}\Gamma_{\theta})}$ for short. The reconstruction of $\theta_{\bullet} \in \mathbb{J}$ (in equal $\theta = \mathrm{U}^{\star}\theta_{\bullet} \in \mathbb{H}$) from its noisy version $\widehat{\theta}_{\bullet} \sim \mathbb{P}^n_{\theta}$ is called a *statistical direct problem*. □

§01|03.04 **Direct empirical mean model**. Consider the reconstruction of $\theta_{\bullet} \in \mathbb{J}$ (in equal $\theta = \mathrm{U}^{\star}\theta_{\bullet} \in \mathbb{H}$) in an Empirical mean model as in §01|02.04. Under Assumption §01|03.01 the observable noisy image has $\mathbb{J}$-mean $\mathrm{U}\theta = \theta_{\bullet} \in \mathbb{J}$, i.e. it is a noisy version of the parameter, and takes the form an Empirical mean model as in §01|02.04, that is $\widehat{\theta}_{\bullet} = \theta_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ with error process $\dot{\varepsilon}_{\bullet} = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_{\bullet}) - \mathbb{P}_{\theta}(\psi_{\bullet})) \in \mathcal{M}(\mathscr{X}^{\otimes n} \otimes \mathscr{J})$ satisfying Assumption §01|01.04. □

§01|03.05 **Direct sequence model (dSM)**. Consider $\mathbb{J} = \ell_2 = \mathbb{L}_2(\nu_{\mathbb{N}})$ as in §01|01.14. Let $\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathbb{N}}$ be a sequence of real-valued random variables with mean zero and let $n \in \mathbb{N}$ be a sample size. The observable noisy version $\widehat{\theta}_{\bullet} = \theta_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} \sim \mathbb{P}^n_{\theta}$ with ℓ_2-mean $\theta_{\bullet} \in \ell_2$ takes the form of a Sequence model as in §01|02.05, that is

$$\widehat{\theta}_j = \theta_j + n^{-1/2}\dot{\varepsilon}_j, \quad j \in \mathbb{N}. \tag{01.06}$$

If $\dot{\varepsilon}_{\bullet}$ admits a covariance function (possibly depending on $\theta_{\bullet}$), say $\mathrm{cov}^{\theta}_{\bullet\bullet} \in \mathcal{M}(2^{\mathbb{N}^2}) = \mathbb{R}^{\mathbb{N}^2}$, then we eventually write $\widehat{\theta}_{\bullet} \sim \mathrm{P}_{(\theta_{\bullet},n^{-1}\mathrm{cov}^{\theta}_{\bullet\bullet})}$ for short. If in addition $\dot{\varepsilon}_{\bullet}$ admits a covariance operator $\Gamma_{\theta_{\bullet}} \in \mathbb{L}^{\geqslant}(\ell_2)$ (an infinite matrix) then we write $\widehat{\theta}_{\bullet} \sim \mathrm{P}_{(\theta_{\bullet},n^{-1}\Gamma_{\theta_{\bullet}})}$. □

§01|03.06 **Gaussian direct sequence model (GdSM)**. Let $\dot{\mathrm{B}}_{\bullet} := (\dot{\mathrm{B}}_j)_{j\in\mathbb{N}} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$ be a Gaussian white noise process. The observable noisy version $\widehat{\theta}_{\bullet} = \theta_{\bullet} + n^{-1/2}\dot{\mathrm{B}}_{\bullet}$ with ℓ_2-mean $\theta_{\bullet} \in \ell_2$ takes the form of a Gaussian sequence model as in §01|02.06, that is

$$\widehat{\theta}_j = \theta_j + n^{-1/2}\dot{\mathrm{B}}_j,\ j \in \mathbb{N} \quad \text{with} \quad (\dot{\mathrm{B}}_j)_{j\in\mathbb{N}} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}. \tag{01.07}$$

We denote by N^n_{θ} the distribution of the stochastic process $\widehat{\theta}_{\bullet}$. □

§01|04 Diagonal statistical inverse problem

§01|04.01 **Notation**. Consider the measure space $(\mathcal{J}, \mathscr{J}, \nu)$ and the Hilbert space $\mathbb{J} = \mathbb{L}_2(\nu)$ as in Notation §01|01.01. For $\mathrm{w}_{\bullet} \in \mathbb{R}^{\mathcal{J}}$ define the multiplication map $\mathrm{M}_{\mathrm{w}} : \mathbb{R}^{\mathcal{J}} \to \mathbb{R}^{\mathcal{J}}$ with $a_{\bullet} \mapsto \mathrm{M}_{\mathrm{w}}a_{\bullet} := \mathrm{w}_{\bullet}a_{\bullet} := (\mathrm{w}_j a_j)_{j\in\mathcal{J}}$. If $\mathrm{w}_{\bullet} \in \mathcal{M}(\mathscr{J})$, i.e. $\mathrm{w}_{\bullet}$ is $\mathscr{J}$-$\mathscr{B}$-measurable, then we have $\mathrm{M}_{\mathrm{w}} : \mathcal{M}(\mathscr{J}) \to \mathcal{M}(\mathscr{J})$ too. If in addition $\mathrm{w}_{\bullet} \in \mathbb{L}_{\infty}(\nu)$ then we have also $\mathrm{M}_{\mathrm{w}} \in \mathbb{L}(\mathbb{J})$ identifying again equivalence classes and representatives. We set

$$\mathbb{L}^{\mathrm{M}}(\mathbb{J}) := \left\{\mathrm{M}_{\mathrm{w}}\colon \mathrm{w}_{\bullet} \in \mathbb{L}_{\infty}(\nu)\right\} \subseteq \mathbb{L}(\mathbb{J})$$

noting that $\|\mathrm{M}_{\mathrm{w}}\|_{\mathbb{L}(\mathbb{J})} = \sup\left\{\|\mathrm{w}_{\bullet}a_{\bullet}\|_{\mathbb{J}}\colon \|a_{\bullet}\|_{\mathbb{J}} \leqslant 1\right\} \leqslant \|\mathrm{w}_{\bullet}\|_{\mathbb{L}_{\infty}(\nu)}$ for each $\mathrm{M}_{\mathrm{w}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$. Finally, given surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \mathbb{J})$ we define

$$\mathbb{L}^{\mathrm{U},\mathrm{V}}(\mathbb{L}^{\mathrm{M}}(\mathbb{J})) := \mathrm{V}^{\star}(\mathbb{L}^{\mathrm{M}}(\mathbb{J}))\mathrm{U} := \left\{\mathrm{V}^{\star}\mathrm{M}_{\mathrm{w}}\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{G})\colon \mathrm{M}_{\mathrm{w}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})\right\}.$$

As a consequence, for each $\mathrm{T} \in \mathbb{L}^{\mathrm{U},\mathrm{V}}(\mathbb{L}^{\mathrm{M}}(\mathbb{J}))$ we have $\mathrm{V}\mathrm{T}\mathrm{U}^{\star} = \mathrm{M}_{\mathrm{w}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$ for some $\mathrm{w}_{\bullet} \in \mathbb{L}_{\infty}(\nu)$. □

§01|04.02 **Notation.** For $A \in \mathscr{J}$ we denote by $\mathbb{1}^A_\bullet = (\mathbb{1}^A_j)_{j\in\mathcal{J}}$ the indicator function where for each $j \in \mathcal{J}$, $\mathbb{1}^A_j = 1$ if $j \in A$ and $\mathbb{1}^A_j = 0$ otherwise. Obviously, $\mathbb{1}^A_\bullet$ is $\mathscr{J}$-$\mathscr{B}$-measurable, i.e. $\mathbb{1}^A_\bullet \in \mathcal{M}(\mathscr{J})$, and it belongs to $\mathbb{L}_\infty(\nu)$, and to $\mathbb{L}_2(\nu)$ whenever $\nu(A) \in \mathbb{R}_{\geqslant 0}$. Since $\{j\} \in \mathscr{J}$ we have $\mathbb{1}^{\{j\}}_\bullet \in \mathscr{J}$ and $\mathbb{1}^{\{j\}}_\bullet \in \mathbb{L}_\infty(\nu)$. In particular, it follows $\mathbb{1}_\bullet = \mathbb{1}^{\mathcal{J}}_\bullet \in \mathbb{L}_\infty(\nu)$ and $\mathrm{M}_{\mathbb{1}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$. For each $\mathrm{w}_\bullet \in \mathbb{L}_\infty(\nu)$ set

$$\mathbb{J}\mathrm{w}_\bullet := \{\{a_\bullet \mathrm{w}_\bullet\}_\nu : a_\bullet \in \mathcal{L}_2(\nu)\} = \{a_\bullet \mathrm{w}_\bullet : a_\bullet \in \mathbb{J} = \mathbb{L}_2(\nu)\}$$

and hence in particular $\mathbb{J}\mathbb{1}^A_\bullet = \{a_\bullet \mathbb{1}^A_\bullet : a_\bullet \in \mathbb{J}\}$. Given $0_\bullet = (0)_{j\in\mathcal{J}}$ for $\mathrm{w}_\bullet \in \mathcal{M}(\mathscr{J})$ we write further

$$\mathcal{N}_{\mathrm{w}} := \{\mathrm{w}_\bullet = 0_\bullet\} := \{j \in \mathcal{J} : \mathrm{w}_j = 0\} \in \mathscr{J},$$

and denote by $\mathrm{dom}(\mathrm{M}_{\mathrm{w}}) = \{a_\bullet \in \mathbb{J} : a_\bullet \mathrm{w}_\bullet \in \mathbb{J}\}$, $\mathrm{ran}(\mathrm{M}_{\mathrm{w}}) = \{a_\bullet \mathrm{w}_\bullet : a_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathrm{w}}) \subseteq \mathbb{J}\}$ and $\ker(\mathrm{M}_{\mathrm{w}}) = \{a_\bullet \in \mathbb{J} : \{a_\bullet \mathrm{w}_\bullet\}_\nu = 0_\bullet\}$, respectively, the domain, range and nullspace of $\mathrm{M}_{\mathrm{w}} : \mathbb{J} \supseteq \mathrm{dom}(\mathrm{M}_{\mathrm{w}}) \to \mathbb{J}$. We write $\mathrm{w}_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$, if $\mathrm{w}_\bullet \in \mathcal{M}(\mathscr{J})$ and $\nu(\mathcal{N}_{\mathrm{w}}) = 0$. Similarly, for $\mathrm{w}_\bullet \in \mathcal{M}(\mathscr{J})$ with $\nu(\{\mathrm{w}_\bullet \leqslant 0_\bullet\}) = 0$ we write $\mathrm{w}_\bullet \in \mathcal{M}_{>0,\nu}(\mathscr{J})$. □

§01|04.03 **Notation.** Consider the special case $(\mathcal{J}, \mathscr{J}, \nu) = (\mathbb{N}, 2^{\mathbb{N}}, \nu_{\mathbb{N}})$ where $(\mathbb{1}^{\{j\}}_\bullet)_{j\in\mathbb{N}}$ forms an orthonormal basis in ℓ_2. For each infinite matrix $\mathrm{A}_{\bullet|\bullet} \in \mathbb{L}(\ell_2) \subseteq \mathbb{R}^{\mathbb{N}^2} = \mathcal{M}(2^{\mathbb{N}^2})$ with

$$\mathrm{A}_{\bullet|\bullet} = (\mathrm{A}_{j|j_\circ} := \langle \mathrm{A}_{\bullet|\bullet}\mathbb{1}^{\{j_\circ\}}_\bullet, \mathbb{1}^{\{j\}}_\bullet\rangle_{\ell_2})_{j,j_\circ\in\mathbb{N}}$$

and for each $j, j_\circ \in \mathbb{N}$ and $a_\bullet \in \ell_2$ we have

$$\mathrm{A}_{j|\bullet} := (\mathrm{A}_{j|j_\circ})_{j_\circ\in\mathbb{N}} = \mathrm{A}^\star_{\bullet|\bullet}\mathbb{1}^{\{j\}}_\bullet \in \ell_2, \quad \mathrm{A}_{\bullet|j_\circ} := (\mathrm{A}_{j|j_\circ})_{j\in\mathbb{N}} = \mathrm{A}_{\bullet|\bullet}\mathbb{1}^{\{j\}}_\bullet \in \ell_2,$$
$$\text{and} \quad \langle \mathrm{A}_{j|\bullet}, a_\bullet\rangle_{\ell_2} = \nu_{\mathbb{N}}(\mathrm{A}_{j|\bullet}a_\bullet) = \sum_{j_\circ\in\mathbb{N}} \mathrm{A}_{j|j_\circ}a_{j_\circ} = \langle \mathrm{A}_{\bullet|\bullet}a_\bullet, \mathbb{1}^{\{j\}}_\bullet\rangle_{\ell_2} \in \mathbb{R}.$$

If $\mathrm{A}_{\bullet|\bullet} \in \mathbb{L}(\ell_2)$ equals a multiplication operator $\mathrm{A}_{\bullet|\bullet} = \mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathrm{M}}(\ell_2)$ for some $\mathfrak{s}_\bullet \in \ell_\infty$ (where $\ell_\infty := \mathbb{L}_\infty(\nu_{\mathbb{N}})$ is the set of all bounded real-valued sequences with respect to the counting measure $\nu_{\mathbb{N}}$ over $\mathbb{N}$) then we call $\mathrm{A}_{\bullet|\bullet} \in \mathbb{L}(\ell_2)$ *diagonal*. Note that $\mathrm{A}_{\bullet|\bullet} \in \mathbb{L}(\ell_2)$ is diagonal if and only if $\mathrm{A}_{j,j_\circ} = 0$ for all $j \in \mathbb{N}$ and $j_\circ \in \mathbb{N}_{\setminus j} = \mathbb{N} \setminus \{j\}$. For each $\mathrm{T} \in \mathbb{L}^{\mathrm{U,V}}(\mathbb{L}^{\mathrm{M}}(\ell_2))$ with $\mathrm{T}_{\bullet|\bullet} = \mathrm{VTU}^\star = \mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathrm{M}}(\ell_2)$, the sequence $\mathfrak{s}_\bullet \in \ell_\infty$ is called singular values of T and $(\mathfrak{s}_\bullet, \mathrm{U}, \mathrm{V})$ singular value decomposition of T. In other words, each $\mathrm{T} \in \mathbb{L}^{\mathrm{U,V}}(\mathbb{L}^{\mathrm{M}}(\mathbb{J}))$ is diagonal wrt. to U and V. □

§01|04.04 **Assumption.** For $\mathbb{J} = \mathbb{L}_2(\nu)$, surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \mathbb{J})$, *fixed* and presumed to be *known in advance*, $\mathrm{T} \in \mathbb{L}^{\mathrm{U,V}}(\mathbb{L}^{\mathrm{M}}(\mathbb{J})) \subseteq \mathbb{L}(\mathbb{H}, \mathbb{G})$ and hence $\mathrm{A} = \mathrm{VT} = \mathrm{M}_{\mathfrak{s}}\mathrm{U}$ or in equal $\mathfrak{s}_\bullet \in \mathbb{L}_\infty(\nu)$ is also presumed to be *known* where $g_\bullet = \mathrm{VT}\theta = \mathrm{M}_{\mathfrak{s}}\mathrm{U}\theta = \mathrm{M}_{\mathfrak{s}}\theta_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathbb{J}$ or in equal $g_\bullet \in \mathbb{J}\mathfrak{s}_\bullet$. □

§01|04.05 **Notation.** Under Assumption §01|04.04 given $\mathfrak{s}_\bullet \in \mathbb{L}_\infty(\nu)$ and $g_\bullet \in \mathbb{J}\mathfrak{s}_\bullet$ we consider the reconstruction of $\theta_\bullet = \mathrm{U}\theta \in \mathbb{J}$ (or in equal $\theta = \mathrm{U}^\star\theta_\bullet \in \mathbb{H}$) from a noisy version of $g_\bullet = \mathrm{V}g = \mathrm{A}\theta = \mathfrak{s}_\bullet\theta_\bullet \in \mathbb{J}$. Keep in mind, that we identify the equivalence class and its representative $g_\bullet$. □

§01|04.06 **Diagonal statistical inverse problem.** Consider as in Definition §01|02.03 a stochastic process $\dot{\varepsilon}_\bullet = (\dot{\varepsilon}_j)_{j\in\mathcal{J}}$ satisfying Assumption §01|01.04 with mean zero and a sample size $n \in \mathbb{N}$. Under Assumption §01|04.04, where $\mathfrak{s}_\bullet \in \mathbb{L}_\infty(\nu)$ is *known in advance*, the observable noisy image has $\mathbb{J}$-mean $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet$ and takes the form $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet = \mathfrak{s}_\bullet\theta_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$. We denote by $\mathbb{P}^n_{\theta|\mathfrak{s}_\bullet}$ the distribution of $\widehat{g}_\bullet$. If $\dot{\varepsilon}_\bullet$ admits (possibly depending on $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet$) a covariance function, say $\mathrm{cov}^{\theta|\mathfrak{s}_\bullet}_{\bullet,\bullet} \in \mathcal{M}(\mathscr{J}^2)$, or a covariance operator, say $\Gamma_{\theta|\mathfrak{s}_\bullet} \in \mathbb{L}^{\geqslant}(\mathbb{J})$, then we eventually write $\dot{\varepsilon}_\bullet \sim \mathrm{P}_{(0_\bullet,\mathrm{cov}^{\theta|\mathfrak{s}_\bullet}_{\bullet,\bullet})}$ and $\widehat{g}_\bullet \sim \mathrm{P}_{(g_\bullet,n^{-1}\mathrm{cov}^{\theta|\mathfrak{s}_\bullet}_{\bullet,\bullet})}$ or $\dot{\varepsilon}_\bullet \sim \mathrm{P}_{(0_\bullet,\Gamma_{\theta|\mathfrak{s}_\bullet})}$ and $\widehat{g}_\bullet \sim \mathrm{P}_{(g_\bullet,n^{-1}\Gamma_{\theta|\mathfrak{s}_\bullet})}$ for short. The reconstruction of $\theta_\bullet \in \mathbb{J}$ (in equal $\theta = \mathrm{U}^\star\theta_\bullet \in \mathbb{H}$) from a noisy version $\widehat{g}_\bullet \sim \mathbb{P}^n_{\theta|\mathfrak{s}_\bullet}$ of the image $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathbb{J}$ is called a *diagonal statistical inverse problem*. □

§01|04.07 **Diagonal inverse empirical mean model (dieMM).** Consider the reconstruction of $\theta_{\bullet} \in \mathbb{J}$ (in equal $\theta = \mathrm{U}^{\star}\theta_{\bullet} \in \mathbb{H}$) in an Empirical mean model as in §01|02.04. Under Assumption §01|04.04, where $\mathfrak{s}_{\bullet} \in \mathbb{L}_{\infty}(\nu)$ is *known in advance*, the observable noisy image has $\mathbb{J}$-mean $\mathrm{V}g = g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathbb{J}\mathfrak{s}_{\bullet} \subseteq \mathbb{J}$ and takes the form of an Empirical mean model as in §01|02.04, that is $\widehat{g}_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ with error process

$$\dot{\varepsilon}_{\bullet} = n^{1/2}(\widehat{\mathbb{P}}_{n} - \mathbb{P}_{g})(\psi_{\bullet}) = n^{1/2}(\widehat{\mathbb{P}}_{n}(\psi_{\bullet}) - \mathbb{P}_{g}(\psi_{\bullet})) \in \mathcal{M}(\mathscr{Z}^{\otimes n} \otimes \mathscr{J})$$

satisfying Assumption §01|01.04. □

§01|04.08 **Diagonal inverse sequence model (diSM).** Consider $\mathbb{J} = \ell_2 = \mathbb{L}_2(\nu_{\mathbb{N}})$ as in §01|01.14. Let $\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathbb{N}}$ be a sequence of real-valued random variables with mean zero and let $n \in \mathbb{N}$ be a sample size. Under Assumption §01|04.04, where $\mathfrak{s}_{\bullet} \in \ell_{\infty}$ is *known in advance*, the observable noisy image has ℓ_2-mean $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \ell_2$ and takes the form of a Sequence model as in §01|02.05, that is $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ or in equal

$$\widehat{g}_{j} = g_{j} + n^{-1/2}\dot{\varepsilon}_{j} = \mathfrak{s}_{j}\theta_{j} + n^{-1/2}\dot{\varepsilon}_{j}, \quad j \in \mathbb{N}. \tag{01.08}$$

We denote by $\mathbb{P}^{n}_{\theta|\mathfrak{s}}$ the distribution of $\widehat{g}_{\bullet}$. □

§01|04.09 **Gaussian diagonal inverse sequence model (GdiSM).** Let $\dot{\mathrm{B}}_{\bullet} := (\dot{\mathrm{B}}_j)_{j\in\mathbb{N}} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$ be a Gaussian white noise process. The observable noisy version $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\mathrm{B}}_{\bullet}$ with ℓ_2-mean $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet}$ takes the form of a Gaussian sequence model as in §01|02.06, that is

$$\widehat{g}_{j} = \mathfrak{s}_{j}\theta_{j} + n^{-1/2}\dot{\mathrm{B}}_{j},\ j \in \mathbb{N} \quad \text{with} \quad (\dot{\mathrm{B}}_j)_{j\in\mathbb{N}} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}. \tag{01.09}$$

We denote by $\mathrm{N}^{n}_{\theta|\mathfrak{s}}$ the distribution of the stochastic process $\widehat{g}_{\bullet}$. □

§01|04|01 Examples of diagonal inverse empirical mean models

§01|04.10 **Diagonal inverse regression with uniform design.** Consider the measure space $([0,1], \mathscr{B}_{[0,1]}, \lambda_{[0,1]})$ and the Hilbert space $\mathbb{L}_2(\lambda_{[0,1]})$ as in Model §01|02.09. Let $\mathrm{T} \in \mathbb{L}^{\mathrm{U},\mathrm{V}}(\mathbb{L}^{\mathrm{M}}(\ell_2)) \subseteq \mathbb{L}(\mathbb{H}, \mathbb{L}_2([0,1]))$ be *known in advance*, i.e. $\mathrm{T}_{\bullet|\bullet} = \mathrm{V}\mathrm{T}\mathrm{U}^{\star} = \mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathrm{M}}(\ell_2)$ and in other words T has a *known* singular value decomposition $(\mathfrak{s}_{\bullet}, \mathrm{U}, \mathrm{V})$ with sequence of singular values $\mathfrak{s}_{\bullet} \in \ell_{\infty}$. Let (X, Y) be a $[0,1] \times \mathbb{R}$-valued random vector. As in Model §01|02.09 we assume in what follows that the regressor X is uniformly distributed on the interval $[0,1]$, i.e. $X \sim \mathrm{U}_{[0,1]} = \lambda_{[0,1]} = \mathbb{P}^{X}$ and that given $\mathrm{T}\theta = g \in \mathbb{L}_2([0,1])$ for some $\theta \in \mathbb{H}$ the joint distribution of (X, Y) is given by $\mathrm{U}_{\theta|\mathrm{T}} := \mathrm{U}_{[0,1]} \odot \mathbb{P}^{Y|X}_{\mathrm{T}\theta}$ without fully specifying the regular conditional distribution $\mathbb{P}^{Y|X}_{\mathrm{T}\theta}$ which however satisfies $\mathbb{P}^{Y|X}_{\mathrm{T}\theta}(\mathrm{id}_{\mathbb{R}}) = \mathbb{P}_{\mathrm{T}\theta}(Y|X) = \mathrm{T}\theta = g \in \mathbb{L}_2([0,1])$. Keep in mind that we tactically identify X and Y with the coordinate map $\Pi_{[0,1]}$ and $\Pi_{\mathbb{R}}$, respectively, and thus (X, Y) with the identity $\mathrm{id}_{[0,1]\times\mathbb{R}}$. Consequently, if $Y \in \mathbb{L}_2(\mathrm{U}_{\theta|\mathrm{T}})$ and $h \in \mathbb{L}_2(\mathbb{P}^{X}) = \mathbb{L}_2(\lambda_{[0,1]})$, hence $h(X) \in \mathbb{L}_2(\mathrm{U}_{\theta|\mathrm{T}})$, then we obtain $Yh(X) \in \mathbb{L}_1(\mathrm{U}_{\theta|\mathrm{T}})$ and

$$\mathrm{U}_{\theta|\mathrm{T}}(Yh(X)) = \mathbb{P}^{X}(\mathbb{P}^{Y|X}_{\mathrm{T}\theta}(Y)h) = \mathbb{P}^{X}\big((\mathrm{T}\theta)h\big) = \lambda_{[0,1]}\big((\mathrm{T}\theta)h\big) = \langle \mathrm{T}\theta, h\rangle_{\mathbb{L}_2(\lambda_{[0,1]})} \in \mathbb{R}$$

identifying again equivalence classes and their representatives. We consider the statistical product experiment $\big(([0,1]\times\mathbb{R})^{n}, (\mathscr{B}_{[0,1]} \otimes \mathscr{B})^{\otimes n}, \mathrm{U}^{\otimes n}_{\Theta\times\{\mathrm{T}\}} := (\mathrm{U}^{\otimes n}_{\theta|\mathrm{T}})_{\theta\in\Theta}\big)$ of size $n \in \mathbb{N}$ and for $\theta \in \Theta$ we denote by $((X_i, Y_i))_{i\in[\![n]\!]} \sim \mathrm{U}^{\otimes n}_{\theta|\mathrm{T}}$ an iid. sample of $(X, Y) \sim \mathrm{U}_{\theta|\mathrm{T}} = \mathrm{U}_{[0,1]} \odot \mathbb{P}^{Y|X}_{\mathrm{T}\theta}$. Keep in mind that $\mathrm{V} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1]}), \ell_2)$ and $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \ell_2)$ are generalised Fourier series transform as in Notation §01|02.07 which are fixed and *known in advanced*. Evidently, for each $\theta \in \Theta \subseteq \mathbb{H}$ the generalised Fourier coefficients $\theta_{\bullet} = (\theta_j)_{j\in\mathbb{N}} = \mathrm{U}\theta$ and $g_{\bullet} = (g_j)_{j\in\mathbb{N}} = \mathrm{V}g = \mathrm{M}_{\mathfrak{s}}\theta_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet}$ satisfy

$$g_j = \mathfrak{s}_j\theta_j = \langle \mathrm{M}_{\mathfrak{s}}\theta_{\bullet}, \mathbb{1}^{\{j\}}_{\bullet}\rangle_{\ell_2} = \langle \mathrm{T}\theta, \mathrm{V}^{\star}\mathbb{1}^{\{j\}}_{\bullet}\rangle_{\mathbb{L}_2(\lambda_{[0,1]})} = \langle \mathrm{T}\theta, \mathrm{v}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1]})} = \mathrm{U}_{\theta|\mathrm{T}}(Y\mathrm{v}_j(X))$$

for each $j \in \mathbb{N}$. The stochastic process $\psi_{\bullet} = (\psi_j(X,Y) := Y \mathrm{v}_j(X))_{j\in\mathbb{N}} \in \mathcal{M}((\mathscr{B}_{[0,1]} \otimes \mathscr{B}) \otimes 2^{\mathbb{N}})$ fulfils Assumption §01|01.04 and $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} = \mathrm{U}_{\theta|\mathrm{T}}(\psi_{\bullet})$. Similar to an Empirical mean model §01|02.04 we define $\widehat{g}_{\bullet} = (\widehat{g}_j := \widehat{\mathbb{P}}_n(\psi_j))_{j\in\mathbb{N}} \in \mathcal{M}((\mathscr{B}_{[0,1]} \otimes \mathscr{B})^{\otimes n} \otimes 2^{\mathbb{N}})$. By construction $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} = \mathrm{U}_{\theta|\mathrm{T}}(\psi_{\bullet}) \in \mathcal{M}(2^{\mathbb{N}})$ is the ℓ_2-mean of $\widehat{g}_{\bullet}$. For each $j \in \mathbb{N}$ the statistic $\dot{\varepsilon}_j := n^{1/2}(\widehat{\mathbb{P}}_n(\psi_j) - \mathrm{U}_{\theta|\mathrm{T}}(\psi_j)) \in \mathcal{M}((\mathscr{B}_{[0,1]} \otimes \mathscr{B})^{\otimes n})$ is centred, i.e. $\dot{\varepsilon}_j \in \mathbb{L}_1(\mathrm{U}_{\theta|\mathrm{T}}^{\otimes n})$ with $\mathrm{U}_{\theta|\mathrm{T}}^{\otimes n}(\dot{\varepsilon}_j) = 0$, and exploiting $\psi_{\bullet} \in (\mathscr{B}_{[0,1]} \otimes \mathscr{B}) \otimes 2^{\mathbb{N}}$ the stochastic process

$$\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathbb{N}} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathrm{U}_{\theta|\mathrm{T}})(\psi_{\bullet}) = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_{\bullet}) - \mathrm{U}_{\theta|\mathrm{T}}(\psi_{\bullet})) \in \mathcal{M}((\mathscr{B}_{[0,1]} \otimes \mathscr{B})^{\otimes n} \otimes 2^{\mathbb{N}})$$

satisfies Assumption §01|01.04, and by construction $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ is a noisy version of $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet}$. □

§01|04.11 **Notation** (Circular additive convolution). Let $\mathbb{q}, \mathbb{p}$ be two Lebesgue densities on $([0,1), \mathscr{B}_{[0,1]})$, then their *circular additive convolution* is given by

$$g(y) := (\mathbb{q} \circledast \mathbb{p})(y) := \int_{[0,1)} \mathbb{q}(y - x - \lfloor y - x \rfloor)\mathbb{p}(x)\lambda_{[0,1)}(dx) \quad \forall y \in [0,1).$$

We note that g is again a density on $([0,1), \mathscr{B}_{[0,1]})$. Consider the *complex* Hilbert spaces $\mathbb{L}_2(\lambda_{[0,1)})$ and $\mathbb{J} := \ell_2(\mathbb{Z})$, the exponentials $\mathrm{e}_{\bullet} := (\mathrm{e}_j)_{j\in\mathbb{Z}}$ given by $\mathrm{e}_j(x) := \exp(-\imath 2\pi j x)$ for $x \in [0,1)$ and $j \in \mathbb{Z}$, and the *Fourier-series transform* $\mathrm{F} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}), \ell_2(\mathbb{Z}))$ with

$$g \mapsto \mathrm{F}g := g_{\bullet} = ((\mathrm{F}g)_j := g_j := \langle g, \mathrm{e}_j \rangle_{\mathbb{L}_2(\lambda_{[0,1)})})_{j\in\mathbb{Z}}$$

(see Notations §01|02.10 and §01|02.12). Let $\varphi \in \mathbb{L}_1(\lambda_{[0,1)})$ and let $\lfloor\cdot\rfloor$ be the floor function, then the *circular additive convolution operator* $\circledast_{\varphi} : \mathbb{L}_2(\lambda_{[0,1)}) \to \mathbb{L}_2(\lambda_{[0,1)})$ with $h \mapsto \circledast_{\varphi} h$ defined by

$$(\circledast_{\varphi} h)(t) := (\varphi \circledast h)(t) := \int_{[0,1)} \varphi(t - s - \lfloor t - s \rfloor)h(s)\lambda_{[0,1)}(ds) \quad \forall t \in [0,1)$$

satisfies $\|\circledast_{\varphi}\|_{\mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}))} \leqslant \|\varphi\|_{\mathbb{L}_1(\lambda_{[0,1)})} = \lambda_{[0,1)}(|\varphi|)$. Since $\varphi \in \mathbb{L}_1(\lambda_{[0,1)})$ and for each $j \in \mathbb{Z}$, $\mathrm{e}_j \in \mathbb{L}_\infty(\lambda_{[0,1)})$ we have $\varphi\bar{\mathrm{e}}_j \in \mathbb{L}_1(\lambda_{[0,1)})$ too. More precisely, for each $j \in \mathbb{Z}$ we have

$$|\lambda_{[0,1)}(\varphi\bar{\mathrm{e}}_j)| \leqslant \|\varphi\bar{\mathrm{e}}_j\|_{\mathbb{L}_1(\lambda_{[0,1)})} \leqslant \|\varphi\|_{\mathbb{L}_1(\lambda_{[0,1)})}\|\bar{\mathrm{e}}_j\|_{\mathbb{L}_\infty(\lambda_{[0,1)})} = \|\varphi\|_{\mathbb{L}_1(\lambda_{[0,1)})}$$

and hence $\varphi_{\bullet} := \lambda_{[0,1)}(\varphi\bar{\mathrm{e}}_{\bullet}) = (\lambda_{[0,1)}(\varphi\bar{\mathrm{e}}_j))_{j\in\mathbb{Z}} \in \ell_\infty(\mathbb{Z})$ with a slight abuse of notation satisfies $\|\varphi_{\bullet}\|_{\ell_\infty(\mathbb{Z})} = \|\lambda_{[0,1)}(\varphi\bar{\mathrm{e}}_{\bullet})\|_{\ell_\infty(\mathbb{Z})} \leqslant \|\varphi\|_{\mathbb{L}_1(\lambda_{[0,1)})}$. Obviously, if $\varphi \in \mathbb{L}_2(\lambda_{[0,1)})$ (implying $\varphi \in \mathbb{L}_1(\lambda_{[0,1)})$) then $\varphi_{\bullet} = \lambda_{[0,1)}(\varphi\bar{\mathrm{e}}_{\bullet}) = (\lambda_{[0,1)}(\varphi\bar{\mathrm{e}}_j) = \langle \varphi, \mathrm{e}_j \rangle_{\mathbb{L}_2(\lambda_{[0,1)})})_{j\in\mathbb{Z}} = \mathrm{F}\varphi \in \ell_2(\mathbb{Z})$. However, for each $\varphi \in \mathbb{L}_1(\lambda_{[0,1)})$ and $h \in \mathbb{L}_2(\lambda_{[0,1)})$ the *circular convolution theorem* states

$$(\circledast_{\varphi} h)_j = \langle \circledast_{\varphi} h, \mathrm{e}_j \rangle_{\mathbb{L}_2(\lambda_{[0,1)})} = \lambda_{[0,1)}(\varphi\bar{\mathrm{e}}_j)\langle h, \mathrm{e}_j \rangle_{\mathbb{L}_2(\lambda_{[0,1)})} = \lambda_{[0,1)}(\varphi\bar{\mathrm{e}}_j)(\mathrm{F}h)_j = \varphi_j h_j \quad \forall j \in \mathbb{Z},$$

or $(\circledast_{\varphi} h)_{\bullet} = \mathrm{F}(\circledast_{\varphi} h) = \lambda_{[0,1)}(\varphi\bar{\mathrm{e}}_{\bullet})(\mathrm{F}h) = \varphi_{\bullet} h_{\bullet}$ in short. Consequently, $(\varphi_{\bullet}, \mathrm{F}, \mathrm{F})$ is a singular value decomposition of $\circledast_{\varphi}$ with $\varphi_{\bullet} \in \ell_\infty(\mathbb{Z})$, and thus $\circledast_{\varphi} \in \mathbb{L}^{\mathrm{F},\mathrm{F}}(\mathbb{L}^{\mathrm{M}}(\ell_2(\mathbb{Z}))) = \mathrm{F}^\star(\mathbb{L}^{\mathrm{M}}(\ell_2(\mathbb{Z})))\mathrm{F}$. □

§01|04.12 **Cicular density deconvolution**. Consider the *complex* Hilbert spaces $\mathbb{L}_2(\lambda_{[0,1)})$ and $\mathbb{J} := \ell_2(\mathbb{Z})$. Let $\mathbb{D}_2$ be a set of square-integrable Lebesgue densities on $([0,1), \mathscr{B}_{[0,1)})$, and hence $\mathbb{D}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)}) \subseteq \mathbb{L}_1(\lambda_{[0,1)})$ (by the usual embedding of real-valued functions) as in Notation §01|02.10. We denote for each density $\mathbb{p} \in \mathbb{L}_1(\lambda_{[0,1)})$ by $\mathbb{P}_{\mathbb{p}} := \mathbb{p}\lambda_{[0,1)} \in \mathscr{W}(\mathscr{B}_{[0,1)})$ the associated probability measure. Given a Lebesgue density $\mathbb{q} \in \mathbb{L}_1(\lambda_{[0,1)})$ presumed to be fixed and *known in advance* for each Lebesgue density $\mathbb{p} \in \mathbb{D}_2$ we consider the Lebesgue density $g = \mathbb{q} \circledast \mathbb{p} \in \mathbb{L}_2(\lambda_{[0,1)})$ (see Notation §01|04.11) and denote by $\mathbb{P}_{\mathbb{p}|\mathbb{q}} := (\mathbb{q} \circledast \mathbb{p})\lambda_{[0,1)} = g\lambda_{[0,1)} \in \mathscr{W}(\mathscr{B}_{[0,1)})$ the associated probability measure. We consider the statistical product experiment $\left([0,1)^n, \mathscr{B}_{[0,1)}^{\otimes n}, \mathbb{P}_{\mathbb{D}_2\times\{\mathbb{q}\}}^{\otimes n} := (\mathbb{P}_{\mathbb{p}|\mathbb{q}}^{\otimes n})_{\mathbb{p}\in\mathbb{D}_2}\right)$. Let $\mathrm{F} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}), \ell_2(\mathbb{Z}))$ be the Fourier-series transform (see Notation §01|02.12). Evidently, for $g \in \mathbb{L}_2(\lambda_{[0,1)}) \subseteq \mathbb{L}_1(\lambda_{[0,1)})$ its Fourier-series $g_{\bullet} = (g_j)_{j\in\mathbb{Z}} = \mathrm{F}g$ satisfies $g_j = \lambda_{[0,1)}(g\overline{\mathrm{e}}_j) = \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\overline{\mathrm{e}}_j)$ for each $j\in\mathbb{Z}$. Moreover, considering the Fourier-series $\mathbb{p}_{\bullet} = (\mathbb{p}_j)_{j\in\mathbb{Z}} = \mathrm{F}\mathbb{p}$ of $\mathbb{p}\in\mathbb{D}_2\subseteq\mathbb{L}_2(\lambda_{[0,1)})$ by the *circular convolution theorem* we have $g_{\bullet} = \mathrm{F}(\mathbb{q}\circledast\mathbb{p}) = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet}$ with $\mathbb{q}_{\bullet} = \lambda_{[0,1)}(\mathbb{q}\overline{\mathrm{e}}_{\bullet}) \in \ell_\infty(\mathbb{Z})$ and $\mathbb{p}_{\bullet} = \mathrm{F}\mathbb{p}\in\ell_2(\mathbb{Z})$ (see Notation §01|04.11). Moreover, the stochastic process $\overline{\mathrm{e}}_{\bullet} = (\overline{\mathrm{e}}_j)_{j\in\mathbb{Z}}$ on $([0,1),\mathscr{B}_{[0,1)})$ is $(\mathscr{B}_{[0,1)}\otimes 2^{\mathbb{Z}})$-$\mathscr{B}$-measurable, i.e. $\overline{\mathrm{e}}_{\bullet}\in\mathcal{M}(\mathscr{B}_{[0,1)}\otimes 2^{\mathbb{Z}})$ for short. We define $\widehat{g}_{\bullet} = (\widehat{g}_j := \widehat{\mathbb{P}}_n(\overline{\mathrm{e}}_j))_{j\in\mathbb{Z}} = \widehat{\mathbb{P}}_n(\overline{\mathrm{e}}_{\bullet}) \in \mathcal{M}(\mathscr{B}_{[0,1)}^{\otimes n}\otimes 2^{\mathbb{Z}})$ similar to an Empirical mean model §01|02.04 where for each $j\in\mathbb{Z}$

$$y=(y_i)_{i\in[\![n]\!]}\mapsto \widehat{g}_j(y) = (\widehat{\mathbb{P}}_n(\overline{\mathrm{e}}_j))(y) = n^{-1}\sum_{i\in[\![n]\!]}\overline{\mathrm{e}}_j(y_i) = n^{-1}\sum_{i\in[\![n]\!]}\exp(\imath 2\pi j y_i).$$

By construction $g_\bullet = \mathbb{q}_\bullet\mathbb{p}_\bullet = \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\overline{\mathrm{e}}_\bullet)\in\ell_2(\mathbb{Z})$ is the mean of $\widehat{g}_\bullet$. For each $j\in\mathbb{Z}$ the statistic $\dot{\varepsilon}_j := n^{1/2}(\widehat{\mathbb{P}}_n(\overline{\mathrm{e}}_j) - \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\overline{\mathrm{e}}_j))\in\mathcal{M}(\mathscr{B}_{[0,1)}^{\otimes n})$ is centred, i.e. $\dot\varepsilon_j\in\mathbb{L}_1(\mathbb{P}_{\mathbb{p}|\mathbb{q}}^{\otimes n})$ with $\mathbb{P}^{\otimes n}_{\mathbb{p}|\mathbb{q}}(\dot\varepsilon_j)=0$, and exploiting $\overline{\mathrm{e}}_\bullet = (\overline{\mathrm{e}}_j)_{j\in\mathbb{Z}}\in\mathcal{M}(\mathscr{B}_{[0,1)}\otimes 2^{\mathbb{Z}})$ the complex valued stochastic process

$$\dot\varepsilon_\bullet = (\dot\varepsilon_j)_{j\in\mathbb{Z}} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_{\mathbb{p}|\mathbb{q}})(\overline{\mathrm{e}}_\bullet) = n^{1/2}(\widehat{\mathbb{P}}_n(\overline{\mathrm{e}}_\bullet) - \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\overline{\mathrm{e}}_\bullet))\in\mathcal{M}(\mathscr{B}^{\otimes n}_{[0,1)}\otimes 2^{\mathbb{Z}})$$

satisfies Assumption §01|02.11. Since $\widehat g_j = g_j + n^{-1/2}\dot\varepsilon_j = \mathbb{q}_j\mathbb{p}_j + n^{-1/2}\dot\varepsilon_j$ for each $j\in\mathbb{R}$ by construction $\widehat g_\bullet = g_\bullet + n^{-1/2}\dot\varepsilon_\bullet = \mathbb{q}_\bullet\mathbb{p}_\bullet + n^{-1/2}\dot\varepsilon_\bullet$ is a noisy version of $g_\bullet = \mathbb{q}_\bullet\mathbb{p}_\bullet$.

§01|04.13 **Cicular regression deconvolution with uniform design**. Consider the measure space $([0,1],\mathscr{B}_{[0,1)},\lambda_{[0,1)})$ and the *complex* Hilbert spaces $\mathbb{L}_2(\lambda_{[0,1)})$ and $\ell_2(\mathbb{Z})$ as in Notation §01|02.10. Let the circular convolution operator $\circledast_\varphi\in\mathbb{L}^{\mathrm{F,F}}(\mathbb{L}^{\mathbb{M}}(\ell_2(\mathbb{Z})))\subseteq\mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}))$ with $\varphi\in\mathbb{L}_1(\lambda_{[0,1)})$ be *known in advance* (see Notation §01|04.11), i.e. $\mathrm{F}\circledast_\varphi\mathrm{F}^\star = \mathrm{M}_{\varphi_\bullet}\in\mathbb{L}^{\mathbb{M}}(\ell_2(\mathbb{Z}))$ and in other words $\circledast_\varphi$ has a *known* singular value decomposition $(\varphi_\bullet,\mathrm{F},\mathrm{F})$ with sequence of singular values $\varphi_\bullet\in\ell_\infty(\mathbb{Z})$. Let (X,Y) be a $[0,1)\times\mathbb{R}$-valued random vector. As in Model §01|02.14 we assume in what follows that the regressor X is uniformly distributed on the interval $[0,1)$, i.e. $X\sim\mathrm{U}_{[0,1)} = \lambda_{[0,1)} = \mathbb{P}^X$ and that given $\circledast_\varphi f =: g\in\mathbb{L}_2(\lambda_{[0,1)})$ for some $f\in\mathbb{F}_2\subseteq\mathbb{L}_2(\lambda_{[0,1)}) = \mathbb{H}$ the joint distribution of (X,Y) is given by $\mathrm{U}_{f|\varphi} := \mathrm{U}_{[0,1)}\odot\mathbb{P}^{Y|X}_{\circledast_\varphi f}$ and without fully specifying the regular conditional distribution $\mathbb{P}^{Y|X}_{\circledast_\varphi f}$ which however satisfies $\mathbb{P}^{Y|X}_{\circledast_\varphi f}(\mathrm{id}_{\mathbb{R}}) = \mathbb{P}_{\circledast_\varphi f}(Y|X) = \circledast_\varphi f = g\in\mathbb{L}_2(\lambda_{[0,1)})$. Keep in mind that we tactically identify X and Y with the coordinate map $\Pi_{[0,1)}$ and $\Pi_{\mathbb{R}}$, respectively, and thus (X,Y) with the identity $\mathrm{id}_{[0,1)\times\mathbb{R}}$. Consequently, if $Y\in\mathbb{L}_2(\mathrm{U}_{f|\varphi})$ and $h\in\mathbb{L}_2(\mathbb{P}^X) = \mathbb{L}_2(\lambda_{[0,1)})$, hence $\overline{h}(X)\in\mathbb{L}_2(\mathrm{U}_{f|\varphi})$, then we obtain $Y\overline{h}(X)\in\mathbb{L}_1(\mathrm{U}_{f|\varphi})$ and

$$\mathrm{U}_{f|\varphi}(Y\overline{h}(X)) = \mathbb{P}^X(\mathbb{P}^{Y|X}_{\circledast_\varphi f}(Y)\overline{h}) = \mathbb{P}^X\big((\circledast_\varphi f)\overline{h}\big) = \lambda_{[0,1)}\big((\circledast_\varphi f)\overline{h}\big) = \langle\circledast_\varphi f, h\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}\in\mathbb{C}$$

identifying again equivalence classes and their representatives. We consider the statistical product experiment $\big(([0,1)\times\mathbb{R})^n, (\mathscr{B}_{[0,1)}\otimes\mathscr{B})^{\otimes n}, \mathrm{U}^{\otimes n}_{\mathbb{F}_2\times\{\varphi\}} := (\mathrm{U}^{\otimes n}_{f|\varphi})_{f\in\mathbb{F}_2}\big)$ of size $n\in\mathbb{N}$ and for $f\in\mathbb{F}_2$ we denote by $((X_i,Y_i))_{i\in[\![n]\!]}\sim\mathrm{U}^{\otimes n}_{f|\varphi}$ an iid. sample of $(X,Y)\sim\mathrm{U}_{f|\varphi} = \mathrm{U}_{[0,1)}\odot\mathbb{P}^{Y|X}_{\circledast_\varphi f}$. Keep in mind that $\mathrm{F}\in\mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}),\ell_2(\mathbb{Z}))$ is the Fourier series transform as in Notation §01|02.12 which is fixed and evidently *known in advanced*. For each $f\in\mathbb{F}_2\subseteq\mathbb{L}_2(\lambda_{[0,1)}) = \mathbb{H}$ the Fourier coefficients $f_\bullet = (f_j)_{j\in\mathbb{Z}} = \mathrm{F}f$ and $g_\bullet = (g_j)_{j\in\mathbb{Z}} = \mathrm{F}g = \mathrm{F}(\circledast_\varphi f) = \mathrm{M}_{\varphi_\bullet}f_\bullet = \varphi_\bullet f_\bullet$ (Notation §01|04.11) satisfy

$$g_j = \varphi_j f_j = \langle\mathrm{M}_{\varphi_\bullet}f_\bullet, \mathbb{1}^{\{j\}}_\bullet\rangle_{\ell_2(\mathbb{Z})} = \langle\circledast_\varphi f, \mathrm{F}^\star\mathbb{1}^{\{j\}}_\bullet\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} = \langle\circledast_\varphi f, \mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} = \mathrm{U}_{f|\varphi}(Y\overline{\mathrm{e}}_j(X))$$

for each $j\in\mathbb{Z}$. The stochastic process $\psi_\bullet = (\psi_j(X,Y) := Y\overline{\mathrm{e}}_j(X))_{j\in\mathbb{Z}}\in\mathcal{M}((\mathscr{B}_{[0,1)}\otimes\mathscr{B})\otimes 2^{\mathbb{Z}})$ fulfils Assumption §01|01.04 and $g_\bullet = \varphi_\bullet f_\bullet = \mathrm{U}_{f|\varphi}(\psi_\bullet)$. Similar to an Empirical mean model §01|02.04 we define $\widehat g_\bullet = (\widehat g_j := \widehat{\mathbb{P}}_n(\psi_j))_{j\in\mathbb{Z}} = \widehat{\mathbb{P}}_n(\psi_\bullet)\in\mathcal{M}((\mathscr{B}_{[0,1)}\otimes\mathscr{B})^{\otimes n}\otimes 2^{\mathbb{Z}})$. By construction $g_\bullet = \varphi_\bullet f_\bullet = \mathrm{U}_{f|\varphi}(\psi_\bullet)\in\ell_2(\mathbb{Z})$ is the mean of $\widehat g_\bullet$. For each $j\in\mathbb{Z}$ the statistic $\dot\varepsilon_j := n^{1/2}(\widehat{\mathbb{P}}_n(\psi_j) - \mathrm{U}_{f|\varphi}(\psi_j))\in\mathcal{M}((\mathscr{B}_{[0,1)}\otimes\mathscr{B})^{\otimes n})$ is centred, i.e. $\dot\varepsilon_j\in\mathbb{L}_1(\mathrm{U}^{\otimes n}_{f|\varphi})$ with $\mathrm{U}^{\otimes n}_{f|\varphi}(\dot\varepsilon_j) = 0$, and exploiting $\psi_\bullet\in\mathcal{M}((\mathscr{B}_{[0,1)}\otimes\mathscr{B})\otimes 2^{\mathbb{Z}})$ the stochastic process

$$\dot\varepsilon_\bullet = (\dot\varepsilon_j)_{j\in\mathbb{Z}} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathrm{U}_{f|\varphi})(\psi_\bullet) = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_\bullet) - \mathrm{U}_{f|\varphi}(\psi_\bullet))\in\mathcal{M}((\mathscr{B}_{[0,1)}\otimes\mathscr{B})^{\otimes n}\otimes 2^{\mathbb{Z}})$$

satisfies Assumption §01|01.04, and by construction $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} = \varphi_{\bullet} f_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ is a noisy version of $g_{\bullet} = \varphi_{\bullet} f_{\bullet}$. □

§01|04.14 **Notation** (Additive convolution on $\mathbb{R}$). Let $\mathbb{q}, \mathbb{p}$ be two Lebesgue densities on $(\mathbb{R}, \mathscr{B})$, then their *additive convolution* is given by

$$g(y) := (\mathbb{q} * \mathbb{p})(y) := \int_{\mathbb{R}} \mathbb{q}(y-x)\mathbb{p}(x)\lambda(dx) = \int_{\mathbb{R}} \mathbb{p}(y-x)\mathbb{q}(x)\lambda(dx) \quad \text{for } \lambda\text{-a.e. } y \in \mathbb{R}.$$

We note that g is again a density on $(\mathbb{R}, \mathscr{B})$ (keep in mind that we identify representatives and equivalence classes). Consider the *complex* Hilbert space $\mathbb{L}_2 = \mathbb{L}_2(\lambda)$, the exponentials $\mathrm{e}_{\bullet} := (\mathrm{e}_j)_{j \in \mathbb{R}} \in \mathcal{M}(\mathscr{B}^2)$ given by $\mathrm{e}_j(x) := \exp(-\imath 2\pi j x)$ for $x, j \in \mathbb{R}$, and the *Fourier-Plancherel transform* $\mathrm{F} \in \mathbb{L}(\mathbb{L}_2)$ satisfying

$$\mathrm{F}h = h_{\bullet} = ((\mathrm{F}h)_j := h_j := \lambda(h\bar{\mathrm{e}}_j))_{j \in \mathbb{R}} = \lambda(h\bar{\mathrm{e}}_{\bullet}), \quad \forall h \in \mathbb{L}_1 \cap \mathbb{L}_2$$

(see Notations §01|02.10 and §01|02.15). Consider $\mathbb{p} \in \mathbb{L}_p$ and $\mathbb{q} \in \mathbb{L}_q$ with conjugate exponents $(1/p + 1/q = 1)$ then the integral $\int_{\mathbb{R}} \mathbb{p}(y-x)\mathbb{q}(x)\lambda(dx)$ exists for all $y \in \mathbb{R}$ and hence $(\mathbb{q} * \mathbb{p})(y)$ is for all $y \in \mathbb{R}$ defined. In the case $\mathbb{q} \in \mathbb{L}_1$ and $\mathbb{p} \in \mathbb{L}_p$ with $p \in \overline{\mathbb{R}}_{\geqslant 1}$ the integral $(\mathbb{q} * \mathbb{p})(y)$ exists for λ-a.e. $y \in \mathbb{R}$ only. However, the λ-a.e.-defined function $\mathbb{q} * \mathbb{p}$ belongs to $\mathbb{L}_p$ and satisfies $\|\mathbb{q} * \mathbb{p}\|_{\mathbb{L}_p} \leqslant \|\mathbb{q}\|_{\mathbb{L}_1} \|\mathbb{p}\|_{\mathbb{L}_p}$. Werner [2011] p.337 for $p = 2$ and general case $p \in \overline{\mathbb{R}}_{\geqslant 1}$ lecture notes P. Maréchal (Analyse pour les problèmes inverses d'imagerie). For $\varphi \in \mathbb{L}_1$ the *additive convolution operator* $*_{\varphi} : \mathbb{L}_2 \to \mathbb{L}_2$ with $h \mapsto *_{\varphi} h$ defined by

$$(*_{\varphi} h)(t) := (\varphi * h)(t) := \int_{\mathbb{R}} \varphi(t-s)h(s)\lambda(ds) \quad \text{for } \lambda\text{-a.e. } y \in \mathbb{R}$$

satisfies $\|*_{\varphi}\|_{\mathbb{L}(\mathbb{L}_2)} \leqslant \|\varphi\|_{\mathbb{L}_1} = \lambda(|\varphi|)$. Since $\varphi \in \mathbb{L}_1$ and for each $j \in \mathbb{R}$, $\mathrm{e}_j \in \mathbb{L}_\infty$ we have $\varphi\bar{\mathrm{e}}_j \in \mathbb{L}_1$ too. More precisely, for each $j \in \mathbb{R}$ we have

$$|\lambda(\varphi\bar{\mathrm{e}}_j)| \leqslant \|\varphi\bar{\mathrm{e}}_j\|_{\mathbb{L}_1} \leqslant \|\varphi\|_{\mathbb{L}_1} \|\bar{\mathrm{e}}_j\|_{\mathbb{L}_\infty} = \|\varphi\|_{\mathbb{L}_1}$$

and hence $\varphi_{\bullet} := \lambda(\varphi\bar{\mathrm{e}}_{\bullet}) = (\varphi_j := \lambda(\varphi\bar{\mathrm{e}}_j))_{j \in \mathbb{R}} \in \mathbb{L}_\infty$ with a slight abuse of notation satisfies $\|\varphi_{\bullet}\|_{\mathbb{L}_\infty} = \|\lambda(\varphi\bar{\mathrm{e}}_{\bullet})\|_{\mathbb{L}_\infty} \leqslant \|\varphi\|_{\mathbb{L}_1}$. Obviously, if $\varphi \in \mathbb{L}_1 \cap \mathbb{L}_2$ then $\varphi_{\bullet} = \lambda(\varphi\bar{\mathrm{e}}_{\bullet}) = \mathrm{F}\varphi \in \mathbb{L}_2$. However, for each $\varphi \in \mathbb{L}_1$ and $h \in \mathbb{L}_1 \cap \mathbb{L}_2$ the *convolution theorem* states

$$(*_{\varphi} h)_j = \lambda((*_{\varphi} h)\bar{\mathrm{e}}_j) = \lambda(\varphi\bar{\mathrm{e}}_j)\lambda(h\bar{\mathrm{e}}_j) = \varphi_j(\mathrm{F}h)_j = \varphi_j h_j \quad \text{for } \lambda\text{-a.e. } j \in \mathbb{R}.$$

or $(*_{\varphi} h)_{\bullet} = \mathrm{F}(*_{\varphi} h) = \lambda(\varphi\bar{\mathrm{e}}_{\bullet})(\mathrm{F}h) = \varphi_{\bullet} h_{\bullet}$ λ-a.s. in short. Consequently, $(\varphi_{\bullet}, \mathrm{F}, \mathrm{F})$ is a singular value decomposition of $*_{\varphi}$ with $\varphi_{\bullet} \in \mathbb{L}_\infty$, and thus $*_{\varphi} \in \mathbb{L}^{\mathrm{F},\mathrm{F}}(\mathbb{L}^{\mathrm{M}}(\mathbb{L}_2)) = \mathrm{F}^{\star}(\mathbb{L}^{\mathrm{M}}(\mathbb{L}_2))\mathrm{F}$. □

§01|04.15 **Density additive deconvolution on $\mathbb{R}$.** Consider the *complex* Hilbert space $\mathbb{L}_2 = \mathbb{L}_2(\lambda)$. Let $\mathbb{D}_2$ be a set of square-integrable Lebesgue densities on $(\mathbb{R}, \mathscr{B})$, and hence $\mathbb{D}_2 \subseteq \mathbb{L}_2 \cap \mathbb{L}_1$ (by the usual embedding of real-valued functions) as in Notation §01|02.10. We denote for each density $\mathbb{p} \in \mathbb{L}_1$ by $\mathbb{P}_{\mathbb{p}} := \mathbb{p}\lambda \in \mathscr{W}(\mathscr{B})$ the associated probability measure. Given a Lebesgue density $\mathbb{q} \in \mathbb{L}_1(\lambda)$ presumed to be fixed and *known in advance* for each Lebesgue density $\mathbb{p} \in \mathbb{D}_2$ we consider the Lebesque density $g = *_{\mathbb{q}}\mathbb{p} = \mathbb{q} * \mathbb{p} \in \mathbb{L}_2 \cap \mathbb{L}_1$ (see Notation §01|04.14) and denote by $\mathbb{P}_{\mathbb{p}|\mathbb{q}} := (\mathbb{q} * \mathbb{p})\lambda = g\lambda \in \mathscr{W}(\mathscr{B})$ the associated probability measure. We consider the statistical product experiment $(\mathbb{R}^n, \mathscr{B}^{\otimes n}, \mathbb{P}^{\otimes n}_{\mathbb{D}_2 \times \{\mathbb{q}\}} := (\mathbb{P}^{\otimes n}_{\mathbb{p}|\mathbb{q}})_{\mathbb{p} \in \mathbb{D}_2})$. Let $\mathrm{F} \in \mathbb{L}(\mathbb{L}_2)$ be the Fourier-Plancherel transform (see Notation §01|02.15). Evidently, for $g \in \mathbb{L}_2 \cap \mathbb{L}_1$ its Fourier-Plancherel transform $g_{\bullet} = (g_j)_{j \in \mathbb{R}} = \mathrm{F}g$ satisfies $g_j = \lambda(g\bar{\mathrm{e}}_j) = \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\bar{\mathrm{e}}_j)$ for all $j \in \mathbb{R}$. Moreover, considering the Fourier-Plancherel transform $\mathbb{p}_{\bullet} = (\mathbb{p}_j)_{j \in \mathbb{R}} = \mathrm{F}\mathbb{p}$ of $\mathbb{p} \in \mathbb{D}_2 \subseteq \mathbb{L}_2 \cap \mathbb{L}_1$ by the *additive convolution theorem* we have $g_{\bullet} = \mathrm{F}(*_{\mathbb{q}}\mathbb{p}) = \lambda(\mathbb{q}\bar{\mathrm{e}}_{\bullet})(\mathrm{F}\mathbb{p}) = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet}$ λ-a.s. with $\mathbb{q}_{\bullet} = \lambda(\mathbb{q}\bar{\mathrm{e}}_{\bullet}) \in \mathbb{L}_\infty$

and $\mathbb{p}_{\bullet} = \mathrm{F}\mathbb{p} \in \mathbb{L}_2$ (see Notation §01|04.14). Moreover, the stochastic process $\bar{\mathrm{e}}_{\bullet} = (\bar{\mathrm{e}}_j)_{j\in\mathbb{R}}$ on $(\mathbb{R}, \mathscr{B})$ is $\mathscr{B}^2$-$\mathscr{B}$-measurable, i.e. $\bar{\mathrm{e}}_{\bullet} \in \mathcal{M}(\mathscr{B}^2)$ for short. We define

$$\widehat{g}_{\bullet} = (\widehat{g}_j := \widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_j))_{j\in\mathbb{R}} = \widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_{\bullet}) \in \mathcal{M}(\mathscr{B}^{\otimes n} \otimes \mathscr{B})$$

similar to an Empirical mean model §01|02.04 where for each $j \in \mathbb{R}$

$$y = (y_i)_{i\in[\![n]\!]} \mapsto \widehat{g}_j(y) = (\widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_j))(y) = n^{-1} \sum_{i\in[\![n]\!]} \bar{\mathrm{e}}_j(y_i) = n^{-1} \sum_{i\in[\![n]\!]} \exp(\imath 2\pi j y_i).$$

By construction $g_{\bullet} = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet} = \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\bar{\mathrm{e}}_{\bullet}) \in \mathbb{L}_2$ is the mean of $\widehat{g}_{\bullet}$. For each $j \in \mathbb{R}$ the statistic $\dot{\varepsilon}_j := n^{1/2}(\widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_j) - \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\bar{\mathrm{e}}_j)) \in \mathcal{M}(\mathscr{B}^{\otimes n})$ is centred, i.e. $\dot{\varepsilon}_j \in \mathbb{L}_1(\mathbb{P}^{\otimes n}_{\mathbb{p}|\mathbb{q}})$ with $\mathbb{P}^{\otimes n}_{\mathbb{p}|\mathbb{q}}(\dot{\varepsilon}_j) = 0$, and exploiting $\bar{\mathrm{e}}_{\bullet} = (\bar{\mathrm{e}}_j)_{j\in\mathbb{R}} \in \mathcal{M}(\mathscr{B}^2)$ the complex valued stochastic process

$$\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathbb{R}} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_{\mathbb{p}|\mathbb{q}})(\bar{\mathrm{e}}_{\bullet}) = n^{1/2}(\widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_{\bullet}) - \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\bar{\mathrm{e}}_{\bullet})) \in \mathcal{M}(\mathscr{B}^{\otimes n} \otimes \mathscr{B})$$

satisfies Assumption §01|02.11, and by construction $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ is a noisy version of $g_{\bullet} = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet}$.

§01|04.16 **Notation** (Multiplicative convolution on $\mathbb{R}_{>0}$). Let $\mathbb{q}, \mathbb{p}$ be two Lebesgue densities on $(\mathbb{R}_{>0}, \mathscr{B}_{>0})$ satisfying , then their *multiplicative convolution* is given by

$$g(y) := (\mathbb{q} \circledast \mathbb{p})(y) := \int_{\mathbb{R}_{>0}} \mathbb{q}(y/x)\mathbb{p}(x)x^{-1}\lambda_{>0}(dx)$$
$$= \int_{\mathbb{R}_{>0}} \mathbb{p}(y/x)\mathbb{q}(x)x^{-1}\lambda_{>0}(dx) \quad \text{for } \lambda_{>0}\text{-a.e. } y \in \mathbb{R}_{>0}.$$

We note that g is again a density on $(\mathbb{R}_{>0}, \mathscr{B}_{>0})$ (keep in mind that we identify representatives and equivalence classes). For $\mathrm{c} \in \mathbb{R}$ fixed and *known in advance* consider the *complex* Hilbert spaces $\mathbb{L}_2(\mathrm{x}^{\mathrm{c}}) = \mathbb{L}_2(\mathbb{R}_{>0}, \mathscr{B}_{>0}, \mathrm{x}^{\mathrm{c}}\lambda_{>0})$ and $\mathbb{L}_2 = \mathbb{L}_2(\mathbb{R}, \mathscr{B}, \lambda)$, the kernel $\mathrm{x}^{\mathrm{c}}\bar{\mathrm{x}}_{\bullet} = (\mathrm{x}^{\mathrm{c}}\bar{\mathrm{x}}_j)_{j\in\mathbb{R}} \in \mathcal{M}(\mathscr{B}_{>0} \otimes \mathscr{B})$ given by $(\mathrm{x}^{\mathrm{c}}\bar{\mathrm{x}}_j)(x) = x^{\mathrm{c}}x^{\imath 2\pi j}$ for $x \in \mathbb{R}_{>0}$, $j \in \mathbb{R}$ and the *Mellin transform* $\mathrm{M}_{\mathrm{c}} \in \mathbb{L}(\mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1}), \mathbb{L}_2)$ (see Notation §01|02.17) satisfying

$$\mathrm{M}_{\mathrm{c}}h = h_{\bullet} = ((\mathrm{M}_{\mathrm{c}}h)_j = \mathrm{x}^{\mathrm{c}-1}\lambda_{>0}(\bar{\mathrm{x}}_j h) = \mathrm{x}^{2\mathrm{c}-1}\lambda_{>0}(\mathrm{x}^{-\mathrm{c}}\bar{\mathrm{x}}_j h))_{j\in\mathbb{R}}, \quad \forall h \in \mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1}) \cap \mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1})$$

(see Notations §01|02.10 and §01|02.17). Consider $\mathbb{p}, \mathbb{q} \in \mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1}) = \mathbb{L}_1(\mathbb{R}_{>0}, \mathscr{B}_{>0}, \mathrm{x}^{\mathrm{c}-1}\lambda_{>0})$ then the integral $\int_{\mathbb{R}_{>0}} \mathbb{p}(y/x)\mathbb{q}(x)x^{-1}\lambda_{>0}(dx)$ exists for $\lambda_{>0}$-a.e. $y \in \mathbb{R}_{>0}$ and hence $(\mathbb{q} \circledast \mathbb{p})(y)$ is for $\lambda_{>0}$-a.e. $y \in \mathbb{R}$ defined and the $\lambda_{>0}$-a.e.-defined function $\mathbb{q} \circledast \mathbb{p}$ belongs to $\mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1})$ and satisfies $\|\mathbb{q} \circledast \mathbb{p}\|_{\mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1})} \leqslant \|\mathbb{q}\|_{\mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1})}\|\mathbb{p}\|_{\mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1})}$. In case $\mathbb{p} \in \mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1}) \cap \mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1})$ and $\mathbb{q} \in \mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1})$ the integral $(\mathbb{q} \circledast \mathbb{p})(y)$ is for $\lambda_{>0}$-a.e. $y \in \mathbb{R}$ defined and the $\lambda_{>0}$-a.e.-defined function $\mathbb{q} \circledast \mathbb{p}$ belongs to $\mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1})$ and satisfies $\|\mathbb{q} \circledast \mathbb{p}\|_{\mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1})} \leqslant \|\mathbb{q}\|_{\mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1})}\|\mathbb{p}\|_{\mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1})}$. (Phd thesis of S. Brenner Miguel [2023]). For $\varphi \in \mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1})$ the *multiplicative convolution operator* $\circledast_{\varphi} : \mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1}) \to \mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1})$ with $h \mapsto \circledast_{\varphi}h$ defined by

$$(\circledast_{\varphi}h)(t) := (\varphi \circledast h)(t) := \int_{\mathbb{R}_{\geqslant 0}} \varphi(t/s)h(s)s^{-1}\lambda_{>0}(ds) \quad \text{for } \lambda_{>0}\text{-a.e. } y \in \mathbb{R}_{>0}$$

satisfies $\|\circledast_{\varphi}\|_{\mathbb{L}(\mathbb{L}_2(\mathrm{x}^{2\mathrm{c}-1}))} \leqslant \|\varphi\|_{\mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1})} = \mathrm{x}^{\mathrm{c}-1}\lambda_{>0}(|\varphi|)$. Since $\varphi \in \mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1})$ and for each $j \in \mathbb{R}$, $\mathrm{x}_j \in \mathbb{L}_\infty(\lambda_{>0})$ we have $\bar{\mathrm{x}}_j\varphi \in \mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1})$ too. More precisely, for each $j \in \mathbb{R}$ we have

$$|\mathrm{x}^{\mathrm{c}-1}\lambda_{>0}(\bar{\mathrm{x}}_j\varphi)| \leqslant \|\bar{\mathrm{x}}_j\varphi\|_{\mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1})} \leqslant \|\varphi\|_{\mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1})}\|\bar{\mathrm{x}}_j\|_{\mathbb{L}_\infty(\lambda_{>0})} = \|\varphi\|_{\mathbb{L}_1(\mathrm{x}^{\mathrm{c}-1})}$$

and hence $\varphi_{\bullet} := x^{c-1}\lambda_{>0}(\overline{x}_{\bullet}\varphi) = (\varphi_j := x^{c-1}\lambda_{>0}(\overline{x}_j\varphi))_{j\in\mathbb{R}} \in \mathbb{L}_\infty$ with a slight abuse of notation satisfies $\|\varphi_{\bullet}\|_{\mathbb{L}_\infty} = \|x^{c-1}\lambda_{>0}(\overline{x}_{\bullet}\varphi)\|_{\mathbb{L}_\infty} \leqslant \|\varphi\|_{\mathbb{L}_1(x^{c-1})} = x^{c-1}\lambda_{>0}(|\varphi|)$. Obviously, if $\varphi \in \mathbb{L}_1(x^{c-1}) \cap \mathbb{L}_2(x^{2c-1})$ then $\varphi_{\bullet} = x^{c-1}\lambda_{>0}(\overline{x}_{\bullet}\varphi) = (x^{2c-1}\lambda_{>0}(x^{-c}\overline{x}_j\varphi))_{j\in\mathbb{R}} = \mathrm{M}_c\varphi \in \mathbb{L}_2$. However, for each $\varphi \in \mathbb{L}_1(x^{c-1})$ and $h \in \mathbb{L}_1(x^{c-1}) \cap \mathbb{L}_2(x^{2c-1})$ the *convolution theorem* states

$$(\circledast_\varphi h)_j = x^{c-1}\lambda_{>0}(\overline{x}_j(\circledast_\varphi h)) = x^{c-1}\lambda_{>0}(\overline{x}_j\varphi)\,x^{c-1}\lambda_{>0}(\overline{x}_j h)$$
$$= x^{c-1}\lambda_{>0}(\overline{x}_j\varphi)(\mathrm{M}_c h)_j = \varphi_j h_j \quad \text{for } \lambda\text{-a.e. } j \in \mathbb{R}.$$

or $(\circledast_\varphi h)_{\bullet} = \mathrm{M}_c(\circledast_\varphi h) = x^{c-1}\lambda_{>0}(\overline{x}_{\bullet}\varphi)(\mathrm{M}_c h) = \varphi_{\bullet}h_{\bullet}$ λ-a.s. in short. Consequently, $(\varphi_{\bullet}, \mathrm{M}_c, \mathrm{M}_c)$ is a singular value decomposition of $\circledast_\varphi$ with $\varphi_{\bullet} \in \mathbb{L}_\infty$, and thus $\circledast_\varphi \in \mathbb{L}^{\mathrm{M}_c,\mathrm{M}_c}(\mathbb{L}(\mathbb{L}_2)) = \mathrm{M}_c^\star(\mathbb{L}^{\mathbb{M}}(\mathbb{L}_2))\mathrm{M}_c$. □

§01|04.17 **Density multiplicative deconvolution on $\mathbb{R}_{>0}$.** Consider the *complex* Hilbert spaces $\mathbb{L}_2(x^{2c-1}) = \mathbb{L}_2(\mathbb{R}_{>0}, \mathscr{B}_{>0}, x^{2c-1}\lambda_{>0})$ and $\mathbb{L}_2 = \mathbb{L}_2(\lambda)$. Let $\mathbb{D}_2 \subseteq \mathbb{L}_1(x^{c-1}) \cap \mathbb{L}_2(x^{2c-1})$ be a set of Lebesgue-densities on $(\mathbb{R}_{>0}, \mathscr{B}_{>0})$ (by the usual embedding of real-valued functions) as in Notation §01|02.10. We denote for each Lebesgue density $\mathbb{p}$ on $(\mathbb{R}_{>0}, \mathscr{B}_{>0})$ by $\mathbb{P}_{\mathbb{p}} := \mathbb{p}\lambda_{>0} \in \mathscr{W}(\mathscr{B}_{>0})$ the associated probability measure. Given a Lebesgue density $\mathbb{q} \in \mathbb{L}_1(x^{c-1})$ presumed to be fixed and *known in advance* for each Lebesgue density $\mathbb{p} \in \mathbb{D}_2$ we consider the Lebesgue density $g = \circledast_{\mathbb{q}}\mathbb{p} = \mathbb{q} \circledast \mathbb{p} \in \mathbb{L}_1(x^{c-1}) \cap \mathbb{L}_2(x^{2c-1})$ (see Notation §01|04.16) and denote by $\mathbb{P}_{\mathbb{p}|\mathbb{q}} := (\mathbb{q} \circledast \mathbb{p})\lambda_{>0} = g\lambda_{>0} \in \mathscr{W}(\mathscr{B}_{>0})$ the associated probability measure. We consider the statistical product experiment $\big(\mathbb{R}^n_{>0}, \mathscr{B}^{\otimes n}_{>0}, \mathbb{P}^{\otimes n}_{\mathbb{D}_2\times\{\mathbb{q}\}} := (\mathbb{P}^{\otimes n}_{\mathbb{p}|\mathbb{q}})_{\mathbb{p}\in\mathbb{D}_2}\big)$. Let $\mathrm{M}_c \in \mathbb{L}(\mathbb{L}_2(x^{2c-1}), \mathbb{L}_2)$ be the Mellin transform (see Notation §01|02.17). Evidently, for $g \in \mathbb{L}_1(x^{c-1}) \cap \mathbb{L}_2(x^{2c-1})$ its Mellin transform $g_{\bullet} = (g_j)_{j\in\mathbb{R}} = \mathrm{M}_c g$ satisfies $g_j = x^{c-1}\lambda_{>0}(\overline{x}_j g) = \mathbb{P}_{\mathbb{p}|\mathbb{q}}(x^{c-1}\overline{x}_j)$ for all $j \in \mathbb{R}$. Moreover, considering the Mellin transform $\mathbb{p}_{\bullet} = (\mathbb{p}_j)_{j\in\mathbb{R}} = \mathrm{M}_c\mathbb{p}$ of $\mathbb{p} \in \mathbb{D}_2 \subseteq \mathbb{L}_1(x^{c-1}) \cap \mathbb{L}_2(x^{2c-1})$ by the *multiplicative convolution theorem* we have $g_{\bullet} = \mathrm{M}_c(\circledast_{\mathbb{q}}\mathbb{p}) = x^{c-1}\lambda_{>0}(\overline{x}_{\bullet}\mathbb{q})(\mathrm{M}_c\mathbb{p}) = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet}$ λ-a.s. with $\mathbb{q}_{\bullet} = x^{c-1}\lambda_{>0}(\overline{x}_{\bullet}\mathbb{q}) \in \mathbb{L}_\infty$ and $\mathbb{p}_{\bullet} = \mathrm{M}_c\mathbb{p} \in \mathbb{L}_2$ (see Notation §01|04.16). Moreover, the complex-valued stochastic process $x^{c-1}\overline{x}_{\bullet} = (x^{c-1}\overline{x}_j)_{j\in\mathbb{R}} \in \mathcal{M}(\mathscr{B}_{>0}\otimes\mathscr{B})$ on $(\mathbb{R}_{>0}, \mathscr{B}_{>0})$ is $\mathscr{B}_{>0} \otimes \mathscr{B}$-$\mathscr{B}$-measurable, i.e. $x^{c-1}\overline{x}_{\bullet} \in \mathcal{M}(\mathscr{B}_{>0}\otimes\mathscr{B})$ for short. We define

$$\widehat{g}_{\bullet} = (\widehat{g}_j := \widehat{\mathbb{P}}_n(x^{c-1}\overline{x}_j))_{j\in\mathbb{R}} = \widehat{\mathbb{P}}_n(x^{c-1}\overline{x}_{\bullet}) \in \mathcal{M}(\mathscr{B}^{\otimes n}_{>0}\otimes\mathscr{B})$$

similar to an Empirical mean model §01|02.04 where for each $j \in \mathbb{R}$

$$y = (y_i)_{i\in[\![n]\!]} \mapsto \widehat{g}_j(y) = (\widehat{\mathbb{P}}_n(x^{c-1}\overline{x}_j))(y) = n^{-1}\sum_{i\in[\![n]\!]}(x^{c-1}\overline{x}_j)(y_i) = n^{-1}\sum_{i\in[\![n]\!]} y_i^{c-1+\imath 2\pi j}.$$

By construction $g_{\bullet} = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet} = \mathbb{P}_{\mathbb{p}|\mathbb{q}}(x^{c-1}\overline{x}_{\bullet}) \in \mathbb{L}_2$ is the mean of $\widehat{g}_{\bullet}$. For each $j \in \mathbb{R}$ the statistic $\dot{\varepsilon}_j := n^{1/2}(\widehat{\mathbb{P}}_n(x^{c-1}\overline{x}_j) - \mathbb{P}_{\mathbb{p}|\mathbb{q}}(x^{c-1}\overline{x}_j)) \in \mathcal{M}(\mathscr{B}^{\otimes n}_{>0})$ is centred, i.e. $\dot{\varepsilon}_j \in \mathbb{L}_1(\mathbb{P}^{\otimes n}_{\mathbb{p}|\mathbb{q}})$ with $\mathbb{P}^{\otimes n}_{\mathbb{p}|\mathbb{q}}(\dot{\varepsilon}_j) = 0$, and exploiting $x^{c-1}\overline{x}_{\bullet} = (x^{c-1}\overline{x}_j)_{j\in\mathbb{R}} \in \mathcal{M}(\mathscr{B}_{>0}\otimes\mathscr{B})$ the complex valued stochastic process

$$\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathbb{R}} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_{\mathbb{p}|\mathbb{q}})(x^{c-1}\overline{x}_{\bullet}) = n^{1/2}(\widehat{\mathbb{P}}_n(x^{c-1}\overline{x}_{\bullet}) - \mathbb{P}_{\mathbb{p}|\mathbb{q}}(x^{c-1}\overline{x}_{\bullet})) \in \mathcal{M}(\mathscr{B}^{\otimes n}_{>0}\otimes\mathscr{B})$$

satisfies Assumption §01|02.11, and by construction $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ is a noisy version of $g_{\bullet} = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet}$.

§01|05 Non-diagonal statistical inverse problem

§01|05.01 **Notation.** Consider the measure space $(\mathcal{J}, \mathscr{J}, \nu)$ and the Hilbert space $\mathbb{J} = \mathbb{L}_2(\nu)$ as in Notation §01|01.01. For $\mathrm{T}_{\bullet|\bullet} \in \mathcal{M}(\mathscr{J}^2)$ denote for each $j, j_o \in \mathcal{J}$ by $\mathrm{T}_{\bullet|j_o} : \mathcal{J} \to \mathbb{R}$ and $\mathrm{T}_{j|\bullet} : \mathcal{J} \to \mathbb{R}$ the map $j \mapsto \mathrm{T}_{j|j_o}$ and $j_o \mapsto \mathrm{T}_{j|j_o}$, respectively. Then we have $\mathrm{T}_{\bullet|j_o}, \mathrm{T}_{j|\bullet} \in \mathcal{M}(\mathscr{J})$ for each $j, j_o \in \mathcal{J}$. If in addition $\mathrm{T}_{j|\bullet} \in \mathbb{J}$ for ν-a.e. $j \in \mathcal{J}$ then for each $a_{\bullet} \in \mathbb{J}$ it follows $(\mathrm{T}_{\bullet|\bullet}a_{\bullet})_j := \langle \mathrm{T}_{j|\bullet}, a_{\bullet}\rangle_{\mathbb{J}} = \nu(\mathrm{T}_{j|\bullet}a_{\bullet}) \in \mathbb{R}$ for ν-a.e. $j \in \mathcal{J}$ and thus $(\mathrm{T}_{\bullet|\bullet}a_{\bullet})_{\bullet} : j \mapsto (\mathrm{T}_{\bullet|\bullet}a_{\bullet})_j$ is ν-a.e. defined and $(\mathrm{T}_{\bullet|\bullet}a_\bullet)_\bullet \in \mathbb{M}(\mathscr{J})$. If for each $a_\bullet \in \mathbb{J}$ in addition $\|(\mathrm{T}_{\bullet|\bullet}a_\bullet)_\bullet\|_{\mathbb{J}}^2 = \nu((\mathrm{T}_{\bullet|\bullet}a_\bullet)_\bullet^2) \in \mathbb{R}_{\geqslant 0}$ and hence $(\mathrm{T}_{\bullet|\bullet}a_\bullet)_\bullet \in \mathbb{J}$, then setting $a_\bullet \mapsto \mathrm{T}a_\bullet := (\mathrm{T}_{\bullet|\bullet}a_\bullet)_\bullet$ defines an *integral operator* $\mathrm{T} : \mathbb{J} \to \mathbb{J}$ which we identify here and subsequently with its kernel $\mathrm{T}_{\bullet|\bullet} \in \mathbb{M}(\mathscr{J}^2)$. Evidently, the operator $\mathrm{T}_{\bullet|\bullet} : \mathbb{J} \to \mathbb{J}$ is bounded, i.e. $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}(\mathbb{J})$, if $\|\mathrm{T}_{\bullet|\bullet}\|_{\mathbb{L}(\mathbb{J})} = \sup\big\{\nu((\mathrm{T}_{\bullet|\bullet}a_\bullet)_\bullet^2) \colon a_\bullet \in \mathbb{J}, \|a_\bullet\|_{\mathbb{J}} \leqslant 1\big\} \in \mathbb{R}_{\geqslant 0}$. We set $\mathbb{L}_{\bullet}(\mathbb{J}) := \big\{\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}(\mathbb{J}) \colon \text{with kernel } \mathrm{T}_{\bullet|\bullet} \in \mathbb{M}(\mathscr{J}^2)\big\}$. Finally, given surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \mathbb{J})$ we define $\mathbb{L}^{\mathrm{U},\mathrm{V}}(\mathbb{L}_{\bullet}(\mathbb{J})) := \mathrm{V}^\star(\mathbb{L}_{\bullet}(\mathbb{J}))\mathrm{U} := \big\{\mathrm{V}^\star\mathrm{T}_{\bullet|\bullet}\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{G}) \colon \mathrm{T}_{\bullet|\bullet} \in \mathbb{L}_{\bullet}(\mathbb{J})\big\}$. As a consequence, for each $\mathrm{T} \in \mathbb{L}^{\mathrm{U},\mathrm{V}}(\mathbb{L}_{\bullet}(\mathbb{J}))$ we have $\mathrm{V}\mathrm{T}\mathrm{U}^\star = \mathrm{T}_{\bullet|\bullet} \in \mathbb{L}_{\bullet}(\mathbb{J})$ for some kernel $\mathrm{T}_{\bullet|\bullet} \in \mathbb{M}(\mathscr{J}^2)$. In the special case $(\mathcal{J}, \mathscr{J}, \nu) = (\mathbb{N}, 2^{\mathbb{N}}, \nu_{\mathbb{N}})$, where $\mathbb{R}^{\mathbb{N}^2} = \mathbb{M}(2^{\mathbb{N}^2})$ is the set of all infinite real-valued matrices, we have $\mathbb{L}_{\bullet}(\ell_2) = \mathbb{L}(\ell_2)$ (compare Notation §01|04.03). □

§01|05.02 **Assumption**. For $\mathbb{J} = \mathbb{L}_2(\nu)$, surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \mathbb{J})$, fixed and presumed to be known in advance, $\mathrm{T} \in \mathbb{L}^{\mathrm{U},\mathrm{V}}(\mathbb{L}_{\bullet}(\mathbb{J}))$ and hence $\mathrm{T}_{\bullet|\bullet} = \mathrm{V}\mathrm{T}\mathrm{U}^\star \in \mathbb{L}_{\bullet}(\mathbb{J})$ with kernel $\mathrm{T}_{\bullet|\bullet} \in \mathbb{M}(\mathscr{J}^2)$ is also known where $g_\bullet = \mathrm{T}_{\bullet|\bullet}\theta_\bullet \in \mathbb{J}$ or inequal $g_\bullet \in \mathrm{ran}(\mathrm{T}_{\bullet|\bullet}) = \big\{\mathrm{T}_{\bullet|\bullet}a_\bullet \colon a_\bullet \in \mathbb{J}\big\}$. □

§01|05.03 **Notation**. Under Assumption §01|05.02 given $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}_{\bullet}(\mathbb{J})$ and $g_\bullet \in \mathrm{ran}(\mathrm{T}_{\bullet|\bullet})$ we consider the reconstruction of $\theta_\bullet = \mathrm{U}\theta \in \mathbb{J}$ (or in equal $\theta = \mathrm{U}^\star\theta_\bullet \in \mathbb{H}$) from a noisy version of the image $g_\bullet = \mathrm{V}\mathrm{T}\mathrm{U}^\star\theta_\bullet = \mathrm{T}_{\bullet|\bullet}\theta_\bullet \in \mathbb{J}$. Keep in mind, that we identify the equivalence class and its representative $g_\bullet$. □

§01|05.04 **Non-diagonal statistical inverse problem**. Consider as in Definition §01|02.03 a stochastic process $\dot{\varepsilon}_\bullet = (\dot{\varepsilon}_j)_{j\in\mathcal{J}}$ satisfying Assumption §01|01.04 with mean zero and a sample size $n \in \mathbb{N}$. Under Assumption §01|05.02, where $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}_{\bullet}(\mathbb{J})$ with kernel $\mathrm{T}_{\bullet|\bullet} \in \mathbb{M}(\mathscr{J}^2)$ is *known in advance*, the observable noisy image has $\mathbb{J}$-mean $g_\bullet = \mathrm{T}_{\bullet|\bullet}\theta_\bullet$ and takes the form $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet = \mathrm{T}_{\bullet|\bullet}\theta_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ or in equal

$$\widehat{g}_j = g_j + n^{-1/2}\dot{\varepsilon}_j = \langle \mathrm{T}_{j|\bullet}, \theta_\bullet\rangle_{\mathbb{J}} + n^{-1/2}\dot{\varepsilon}_j, \quad \nu\text{-a.e. } j \in \mathcal{J}. \tag{01.10}$$

We denote by $\mathbb{P}^n_{\theta_\bullet|\mathrm{T}_{\bullet|\bullet}}$ the distribution of $\widehat{g}_\bullet$. If $\dot{\varepsilon}_\bullet$ admits (possibly depending on $g = \mathrm{T}\theta$) a covariance function, say $\mathrm{cov}^{\mathrm{T}\theta}_{\bullet,\bullet} \in \mathbb{M}(\mathscr{J}^2)$, or a covariance operator, say $\Gamma_{\mathrm{T}\theta} \in \mathbb{L}^{\geqslant}(\mathbb{J})$, then we eventually write $\dot{\varepsilon}_\bullet \sim \mathrm{P}_{(0_\bullet, \mathrm{cov}^{\mathrm{T}\theta}_{\bullet,\bullet})}$ and $\widehat{g}_\bullet \sim \mathrm{P}_{(g_\bullet, n^{-1}\mathrm{cov}^{\mathrm{T}\theta}_{\bullet,\bullet})}$ or $\dot{\varepsilon}_\bullet \sim \mathrm{P}_{(0_\bullet, \Gamma_{\mathrm{T}\theta})}$ and $\widehat{g}_\bullet \sim \mathrm{P}_{(g_\bullet, n^{-1}\Gamma_{\mathrm{T}\theta})}$ for short. The reconstruction of $\theta_\bullet \in \mathbb{J}$ (in equal $\theta = \mathrm{U}^\star\theta_\bullet \in \mathbb{H}$) from a noisy version $\widehat{g}_\bullet \sim \mathbb{P}^n_{g}$ of the image $g_\bullet = \mathrm{T}_{\bullet|\bullet}\theta_\bullet \in \mathbb{J}$ is called a *non-diagonal statistical inverse problem*. □

§01|05.05 **Non-diagonal inverse empirical mean model (nieMM)**. Consider the reconstruction of $\theta_\bullet \in \mathbb{J}$ (in equal $\theta = \mathrm{U}^\star\theta_\bullet \in \mathbb{H}$) in an Empirical mean model as in §01|02.04. Under Assumption §01|05.02, where $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}_{\bullet}(\mathbb{J})$ with kernel $\mathrm{T}_{\bullet|\bullet} \in \mathbb{M}(\mathscr{J}^2)$ is known in advance, the observable noisy image has $\mathbb{J}$-mean $\mathrm{V}g = g_\bullet = \mathrm{T}_{\bullet|\bullet}\theta_\bullet \in \mathbb{J}$ and takes the form of an Empirical mean model as in §01|02.04, that is $\widehat{g}_\bullet = \mathrm{T}_{\bullet|\bullet}\theta_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ or in equal (01.10) with error process

$$\dot{\varepsilon}_\bullet = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_g)(\psi_\bullet) = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_\bullet) - \mathbb{P}_g(\psi_\bullet)) \in \mathbb{M}(\mathscr{X}^{\otimes n} \otimes \mathscr{J})$$

satisfying Assumption §01|01.04. □

§01|05.06 **Comment**. In the special case $\mathbb{J} = \ell_2$ (compare Notation §01|04.03) a Diagonal statistical inverse problem §01|04.06 with multiplication operator, i.e. $\mathrm{T}_{\bullet|\bullet} = \mathrm{M}_{s_\bullet}$, is indeed the diagonal case of the statistical inverse problem given in Assumption §01|05.02. Moreover, introducing $\mathbb{1}_\bullet := (1)_{j\in\mathbb{N}}$ the multiplication operator $\mathrm{id}_{\bullet|\bullet} := \mathrm{M}_{\mathbb{1}_\bullet} \in \mathbb{L}^{\mathbb{M}}(\ell_2)$ with diagonal kernel $\mathrm{id}_{\bullet|\bullet} \in \mathbb{M}(2^{\mathbb{N}^2})$ equals the identity on ℓ_2, i.e. $\mathrm{id}_{\ell_2} = \mathrm{id}_{\bullet|\bullet}$. As a consequence for $\mathbb{J} = \ell_2$ a statistical direct problem as in Definition §01|03.03 is also a statistical inverse problem with known identity operator, i.e. $\mathrm{T}_{\bullet|\bullet} = \mathrm{id}_{\bullet|\bullet}$. □

§01|05.07 **Non-diagonal inverse sequence model (niSM).** Consider $\mathbb{J} = \ell_2 = \mathbb{L}_2(\nu_{\mathbb{N}})$ as in §01|01.14. Let $\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathbb{N}}$ be a sequence of real-valued random variables with mean zero and let $n \in \mathbb{N}$ be a sample size. Under Assumption §01|05.02, where $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}(\ell_2)$ with kernel $\mathrm{T}_{\bullet|\bullet} \in \mathcal{M}(2^{\mathbb{N}^2})$ is known in advance, the observable noisy image has ℓ_2-mean $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet}$ and takes the form of a Sequence model as in §01|02.05, that is $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ or in equal

$$\widehat{g}_j = g_j + n^{-1/2}\dot{\varepsilon}_j = \langle \mathrm{T}_{j|\bullet}, \theta_{\bullet}\rangle_{\ell_2} + n^{-1/2}\dot{\varepsilon}_j, \quad j \in \mathbb{N}. \tag{01.11}$$

We denote by $\mathbb{P}^n_{\theta|\mathrm{T}}$ the distribution of $\widehat{g}_{\bullet}$. □

§01|05.08 **Gaussian non-diagonal inverse sequence model (GniSM).** Let $\dot{\mathrm{B}}_{\bullet} := (\dot{\mathrm{B}}_j)_{j\in\mathbb{N}} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$ be a Gaussian white noise process. The observable noisy version $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\mathrm{B}}_{\bullet}$ with ℓ_2-mean $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet}$ takes the form of a Gaussian sequence model as in §01|02.06, that is

$$\widehat{g}_j = \langle \mathrm{T}_{j|\bullet}, \theta_{\bullet}\rangle_{\ell_2} + n^{-1/2}\dot{\mathrm{B}}_j,\ j \in \mathbb{N} \quad \text{with} \quad (\dot{\mathrm{B}}_j)_{j\in\mathbb{N}} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}. \tag{01.12}$$

We denote by $\mathrm{N}^n_{\theta|\mathrm{T}}$ the distribution of the stochastic process $\widehat{g}_{\bullet}$. □

§01|05|01 Examples of non-diagonal inverse empirical mean models

§01|05.09 **Non-diagonal inverse regression with uniform design.** Consider the measure space $([0,1], \mathscr{B}_{[0,1]}, \lambda_{[0,1]})$ and the *real* Hilbert space $\mathbb{L}_2(\lambda_{[0,1]})$ as in Model §01|02.09. Let $\mathrm{T} \in \mathbb{L}^{\mathrm{U,V}}(\mathbb{L}(\ell_2))$ and hence $\mathrm{T}_{\bullet|\bullet} = \mathrm{VTU}^{\star} \in \mathbb{L}(\ell_2)$ with kernel $\mathrm{T}_{\bullet|\bullet} \in \mathcal{M}(2^{\mathbb{N}^2})$ be *known in advance*. Let (X, Y) be a $[0,1] \times \mathbb{R}$-valued random vector. As in Model §01|02.09 we assume in what follows that the regressor X is uniformly distributed on the interval $[0,1]$, i.e. $X \sim \mathrm{U}_{[0,1]} = \lambda_{[0,1]} = \mathbb{P}^X$ and that given $\mathrm{T}\theta = g \in \mathbb{L}_2([0,1])$ for some $\theta \in \mathbb{H}$ the joint distribution of (X, Y) is given by $\mathrm{U}_{\theta|\mathrm{T}} := \mathrm{U}_{[0,1]} \odot \mathbb{P}^{Y|X}_{\mathrm{T}\theta}$ without fully specifying the regular conditional distribution $\mathbb{P}^{Y|X}_{\mathrm{T}\theta}$ which however satisfies $\mathbb{P}^{Y|X}_{\mathrm{T}\theta}(\mathrm{id}_{\mathbb{R}}) = \mathbb{P}_{\mathrm{T}\theta}(Y|X) = \mathrm{T}\theta = g \in \mathbb{L}_2([0,1])$. Keep in mind that we tactically identify X and Y with the coordinate map $\Pi_{[0,1]}$ and $\Pi_{\mathbb{R}}$, respectively, and thus (X, Y) with the identity $\mathrm{id}_{[0,1]\times\mathbb{R}}$. Consequently, if $Y \in \mathbb{L}_2(\mathrm{U}_{\theta|\mathrm{T}})$ and $h \in \mathbb{L}_2(\mathbb{P}^X) = \mathbb{L}_2(\lambda_{[0,1]})$, hence $h(X) \in \mathbb{L}_2(\mathrm{U}_{\theta|\mathrm{T}})$, then we obtain $Yh(X) \in \mathbb{L}_1(\mathrm{U}_{\theta|\mathrm{T}})$ and

$$\mathrm{U}_{\theta|\mathrm{T}}(Yh(X)) = \mathbb{P}^X(\mathbb{P}^{Y|X}_{\mathrm{T}\theta}(Y)h) = \mathbb{P}^X\big((\mathrm{T}\theta)h\big) = \lambda_{[0,1]}\big((\mathrm{T}\theta)h\big) = \langle \mathrm{T}\theta, h\rangle_{\mathbb{L}_2(\lambda_{[0,1]})} \in \mathbb{R}$$

identifying again equivalence classes and their representatives. We consider the statistical product experiment $\big(([0,1]\times\mathbb{R})^n, (\mathscr{B}_{[0,1]} \otimes \mathscr{B})^{\otimes n}, \mathrm{U}^{\otimes n}_{\Theta\times\{\mathrm{T}\}} := (\mathrm{U}^{\otimes n}_{\theta|\mathrm{T}})_{\theta\in\Theta}\big)$ of size $n \in \mathbb{N}$ and for $\theta \in \Theta$ we denote by $((X_i, Y_i))_{i\in[\![n]\!]} \sim \mathrm{U}^{\otimes n}_{\theta|\mathrm{T}}$ an iid. sample of $(X, Y) \sim \mathrm{U}_{\theta|\mathrm{T}} = \mathrm{U}_{[0,1]} \odot \mathbb{P}^{Y|X}_{\mathrm{T}\theta}$. Keep in mind that $\mathrm{V} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1]}), \ell_2)$ and $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \ell_2)$ are generalised Fourier series transform as in Notation §01|02.07 which are fixed and *known in advance*. Evidently, for each $\theta \in \Theta \subseteq \mathbb{H}$ the generalised Fourier coefficients $\theta_{\bullet} = (\theta_j)_{j\in\mathbb{N}} = \mathrm{U}\theta$ and $g_{\bullet} = (g_j)_{j\in\mathbb{N}} = \mathrm{V}g = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet}$ satisfy

$$g_j = \langle \mathrm{T}_{j|\bullet}, \theta_{\bullet}\rangle_{\ell_2} = \langle \mathrm{T}_{\bullet|\bullet}\theta_{\bullet}, \mathbb{1}^{\{j\}}_{\bullet}\rangle_{\ell_2} = \langle \mathrm{T}\theta, \mathrm{V}^{\star}\mathbb{1}^{\{j\}}_{\bullet}\rangle_{\mathbb{L}_2(\lambda_{[0,1]})} = \lambda_{[0,1]}\big((\mathrm{T}\theta)\mathrm{v}_j\big) = \mathrm{U}_{\theta|\mathrm{T}}(Y\mathrm{v}_j(X))$$

for each $j \in \mathbb{N}$. The stochastic process $\psi_{\bullet} = (\psi_j(X, Y) := Y\mathrm{v}_j(X))_{j\in\mathbb{N}} \in \mathcal{M}((\mathscr{B}_{[0,1]} \otimes \mathscr{B}) \otimes 2^{\mathbb{N}})$ fulfils Assumption §01|01.04 and $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} = \mathrm{U}_{\theta|\mathrm{T}}(\psi_{\bullet})$. Similar to an Empirical mean model §01|02.04 we define $\widehat{g}_{\bullet} = (\widehat{g}_j := \widehat{\mathbb{P}}_n(\psi_j))_{j\in\mathbb{N}} \in \mathcal{M}((\mathscr{B}_{[0,1]} \otimes \mathscr{B})^{\otimes n} \otimes 2^{\mathbb{N}})$. By construction $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} = \mathrm{U}_{\theta|\mathrm{T}}(\psi_{\bullet}) \in \mathcal{M}(2^{\mathbb{N}})$ is the ℓ_2-mean of $\widehat{g}_{\bullet}$. For each $j \in \mathbb{N}$ the statistic $\dot{\varepsilon}_j := n^{1/2}(\widehat{\mathbb{P}}_n(\psi_j) - \mathrm{U}_{\theta|\mathrm{T}}(\psi_j)) \in \mathcal{M}((\mathscr{B}_{[0,1]} \otimes \mathscr{B})^{\otimes n})$ is centred, i.e. $\dot{\varepsilon}_j \in \mathbb{L}_1(\mathrm{U}^{\otimes n}_{\theta|\mathrm{T}})$ with $\mathrm{U}^{\otimes n}_{\theta|\mathrm{T}}(\dot{\varepsilon}_j) = 0$, and exploiting $\psi_{\bullet} \in \mathcal{M}((\mathscr{B}_{[0,1]} \otimes \mathscr{B}) \otimes 2^{\mathbb{N}})$ the stochastic process

$$\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathbb{N}} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathrm{U}_{\theta|\mathrm{T}})(\psi_{\bullet}) = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_{\bullet}) - \mathrm{U}_{\theta|\mathrm{T}}(\psi_{\bullet})) \in \mathcal{M}((\mathscr{B}_{[0,1]} \otimes \mathscr{B})^{\otimes n} \otimes 2^{\mathbb{N}})$$

satisfies Assumption §01|01.04. Since $\widehat{g}_j = g_j + n^{-1/2}\dot{\varepsilon}_j = \langle \mathrm{T}_{j|\bullet}, \theta_{\bullet}\rangle_{\ell_2} + n^{-1/2}\dot{\varepsilon}_j$ for each $j \in \mathbb{N}$ by construction $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ is a noisy version of $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet}$. □

§01|05.10 **Regression with known design**. Consider the measure space $(\mathcal{D},\mathscr{B}_{\mathcal{D}},\lambda_{\mathcal{D}})$ where $\lambda_{\mathcal{D}}$ denotes the restriction of the Lebesgue measure to the Borel-σ-algebra $\mathscr{B}_{\mathcal{D}}$ over $\mathcal{D}\in\mathscr{B}$, and the *real* Hilbert space $\mathbb{L}_2(\lambda_{\mathcal{D}}) := \mathbb{L}_2(\mathcal{D},\mathscr{B}_{\mathcal{D}},\lambda_{\mathcal{D}})$ of square Lebesgue-integrable real-valued functions. Let (X,Y) be a $\mathcal{D}\times\mathbb{R}$-valued random vector. We assume in what follows that the marginal distribution $\mathbb{P}^X\in\mathscr{W}(\mathscr{B}_{\mathcal{D}})$ of the regressor X admits a Lebesgue density $\varphi\in\mathbb{L}_1(\lambda_{\mathcal{D}})$ presumed to be fixed and *known in advance*, that is $\mathbb{P}^X=\varphi\lambda_{\mathcal{D}}$. For a real random variable $\xi\sim\mathbb{P}^\xi\in\mathscr{W}(\mathscr{B})$ and $a\in\mathbb{R}$ we denote by $\mathbb{P}^\xi_a\in\mathscr{W}(\mathscr{B})$ the distribution of $a+\xi$. We assume that for each $B\in\mathscr{B}$ the map $\mathbb{P}^\xi_\cdot(B): a\mapsto \mathbb{P}^\xi_a(B)$ is $\mathscr{B}$-$\mathscr{B}_{[0,1]}$-*measurable*. Then $\mathbb{P}^\xi_\cdot:\mathbb{R}\times\mathscr{B}\to[0,1]$ with $(a,B)\mapsto\mathbb{P}^\xi_a(B)$ is a Markov kernel from $(\mathbb{R},\mathscr{B})$ to $(\mathbb{R},\mathscr{B})$. In this situation, for any $f\in\mathcal{M}(\mathscr{B}_{\mathcal{D}})$ the map $\mathbb{P}^\xi_{f(X)}:\mathcal{D}\times\mathscr{B}\to[0,1]$ with $(x,B)\mapsto\mathbb{P}^\xi_{f(x)}(B)$ is a Markov kernel from $(\mathcal{D},\mathscr{B}_{\mathcal{D}})$ to $(\mathbb{R},\mathscr{B})$. If ξ and X are *independent* and $Y=f(X)+\xi$ for some $f\in\mathcal{M}(\mathscr{B}_{\mathcal{D}})$, which is assumed throughtout this model, then $\mathbb{P}^\xi_{f(X)}$ is a regular version of the conditional distribution of Y given X, in symbols $\mathbb{P}^{Y|X}_f=\mathbb{P}^\xi_{f(X)}$. In other words there exists a $\mathbb{P}^X$-null set $\mathcal{N}\in\mathscr{B}_{\mathcal{D}}$ such that $\mathbb{P}^{Y|X=x}_f(B)=\mathbb{P}^\xi_{f(x)}(B)$ for all $B\in\mathscr{B}$ and $x\in\mathcal{N}^c$ (Witting [1985], Satz 129, p.130). In summary the joint distribution of (X,Y) is given by $\mathbb{P}^{X,Y}_{f|\varphi}:=\varphi\lambda_{\mathcal{D}}\odot\mathbb{P}^\xi_{f(X)}$ without fully specifying the error distribution $\mathbb{P}^\xi\in\mathscr{W}(\mathscr{B})$ and thus the regular conditional distribution $\mathbb{P}^{Y|X}_f=\mathbb{P}^\xi_{f(X)}$. (Since $\lambda_{\mathcal{D}}$ dominates $\varphi\lambda_{\mathcal{D}}$ each representative of $\{f\}_{\lambda_{\mathcal{D}}}$ induces the same joint distribution $\mathbb{P}^{X,Y}_{\{f\}_{\lambda_{\mathcal{D}}}|\varphi}=\mathbb{P}^{X,Y}_{f|\varphi}\in\mathscr{W}(\mathscr{B}_{\mathcal{D}}\otimes\mathscr{B})$.) We tactically identify X and Y with the coordinate map $\Pi_{\mathcal{D}}$ and $\Pi_{\mathbb{R}}$, respectively, and thus (X,Y) with the identity $\mathrm{id}_{\mathcal{D}\times\mathbb{R}}$ such that $\mathbb{P}_{f|\varphi}=\mathbb{P}^{X,Y}_{f|\varphi}\in\mathscr{W}(\mathscr{B}_{\mathcal{D}}\otimes\mathscr{B})$. Let in addition $\varphi\in\mathbb{L}_\infty(\lambda_{\mathcal{D}})$, then $\mathrm{M}_\varphi\in\mathbb{L}^{\mathrm{M}}(\mathbb{L}_2(\lambda_{\mathcal{D}}))$ with $h\mapsto\mathrm{M}_\varphi h:=\varphi h$. Note that then for each $h\in\mathbb{L}_2(\lambda_{\mathcal{D}})$ we have $\mathrm{M}_\varphi h\in\mathbb{L}_2(\lambda_{\mathcal{D}})$ and hence for each representative $h\in\mathcal{L}_2(\varphi\lambda_{\mathcal{D}})$. (Since $\lambda_{\mathcal{D}}$ dominates $\varphi\lambda_{\mathcal{D}}$ for each $h\in\overline{\mathcal{M}}(\mathscr{B}_{[0,1]})$ we have $\{h\}_{\lambda_{\mathcal{D}}}\subseteq\{h\}_{\varphi\lambda_{\mathcal{D}}}$. If $\lambda_{\mathcal{D}}$ and $\varphi\lambda_{\mathcal{D}}$ dominate mutually each other, i.e. they share the same null sets, then $\{h\}_{\varphi\lambda_{\mathcal{D}}}=\{h\}_{\lambda_{\mathcal{D}}}$ and hence $\mathbb{L}_2(\lambda_{\mathcal{D}})\subseteq\mathbb{L}_2(\varphi\lambda_{\mathcal{D}})$.) If in addition $\mathbb{P}^\xi\in\mathrm{P}_{\{0\}\times\mathbb{R}_{>0}}\subseteq\mathscr{W}_2(\mathscr{B})$, i.e. ξ has mean zero and a finite second moment, and $f\in\mathbb{L}_2(\lambda_{\mathcal{D}})$, then for each representative $f\in\mathcal{L}_2(\varphi\lambda_{\mathcal{D}})$, $f(X)\in\mathcal{L}_2(\mathbb{P}_{f|\varphi})$ and $Y\in\mathcal{L}_2(\mathbb{P}_{f|\varphi})$ too. In particular it follows $\mathbb{P}^{Y|X}_f(\mathrm{id}_{\mathbb{R}})=\mathbb{P}_f(Y|X)=\{f\}_{\varphi\lambda_{\mathcal{D}}}\in\mathbb{L}_2(\mathbb{P}^X)=\mathbb{L}_2(\varphi\lambda_{\mathcal{D}})$. Consequently, for each $h\in\mathbb{L}_2(\lambda_{\mathcal{D}})$, hence $h(X)\in\mathcal{L}_2(\mathbb{P}_{f|\varphi})$ we obtain $Yh(X)\in\mathcal{L}_1(\mathbb{P}_{f|\varphi})$ and

$$\mathbb{P}_{f|\varphi}(Yh(X))=\mathbb{P}^X(\mathbb{P}^{Y|X}_f(Y)h)=\varphi\lambda_{\mathcal{D}}(fh)=\lambda_{\mathcal{D}}(\varphi fh)=\langle\mathrm{M}_\varphi f,h\rangle_{\mathbb{L}_2(\lambda_{\mathcal{D}})}\in\mathbb{R}$$

identifying again equivalence classes and their representatives. We note that $\mathrm{M}_\varphi\in\mathbb{L}^{\mathrm{M}}(\mathbb{L}_2(\lambda_{\mathcal{D}}))$ with density $\varphi\in\mathbb{L}_\infty(\lambda_{\mathcal{D}})$ is positive semi-definite, i.e. $\mathrm{M}_\varphi\in\mathbb{L}^{\geq}(\mathbb{L}_2(\lambda_{\mathcal{D}}))$ and if in addition $\varphi\in\mathcal{M}_{>0,\lambda_{\mathcal{D}}}(\mathscr{B}_{\mathcal{D}})$ (i.e. $\varphi\in\mathcal{M}_{\geqslant0}(\mathscr{B}_{\mathcal{D}})$ and $\lambda_{\mathcal{D}}(\mathcal{N}_\varphi)=0$) then it is stictly positive definite, i.e. $\mathrm{M}_\varphi\in\mathbb{L}^{\gneq}(\mathbb{L}_2(\lambda_{\mathcal{D}}))$. Keep in mind that $\mathrm{U}\in\mathbb{L}(\mathbb{L}_2(\lambda_{\mathcal{D}}),\ell_2)$ is generalised Fourier series transform as in Notation §01|02.07 which is fixed and *known in advance*. Evidently, we have $\mathrm{M}^\varphi_{\bullet|\bullet}:=\mathrm{U}\mathrm{M}_\varphi\mathrm{U}^\star\in\mathbb{L}^{\gneq}_{\bullet}(\ell_2)\subseteq\mathbb{L}_{\bullet}(\ell_2)=\mathbb{L}(\ell_2)$ and for each $f\in\mathbb{F}_2\subseteq\mathbb{L}_2(\lambda_{\mathcal{D}})$ and $g:=\mathrm{M}_\varphi f\in\mathbb{L}_2(\lambda_{\mathcal{D}})$ the generalised Fourier coefficients $f_\bullet=(f_j)_{j\in\mathbb{N}}=\mathrm{U}f$ and $g_\bullet=(g_j)_{j\in\mathbb{N}}=\mathrm{U}g=\mathrm{M}^\varphi_{\bullet|\bullet}f_\bullet$ for each $j\in\mathbb{N}$ satisfy

$$g_j=\langle\mathrm{M}^\varphi_{j|\bullet},f_\bullet\rangle_{\ell_2}=\langle\mathrm{M}^\varphi_{\bullet|\bullet}f_\bullet,\mathbb{1}^{\{j\}}_\bullet\rangle_{\ell_2}=\langle\mathrm{M}_\varphi f,\mathrm{U}^\star\mathbb{1}^{\{j\}}_\bullet\rangle_{\mathbb{L}_2(\lambda_{\mathcal{D}})}=\langle\mathrm{M}_\varphi f,\mathrm{u}_j\rangle_{\mathbb{L}_2(\lambda_{\mathcal{D}})}=\mathbb{P}_{f|\varphi}(Y\mathrm{u}_j(X))\in\mathbb{R}$$

The stochastic process $\psi_\bullet=(\psi_j(X,Y):=Y\mathrm{u}_j(X))_{j\in\mathbb{N}}\in\mathcal{M}((\mathscr{B}_{\mathcal{D}}\otimes\mathscr{B})\otimes 2^{\mathbb{N}})$ fulfils Assumption §01|01.04 and $g_\bullet=\mathrm{M}^\varphi_{\bullet|\bullet}f_\bullet=\mathbb{P}_{f|\varphi}(\psi_\bullet)$. Similar to an Empirical mean model §01|02.04 we consider the statistical product experiment $((\mathcal{D}\times\mathbb{R})^n,(\mathscr{B}_{\mathcal{D}}\otimes\mathscr{B})^{\otimes n},\mathbb{P}^{\otimes n}_{\mathbb{F}_2\times\{\varphi\}}:=(\mathbb{P}^{\otimes n}_{f|\varphi})_{f\in\mathbb{F}_2})$ of size $n\in\mathbb{N}$ and for $f\in\mathbb{F}_2$ we denote by $((X_i,Y_i))_{i\in[\![n]\!]}\sim\mathbb{P}^{\otimes n}_{f|\varphi}$ an iid. sample of $(X,Y)\sim\mathbb{P}_{f|\varphi}=\varphi\lambda_{\mathcal{D}}\odot\mathbb{P}^{Y|X}_f$. We define $\widehat{g}_\bullet=(\widehat{g}_j:=\widehat{\mathbb{P}}_n(\psi_j))_{j\in\mathbb{N}}=\widehat{\mathbb{P}}_n(\psi_\bullet)\in\mathcal{M}((\mathscr{B}_{\mathcal{D}}\otimes\mathscr{B})^{\otimes n}\otimes2^{\mathbb{N}})$. By construction $g_\bullet=\mathrm{M}^\varphi_{\bullet|\bullet}f_\bullet=\mathbb{P}_{f|\varphi}(\psi_\bullet)\in\mathcal{M}(2^{\mathbb{N}})$ is the ℓ_2-mean of $\widehat{g}_\bullet$. For each $j\in\mathbb{N}$ the statistic $\dot{\varepsilon}_j:=n^{1/2}(\widehat{\mathbb{P}}_n(\psi_j)-\mathbb{P}_{f|\varphi}(\psi_j))\in\mathcal{M}((\mathscr{B}_{\mathcal{D}}\otimes\mathscr{B})^{\otimes n})$ is centred, i.e. $\dot{\varepsilon}_j\in\mathbb{L}_1(\mathbb{P}^{\otimes n}_{f|\varphi})$ with $\mathbb{P}^{\otimes n}_{f|\varphi}(\dot{\varepsilon}_j)=0$, and exploiting $\psi_\bullet\in\mathcal{M}((\mathscr{B}_{\mathcal{D}}\otimes\mathscr{B})\otimes2^{\mathbb{N}})$ the stochastic process

$$\dot{\varepsilon}_\bullet=(\dot{\varepsilon}_j)_{j\in\mathbb{N}}=n^{1/2}(\widehat{\mathbb{P}}_n-\mathbb{P}_{f|\varphi})(\psi_\bullet)=n^{1/2}(\widehat{\mathbb{P}}_n(\psi_\bullet)-\mathbb{P}_{f|\varphi}(\psi_\bullet))\in\mathcal{M}((\mathscr{B}_{\mathcal{D}}\otimes\mathscr{B})^{\otimes n}\otimes2^{\mathbb{N}})$$

satisfies Assumption §01|01.04. Since $\widehat{g}_j=g_j+n^{-1/2}\dot{\varepsilon}_j=\langle\mathrm{M}^\varphi_{j|\bullet},f_\bullet\rangle_{\ell_2}+n^{-1/2}\dot{\varepsilon}_j$ for each $j\in\mathbb{N}$ by construction $\widehat{g}_\bullet=g_\bullet+n^{-1/2}\dot{\varepsilon}_\bullet=\mathrm{M}^\varphi_{\bullet|\bullet}f_\bullet+n^{-1/2}\dot{\varepsilon}_\bullet$ is a noisy version of $g_\bullet=\mathrm{M}^\varphi_{\bullet|\bullet}f_\bullet$. □

§02 Noisy image and noisy operator

§02|00.01 **Assumption**. The Hilbert space $\mathbb{J} = \mathbb{L}_2(\mathcal{J}, \mathscr{J}, \nu)$ with σ-finite measure $\nu \in \mathscr{M}_\sigma(\mathscr{J})$, σ-algebra $\mathscr{J}$ over $\mathcal{J}$ containing all elementary events $\{j\}$, $j \in \mathcal{J}$, and the surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \mathbb{J})$, i.e. $\mathrm{U}\mathrm{U}^\star = \mathrm{id}_{\mathbb{J}} = \mathrm{V}\mathrm{V}^\star$, are fixed and presumed to be known in advance. □

§02|01 Noisy non-diagonal operator

§02|01.01 **Notation**. Under Assumption §02|00.01 we consider the reconstruction of $\theta_{\bullet} = \mathrm{U}\theta \in \mathbb{J}$ (or in equal $\theta = \mathrm{U}^\star\theta_{\bullet} \in \mathbb{H}$) from noisy versions of $\mathrm{V}g = g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathbb{J}$ and $\mathrm{T}_{\bullet|\bullet} = \mathrm{V}\mathrm{T}\mathrm{U}^\star \in \mathbb{L}_{\bullet}(\mathbb{J})$. □

§02|01.02 **Assumption**. The real-valued stochastic process $Y_{\bullet|\bullet} = (Y_{j|j_\circ})_{j,j_\circ \in \mathcal{J}}$ on a common measurable space $(\Omega, \mathscr{A})$ as a function $\Omega \times \mathcal{J}^2 \to \mathbb{R}$ with $(\omega, j, j_\circ) \mapsto Y_{j|j_\circ}(\omega)$ is $\mathscr{A} \otimes \mathscr{J}^2$-$\mathscr{B}$-measurable, $Y_{\bullet|\bullet} \in \mathcal{M}(\mathscr{A} \otimes \mathscr{J}^2)$ for short. □

§02|01.03 **Noisy non-diagonal operator**. Let $\dot{\boldsymbol{\eta}}_{\bullet|\bullet} = (\dot{\boldsymbol{\eta}}_{j|j_\circ})_{j,j_\circ \in \mathcal{J}}$ be a stochastic process satisfying Assumption §02|01.02 with mean zero and let $k \in \mathbb{N}$ be a sample size. The stochastic process $\widehat{\mathrm{T}}_{\bullet|\bullet} = \mathrm{T}_{\bullet|\bullet} + k^{-1/2}\dot{\boldsymbol{\eta}}_{\bullet|\bullet}$ with mean kernel $\mathrm{T}_{\bullet|\bullet} \in \mathcal{M}(\mathscr{J}^2)$ is called a *noisy version* of the non-diagonal operator $\mathrm{T}_{\bullet|\bullet} = \mathrm{V}\mathrm{T}\mathrm{U}^\star \in \mathbb{L}_{\bullet}(\mathbb{J})$, or *noisy non-diagonal operator* for short. We denote by $\mathbb{P}^k_{\mathrm{T}}$ the distribution of $\widehat{\mathrm{T}}_{\bullet|\bullet}$. If $\dot{\boldsymbol{\eta}}_{\bullet|\bullet}$ admits a covariance function (possibly depending on T), say $\mathrm{cov} \in \mathcal{M}(\mathscr{J}^4)$, then we eventually write $\dot{\boldsymbol{\eta}}_{\bullet|\bullet} \sim \mathrm{P}_{(0_{\bullet|\bullet}, \mathrm{cov})}$ and $\widehat{\mathrm{T}}_{\bullet|\bullet} \sim \mathrm{P}_{(\mathrm{T}_{\bullet|\bullet}, k^{-1}\mathrm{cov})}$ for short. □

§02|01.04 **Empirical mean model**. For each $\mathrm{T} \in \mathbb{T} \subseteq \mathbb{L}^{\mathrm{U},\mathrm{V}}(\mathbb{L}_{\bullet}(\mathbb{J}))$ let $\mathbb{P}_{\mathrm{T}} \in \mathscr{W}(\mathscr{Z})$ be a probability measure on a measurable space $(\mathcal{Z}, \mathscr{Z})$. Similar to an Empirical mean function §01|01.10 consider a stochastic process $\psi_{\bullet|\bullet} = (\psi_{j|j_\circ})_{j,j_\circ \in \mathcal{J}} \in \mathcal{M}(\mathscr{Z} \otimes \mathscr{J}^2)$ which in addition for all $\mathrm{T} \in \mathbb{T}$ with $\mathrm{T}_{\bullet|\bullet} = \mathrm{V}\mathrm{T}\mathrm{U}^\star \in \mathbb{L}_{\bullet}(\mathbb{J})$ and kernel $\mathrm{T}_{\bullet|\bullet} \in \mathcal{M}(\mathscr{J}^2)$ satisfies $\psi_{j|j_\circ} \in \mathcal{L}_1(\mathbb{P}_{\mathrm{T}}) := \mathcal{L}_1(\mathcal{Z}, \mathscr{Z}, \mathbb{P}_{\mathrm{T}})$ for each $j, j_\circ \in \mathcal{J}$ and $\mathbb{P}_{\mathrm{T}}(\psi_{\bullet|\bullet}) = (\mathrm{T}_{j|j_\circ} = \mathbb{P}_{\mathrm{T}}(\psi_{j|j_\circ}))_{j,j_\circ \in \mathcal{J}} = \mathrm{T}_{\bullet|\bullet}$. Considering a statistical product experiment $(\mathcal{Z}^k, \mathscr{Z}^{\otimes k}, \mathbb{P}^{\otimes k}_{\mathbb{T}} = (\mathbb{P}^{\otimes k}_{\mathrm{T}})_{\mathrm{T} \in \mathbb{T}})$ similar to an Empirical mean function §01|01.10 we define $\widehat{\mathrm{T}}_{\bullet|\bullet} = (\widehat{\mathrm{T}}_{j|j_\circ} := \widehat{\mathbb{P}}_k(\psi_{j|j_\circ}))_{j,j_\circ \in \mathcal{J}} = \widehat{\mathbb{P}}_k(\psi_{\bullet|\bullet}) \in \mathcal{M}(\mathscr{Z}^{\otimes k} \otimes \mathscr{J}^2)$. For $\mathrm{T} \in \mathbb{T}$ assuming a $\mathbb{P}^{\otimes k}_{\mathrm{T}}$-sample the mean kernel of $\widehat{\mathrm{T}}_{\bullet|\bullet}$ is by construction $\mathbb{P}_{\mathrm{T}}(\psi_{\bullet|\bullet}) = \mathrm{T}_{\bullet|\bullet} = \mathrm{V}\mathrm{T}\mathrm{U}^\star \in \mathbb{L}_{\bullet}(\mathbb{J})$. Moreover for each $j, j_\circ \in \mathcal{J}$ the statistic $\dot{\boldsymbol{\eta}}_{j|j_\circ} := k^{1/2}(\widehat{\mathbb{P}}_k(\psi_{j|j_\circ}) - \mathbb{P}_{\mathrm{T}}(\psi_{j|j_\circ})) \in \mathcal{M}(\mathscr{Z}^{\otimes k})$ is centred, i.e. $\dot{\boldsymbol{\eta}}_{j|j_\circ} \in \mathbb{L}_1(\mathbb{P}^{\otimes k}_{\mathrm{T}}) = \mathbb{L}_1(\mathcal{Z}^k, \mathscr{Z}^{\otimes k}, \mathbb{P}^{\otimes k}_{\mathrm{T}})$ with $\mathbb{P}^{\otimes k}_{\mathrm{T}}(\dot{\boldsymbol{\eta}}_{j|j_\circ}) = 0$, and exploiting $\psi_{\bullet|\bullet} \in \mathcal{M}(\mathscr{Z} \otimes \mathscr{J}^2)$ the stochastic process

$$\dot{\boldsymbol{\eta}}_{\bullet|\bullet} = (\dot{\boldsymbol{\eta}}_{j|j_\circ})_{j,j_\circ \in \mathcal{J}} = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_{\mathrm{T}})(\psi_{\bullet|\bullet}) = k^{1/2}(\widehat{\mathbb{P}}_k(\psi_{\bullet|\bullet}) - \mathbb{P}_{\mathrm{T}}(\psi_{\bullet|\bullet})) \in \mathcal{M}(\mathscr{Z}^{\otimes k} \otimes \mathscr{J}^2)$$

satisfies Assumption §01|01.04. Since $\widehat{\mathrm{T}}_{j|j_\circ} = \mathrm{T}_{j|j_\circ} + k^{-1/2}\dot{\boldsymbol{\eta}}_{j|j_\circ}$ for each $j, j_\circ \in \mathcal{J}$ the stochastic process $\widehat{\mathrm{T}}_{\bullet|\bullet} = \mathrm{T}_{\bullet|\bullet} + k^{-1/2}\dot{\boldsymbol{\eta}}_{\bullet|\bullet}$ is a noisy version of the operator $\mathrm{T}_{\bullet|\bullet} = \mathrm{V}\mathrm{T}\mathrm{U}^\star \in \mathbb{L}_{\bullet}(\mathbb{J})$. □

§02|01.05 **Bivariate sequence model**. Consider the measure space $(\mathcal{J}, \mathscr{J}, \nu) = (\mathbb{N}, 2^{\mathbb{N}}, \nu_{\mathbb{N}})$ as in §01|01.14. Let $\dot{\boldsymbol{\eta}}_{\bullet|\bullet} = (\dot{\boldsymbol{\eta}}_{j|j_\circ})_{j,j_\circ \in \mathbb{N}}$ be a real-valued stochastic process satisfying Assumption §01|01.04 with mean zero and let $k \in \mathbb{N}$ be a sample size. The observable noisy version $\widehat{\mathrm{T}}_{\bullet|\bullet} = \mathrm{T}_{\bullet|\bullet} + k^{-1/2}\dot{\boldsymbol{\eta}}_{\bullet|\bullet} \sim \mathbb{P}^k_{\mathrm{T}}$ with mean kernel $\mathrm{T}_{\bullet|\bullet} \in \mathcal{M}(2^{\mathbb{N}^2}) = \mathbb{R}^{\mathbb{N}^2}$ takes the form of a *bivariate sequence model*

$$\widehat{\mathrm{T}}_{j|j_\circ} = \mathrm{T}_{j|j_\circ} + k^{-1/2}\dot{\boldsymbol{\eta}}_{j|j_\circ}, \quad j, j_\circ \in \mathbb{N}. \tag{02.01}$$

If $\dot{\boldsymbol{\eta}}_{\bullet|\bullet}$ admits a covariance function (possibly depending on $\mathrm{T}_{\bullet|\bullet}$), say $\mathrm{cov} \in \mathcal{M}(2^{\mathbb{N}^4})$, then we eventually write $\widehat{\mathrm{T}}_{\bullet|\bullet} \sim \mathrm{P}_{(\mathrm{T}_{\bullet|\bullet}, k^{-1}\mathrm{cov})}$ for short. □

§02|01.06 **Gaussian bivariate sequence model**. Let $\dot{\mathrm{W}}_{\bullet|\bullet} := (\dot{\mathrm{W}}_{j|j_\circ})_{j,j_\circ\in\mathbb{N}} \sim \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}^2}$ be a Gaussian white noise process. The observable noisy version $\widehat{\mathrm{T}}_{\bullet|\bullet} = \mathrm{T}_{\bullet|\bullet} + k^{-1/2}\dot{\mathrm{W}}_{\bullet|\bullet}$ with mean kernel $\mathrm{T}_{\bullet|\bullet} \in \mathcal{M}(2^{\mathbb{N}^2})$ takes the form of a *Gaussian bivariate sequence model*

$$\widehat{\mathrm{T}}_{j|j_\circ} = \mathrm{T}_{j|j_\circ} + k^{-1/2}\dot{\mathrm{W}}_{j|j_\circ},\ j, j_\circ \in \mathbb{N} \quad \text{with} \quad (\dot{\mathrm{W}}_{j|j_\circ})_{j,j_\circ\in\mathbb{N}} \sim \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}^2} \tag{02.02}$$

and we denote by $\mathrm{N}_{\mathrm{T}}^{k}$ the distribution of the stochastic process $\widehat{\mathrm{T}}_{\bullet|\bullet}$. □

§02|01|01 Examples of empirical mean models

§02|01.07 **Conditional expectation operator**. Consider the Borel-measurable spaces $(\mathcal{X}, \mathscr{B}_x)$ and $(\mathcal{Z}, \mathscr{B}_z)$ for $\mathcal{X}, \mathcal{Z} \in \mathscr{B}$. Let (Z, X) be a $\mathcal{Z}\times\mathcal{X}$-valued random vector. We denote by $\mathbb{P}^Z \in \mathscr{W}(\mathscr{B}_z)$ and $\mathbb{P}^X \in \mathscr{W}(\mathscr{B}_x)$ the marginal distribution of Z and X, respectively, by $\mathbb{P}^{X|Z}$ a regular conditional distribution of X given Z, and by $\mathbb{P}^{Z,X} = \mathbb{P}^Z \odot \mathbb{P}^{X|Z} \in \mathscr{W}(\mathscr{B}_z\otimes\mathscr{B}_x)$ the joint distribution of (Z, X). We tactically identify Z and X with the coordinate map Π_z and Π_x, respectively, and thus (Z, X) with the identity $\mathrm{id}_{\mathcal{Z}\times\mathcal{X}}$ such that $\mathbb{P} = \mathbb{P}^{Z,X} \in \mathscr{W}(\mathscr{B}_z\otimes\mathscr{B}_x)$. Introduce further the Hilbert spaces $\mathbb{L}_2(\mathbb{P}^X) := \mathbb{L}_2(\mathcal{X}, \mathscr{B}_x, \mathbb{P}^X) =: \mathbb{H}$, $\mathbb{L}_2(\mathbb{P}^Z) := \mathbb{L}_2(\mathcal{Z}, \mathscr{B}_z, \mathbb{P}^Z) =: \mathbb{G}$ and $\mathbb{L}_2(\mathbb{P}^{Z,X}) := \mathbb{L}_2(\mathcal{Z}\times\mathcal{X}, \mathscr{B}_z\otimes\mathscr{B}_x, \mathbb{P}^{Z,X})$. For each $h \in \mathbb{H} = \mathbb{L}_2(\mathbb{P}^X)$, and hence $h(X) \in \mathbb{L}_2(\mathbb{P}^{Z,X})$ we have $\mathbb{P}^{X|Z}h := \mathbb{P}^{X|Z}(h) = \mathbb{P}\big(h(X)\big|Z\big) \in \mathbb{L}_2(\mathbb{P}^Z) = \mathbb{G}$. We call $\mathbb{P}^{X|Z} : \mathbb{H} \to \mathbb{G}$ with $h \mapsto \mathbb{P}^{X|Z}h$ *conditional expectation operator*. Since by exploiting Jensens inequality for each $h \in \mathbb{H} = \mathbb{L}_2(\mathbb{P}^X)$ we have

$$\|\mathbb{P}^{X|Z}h\|_{\mathbb{G}}^2 = \mathbb{P}^Z\big(|\mathbb{P}^{X|Z}(h)|^2\big) = \mathbb{P}^Z\big(|\mathbb{P}\big(h(X)\big|Z\big)|^2\big) \leqslant \mathbb{P}^Z\big(\mathbb{P}\big(h^2(X)\big|Z\big)\big) = \mathbb{P}^X(h^2) = \|h\|_{\mathbb{H}}^2$$

it follows $\mathbb{P}^{X|Z} \in \mathbb{L}(\mathbb{H}, \mathbb{G})$ with $\|\mathbb{P}^{X|Z}\|_{\mathbb{L}(\mathbb{H},\mathbb{G})} \leqslant 1$. Its adjoint $(\mathbb{P}^{X|Z})^\star \in \mathbb{L}(\mathbb{G}, \mathbb{H})$ satisfies $(\mathbb{P}^{X|Z})^\star = \mathbb{P}^{Z|X}$. Moreover, for each $h \in \mathbb{H} = \mathbb{L}_2(\mathbb{P}^X)$ and $g \in \mathbb{G} = \mathbb{L}_2(\mathbb{P}^Z)$, hence $h(X), g(Z) \in \mathbb{L}_2(\mathbb{P}^{Z,X})$, we have

$$\langle\mathbb{P}^{X|Z}h, g\rangle_{\mathbb{G}} = \mathbb{P}^{Z,X}\big(g(Z)\mathbb{P}\big(h(X)\big|Z\big)\big) = \mathbb{P}^{Z,X}\big(g(Z)h(X)\big) = \langle h, \mathbb{P}^{Z|X}g\rangle_{\mathbb{H}}.$$

Evidently, the conditional expectation operator $\mathbb{P}^{X|Z}$ determines fully (and vice versa) the regular conditional distribution $\mathbb{P}^{X|Z}$ of X given Z. However, in general the marginal distributions $\mathbb{P}^X$ and $\mathbb{P}^Z$, and hence the Hilbert spaces $\mathbb{H} = \mathbb{L}_2(\mathbb{P}^X)$ and $\mathbb{G} = \mathbb{L}_2(\mathbb{P}^Z)$ are not known in advance. We assume in what follows that $\mathcal{X} = \mathcal{Z} = [0, 1]$ and that X and Z is uniformly distributed on the interval $[0, 1]$, i.e. $X \sim \mathrm{U}_{[0,1]} = \lambda_{[0,1]} = \mathbb{P}^X$ and $Z \sim \mathrm{U}_{[0,1]} = \lambda_{[0,1]} = \mathbb{P}^Z$. We denote by $\mathrm{U}_{\mathbb{P}^{X|Z}} := \mathrm{U}_{[0,1]} \odot \mathbb{P}^{X|Z}$ the joint distribution of (Z, X) which is now fully specified once the conditional expectation operator $\mathbb{P}^{X|Z} \in \mathbb{T} \subseteq \mathbb{L}(\mathbb{H}, \mathbb{G})$ is known. We consider the statistical product experiment $\big([0, 1]^{2k}, \mathscr{B}_{[0,1]}^{\otimes 2k}, \mathrm{U}_{\mathrm{T}}^{\otimes k} := (\mathrm{U}_{\mathbb{P}^{X|Z}}^{\otimes k})_{\mathbb{P}^{X|Z}\in\mathbb{T}}\big)$ of size $k \in \mathbb{N}$ and for $\mathbb{P}^{X|Z} \in \mathbb{T}$ we denote by $((Z_i, X_i))_{i\in[\![k]\!]} \sim \mathrm{U}_{\mathbb{P}^{X|Z}}^{\otimes k}$ an iid. sample of $(Z, X) \sim \mathrm{U}_{\mathbb{P}^{X|Z}} = \mathrm{U}_{[0,1]} \odot \mathbb{P}^{X|Z}$. Let $\mathrm{U}, \mathrm{V} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1]}), \ell_2)$ be generalised Fourier series transforms as in Notation §01|02.07 which are fixed and *known* in advanced. Then $\mathbb{P}_{\bullet|\bullet}^{X|Z} := \mathrm{V}\mathbb{P}^{X|Z}\mathrm{U}^\star \in \mathbb{L}_\bullet(\ell_2)$ is an operator with kernel (infinite matrix) $\mathbb{P}_{\bullet|\bullet}^{X|Z} \in \mathcal{M}(2^{\mathbb{N}^2})$ satisfying $\mathbb{P}_{\bullet|\bullet}^{X|Z} = (\mathbb{P}_{j|j_\circ}^{X|Z} = \langle\mathbb{P}^{X|Z}\mathrm{u}_{j_\circ}, \mathrm{v}_j\rangle_{\mathbb{G}} = \mathrm{U}_{\mathbb{P}^{X|Z}}\big(\mathrm{u}_{j_\circ}(X)\mathrm{v}_j(Z)\big))_{j,j_\circ\in\mathbb{N}}$. Therefore the stochastic process $\psi_{\bullet|\bullet} = (\psi_{j|j_\circ}(Z, X) := \mathrm{u}_{j_\circ}(X)\mathrm{v}_j(Z))_{j,j_\circ\in\mathbb{N}} \in \mathcal{M}(\mathscr{B}_{[0,1]}^2 \otimes 2^{\mathbb{N}^2})$ fulfils Assumption §02|01.02 and $\mathbb{P}_{\bullet|\bullet}^{X|Z} = \mathrm{U}_{\mathbb{P}^{X|Z}}(\psi_{\bullet|\bullet})$. Similar to an Empirical mean model §02|01.04 we define $\widehat{\mathbb{P}}_{\bullet|\bullet}^{X|Z} = (\widehat{\mathbb{P}}_{j|j_\circ}^{X|Z} := \widehat{\mathbb{P}}_n(\psi_{j|j_\circ}))_{j,j_\circ\in\mathbb{N}} \in \mathcal{M}(\mathscr{B}_{[0,1]}^{\otimes 2k} \otimes 2^{\mathbb{N}^2})$. By construction $\mathbb{P}_{\bullet|\bullet}^{X|Z} = \mathrm{U}_{\mathbb{P}^{X|Z}}(\psi_{\bullet|\bullet}) \in \mathcal{M}(2^{\mathbb{N}^2})$ is the mean kernel of $\widehat{\mathbb{P}}_{\bullet|\bullet}^{X|Z}$. For each $j, j_\circ \in \mathbb{N}$ the statistic $\dot{\boldsymbol{\eta}}_{j|j_\circ} := k^{1/2}(\widehat{\mathbb{P}}_k(\psi_{j|j_\circ}) - \mathrm{U}_{\mathbb{P}^{X|Z}}(\psi_{j|j_\circ})) \in \mathcal{M}(\mathscr{B}_{[0,1]}^{\otimes 2k})$ is centred, i.e. $\dot{\boldsymbol{\eta}}_{j|j_\circ} \in \mathbb{L}_1(\mathrm{U}_{\mathbb{P}^{X|Z}}^{\otimes k})$ with $\mathrm{U}_{\mathbb{P}^{X|Z}}^{\otimes k}(\dot{\boldsymbol{\eta}}_{j|j_\circ}) = 0$, and exploiting $\psi_{\bullet|\bullet} \in \mathcal{M}(\mathscr{B}_{[0,1]}^2 \otimes 2^{\mathbb{N}^2})$ the stochastic process

$$\dot{\boldsymbol{\eta}}_{\bullet|\bullet} = (\dot{\boldsymbol{\eta}}_{j|j_\circ})_{j,j_\circ\in\mathbb{N}} = k^{1/2}(\widehat{\mathbb{P}}_k - \mathrm{U}_{\mathbb{P}^{X|Z}})(\psi_{\bullet|\bullet}) = k^{1/2}(\widehat{\mathbb{P}}_k(\psi_{\bullet|\bullet}) - \mathrm{U}_{\mathbb{P}^{X|Z}}(\psi_{\bullet|\bullet})) \in \mathcal{M}(\mathscr{B}_{[0,1]}^{\otimes 2k} \otimes 2^{\mathbb{N}^2})$$

satisfies Assumption §02|01.02. Since $\widehat{\mathbb{P}}^{X|Z}_{j|j_\circ} = \mathbb{P}^{X|Z}_{j|j_\circ} + k^{-1/2}\dot{\boldsymbol{\eta}}_{j|j_\circ}$ for each $j, j_\circ \in \mathbb{N}$ by construction $\widehat{\mathbb{P}}^{X|Z}_{\bullet|\bullet} = \mathbb{P}^{X|Z}_{\bullet|\bullet} + k^{-1/2}\dot{\boldsymbol{\eta}}_{\bullet|\bullet}$ is a noisy version of $\mathbb{P}^{X|Z}_{\bullet|\bullet}$. □

§02|01.08 **Covariance operator**. Let $(\mathbb{H}, \langle\cdot,\cdot\rangle_{\mathbb{H}})$ be a separable Hilbert space equipped with its Borel-σ-algebra $\mathscr{B}_{\mathbb{H}}$ and X be an $\mathbb{H}$-valued *random function*. We tactically identify X with the identity $\mathrm{id}_{\mathbb{H}}$ on $\mathbb{H}$ such that X is defined on the measure space $(\mathbb{H}, \mathscr{B}_{\mathbb{H}}, \mathbb{P})$ and $X \sim \mathbb{P} = \mathbb{P}^X \in \mathscr{W}(\mathscr{B}_{\mathbb{H}})$. Here and subsequently, we assume that $\|X\|^2_{\mathbb{H}} \in \mathbb{L}_2(\mathbb{P})$ and $\mathbb{P}(\langle x, X\rangle_{\mathbb{H}}) = 0$ for all $x \in \mathbb{H}$. In this situation X admits a *covariance operator* $\Gamma^X \in \mathbb{L}^{\geqslant}(\mathbb{H})$ (see Remark §01|01.07). Let us denote by $\mathbb{P}_{\Gamma^X} \in \mathscr{W}(\mathscr{B}_{\mathbb{H}})$ the destribution of X which is not fully specified given $\Gamma^X \in \mathbb{T} \subseteq \mathbb{L}^{\geqslant}(\mathbb{H})$. We consider the statistical product experiment $(\mathbb{H}^k, \mathscr{B}^{\otimes k}_{\mathbb{H}}, \mathbb{P}^{\otimes k}_{\mathbb{T}} = (\mathbb{P}^{\otimes k}_{\Gamma^X})_{\Gamma^X\in\mathbb{T}})$. Let $(\mathrm{u}_j)_{j\in\mathbb{N}}$ be an *orthonormal system* in $\mathbb{H}$ and denote by $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \ell_2)$ its associated generalised Fourier series transform (see Notation §01|02.07). Then $\Gamma^X_{\bullet|\bullet} = \mathrm{U}\Gamma^X\mathrm{U}^\star \in \mathbb{L}^{\geqslant}(\ell_2)$ is a positive semi-definite operator with kernel (infinite matrix) $\Gamma^X_{\bullet|\bullet} \in \mathcal{M}(2^{\mathbb{N}^2})$ which satisfies $\Gamma^X_{\bullet|\bullet} = (\Gamma^X_{j|j_\circ} = \mathbb{P}_{\Gamma^X}(\langle \mathrm{u}_j, X\rangle_{\mathbb{H}}\langle X, \mathrm{u}_{j_\circ}\rangle_{\mathbb{H}}))_{j,j_\circ\in\mathbb{N}}$. The process $\psi_{\bullet|\bullet} = (\psi_{j|j_\circ}(X) := \langle \mathrm{u}_j, X\rangle_{\mathbb{H}}\langle X, \mathrm{u}_{j_\circ}\rangle_{\mathbb{H}})_{j,j_\circ\in\mathbb{N}} \in \mathcal{M}(\mathscr{B}_{\mathbb{H}} \otimes 2^{\mathbb{N}^2})$ fulfils Assumption §02|01.02 and $\Gamma^X_{\bullet|\bullet} = \mathbb{P}_{\Gamma^X}(\psi_{\bullet|\bullet})$. Similar to an Empirical mean model §02|01.04 we define $\widehat{\Gamma}^X_{\bullet|\bullet} = (\widehat{\Gamma}^X_{j|j_\circ} := \widehat{\mathbb{P}}_k(\psi_{j|j_\circ}))_{j,j_\circ\in\mathbb{N}} \in \mathcal{M}(\mathscr{B}^{\otimes k}_{\mathbb{H}} \otimes 2^{\mathbb{N}^2})$. By construction $\Gamma^X_{\bullet|\bullet} = \mathbb{P}_{\Gamma^X}(\psi_{\bullet|\bullet}) \in \mathcal{M}(2^{\mathbb{N}^2})$ is the mean kernel of $\widehat{\Gamma}^X_{\bullet|\bullet}$. For each $j, j_\circ \in \mathbb{N}$ the statistic $\dot{\boldsymbol{\eta}}_{j|j_\circ} := n^{1/2}(\widehat{\mathbb{P}}_k(\psi_{j|j_\circ}) - \mathbb{P}_{\Gamma^X}(\psi_{j|j_\circ})) \in \mathcal{M}(\mathscr{B}^{\otimes k}_{\mathbb{H}})$ is centred, i.e. $\dot{\boldsymbol{\eta}}_{j|j_\circ} \in \mathbb{L}_1(\mathbb{P}^{\otimes k}_{\Gamma^X})$ with $\mathbb{P}^{\otimes k}_{\Gamma^X}(\dot{\boldsymbol{\eta}}_{j|j_\circ}) = 0$, and exploiting $\psi_{\bullet|\bullet} \in \mathcal{M}(\mathscr{B}_{\mathbb{H}} \otimes 2^{\mathbb{N}^2})$ the stochastic process

$$\dot{\boldsymbol{\eta}}_{\bullet|\bullet} = (\dot{\boldsymbol{\eta}}_{j|j_\circ})_{j,j_\circ\in\mathbb{N}} = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_{\Gamma^X})(\psi_{\bullet|\bullet}) = k^{1/2}(\widehat{\mathbb{P}}_k(\psi_{\bullet|\bullet}) - \mathbb{P}_{\Gamma^X}(\psi_{\bullet|\bullet})) \in \mathcal{M}(\mathscr{B}^{\otimes k}_{\mathbb{H}} \otimes 2^{\mathbb{N}^2})$$

satisfies Assumption §02|01.02. Since $\widehat{\Gamma}^X_{j|j_\circ} = \Gamma^X_{j|j_\circ} + k^{-1/2}\dot{\boldsymbol{\eta}}_{j|j_\circ}$ for each $j, j_\circ \in \mathbb{N}$ by construction $\widehat{\Gamma}^X_{\bullet|\bullet} = \Gamma^X_{\bullet|\bullet} + k^{-1/2}\dot{\boldsymbol{\eta}}_{\bullet|\bullet}$ is a noisy version of $\Gamma^X_{\bullet|\bullet}$. □

§02|01.09 **Cross-covariance operator**. Let $(\mathbb{H}, \langle\cdot,\cdot\rangle_{\mathbb{H}})$ and $(\mathbb{G}, \langle\cdot,\cdot\rangle_{\mathbb{G}})$ be separable Hilbert space equipped with its Borel-σ-algebra $\mathscr{B}_{\mathbb{H}}$ and $\mathscr{B}_{\mathbb{G}}$, respectively. Consider an $\mathbb{H}$-valued *random function* X and an $\mathbb{G}$-valued *random function* Z. Then (Z, X) is an $(\mathbb{G}\times\mathbb{H}, \mathscr{B}_{\mathbb{G}} \otimes \mathscr{B}_{\mathbb{H}})$-valued *random function*. We denote by $\mathbb{P}^Z \in \mathscr{W}(\mathscr{B}_{\mathbb{G}})$ and $\mathbb{P}^X \in \mathscr{W}(\mathscr{B}_{\mathbb{H}})$ the marginal distribution of Z and X, respectively, and by $\mathbb{P}^{Z,X} \in \mathscr{W}(\mathscr{B}_{\mathbb{G}} \otimes \mathscr{B}_{\mathbb{H}})$ the joint distribution of (Z, X). We tactically identify Z and X with the coordinate map $\Pi_{\mathbb{G}}$ and $\Pi_{\mathbb{H}}$, respectively, and thus (Z, X) with the identity $\mathrm{id}_{\mathbb{G}\times\mathbb{H}}$ such that $\mathbb{P} = \mathbb{P}^{Z,X} \in \mathscr{W}(\mathscr{B}_{\mathbb{G}} \otimes \mathscr{B}_{\mathbb{H}})$. Here and subsequently, we assume that $\|Z\|^2_{\mathbb{G}} \in \mathcal{L}_2(\mathbb{P})$, $\|X\|^2_{\mathbb{H}} \in \mathcal{L}_2(\mathbb{P})$, $\mathbb{P}(\langle z, Z\rangle_{\mathbb{G}}) = 0$ and $\mathbb{P}(\langle x, X\rangle_{\mathbb{H}}) = 0$ for all $z \in \mathbb{G}$ and $x \in \mathbb{H}$. In this situation Z and X admits a *covariance operator* $\Gamma^Z \in \mathbb{L}^{\geqslant}(\mathbb{G})$ and $\Gamma^X \in \mathbb{L}^{\geqslant}(\mathbb{H})$, respectively (see Remark §01|01.07), and (Z, X) admits a *cross-covariance operator* $\Gamma^{ZX} \in \mathbb{L}(\mathbb{H}, \mathbb{G})$ satisfying

$$\langle \Gamma^{ZX} x, z\rangle_{\mathbb{G}} = \mathbb{P}^{X,Z}(\langle z, Z\rangle_{\mathbb{G}}\langle X, x\rangle_{\mathbb{H}}) \quad \forall x \in \mathbb{H}, z \in \mathbb{G}.$$

where $\|\Gamma^{ZX}\|_{\mathbb{L}(\mathbb{H},\mathbb{G})} \leqslant \|\Gamma^Z\|^{1/2}_{\mathbb{L}(\mathbb{H})}\|\Gamma^X\|^{1/2}_{\mathbb{L}(\mathbb{G})}$ (Baker [1973] p.275). Let us denote by $\mathbb{P}_{\Gamma^{ZX}} \in \mathscr{W}(\mathscr{B}_{\mathbb{G}} \otimes \mathscr{B}_{\mathbb{H}})$ the destribution of (Z, X) which is not fully specified given $\Gamma^{ZX} \in \mathbb{T} \subseteq \mathbb{L}(\mathbb{H}, \mathbb{G})$. We consider the statistical product experiment $((\mathbb{G}\times\mathbb{H})^k, (\mathscr{B}_{\mathbb{G}} \otimes \mathscr{B}_{\mathbb{H}})^{\otimes k}, \mathbb{P}^{\otimes k}_{\mathbb{T}} = (\mathbb{P}^{\otimes k}_{\Gamma^{ZX}})_{\Gamma^{ZX}\in\mathbb{T}})$. Let $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \ell_2)$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \ell_2)$ be generalised Fourier series transforms as in Notation §01|02.07 which are fixed and known in advanced. Then $\Gamma^{ZX}_{\bullet|\bullet} := \mathrm{V}\Gamma^{ZX}\mathrm{U}^\star \in \mathbb{L}^{\geqslant}(\ell_2)$ is an operator with kernel (infinite matrix) $\Gamma^{ZX}_{\bullet|\bullet} \in \mathcal{M}(2^{\mathbb{N}^2})$ satisfying $\Gamma^{ZX}_{\bullet|\bullet} = (\Gamma^{ZX}_{j|j_\circ} = \langle \Gamma^{ZX}\mathrm{u}_{j_\circ}, \mathrm{v}_j\rangle_{\mathbb{G}} = \mathbb{P}_{\Gamma^{ZX}}(\langle \mathrm{v}_j, Z\rangle_{\mathbb{G}}\langle X, \mathrm{u}_{j_\circ}\rangle_{\mathbb{H}}))_{j,j_\circ\in\mathbb{N}}$. Therefore the stochastic process $\psi_{\bullet|\bullet} = (\psi_{j|j_\circ}(Z, X) := \langle \mathrm{v}_j, Z\rangle_{\mathbb{G}}\langle X, \mathrm{u}_{j_\circ}\rangle_{\mathbb{H}})_{j,j_\circ\in\mathbb{N}} \in \mathcal{M}((\mathscr{B}_{\mathbb{G}} \otimes \mathscr{B}_{\mathbb{H}}) \otimes 2^{\mathbb{N}^2})$ fulfils Assumption §02|01.02 and $\Gamma^{ZX}_{\bullet|\bullet} = \mathbb{P}_{\Gamma^{ZX}}(\psi_{\bullet|\bullet})$. Similar to an Empirical mean model §02|01.04 we define $\widehat{\Gamma}^{ZX}_{\bullet|\bullet} = (\widehat{\Gamma}^{ZX}_{j|j_\circ} := \widehat{\mathbb{P}}_n(\psi_{j|j_\circ}))_{j,j_\circ\in\mathbb{N}} \in \mathcal{M}((\mathscr{B}_{\mathbb{G}} \otimes \mathscr{B}_{\mathbb{H}})^{\otimes k} \otimes 2^{\mathbb{N}^2})$. By construction $\Gamma^{ZX}_{\bullet|\bullet} = \mathbb{P}_{\Gamma^{ZX}}(\psi_{\bullet|\bullet}) \in$

$\mathcal{M}(2^{\mathbb{N}^2})$ is the mean kernel of $\widehat{\Gamma}^{ZX}_{\bullet|\bullet}$. For each $j, j_\circ \in \mathbb{N}$ the statistic $\dot{\eta}_{j|j_\circ} := k^{1/2}(\widehat{\mathbb{P}}_k(\psi_{j|j_\circ}) - \mathbb{P}_{\Gamma^{zx}}(\psi_{j|j_\circ})) \in \mathcal{M}((\mathscr{B}_{\mathbb{G}} \otimes \mathscr{B}_{\mathbb{H}})^{\otimes k})$ is centred, i.e. $\dot{\eta}_{j|j_\circ} \in \mathbb{L}_1(\mathbb{P}_{\Gamma^{zx}} \otimes k)$ with $\mathbb{P}^{\otimes k}_{\Gamma^{zx}}(\dot{\eta}_{j|j_\circ}) = 0$, and the stochastic process

$$\dot{\eta}_{\bullet|\bullet} = (\dot{\eta}_{j|j_\circ})_{j,j_\circ \in \mathbb{N}} = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_{\Gamma^{zx}})(\psi_{\bullet|\bullet}) = k^{1/2}(\widehat{\mathbb{P}}_k(\psi_{\bullet|\bullet}) - \mathbb{P}_{\Gamma^{zx}}(\psi_{\bullet|\bullet})) \in \mathcal{M}((\mathscr{B}_{\mathbb{G}} \otimes \mathscr{B}_{\mathbb{H}})^{\otimes k} \otimes 2^{\mathbb{N}^2})$$

satisfies Assumption §02|01.02 exploiting $\psi_{\bullet|\bullet} \in \mathcal{M}((\mathscr{B}_{\mathbb{G}} \otimes \mathscr{B}_{\mathbb{H}})^{\otimes k} \otimes 2^{\mathbb{N}^2})$. Since $\widehat{\Gamma}^{ZX}_{j|j_\circ} = \Gamma^{ZX}_{j|j_\circ} + k^{-1/2}\dot{\eta}_{j|j_\circ}$ for each $j, j_\circ \in \mathbb{N}$ by construction $\widehat{\Gamma}^{ZX}_{\bullet|\bullet} = \Gamma^{ZX}_{\bullet|\bullet} + k^{-1/2}\dot{\eta}_{\bullet|\bullet}$ is a noisy version of $\Gamma^{ZX}_{\bullet|\bullet}$. □

§02|01.10 **Design operator**. Consider the measure space $(\mathcal{D}, \mathscr{B}_{\mathcal{D}}, \lambda_{\mathcal{D}})$ where $\lambda_{\mathcal{D}}$ denotes the restriction of the Lebesgue measure to the Borel-σ-algebra $\mathscr{B}_{\mathcal{D}}$ over $\mathcal{D} \in \mathscr{B}$, and the *real* Hilbert space $\mathbb{L}_2(\lambda_{\mathcal{D}}) := \mathbb{L}_2(\mathcal{D}, \mathscr{B}_{\mathcal{D}}, \lambda_{\mathcal{D}})$ of square Lebesgue-integrable real-valued functions. Let $\mathbb{P}_\varphi \in \mathscr{W}(\mathscr{B}_{\mathcal{D}})$ admit a Lebesgue density $\varphi \in \mathbb{L}_1(\lambda_{\mathcal{D}})$, that is $\mathbb{P}_\varphi = \varphi\lambda_{\mathcal{D}}$ (compare Regression with known design §01|05.10). Let in addition $\varphi \in \mathbb{L}_\infty(\lambda_{\mathcal{D}})$, then $\mathrm{M}_\varphi \in \mathbb{L}^{\mathrm{M}}(\mathbb{L}_2(\lambda_{\mathcal{D}}))$ with $h \mapsto \mathrm{M}_\varphi h := \varphi h$. Note that then for each $h \in \mathbb{L}_2(\lambda_{\mathcal{D}})$ we have $\mathrm{M}_\varphi h \in \mathbb{L}_2(\lambda_{\mathcal{D}})$. Consequently, for each $g, h \in \mathbb{L}_2(\lambda_{\mathcal{D}})$, hence $g, h \in \mathcal{L}_2(\mathbb{P}_\varphi)$ we obtain $gh \in \mathcal{L}_1(\mathbb{P}_\varphi)$ and

$$\mathbb{P}_\varphi(gh) = \varphi\lambda_{\mathcal{D}}(gh) = \lambda_{\mathcal{D}}(\varphi gh) = \langle \mathrm{M}_\varphi g, h\rangle_{\mathbb{L}_2(\lambda_{\mathcal{D}})} \in \mathbb{R}$$

identifying again equivalence classes and their representatives. We note that $\mathrm{M}_\varphi \in \mathbb{L}^{\mathrm{M}}(\mathbb{L}_2(\lambda_{\mathcal{D}}))$ with density $\varphi \in \mathbb{L}_\infty(\lambda_{\mathcal{D}})$ is positive semi-definite, i.e. $\mathrm{M}_\varphi \in \mathbb{L}^{\geqslant}(\mathbb{L}_2(\lambda_{\mathcal{D}}))$ and if in addition $\varphi \in \mathcal{M}_{>0,\lambda_{\mathcal{D}}}(\mathscr{B}_{\mathcal{D}})$ (i.e. $\varphi \in \mathcal{M}_{\geqslant 0}(\mathscr{B}_{\mathcal{D}})$ and $\lambda_{\mathcal{D}}(\mathcal{N}_\varphi) = 0$) then it is stictly positive definite, i.e. $\mathrm{M}_\varphi \in \mathbb{L}^{\gneq}(\mathbb{L}_2(\lambda_{\mathcal{D}}))$. Keep in mind that $\mathrm{U} \in \mathbb{L}(\mathbb{L}_2(\lambda_{\mathcal{D}}), \ell_2)$ is generalised Fourier series transform as in Notation §01|02.07 which is fixed and *known in advance*. Evidently, we have $\mathrm{M}^\varphi_{\bullet|\bullet} := \mathrm{U}\mathrm{M}_\varphi\mathrm{U}^\star \in \mathbb{L}^{\gneq}_{\bullet}(\ell_2) \subseteq \mathbb{L}_{\bullet}(\ell_2) = \mathbb{L}(\ell_2)$ satisfying $\mathrm{M}^\varphi_{\bullet|\bullet} = (\mathrm{M}^\varphi_{j|j_\circ} = \langle \mathrm{M}^\varphi_{\bullet|\bullet}\mathrm{u}_{j_\circ}, \mathrm{u}_j\rangle_{\mathbb{G}} = \mathbb{P}_\varphi(\mathrm{u}_{j_\circ}\mathrm{u}_j))_{j,j_\circ \in \mathbb{N}}$. Therefore the stochastic process $\psi_{\bullet|\bullet} = (\psi_{j|j_\circ} := \mathrm{u}_{j_\circ}\mathrm{u}_j)_{j,j_\circ \in \mathbb{N}} \in \mathcal{M}(\mathscr{B}_{\mathcal{D}} \otimes 2^{\mathbb{N}^2})$ fulfils Assumption §02|01.02 and $\mathrm{M}^\varphi_{\bullet|\bullet} = \mathbb{P}_\varphi(\psi_{\bullet|\bullet})$. Similar to an Empirical mean model §02|01.04 we define $\widehat{\mathrm{M}}_{\bullet|\bullet} = (\widehat{\mathrm{M}}_{j|j_\circ} := \widehat{\mathbb{P}}_k(\psi_{j|j_\circ}))_{j,j_\circ \in \mathbb{N}} \in \mathcal{M}(\mathscr{B}^{\otimes k}_{\mathcal{D}} \otimes 2^{\mathbb{N}^2})$. By construction $\mathrm{M}^\varphi_{\bullet|\bullet} = \mathbb{P}_\varphi(\psi_{\bullet|\bullet}) \in \mathcal{M}(2^{\mathbb{N}^2})$ is the mean kernel of $\widehat{\mathrm{M}}_{\bullet|\bullet}$. For each $j, j_\circ \in \mathbb{N}$ the statistic $\dot{\eta}_{j|j_\circ} := k^{1/2}(\widehat{\mathbb{P}}_k(\psi_{j|j_\circ}) - \mathbb{P}_\varphi(\psi_{j|j_\circ})) \in \mathcal{M}(\mathscr{B}^{\otimes k}_{\mathcal{D}})$ is centred, i.e. $\dot{\eta}_{j|j_\circ} \in \mathbb{L}_1(\mathbb{P}^{\otimes k}_\varphi)$ with $\mathbb{P}^{\otimes k}_\varphi(\dot{\eta}_{j|j_\circ}) = 0$, and exploiting $\psi_{\bullet|\bullet} \in \mathcal{M}(\mathscr{B}_{\mathcal{D}} \otimes 2^{\mathbb{N}^2})$ the stochastic process

$$\dot{\eta}_{\bullet|\bullet} = (\dot{\eta}_{j|j_\circ})_{j,j_\circ \in \mathbb{N}} = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_\varphi)(\psi_{\bullet|\bullet}) = k^{1/2}(\widehat{\mathbb{P}}_k(\psi_{\bullet|\bullet}) - \mathbb{P}_\varphi(\psi_{\bullet|\bullet})) \in \mathcal{M}(\mathscr{B}^{\otimes k}_{\mathcal{D}} \otimes 2^{\mathbb{N}^2})$$

satisfies Assumption §02|01.02. Since $\widehat{\mathrm{M}}_{j|j_\circ} = \mathrm{M}^\varphi_{j|j_\circ} + k^{-1/2}\dot{\eta}_{j|j_\circ}$ for each $j, j_\circ \in \mathbb{N}$ by construction $\widehat{\mathrm{M}}_{\bullet|\bullet} = \mathrm{M}^\varphi_{\bullet|\bullet} + k^{-1/2}\dot{\eta}_{\bullet|\bullet}$ is a noisy version of $\mathrm{M}^\varphi_{\bullet|\bullet}$. □

§02|02 Non-diagonal statistical inverse problem with noisy operator

§02|02.01 **Assumption**. For $\mathbb{J} = \mathbb{L}_2(\nu)$, surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \mathbb{J})$, fixed and presumed to be known in advance, the operator $\mathrm{T} \in \mathbb{L}^{\mathrm{U,V}}(\mathbb{L}_{\bullet}(\mathbb{J}))$ and hence $\mathrm{T}_{\bullet|\bullet} = \mathrm{V}\mathrm{T}\mathrm{U}^\star \in \mathbb{L}_{\bullet}(\mathbb{J})$ with kernel $\mathrm{T}_{\bullet|\bullet} \in \mathscr{J}^2$ is *not known* in advance where $g_\bullet = \mathrm{T}_{\bullet|\bullet}\theta_\bullet \in \mathbb{J}$ or inequal $g_\bullet \in \mathrm{ran}(\mathrm{T}_{\bullet|\bullet}) = \{\mathrm{T}_{\bullet|\bullet}a_\bullet \colon a_\bullet \in \mathbb{J}\}$. □

§02|02.02 **Notation**. Under Assumption §02|02.01 given $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}_{\bullet}(\mathbb{J})$ and $g_\bullet \in \mathrm{ran}(\mathrm{T}_{\bullet|\bullet})$ we consider the reconstruction of $\theta_\bullet = \mathrm{U}\theta \in \mathbb{J}$ (or in equal $\theta = \mathrm{U}^\star\theta_\bullet \in \mathbb{H}$) from a noisy version of the image $g_\bullet = \mathrm{V}\mathrm{T}\mathrm{U}^\star\theta_\bullet = \mathrm{T}_{\bullet|\bullet}\theta_\bullet \in \mathbb{J}$ and a noisy version of the operator $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}_{\bullet}(\mathbb{J})$. Keep in mind, that we identify the equivalence class and its representative $g_\bullet$. □

§02|02.03 **Non-diagonal statistical inverse problem with noisy operator**. As in Definition §01|02.03 consider a stochastic process $\dot{\varepsilon}_\bullet = (\dot{\varepsilon}_j)_{j \in \mathcal{J}}$ satisfying Assumption §01|01.04 with mean zero and a sample size $n \in \mathbb{N}$, and in addition as in Definition §02|01.03 a stochastic process $\dot{\eta}_{\bullet|\bullet} =$

$(\dot{\eta}_{j|j_\circ})_{j,j_\circ\in\mathcal{J}}$ satisfying Assumption §02|01.02 with mean zero and a sample size $k\in\mathbb{N}$. Under Assumption §02|02.01 where $\mathrm{T}_{\bullet|\bullet}\in\mathbb{L}_\bullet(\mathbb{J})$ with kernel $\mathrm{T}_{\bullet|\bullet}\in\mathcal{M}(\mathscr{J}^2)$ is *not known* anymore, the observable noisy image (Definition §01|02.03) has $\mathbb{J}$-mean $g_\bullet=\mathrm{T}_{\bullet|\bullet}\theta_\bullet$ and the observable noisy non-diagonal operator (Definition §02|01.03) has mean kernel $\mathrm{T}_{\bullet|\bullet}\in\mathcal{M}(\mathscr{J}^2)$, and take the form $\widehat{g}_\bullet=g_\bullet+n^{-1/2}\dot{\varepsilon}_\bullet$ and $\widehat{\mathrm{T}}_{\bullet|\bullet}=\mathrm{T}_{\bullet|\bullet}+k^{-1/2}\dot{\eta}_{\bullet|\bullet}$, respectively, or in equal

$$\widehat{g}_j=\langle\mathrm{T}_{j|\bullet},\theta_\bullet\rangle_{\mathbb{J}}+n^{-1/2}\dot{\varepsilon}_j\quad\text{and}\quad\widehat{\mathrm{T}}_{j|j_\circ}=\mathrm{T}_{j|j_\circ}+k^{-1/2}\dot{\eta}_{j|j_\circ},\quad j,j_\circ\in\mathcal{J}.\tag{02.03}$$

We denote by $\mathbb{P}^{n,k}_{\theta|\mathrm{T}}$ the joint distribution of $(\widehat{g}_\bullet,\widehat{\mathrm{T}}_{\bullet|\bullet})$. The reconstruction of $\theta_\bullet\in\mathbb{J}$ (or in equal $\theta=\mathrm{U}^\star\theta_\bullet\in\mathbb{H}$) from a noisy version $(\widehat{g}_\bullet,\widehat{\mathrm{T}}_{\bullet|\bullet})\sim\mathbb{P}^{n,k}_{\theta|\mathrm{T}}$ of the image $g_\bullet=\mathrm{V}\mathrm{T}\mathrm{U}^\star\theta_\bullet=\mathrm{T}_{\bullet|\bullet}\theta_\bullet\in\mathbb{J}$ and the operator $\mathrm{T}\in\mathbb{L}^{\mathrm{U},\mathrm{V}}(\mathbb{L}_\bullet(\mathbb{J}))$ is called a *non-diagonal statistical inverse problem with noisy operator*. □

§02|02.04 **Non-diagonal inverse empirical mean model (nieMM) with noisy operator**. Consider the reconstruction of $\theta_\bullet\in\mathbb{J}$ (in equal $\theta=\mathrm{U}^\star\theta_\bullet\in\mathbb{H}$) in an Empirical mean model as in §01|02.04. Under Assumption §02|02.01, where $\mathrm{T}_{\bullet|\bullet}\in\mathbb{L}_\bullet(\mathbb{J})$ with kernel $\mathrm{T}_{\bullet|\bullet}\in\mathcal{M}(\mathscr{J}^2)$ is *not known* in advance, the observable noisy image has $\mathbb{J}$-mean $\mathrm{V}g=g_\bullet=\mathrm{T}_{\bullet|\bullet}\theta_\bullet\in\mathbb{J}$ and the observable noisy non-diagonal operator (Definition §02|01.03) has mean kernel $\mathrm{T}_{\bullet|\bullet}\in\mathscr{J}^2$, and take, respectively, the form of an Empirical mean model as in §01|02.04 and Empirical mean model as in §02|01.04. More precisely, for each $\theta\in\Theta\subseteq\mathbb{H}$ and $\mathrm{T}\in\mathbb{T}\subseteq\mathbb{L}^{\mathrm{U},\mathrm{V}}(\mathbb{L}_\bullet(\mathbb{J}))$ let $\mathbb{P}_{\theta|\mathrm{T}}\in\mathscr{W}(\mathscr{Z})$ be a probability measure on a measurable space $(\mathcal{Z},\mathscr{Z})$. Similar to §01|02.04 and §02|01.04 consider stochastic processes $\psi_\bullet^{\Theta|\mathbb{T}}\in\mathscr{Z}\otimes\mathscr{J}$ and $\psi_{\bullet|\bullet}^{\mathbb{T}}\in\mathscr{Z}\otimes\mathscr{J}^2$ which in addition for all $\theta\in\Theta$ and $\mathrm{T}\in\mathbb{T}$ satisfy $\psi_j^{\Theta|\mathbb{T}},\psi_{j,j_\circ}^{\mathbb{T}}\in\mathcal{L}_1(\mathbb{P}_{\theta|\mathrm{T}})$ for each $j,j_\circ\in\mathcal{J}$ and $\mathbb{P}_{\theta|\mathrm{T}}(\psi_\bullet^{\Theta|\mathbb{T}})=g_\bullet=\mathrm{T}_{\bullet|\bullet}\theta_\bullet$ and $\mathbb{P}_{\theta|\mathrm{T}}(\psi_{\bullet|\bullet}^{\mathbb{T}})=\mathrm{T}_{\bullet|\bullet}$. The observable noisy versions take the form $\widehat{g}_\bullet=\mathrm{T}_{\bullet|\bullet}\theta_\bullet+n^{-1/2}\dot{\varepsilon}_\bullet$ and $\widehat{\mathrm{T}}_{\bullet|\bullet}=\mathrm{T}_{\bullet|\bullet}+k^{-1/2}\dot{\eta}_{\bullet|\bullet}$, or in equal (02.03) with error processes $\dot{\varepsilon}_\bullet=n^{1/2}(\widehat{\mathbb{P}}_n(\psi_\bullet^{\Theta|\mathbb{T}})-\mathbb{P}_{\theta|\mathrm{T}}(\psi_\bullet^{\Theta|\mathbb{T}}))\in\mathcal{M}(\mathscr{Z}^{\otimes n}\otimes\mathscr{J})$ and $\dot{\eta}_{\bullet|\bullet}=k^{1/2}(\widehat{\mathbb{P}}_k(\psi_{\bullet|\bullet}^{\mathbb{T}})-\mathbb{P}_{\theta|\mathrm{T}}(\psi_{\bullet|\bullet}^{\mathbb{T}}))\in\mathcal{M}(\mathscr{Z}^{\otimes k}\otimes\mathscr{J}^2)$ satisfying Assumption §01|01.04 and Assumption §02|01.02. □

§02|02.05 **Non-diagonal inverse sequence model (niSM) with noisy operator**. Consider $\mathbb{J}=\ell_2=\mathbb{L}_2(\nu_{\mathbb{N}})$ as in §01|01.14. Let $\dot{\varepsilon}_\bullet=(\dot{\varepsilon}_j)_{j\in\mathbb{N}}$ and $\dot{\eta}_{\bullet|\bullet}=(\dot{\eta}_{j|j_\circ})_{j,j_\circ\in\mathbb{N}}$ be real-valued stochastic processes satisfying Assumption §01|01.04 and Assumption §02|01.02 with mean zero and let $n,k\in\mathbb{N}$ be sample sizes. Under Assumption §02|02.01, where $\mathrm{T}_{\bullet|\bullet}\in\mathbb{L}_\bullet(\ell_2)$ with kernel $\mathrm{T}_{\bullet|\bullet}\in\mathcal{M}(2^{\mathbb{N}^2})$ is not known in advance, the observable noisy image has ℓ_2-mean $g_\bullet=\mathrm{T}_{\bullet|\bullet}\theta_\bullet$ and the observable noisy operator has mean kernel $\mathrm{T}_{\bullet|\bullet}\in\mathcal{M}(2^{\mathbb{N}^2})$, and take the form of a Sequence model as in §01|02.05 and Bivariate sequence model as in §02|01.05, that is $\widehat{g}_\bullet=\mathrm{T}_{\bullet|\bullet}\theta_\bullet+n^{-1/2}\dot{\varepsilon}_\bullet$ and $\widehat{\mathrm{T}}_{\bullet|\bullet}=\mathrm{T}_{\bullet|\bullet}+k^{-1/2}\dot{\eta}_{\bullet|\bullet}$ or in equal

$$\widehat{g}_j=\langle\mathrm{T}_{j|\bullet},\theta_\bullet\rangle_{\ell_2}+n^{-1/2}\dot{\varepsilon}_j\quad\text{and}\quad\widehat{\mathrm{T}}_{j|j_\circ}=\mathrm{T}_{j|j_\circ}+k^{-1/2}\dot{\eta}_{j|j_\circ},\quad j,j_\circ\in\mathbb{N}.\tag{02.04}$$

We denote by $\mathbb{P}^{n,k}_{\theta|\mathrm{T}}$ the joint distribution of $(\widehat{g}_\bullet,\widehat{\mathrm{T}}_{\bullet|\bullet})$. □

§02|02.06 **Gaussian non-diagonal inverse sequence model (GniSM) with noisy operator**. Consider $\mathbb{J}=\ell_2$ as in §01|01.14. Let $\dot{\mathrm{B}}_\bullet:=(\dot{\mathrm{B}}_j)_{j\in\mathbb{N}}\sim\mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}$ and $\dot{\mathrm{W}}_{\bullet|\bullet}:=(\dot{\mathrm{W}}_{j|j_\circ})_{j,j_\circ\in\mathbb{N}}\sim\mathrm{N}_{(0,1)}^{\otimes\mathbb{N}^2}$ be Gaussian white noise process. The observable noisy versions $\widehat{g}_\bullet=g_\bullet+n^{-1/2}\dot{\mathrm{B}}_\bullet$ with ℓ_2-mean $g_\bullet=\mathrm{T}_{\bullet|\bullet}\theta_\bullet$ and $\widehat{\mathrm{T}}_{\bullet|\bullet}=\mathrm{T}_{\bullet|\bullet}+k^{-1/2}\dot{\mathrm{W}}_{\bullet|\bullet}$ with mean kernel $\mathrm{T}_{\bullet|\bullet}\in\mathcal{M}(2^{\mathbb{N}^2})$ take the form of a Gaussian sequence model as in §01|02.06 and Gaussian bivariate sequence model as in §02|01.06, that is

$$\widehat{g}_j=\langle\mathrm{T}_{j|\bullet},\theta_\bullet\rangle_{\ell_2}+n^{-1/2}\dot{\mathrm{B}}_j\quad\text{and}\quad\widehat{\mathrm{T}}_{j|j_\circ}=\mathrm{T}_{j|j_\circ}+k^{-1/2}\dot{\mathrm{W}}_{j|j_\circ},\quad j,j_\circ\in\mathbb{N}$$
$$\text{with}\quad(\dot{\mathrm{B}}_j)_{j\in\mathbb{N}}\sim\mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}\quad\text{and}\quad(\dot{\mathrm{W}}_{j|j_\circ})_{j,j_\circ\in\mathbb{N}}\sim\mathrm{N}_{(0,1)}^{\otimes\mathbb{N}^2}.\tag{02.05}$$

We denote by $\mathrm{N}^{n,k}_{\theta|\mathrm{T}}$ the joint distribution of the stochastic process $(\widehat{g}_\bullet,\widehat{\mathrm{T}}_{\bullet|\bullet})$. □

§02|02|01 Examples of non-diagonal inverse empirical mean models with noisy operator

§02|02.07 **Instrumental regression.** Consider for $\mathcal{Z},\mathcal{X}\in\mathscr{B}$ the Borel-measurable spaces $(\mathcal{Z},\mathscr{B}_z)$, $(\mathcal{X},\mathscr{B}_x)$ and $(\mathbb{R},\mathscr{B})$. Let (Z,X,Y) be a $\mathcal{Z}\times\mathcal{X}\times\mathbb{R}$-valued random vector with joint distribution $\mathbb{P}^{Z,X,Y}\in\mathscr{W}(\mathscr{B}_z\otimes\mathscr{B}_x\otimes\mathscr{B})$. We denote by $\mathbb{P}^Z\in\mathscr{W}(\mathscr{B}_z)$ the marginal distribution of Z, by $\mathbb{P}^{X|Z}$ and $\mathbb{P}^{Y|Z}$ a regular conditional distribution of X given Z and Y given Z, respectively, and by $\mathbb{P}^{Z,X}=\mathbb{P}^Z\odot\mathbb{P}^{X|Z}\in\mathscr{W}(\mathscr{B}_z\otimes\mathscr{B}_x)$ and $\mathbb{P}^{Z,Y}=\mathbb{P}^Z\odot\mathbb{P}^{Y|Z}\in\mathscr{W}(\mathscr{B}_z\otimes\mathscr{B})$ the marginal distributions of (Z,X) and (Z,Y). We tactically identify Z, X and Y with the coordinate map Π_z, Π_x and $\Pi_{\mathbb{R}}$, respectively, and thus (Z,X,Y) with the identity $\mathrm{id}_{\mathcal{Z}\times\mathcal{X}\times\mathbb{R}}$ such that $\mathbb{P}=\mathbb{P}^{Z,X,Y}\in\mathscr{W}(\mathscr{B}_z\otimes\mathscr{B}_x\otimes\mathscr{B})$. If in addition $Y\in\mathbb{L}_1(\mathbb{P})=\mathbb{L}_1(\mathcal{Z}\times\mathcal{X}\times\mathbb{R},\mathscr{B}_z\otimes\mathscr{B}_x\otimes\mathscr{B},\mathbb{P})$ then $\mathbb{P}^{Y|Z}(\mathrm{id}_{\mathscr{B}_y})=\mathbb{P}\big(Y\big|Z\big)=:g\in\mathbb{L}_1(\mathbb{P}^Z)$ is unique up to $\mathbb{P}^Z$-a.s. equality (compare Regression with uniform design §01|02.09). Introduce further the Hilbert spaces $\mathbb{L}_2(\mathbb{P}^X):=\mathbb{L}_2(\mathcal{X},\mathscr{B}_x,\mathbb{P}^X)$, $\mathbb{L}_2(\mathbb{P}^Z):=\mathbb{L}_2(\mathcal{Z},\mathscr{B}_z,\mathbb{P}^Z)$ and as in §02|01.07 the *conditional expectation operator* $\mathbb{P}^{X|Z}\in\mathbb{L}(\mathbb{L}_2(\mathbb{P}^X),\mathbb{L}_2(\mathbb{P}^Z))$ with $h\mapsto\mathbb{P}^{X|Z}h:=\mathbb{P}^{X|Z}(h)=\mathbb{P}\big(h(X)\big|Z\big)$. In what follows we assume that $Y\in\mathbb{L}_2(\mathbb{P})$ and hence $g\in\mathbb{L}_2(\mathbb{P}^Z)$, and that in addition $g\in\mathrm{ran}(\mathbb{P}^{X|Z})\subseteq\mathbb{L}_2(\mathbb{P}^Z)$. In this situation there exists $f\in\mathbb{L}_2(\mathbb{P}^X)$ such that for any $h\in\mathbb{L}_2(\mathbb{P}^Z)$

$$\langle g,h\rangle_{\mathbb{L}_2(\mathbb{P}^Z)}=\mathbb{P}^Z\big(\mathbb{P}\big(Y\big|Z\big)h(Z)\big)=\mathbb{P}^Z\big(\mathbb{P}\big(f(X)\big|Z\big)h(Z)\big)=\langle\mathbb{P}^{X|Z}f,h\rangle_{\mathbb{L}_2(\mathbb{P}^Z)}$$

or in equal $\mathbb{P}$-a.s. we have $Y=f(X)+\xi$ with $\mathbb{P}\big(\xi\big|Z\big)=0$. We note that for all $h\in\mathbb{L}_2(\mathbb{P}^Z)$ we have $\langle g,h\rangle_{\mathbb{L}_2(\mathbb{P}^Z)}=\mathbb{P}(Yh(Z))$. We assume moreover that $\mathcal{X}=\mathcal{Z}=[0,1]$ and that X and Z is uniformly distributed on the interval $[0,1]$, i.e. $X\sim\mathrm{U}_{[0,1]}=\lambda_{[0,1]}=\mathbb{P}^X$ and $Z\sim\mathrm{U}_{[0,1]}=\lambda_{[0,1]}=\mathbb{P}^Z$. Consequently, we set $\mathbb{H}:=\mathbb{L}_2(\mathbb{P}^X)=\mathbb{L}_2(\lambda_{[0,1]})$ and $\mathbb{G}:=\mathbb{L}_2(\mathbb{P}^Z)=\mathbb{L}_2(\lambda_{[0,1]})$. We denote by $\mathrm{U}_{\mathbb{P}^{X|Z}}:=\mathrm{U}_{[0,1]}\odot\mathbb{P}^{X|Z}$ the joint distribution of (Z,X) which is now fully specified once the conditional expectation operator $\mathbb{P}^{X|Z}\in\mathbb{T}\subseteq\mathbb{L}(\mathbb{L}_2(\lambda_{[0,1]}))$ is given (see Model §02|01.07). Moreover, for $\mathbb{P}^{X|Z}\in\mathbb{T}\subseteq\mathbb{L}(\mathbb{L}_2(\lambda_{[0,1]}))$ and $f\in\mathbb{F}_2\subseteq\mathbb{L}_2(\lambda_{[0,1]})=\mathbb{H}$, and hence $g:=\mathbb{P}^{X|Z}f\in\mathbb{L}_2(\lambda_{[0,1]})=\mathbb{G}$, we denote by $\mathrm{U}_{f|\mathbb{P}^{X|Z}}:=\mathrm{U}_{[0,1]}\odot\mathbb{P}_g^{Y|Z}$ the joint distribution of (Z,Y) without fully specifying the regular conditional distribution $\mathbb{P}_g^{Y|Z}$ which however satisfies $\mathbb{P}_g^{Y|Z}(\mathrm{id}_{\mathscr{B}_y})=\mathbb{P}_g\big(Y\big|Z\big)=g=\mathbb{P}^{X|Z}f$ (see Model §01|05.09). Let $\mathrm{U},\mathrm{V}\in\mathbb{L}(\mathbb{L}_2(\lambda_{[0,1]}),\ell_2)$ be generalised Fourier series transforms as in Notation §01|02.07 which are fixed and *known* in advance. Following Model §02|01.07 $\mathbb{P}^{X|Z}_{\cdot|\cdot}:=\mathrm{V}\mathbb{P}^{X|Z}\mathrm{U}^\star\in\mathbb{L}_{\bullet}(\ell_2)$ is an operator with kernel (infinite matrix) $\mathbb{P}^{X|Z}_{\cdot|\cdot}\in\mathcal{M}(2^{\mathbb{N}^2})$ satisfying $\mathbb{P}^{X|Z}_{\cdot|\cdot}=(\mathbb{P}^{X|Z}_{j|j_\circ}=\langle\mathbb{P}^{X|Z}\mathrm{u}_{j_\circ},\mathrm{v}_j\rangle_{\mathbb{G}}=\mathrm{U}_{\mathbb{P}^{X|Z}}(\mathrm{u}_{j_\circ}(X)\mathrm{v}_j(Z)))_{j,j_\circ\in\mathbb{N}}$. Therefore the stochastic process $\psi_{\cdot|\cdot}=(\psi_{j|j_\circ}(Z,X):=\mathrm{u}_{j_\circ}(X)\mathrm{v}_j(Z))_{j,j_\circ\in\mathbb{N}}\in\mathcal{M}(\mathscr{B}^2_{[0,1]}\otimes 2^{\mathbb{N}^2})$ fulfils Assumption §02|01.02 and $\mathbb{P}^{X|Z}_{\cdot|\cdot}=\mathrm{U}_{\mathbb{P}^{X|Z}}(\psi_{\cdot|\cdot})$. Moreover, similar to Model §01|05.09 for each $f\in\mathbb{F}_2\subseteq\mathbb{L}_2(\lambda_{[0,1]})=\mathbb{H}$ the generalised Fourier coefficients $g_{\cdot}=(g_j)_{j\in\mathbb{N}}=\mathrm{V}g=\mathrm{V}\mathbb{P}^{X|Z}\mathrm{U}^\star\mathrm{U}f=\mathbb{P}^{X|Z}_{\cdot|\cdot}f_{\cdot}$, satisfy $g_{\cdot}=\mathrm{U}_{f|\mathbb{P}^{X|Z}}(Y\mathrm{v}_{\cdot}(Z))$. The stochastic process $\psi_{\cdot}=(\psi_j:=Y\mathrm{v}_j(X))_{j\in\mathbb{N}}\in\mathcal{M}((\mathscr{B}_{[0,1]}\otimes\mathscr{B})\otimes 2^{\mathbb{N}})$ fulfils Assumption §01|01.04 and $g_{\cdot}=\mathbb{P}^{X|Z}_{\cdot|\cdot}f_{\cdot}=\mathrm{U}_{f|\mathbb{P}^{X|Z}}(\psi_{\cdot}^{\mathbb{F}_2|\mathbb{T}})$. The observable noisy versions take the form $\widehat{g}_{\cdot}=\mathbb{P}^{X|Z}_{\cdot|\cdot}f_{\cdot}+n^{-1/2}\dot{\varepsilon}_{\cdot}$ and $\widehat{\mathbb{P}}^{X|Z}_{\cdot|\cdot}=\mathbb{P}^{X|Z}_{\cdot|\cdot}+k^{-1/2}\dot{\eta}_{\cdot|\cdot}$, or in equal (02.03) with error processes

$$\dot{\varepsilon}_{\cdot}=n^{1/2}(\widehat{\mathbb{P}}_n-\mathrm{U}_{f|\mathbb{P}^{X|Z}})(\psi_{\cdot})=n^{1/2}(\widehat{\mathbb{P}}_n(\psi_{\cdot})-\mathrm{U}_{f|\mathbb{P}^{X|Z}}(\psi_{\cdot}))\in\mathcal{M}((\mathscr{B}_{[0,1]}\otimes\mathscr{B})^{\otimes n}\otimes 2^{\mathbb{N}})\quad\text{and}$$
$$\dot{\eta}_{\cdot|\cdot}=k^{1/2}(\widehat{\mathbb{P}}_k-\mathrm{U}_{\mathbb{P}^{X|Z}})(\psi_{\cdot|\cdot})=k^{1/2}(\widehat{\mathbb{P}}_k(\psi_{\cdot|\cdot})-\mathrm{U}_{\mathbb{P}^{X|Z}}(\psi_{\cdot|\cdot}))\in\mathcal{M}(\mathscr{B}^{\otimes 2k}_{[0,1]}\otimes 2^{\mathbb{N}^2})$$

satisfying Assumption §01|01.04 and Assumption §02|01.02. □

§02|02.08 **Functional linear regression.** Let $(\mathbb{H},\langle\cdot,\cdot\rangle_{\mathbb{H}})$ be a separable Hilbert space equipped with its Borel-σ-algebra $\mathscr{B}_{\mathbb{H}}$ and let (X,Y) be an $\mathbb{H}\times\mathbb{R}$-valued random vector with joint distribution $\mathbb{P}^{X,Y}\in\mathscr{W}(\mathscr{B}_{\mathbb{H}}\otimes\mathscr{B})$. We denote by $\mathbb{P}^X\in\mathscr{W}(\mathscr{B}_{\mathbb{H}})$ the marginal distribution of X. We tactically identify X and Y with the coordinate map $\Pi_{\mathbb{H}}$ and $\Pi_{\mathbb{R}}$, respectively, and thus (X,Y) with the identity $\mathrm{id}_{\mathbb{H}\times\mathbb{R}}$ such that $\mathbb{P} = \mathbb{P}^{X,Y} \in \mathscr{W}(\mathscr{B}_{\mathbb{H}} \otimes \mathscr{B})$. Here and subsequently, we assume that $Y, \|X\|_{\mathbb{H}}^2 \in \mathbb{L}_2(\mathbb{P}) = \mathbb{L}_2(\mathbb{H}\times\mathbb{R}, \mathscr{B}_{\mathbb{H}} \otimes \mathscr{B}, \mathbb{P})$ and $\mathbb{P}(\langle x, X\rangle_{\mathbb{H}}) = 0$ for all $x \in \mathbb{H}$. In this situation X admits a *covariance operator* $\Gamma^X \in \mathbb{L}^{\geq}(\mathbb{H})$ (see Remark §01|01.07) and there is $g \in \mathbb{H}$ satisfying $\langle g, x\rangle_{\mathbb{H}} = \mathbb{P}(Y\langle X, x\rangle_{\mathbb{H}})$ for all $x \in \mathbb{H}$. In what follows we assume that in addition $g \in \mathrm{ran}(\Gamma^X) \subseteq \mathbb{H}$. In this situation there exists $f \in \mathbb{H}$ such that

$$\langle g, x\rangle_{\mathbb{H}} = \mathbb{P}(Y\langle X, x\rangle_{\mathbb{H}}) = \mathbb{P}(\langle X, f\rangle_{\mathbb{H}}\langle X, x\rangle_{\mathbb{H}}) = \langle \Gamma^X f, x\rangle_{\mathbb{H}} \quad \forall x \in \mathbb{H}$$

or in equal $\mathbb{P}$-a.s. we have $Y = \langle X, f\rangle_{\mathbb{H}} + \xi$ with $\mathbb{P}(\xi\langle X, x\rangle_{\mathbb{H}}) = 0$ for all $x \in \mathbb{H}$. Let us denote by $\mathbb{P}_{\Gamma^X} \in \mathscr{W}(\mathscr{B}_{\mathbb{H}})$ the marginal destribution of X which is not fully specified given $\Gamma^X \in \mathbb{T} \subseteq \mathbb{L}^{\geq}(\mathbb{H})$ (see Model §02|01.08). Moreover, for $\Gamma^X \in \mathbb{T} \subseteq \mathbb{L}^{\geq}(\mathbb{H})$ and $f \in \mathbb{F}_2 \subseteq \mathbb{H}$, and hence $g := \Gamma^X f \in \mathbb{H}$, we denote by $\mathbb{P}_{f|\Gamma^X}$ the joint distribution of (X, Y) without fully specifying the distribution which however is assumed to satisfy $\mathbb{P}_{f|\Gamma^X}(Y\langle X, x\rangle_{\mathbb{H}}) = \mathbb{P}_{\Gamma^X}(\langle X, f\rangle_{\mathbb{H}}\langle X, x\rangle_{\mathbb{H}})$ for all $x \in \mathbb{H}$. Let $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \ell_2)$ be a generalised Fourier series transform as in Notation §01|02.07 which is fixed and *known* in advance. Following Model §02|01.08 $\Gamma^X_{\bullet|\bullet} := \mathrm{U}\Gamma^X\mathrm{U}^\star \in \mathbb{L}^{\geq}(\ell_2)$ is an operator with kernel (infinite matrix) $\Gamma^X_{\bullet|\bullet} \in \mathcal{M}(2^{\mathbb{N}^2})$ satisfying $\Gamma^X_{\bullet|\bullet} = (\Gamma^X_{j|j_\circ} = \langle \Gamma^X \mathrm{u}_{j_\circ}, \mathrm{u}_j\rangle_{\mathbb{H}} = \mathbb{P}_{\Gamma^X}(\langle X, \mathrm{u}_{j_\circ}\rangle_{\mathbb{H}}\langle X, \mathrm{u}_j\rangle_{\mathbb{H}}))_{j,j_\circ\in\mathbb{N}}$. Therefore the stochastic process $\psi_{\bullet|\bullet} = (\psi_{j|j_\circ}(X) := \langle X, \mathrm{u}_{j_\circ}\rangle_{\mathbb{H}}\langle X, \mathrm{u}_j\rangle_{\mathbb{H}})_{j,j_\circ\in\mathbb{N}} \in \mathcal{M}(\mathscr{B}_{\mathbb{H}} \otimes 2^{\mathbb{N}^2})$ fulfils Assumption §02|01.02 and $\Gamma^X_{\bullet|\bullet} = \mathbb{P}_{\Gamma^X}(\psi_{\bullet|\bullet})$. Moreover, for each $f \in \mathbb{F}_2 \subseteq \mathbb{H}$ the generalised Fourier coefficients $g_\bullet = (g_j)_{j\in\mathbb{N}} = \mathrm{U}g = \mathrm{U}\Gamma^X\mathrm{U}^\star\mathrm{U}f = \Gamma^X_{\bullet|\bullet} f_\bullet$, satisfy $g_\bullet = \mathbb{P}_{f|\Gamma^X}(Y\langle X, \mathrm{u}_\bullet\rangle_{\mathbb{H}})$. The stochastic process $\psi_\bullet = (\psi_j := Y\langle X, \mathrm{u}_j\rangle_{\mathbb{H}})_{j\in\mathbb{N}} \in \mathcal{M}((\mathscr{B}_{\mathbb{H}} \otimes \mathscr{B}) \otimes 2^{\mathbb{N}})$ fulfils Assumption §01|01.04 and $g_\bullet = \Gamma^X_{\bullet|\bullet} f_\bullet = \mathbb{P}_{f|\Gamma^X}(\psi_\bullet)$. The observable noisy versions take the form $\widehat{g}_\bullet = \Gamma^X_{\bullet|\bullet} f_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ and $\widehat{\Gamma}^X_{\bullet|\bullet} = \Gamma^X_{\bullet|\bullet} + k^{-1/2}\dot{\eta}_{\bullet|\bullet}$, or in equal (02.03) with error processes

$$\dot{\varepsilon}_\bullet = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_{f|\Gamma^X})(\psi_\bullet) = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_\bullet) - \mathbb{P}_{f|\Gamma^X}(\psi_\bullet)) \in \mathcal{M}((\mathscr{B}_{\mathbb{H}} \otimes \mathscr{B})^{\otimes n} \otimes 2^{\mathbb{N}}) \quad \text{and}$$
$$\dot{\eta}_{\bullet|\bullet} = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_{\Gamma^X})(\psi_{\bullet|\bullet}) = k^{1/2}(\widehat{\mathbb{P}}_k(\psi_{\bullet|\bullet}) - \mathbb{P}_{\Gamma^X}(\psi_{\bullet|\bullet})) \in \mathcal{M}(\mathscr{B}_{\mathbb{H}}^{\otimes k} \otimes 2^{\mathbb{N}^2})$$

satisfying Assumption §01|01.04 and Assumption §02|01.02. □

§02|02.09 **Functional linear instrumental regression**. Let $(\mathbb{H}, \langle\cdot,\cdot\rangle_{\mathbb{H}})$ and $(\mathbb{G}, \langle\cdot,\cdot\rangle_{\mathbb{G}})$ be separable Hilbert space equipped with its Borel-σ-algebra $\mathscr{B}_{\mathbb{H}}$ and $\mathscr{B}_{\mathbb{G}}$, respectively, and let (Z, X, Y) be an $\mathbb{G}\times\mathbb{H}\times\mathbb{R}$-valued random vector with joint distribution $\mathbb{P}^{Z,X,Y} \in \mathscr{W}(\mathscr{B}_{\mathbb{G}} \otimes \mathscr{B}_{\mathbb{H}} \otimes \mathscr{B})$. We denote by $\mathbb{P}^Z \in \mathscr{W}(\mathscr{B}_{\mathbb{Z}})$, $\mathbb{P}^{Z,X} \in \mathscr{W}(\mathscr{B}_{\mathbb{G}} \otimes \mathscr{B}_{\mathbb{H}})$, and $\mathbb{P}^{Z,Y} \in \mathscr{W}(\mathscr{B}_{\mathbb{G}} \otimes \mathscr{B})$ the marginal distribution of Z, (Z, X) and (Z, Y), respectively. We tactically identify Z, X and Y with the coordinate map $\Pi_{\mathbb{G}}$, $\Pi_{\mathbb{H}}$ and $\Pi_{\mathbb{R}}$, respectively, and thus (Z, X, Y) with the identity $\mathrm{id}_{\mathbb{G}\times\mathbb{H}\times\mathbb{R}}$ such that $\mathbb{P} = \mathbb{P}^{Z,X,Y} \in \mathscr{W}(\mathscr{B}_{\mathbb{G}} \otimes \mathscr{B}_{\mathbb{H}} \otimes \mathscr{B})$. Here and subsequently, we assume that $Y, \|Z\|_{\mathbb{G}}^2, \|X\|_{\mathbb{H}}^2 \in \mathbb{L}_2(\mathbb{P}) = \mathbb{L}_2(\mathbb{G}\times\mathbb{H}\times\mathbb{R}, \mathscr{B}_{\mathbb{G}} \otimes \mathscr{B}_{\mathbb{H}} \otimes \mathscr{B}, \mathbb{P})$, $\mathbb{P}(\langle z, Z\rangle_{\mathbb{G}}) = 0$ and $\mathbb{P}(\langle x, X\rangle_{\mathbb{H}}) = 0$ for all $z \in \mathbb{G}$ and $x \in \mathbb{H}$. In this situation (Z, X) admits a *cross-covariance operator* $\Gamma^{ZX} \in \mathbb{L}(\mathbb{H}, \mathbb{G})$ (see Model §02|01.09) and there is $g \in \mathbb{G}$ satisfying $\langle g, z\rangle_{\mathbb{G}} = \mathbb{P}(Y\langle Z, z\rangle_{\mathbb{G}})$ for all $z \in \mathbb{G}$. In what follows we assume that in addition $g \in \mathrm{ran}(\Gamma^{ZX}) \subseteq \mathbb{G}$. In this situation there exists $f \in \mathbb{H}$ such that

$$\langle g, z\rangle_{\mathbb{G}} = \mathbb{P}(Y\langle Z, z\rangle_{\mathbb{G}}) = \mathbb{P}(\langle X, f\rangle_{\mathbb{H}}\langle Z, z\rangle_{\mathbb{G}}) = \langle \Gamma^{ZX} f, z\rangle_{\mathbb{G}}$$

or in equal $\mathbb{P}$-a.s. we have $Y = \langle X, f\rangle_{\mathbb{H}} + \xi$ with $\mathbb{P}(\xi\langle Z, z\rangle_{\mathbb{G}}) = 0$ for all $z \in \mathbb{G}$. Let us denote by $\mathbb{P}_{\Gamma^{ZX}} \in \mathscr{W}(\mathscr{B}_{\mathbb{G}} \otimes \mathscr{B}_{\mathbb{H}})$ the marginal destribution of (Z, X) which is not fully specified given $\Gamma^{ZX} \in \mathbb{T} \subseteq \mathbb{L}(\mathbb{H}, \mathbb{G})$ (see Model §02|01.09). Moreover, for $\Gamma^{ZX} \in \mathbb{T} \subseteq \mathbb{L}(\mathbb{H}, \mathbb{G})$ and $f \in \mathbb{F}_2 \subseteq \mathbb{H}$, and hence $g := \Gamma^{ZX} f \in \mathbb{G}$, we denote by $\mathbb{P}_{f|\Gamma^{ZX}}$ the joint distribution of (Z, X, Y) without fully specifying the distribution which however is assumed to satisfy $\mathbb{P}_{f|\Gamma^{ZX}}(Y\langle Z, z\rangle_{\mathbb{G}}) = \mathbb{P}_{\Gamma^{ZX}}(\langle X, f\rangle_{\mathbb{H}}\langle Z, z\rangle_{\mathbb{G}})$

for all $z \in \mathbb{G}$. Let $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \ell_2)$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \ell_2)$ be generalised Fourier series transforms as in Notation §01|02.07 which are fixed and *known* in advance. Following Model §02|01.09 $\Gamma^{ZX}_{\cdot|\cdot} := \mathrm{V}\Gamma^{ZX}\mathrm{U}^\star \in \mathbb{L}_\cdot(\ell_2)$ is an operator with kernel (infinite matrix) $\Gamma^{ZX}_{\cdot|\cdot} \in \mathcal{M}(2^{\mathbb{N}^2})$ satisfying $\Gamma^{ZX}_{\cdot|\cdot} = (\Gamma^{ZX}_{j|j_\circ} = \langle \Gamma^{ZX}\mathrm{u}_{j_\circ}, \mathrm{v}_j\rangle_{\mathbb{G}} = \mathbb{P}_{\Gamma^{zx}}(\langle X, \mathrm{u}_{j_\circ}\rangle_{\mathbb{H}}\langle Z, \mathrm{v}_j\rangle_{\mathbb{G}}))_{j,j_\circ\in\mathbb{N}}$. Therefore the stochastic process $\psi_{\cdot|\cdot} = (\psi_{j|j_\circ}(Z,X) := \langle X, \mathrm{u}_{j_\circ}\rangle_{\mathbb{H}}\langle Z, \mathrm{v}_j\rangle_{\mathbb{G}})_{j,j_\circ\in\mathbb{N}} \in \mathcal{M}(\mathscr{B}_{\mathbb{G}}\otimes\mathscr{B}_{\mathbb{H}}\otimes 2^{\mathbb{N}^2})$ fulfils Assumption §02|01.02 and $\Gamma^{ZX}_{\cdot|\cdot} = \mathbb{P}_{\Gamma^{zx}}(\psi_{\cdot|\cdot})$. Moreover, for each $f \in \mathbb{F}_2 \subseteq \mathbb{H}$ the generalised Fourier coefficients $g_\cdot = (g_j)_{j\in\mathbb{N}} = \mathrm{V}g = \mathrm{V}\Gamma^{ZX}\mathrm{U}^\star\mathrm{U}f = \Gamma^{ZX}_{\cdot|\cdot}f_\cdot$, satisfy $g_\cdot = \mathbb{P}_{f|\Gamma^{zx}}(Y\langle Z, \mathrm{v}_\cdot\rangle_{\mathbb{G}})$. The stochastic process $\psi_\cdot = (\psi_j(Z,Y) := Y\langle Z, \mathrm{v}_j\rangle_{\mathbb{G}})_{j\in\mathbb{N}} \in \mathcal{M}((\mathscr{B}_{\mathbb{G}}\otimes\mathscr{B})\otimes 2^{\mathbb{N}})$ fulfils Assumption §01|01.04 and $g_\cdot = \Gamma^{ZX}_{\cdot|\cdot}f_\cdot = \mathbb{P}_{f|\Gamma^{zx}}(\psi_\cdot)$. The observable noisy versions take the form $\widehat{g}_\cdot = \Gamma^{ZX}_{\cdot|\cdot}f_\cdot + n^{-1/2}\dot{\varepsilon}_\cdot$ and $\widehat{\Gamma}^{ZX}_{\cdot|\cdot} = \Gamma^{ZX}_{\cdot|\cdot} + k^{-1/2}\dot{\eta}_{\cdot|\cdot}$, or in equal (02.03) with error processes

$$\dot{\varepsilon}_\cdot = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_{f|\Gamma^{zx}})(\psi_\cdot) = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_\cdot) - \mathbb{P}_{f|\Gamma^{zx}}(\psi_\cdot)) \in \mathcal{M}((\mathscr{B}_{\mathbb{G}}\otimes\mathscr{B})^{\otimes n}\otimes 2^{\mathbb{N}}) \quad \text{and}$$
$$\dot{\eta}_{\cdot|\cdot} = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_{\Gamma^{zx}})(\psi_{\cdot|\cdot}) = k^{1/2}(\widehat{\mathbb{P}}_k(\psi_{\cdot|\cdot}) - \mathbb{P}_{\Gamma^{zx}}(\psi_{\cdot|\cdot})) \in \mathcal{M}((\mathscr{B}_{\mathbb{G}}\otimes\mathscr{B}_{\mathbb{H}})^{\otimes k}\otimes 2^{\mathbb{N}^2})$$

satisfying Assumption §01|01.04 and Assumption §02|01.02. □

§02|02.10 **Regression with unknown design.** Consider the measure space $(\mathcal{D}, \mathscr{B}_{\mathcal{D}}, \lambda_{\mathcal{D}})$ where $\lambda_{\mathcal{D}}$ denotes the restriction of the Lebesgue measure to the Borel-σ-algebra $\mathscr{B}_{\mathcal{D}}$ over $\mathcal{D} \in \mathscr{B}$, and the *real* Hilbert space $\mathbb{L}_2(\lambda_{\mathcal{D}}) := \mathbb{L}_2(\mathcal{D}, \mathscr{B}_{\mathcal{D}}, \lambda_{\mathcal{D}})$ of square Lebesgue-integrable real-valued functions. Let (X,Y) be a $\mathcal{D}\times\mathbb{R}$-valued random vector. We assume in what follows that the marginal distribution $\mathbb{P}^X \in \mathscr{W}(\mathscr{B}_{\mathcal{D}})$ of the regressor X admits a Lebesgue density $\varphi \in \mathbb{L}_1(\lambda_{\mathcal{D}})$, that is $\mathbb{P}^X = \varphi\lambda_{\mathcal{D}}$, which is *not known* in advance. Moreover, let the joint distribution of (X,Y) be given by $\mathbb{P}^{X,Y}_{f|\varphi} := \varphi\lambda_{\mathcal{D}} \odot \mathbb{P}^{\xi}_{f(X)}$ without fully specifying the error distribution $\mathbb{P}^{\xi} \in \mathscr{W}(\mathscr{B})$ and thus the regular conditional distribution $\mathbb{P}^{Y|X}_f = \mathbb{P}^{\xi}_{f(X)}$ (compare Model §01|05.10). We tactically identify X and Y with the coordinate map $\Pi_{\mathcal{D}}$ and $\Pi_{\mathbb{R}}$, respectively, and thus (X,Y) with the identity $\mathrm{id}_{\mathcal{D}\times\mathbb{R}}$ such that $\mathbb{P}_{f|\varphi} = \mathbb{P}^{X,Y}_{f|\varphi} \in \mathscr{W}(\mathscr{B}_{\mathcal{D}}\otimes\mathscr{B})$. In addition we assume that $\varphi \in \mathbb{L}_\infty(\lambda_{\mathcal{D}})$, and hence $\mathrm{M}_\varphi \in \mathbb{L}^{\mathrm{M}}(\mathbb{L}_2(\lambda_{\mathcal{D}}))$, $\mathbb{P}^{\xi} \in \mathrm{P}_{\{0\}\times\mathbb{R}_{\geq 0}} \subseteq \mathscr{W}_2(\mathscr{B})$, i.e. ξ has mean zero and a finite second moment, and $f \in \mathbb{L}_2(\lambda_{\mathcal{D}})$, then $g = \mathrm{M}_\varphi f \in \mathbb{L}_2(\lambda_{\mathcal{D}})$ for each $h \in \mathbb{L}_2(\lambda_{\mathcal{D}})$ satisfies

$$\mathbb{P}_{f|\varphi}(Yh(X)) = \mathbb{P}^X(\mathbb{P}^{Y|X}_f(Y)h) = \varphi\lambda_{\mathcal{D}}(fh) = \lambda_{\mathcal{D}}(\varphi fh) = \langle \mathrm{M}_\varphi f, h\rangle_{\mathbb{L}_2(\lambda_{\mathcal{D}})} = \langle g, h\rangle_{\mathbb{L}_2(\lambda_{\mathcal{D}})} \in \mathbb{R}.$$

Let $\mathrm{U} \in \mathbb{L}(\mathbb{L}_2(\lambda_{\mathcal{D}}), \ell_2)$ be a generalised Fourier series transform as in Notation §01|02.07 which is fixed and *known in advance*. Evidently, we have $\mathrm{M}^{\varphi}_{\cdot|\cdot} := \mathrm{U}\mathrm{M}_\varphi\mathrm{U}^\star \in \mathbb{L}^{\geq}_\cdot(\ell_2) \subseteq \mathbb{L}_\cdot(\ell_2) = \mathbb{L}(\ell_2)$ and for each $f \in \mathbb{F}_2 \subseteq \mathbb{L}_2(\lambda_{\mathcal{D}})$ and $g := \mathrm{M}_\varphi f \in \mathbb{L}_2(\lambda_{\mathcal{D}})$ the generalised Fourier coefficients $f_\cdot = (f_j)_{j\in\mathbb{N}} = \mathrm{U}f$ and $g_\cdot = (g_j)_{j\in\mathbb{N}} = \mathrm{U}g = \mathrm{M}^{\varphi}_{\cdot|\cdot}f_\cdot$ for each $j \in \mathbb{N}$ satisfy

$$g_j = \langle \mathrm{M}^{\varphi}_{j|\cdot}, f_\cdot\rangle_{\ell_2} = \langle \mathrm{M}^{\varphi}_{\cdot|\cdot}f_\cdot, \mathbb{1}^{\{j\}}_\cdot\rangle_{\ell_2} = \langle \mathrm{M}_\varphi f, \mathrm{U}^\star\mathbb{1}^{\{j\}}_\cdot\rangle_{\mathbb{L}_2(\lambda_{\mathcal{D}})} = \langle \mathrm{M}_\varphi f, \mathrm{u}_j\rangle_{\mathbb{L}_2(\lambda_{\mathcal{D}})} = \mathbb{P}_{f|\varphi}(Y\mathrm{u}_j(X)) \in \mathbb{R}$$

The stochastic process $\psi_\cdot = (\psi_j(X,Y) := Y\mathrm{u}_j(X))_{j\in\mathbb{N}} \in \mathcal{M}((\mathscr{B}_{\mathcal{D}}\otimes\mathscr{B})\otimes 2^{\mathbb{N}})$ fulfils Assumption §01|01.04 and $g_\cdot = \mathrm{M}^{\varphi}_{\cdot|\cdot}f_\cdot = \mathbb{P}_{f|\varphi}(\psi_\cdot)$ (compare Model §01|05.10). Moreover, considering the marginal distribution $\mathbb{P}_\varphi = \varphi\lambda_{\mathcal{D}}$ of X we have $\mathrm{M}^{\varphi}_{\cdot|\cdot} := \mathrm{U}\mathrm{M}_\varphi\mathrm{U}^\star \in \mathbb{L}^{\geq}_\cdot(\ell_2) \subseteq \mathbb{L}_\cdot(\ell_2) = \mathbb{L}(\ell_2)$ satisfying $\mathrm{M}^{\varphi}_{\cdot|\cdot} = (\mathrm{M}^{\varphi}_{j|j_\circ} = \langle \mathrm{M}^{\varphi}_{\cdot|\cdot}\mathrm{u}_{j_\circ}, \mathrm{u}_j\rangle_{\mathbb{G}} = \mathbb{P}_\varphi(\mathrm{u}_{j_\circ}\mathrm{u}_j))_{j,j_\circ\in\mathbb{N}}$. Therefore the stochastic process $\psi_{\cdot|\cdot} = (\psi_{j|j_\circ} := \mathrm{u}_{j_\circ}(X)\mathrm{u}_j(X))_{j,j_\circ\in\mathbb{N}} \in \mathcal{M}(\mathscr{B}_{\mathcal{D}}\otimes 2^{\mathbb{N}^2})$ fulfils Assumption §02|01.02 and $\mathrm{M}^{\varphi}_{\cdot|\cdot} = \mathbb{P}_\varphi(\psi_{\cdot|\cdot})$ (compare Model §02|01.10). The observable noisy versions take the form $\widehat{g}_\cdot = \mathrm{M}^{\varphi}_{\cdot|\cdot}f_\cdot + n^{-1/2}\dot{\varepsilon}_\cdot$ and $\widehat{\mathrm{M}}_{\cdot|\cdot} = \mathrm{M}^{\varphi}_{\cdot|\cdot} + n^{-1/2}\dot{\eta}_{\cdot|\cdot}$, or in equal (02.03) with error processes

$$\dot{\varepsilon}_\cdot = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_{f|\varphi})(\psi_\cdot) = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_\cdot) - \mathbb{P}_{f|\varphi}(\psi_\cdot)) \in \mathcal{M}((\mathscr{B}_{\mathcal{D}}\otimes\mathscr{B})^{\otimes n}\otimes 2^{\mathbb{N}}) \quad \text{and}$$
$$\dot{\eta}_{\cdot|\cdot} = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_\varphi)(\psi_{\cdot|\cdot}) = k^{1/2}(\widehat{\mathbb{P}}_k(\psi_{\cdot|\cdot}) - \mathbb{P}_\varphi(\psi_{\cdot|\cdot})) \in \mathcal{M}(\mathscr{B}_{\mathcal{D}}^{\otimes n}\otimes 2^{\mathbb{N}^2})$$

satisfying Assumption §01|01.04 and Assumption §02|01.02. □

§02|03 Noisy diagonal operator

§02|03.01 **Notation.** Under Assumption §02|00.01 we consider the reconstruction of $\theta_{\bullet} = \mathrm{U}\theta \in \mathbb{J}$ (or in equal $\theta = \mathrm{U}^{\star}\theta_{\bullet} \in \mathbb{H}$) from noisy versions of $\mathrm{V}g = g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathbb{J}$ and $\mathfrak{s}_{\bullet} \in \mathbb{L}_{\infty}(\nu)$ or in equal $\mathrm{M}_{\mathfrak{s}} = \mathrm{VTU}^{\star} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$. □

§02|03.02 **Noisy diagonal operator.** Let $\dot{\eta}_{\bullet} = (\dot{\eta}_j)_{j\in\mathcal{J}}$ be a stochastic process satisfying Assumption §01|01.04 with mean zero and let $k \in \mathbb{N}$ be a sample size. The stochastic process $\widehat{\mathfrak{s}}_{\bullet} = \mathfrak{s}_{\bullet} + k^{-1/2}\dot{\eta}_{\bullet}$ with mean function $\mathfrak{s}_{\bullet} \in \mathscr{J}$ is called a *noisy version* of $\mathfrak{s}_{\bullet} \in \mathbb{L}_{\infty}(\nu)$ and hence the diagonal operator $\mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$, or *noisy diagonal operator* for short. We denote by $\mathbb{P}^k_{\mathfrak{s}}$ the distribution of $\widehat{\mathfrak{s}}_{\bullet}$. If $\dot{\eta}_{\bullet}$ admits a covariance function (possibly depending on $\mathfrak{s}$), say $\mathrm{cov}^{\mathfrak{s}}_{\bullet,\bullet} \in \mathscr{J}^2$, then we eventually write $\dot{\eta}_{\bullet} \sim \mathrm{P}_{(0_{\bullet},\mathrm{cov}^{\mathfrak{s}}_{\bullet,\bullet})}$ and $\widehat{\mathfrak{s}}_{\bullet} \sim \mathrm{P}_{(\mathfrak{s}_{\bullet},k^{-1}\mathrm{cov}^{\mathfrak{s}}_{\bullet,\bullet})}$ for short. □

§02|03.03 **Comment.** Similar to a noisy image (Definition §01|02.03) we consider Empirical mean model §01|02.04, Sequence model §01|02.05 or Gaussian sequence model §01|02.06. Examples are provided in Subsubsection §01|02|01. □

§02|03|01 Examples of empirical mean models

§02|03.04 **Covariance operator under second order stationarity.** Consider the *complex* Hilbert spaces $\mathbb{L}_2(\lambda_{[0,1)})$ and $\mathbb{J} := \ell_2(\mathbb{Z})$. Let $(\mathbb{L}_2(\lambda_{[0,1)}), \langle\cdot,\cdot\rangle_{\mathbb{L}_2(\lambda_{[0,1)})})$ be equipped with its Borel-σ-algebra $\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})}$ and X be a $\mathbb{L}_2(\lambda_{[0,1)})$-valued *real random function* (by the usual embedding of real-valued functions as in Notation §01|02.10). We tactically identify X with the identity $\mathrm{id}_{\mathbb{L}_2(\lambda_{[0,1)})}$ on $\mathbb{L}_2(\lambda_{[0,1)})$ such that X is defined on the measure space $(\mathbb{L}_2(\lambda_{[0,1)}), \mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})}, \mathbb{P})$ and $X \sim \mathbb{P} = \mathbb{P}^X \in \mathscr{W}(\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})})$. Here and subsequently, we assume that $\|X\|^2_{\mathbb{L}_2(\lambda_{[0,1)})} \in \mathbb{L}_2(\mathbb{P})$ and $\mathbb{P}(\langle X, x\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = 0$ for all $x \in \mathbb{L}_2(\lambda_{[0,1)})$. In this situation X admits a *covariance operator* $\Gamma^X \in \mathbb{L}^{\geqslant}(\mathbb{L}_2(\lambda_{[0,1)}))$ (see Remark §01|01.07). Moreover, let X be *second order stationary*, i.e. there exists $\mathrm{c}^X \in \mathcal{M}(\mathscr{B}_{[0,1]})$ such that

$$\mathrm{cov}_{t,s} = \mathbb{C}\mathrm{ov}(X(t), X(s)) = \mathrm{c}^X(t - s - \lfloor t - s\rfloor), \quad \forall s,t \in [0,1).$$

Evidently, since $\|X\|^2_{\mathbb{L}_2(\lambda_{[0,1)})} \in \mathbb{L}_2(\mathbb{P})$ we have

$$\begin{aligned}\|\mathrm{c}^X\|^2_{\mathbb{L}_2(\lambda_{[0,1)})} = \lambda_{[0,1)}(|\mathrm{c}^X|^2) &= \int_{[0,1)} |\mathbb{C}\mathrm{ov}(X(0), X(t))|^2 \lambda_{[0,1)}(dt)\\ &\leqslant \mathbb{P}(|X(0)|^2)\int_{[0,1)} \mathbb{P}(|X(t)|^2)\lambda_{[0,1)}(dt) = \mathbb{P}(|X(0)|^2)\mathbb{P}(\|X\|^2_{\mathbb{L}_2(\lambda_{[0,1)})}) \in \mathbb{R}_{\geqslant 0}\end{aligned}$$

and hence $\mathrm{c}^X \in \mathbb{L}_2(\lambda_{[0,1)})$ too. Furthermore, the *covariance operator* $\Gamma^X \in \mathbb{L}^{\geqslant}(\mathbb{L}_2(\lambda_{[0,1)}))$ equals a *circular additive convolution* (see Notation §01|04.11), since for all $x, y \in \mathbb{L}_2(\lambda_{[0,1)})$

$$\begin{aligned}\langle\Gamma^X x, y\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} &= \mathbb{C}\mathrm{ov}(\langle X, y\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}, \langle X, x\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = \mathbb{P}(\lambda_{[0,1)}(X\overline{y})\overline{\lambda_{[0,1)}(X\overline{x})})\\ &= \int_{[0,1)}\int_{[0,1)} \overline{y}(t)\,\mathbb{C}\mathrm{ov}(X(t), X(s))x(s)\lambda_{[0,1)}(ds)\lambda_{[0,1)}(dt)\\ &= \int_{[0,1)} \overline{y}(t)\int_{[0,1)} \mathrm{c}^X(t - s - \lfloor t-s\rfloor)x(s)\lambda_{[0,1)}(ds)\lambda_{[0,1)}(dt)\\ &= \int_{[0,1)} (\mathrm{c}^X \circledast x)(t)\overline{y}(t)\lambda_{[0,1)}(dt) = \langle\circledast_{\mathrm{c}^X} x, y\rangle_{\mathbb{L}_2(\lambda_{[0,1)})},\end{aligned}$$

hence $\circledast_{\mathrm{c}^X} = \Gamma^X \in \mathbb{L}^{\geqslant}(\mathbb{L}_2(\lambda_{[0,1)}))$ in short. Let $\mathrm{F} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}), \ell_2(\mathbb{Z}))$ be the *Fourier-series transform* with $h \mapsto \mathrm{F}h := h_{\bullet} = \lambda_{[0,1)}(h\overline{\mathrm{e}}_{\bullet})$ and exponential basis $\mathrm{e}_{\bullet} := (\mathrm{e}_j)_{j\in\mathbb{Z}}$ (see Notations §01|02.10

and §01|02.12). Since $\mathrm{c}^X \in \mathbb{L}_2(\lambda_{[0,1)})$ we denote $\mathrm{c}^X_\bullet = \mathrm{F}\mathrm{c}^X$. Then the *circular convolution theorem* states $(\circledast_{\mathrm{c}^X} h)_\bullet = \mathrm{F}(\circledast_{\mathrm{c}^X} h) = (\mathrm{F}\mathrm{c}^X)(\mathrm{F}h) = \mathrm{c}^X_\bullet h_\bullet$. Consequently, $(\mathrm{c}^X_\bullet, \mathrm{F}, \mathrm{F})$ is an eigen value decomposition of $\circledast_{\mathrm{c}^X} \in \mathbb{L}^{\geqslant}(\mathbb{L}_2(\lambda_{[0,1)}))$ with $\mathrm{c}^X_\bullet \in \ell_2(\mathbb{Z}) \subseteq \ell_\infty(\mathbb{Z})$, and thus $\Gamma^X = \circledast_{\mathrm{c}^X} \in \mathbb{L}^{\mathrm{F},\mathrm{F}}(\mathbb{L}^{\mathrm{M}}(\ell_2(\mathbb{Z}))) = \mathrm{F}^\star(\mathbb{L}^{\mathrm{M}}(\ell_2(\mathbb{Z})))\mathrm{F}$. Let us denote by $\mathbb{P}_{\mathrm{c}^X} \in \mathscr{W}(\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})})$ the destribution of X which is not fully specified given $\mathrm{c}^X \in \mathbb{D}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)})$. We consider the statistical product experiment

$$((\mathbb{L}_2(\lambda_{[0,1)}))^k, \mathscr{B}^{\otimes k}_{\mathbb{L}_2(\lambda_{[0,1)})}, \mathbb{P}^{\otimes k}_{\mathbb{D}_2} = (\mathbb{P}^{\otimes k}_{\mathrm{c}^X})_{\mathrm{c}^X \in \mathbb{D}_2}).$$

The stochastic process $\psi_\bullet = (\psi_j(X) := |\langle X, \mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}|^2)_{j\in\mathbb{Z}} \in \mathcal{M}(\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes 2^{\mathbb{Z}})$ fulfils Assumption §01|01.04 and $\mathrm{c}^X_\bullet = \mathbb{P}_{\mathrm{c}^X}(\psi_\bullet)$ since for each $j \in \mathbb{Z}$ we have

$$\mathbb{P}_{\mathrm{c}^X}(|\langle X, \mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}|^2) = \langle \Gamma^X \mathrm{e}_j, \mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} = \langle \circledast_{\mathrm{c}^X} \mathrm{e}_j, \mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} = \langle \mathrm{c}^X_\bullet \mathrm{F}\mathrm{e}_j, \mathrm{F}\mathrm{e}_j\rangle_{\ell_2(\mathbb{Z})}$$
$$= \langle \mathrm{c}^X_\bullet \mathbb{1}^{\{j\}}_\bullet, \mathbb{1}^{\{j\}}_\bullet\rangle_{\ell_2(\mathbb{Z})} = \mathrm{c}^X_j.$$

Similar to an Empirical mean model §01|02.04 we define $\widehat{\mathrm{c}}^X_\bullet = (\widehat{\mathrm{c}}^X_j := \widehat{\mathbb{P}}_k(\psi_j))_{j\in\mathbb{Z}} \in \mathcal{M}(\mathscr{B}^{\otimes k}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes 2^{\mathbb{Z}})$. By construction $\mathrm{c}^X_\bullet = \mathbb{P}_{\mathrm{c}^X}(\psi_\bullet) \in \mathcal{M}(2^{\mathbb{Z}})$ is the mean sequence of $\widehat{\mathrm{c}}^X_\bullet$. For each $j \in \mathbb{Z}$ the statistic $\dot{\boldsymbol{\eta}}_j := k^{1/2}(\widehat{\mathbb{P}}_k(\psi_j) - \mathbb{P}_{\mathrm{c}^X}(\psi_j)) \in \mathcal{M}(\mathscr{B}^{\otimes k}_{\mathbb{L}_2(\lambda_{[0,1)})})$ is centred, i.e. $\dot{\boldsymbol{\eta}}_j \in \mathbb{L}_1(\mathbb{P}^{\otimes k}_{\mathrm{c}^X})$ with $\mathbb{P}^{\otimes k}_{\mathrm{c}^X}(\dot{\boldsymbol{\eta}}_j) = 0$, and exploiting $\psi_\bullet \in \mathcal{M}(\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes 2^{\mathbb{Z}})$ the stochastic process

$$\dot{\boldsymbol{\eta}}_\bullet = (\dot{\boldsymbol{\eta}}_j)_{j\in\mathbb{Z}} = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_{\mathrm{c}^X})(\psi_\bullet) = k^{1/2}(\widehat{\mathbb{P}}_k(\psi_\bullet) - \mathbb{P}_{\mathrm{c}^X}(\psi_\bullet)) \in \mathcal{M}(\mathscr{B}^{\otimes k}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes 2^{\mathbb{Z}})$$

satisfies Assumption §01|01.04 and by construction $\widehat{\mathrm{c}}^X_\bullet = \mathrm{c}^X_\bullet + k^{-1/2}\dot{\boldsymbol{\eta}}_\bullet$ is a noisy version of $\mathrm{c}^X_\bullet$. □

§02|03.05 **Cross-covariance operator under second order stationarity**. Consider the *complex* Hilbert spaces $\mathbb{L}_2(\lambda_{[0,1)})$ and $\mathbb{J} := \ell_2(\mathbb{Z})$. Let $(\mathbb{L}_2(\lambda_{[0,1)}), \langle\cdot,\cdot\rangle_{\mathbb{L}_2(\lambda_{[0,1)})})$ be equipped with its Borel-σ-algebra $\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})}$ and let X and Z be a $\mathbb{L}_2(\lambda_{[0,1)})$-valued *real random function* (by the usual embedding of real-valued functions as in Notation §01|02.10). Then (Z, X) is an $((\mathbb{L}_2(\lambda_{[0,1)}))^2, \mathscr{B}^{\otimes 2}_{\mathbb{L}_2(\lambda_{[0,1)})})$-valued *random function*. We denote by $\mathbb{P}^Z, \mathbb{P}^X \in \mathscr{W}(\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})})$ the marginal distribution of Z and X, respectively, and by $\mathbb{P}^{Z,X} \in \mathscr{W}(\mathscr{B}^{\otimes 2}_{\mathbb{L}_2(\lambda_{[0,1)})})$ the joint distribution of (Z, X). We tactically take Z and X as coordinate map, and thus identify (Z, X) with the identity $\mathrm{id}_{(\mathbb{L}_2(\lambda_{[0,1)}))^2}$ such that $\mathbb{P} = \mathbb{P}^{Z,X} \in \mathscr{W}(\mathscr{B}^{\otimes 2}_{\mathbb{L}_2(\lambda_{[0,1)})})$. Here and subsequently, we assume that $\|Z\|^2_{\mathbb{L}_2(\lambda_{[0,1)})} \in \mathcal{L}_2(\mathbb{P})$, $\|X\|^2_{\mathbb{L}_2(\lambda_{[0,1)})} \in \mathcal{L}_2(\mathbb{P})$, $\mathbb{P}(\langle Z, z\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = 0$ and $\mathbb{P}(\langle X, x\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = 0$ for all $z, x \in \mathbb{L}_2(\lambda_{[0,1)})$. In this situation (Z, X) admits a *cross-covariance operator* $\Gamma^{ZX} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}))$ (see Model §02|01.09). Moreover, let (Z, X) be *second order stationary*, i.e. there exists $\mathrm{c}^{ZX} \in \mathcal{M}(\mathscr{B}_{[0,1)})$ such that

$$\mathrm{cov}^{ZX}_{t,s} = \mathbb{C}\mathrm{ov}(Z(t), X(s)) = \mathrm{c}^{ZX}(t - s - \lfloor t - s\rfloor), \quad \forall s, t \in [0,1).$$

Evidently, since $\|X\|^2_{\mathbb{L}_2(\lambda_{[0,1)})} \in \mathbb{L}_2(\mathbb{P})$ we have

$$\|\mathrm{c}^{ZX}\|^2_{\mathbb{L}_2(\lambda_{[0,1)})} = \lambda_{[0,1)}(|\mathrm{c}^{ZX}|^2) = \int_{[0,1)} |\mathbb{C}\mathrm{ov}(Z(0), X(t))|^2 \lambda_{[0,1)}(dt)$$
$$\leqslant \mathbb{P}(|Z(0)|^2) \int_{[0,1)} \mathbb{P}(|X(t)|^2)\lambda_{[0,1)}(dt) = \mathbb{P}(|Z(0)|^2)\mathbb{P}(\|X\|^2_{\mathbb{L}_2(\lambda_{[0,1)})}) \in \mathbb{R}_{\geqslant 0}$$

and hence $\mathrm{c}^{ZX} \in \mathbb{L}_2(\lambda_{[0,1)})$ too. Furthermore, the *cross-covariance operator* $\Gamma^{ZX} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}))$ equals a *circular additive convolution* (see Notation §01|04.11), since for all $x, y \in \mathbb{L}_2(\lambda_{[0,1)})$

$$\begin{aligned}\langle \Gamma^{ZX} x, z\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} &= \mathbb{Cov}(\langle Z, z\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}, \langle X, x\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = \mathbb{P}(\lambda_{[0,1)}(Z\overline{z})\overline{\lambda_{[0,1)}(X\overline{x})}) \\ &= \int_{[0,1)}\int_{[0,1)} \overline{z}(t)\, \mathbb{Cov}(Z(t), X(s)) x(s)\lambda_{[0,1)}(ds)\lambda_{[0,1)}(dt) \\ &= \int_{[0,1)} \overline{z}(t) \int_{[0,1)} c^{ZX}(t-s-\lfloor t-s\rfloor)x(s)\lambda_{[0,1)}(ds)\lambda_{[0,1)}(dt) \\ &= \int_{[0,1)} (c^{ZX} \circledast x)(t)\overline{z}(t)\lambda_{[0,1)}(dt) = \langle \circledast_{c^{ZX}} x, z\rangle_{\mathbb{L}_2(\lambda_{[0,1)})},\end{aligned}$$

hence $\circledast_{c^{ZX}} = \Gamma^{ZX} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}))$ in short. Let $\mathrm{F} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}), \ell_2(\mathbb{Z}))$ be the *Fourier-series transform* with $h \mapsto \mathrm{F}h := h_{\bullet} = \lambda_{[0,1)}(h\overline{e}_{\bullet})$ and exponential basis $e_{\bullet} := (e_j)_{j\in\mathbb{Z}}$ (see Notations §01|02.10 and §01|02.12). Since $c^{ZX} \in \mathbb{L}_2(\lambda_{[0,1)})$ we denote $c_{\bullet}^{ZX} = \mathrm{F}c^{ZX}$. Then the *circular convolution theorem* states $(\circledast_{c^{ZX}} h)_{\bullet} = \mathrm{F}(\circledast_{c^{ZX}} h) = (\mathrm{F}c^{ZX})(\mathrm{F}h) = c_{\bullet}^{ZX} h_{\bullet}$. Consequently, $(c_{\bullet}^{X}, \mathrm{F}, \mathrm{F})$ is a singular value decomposition of $\circledast_{c^{ZX}} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}))$ with $c_{\bullet}^{ZX} \in \ell_2(\mathbb{Z}) \subseteq \ell_\infty(\mathbb{Z})$, and thus $\Gamma^{ZX} = \circledast_{c^{ZX}} \in \mathbb{L}^{\mathrm{F},\mathrm{F}}(\mathbb{L}^{\mathrm{M}}(\ell_2(\mathbb{Z}))) = \mathrm{F}^\star(\mathbb{L}^{\mathrm{M}}(\ell_2(\mathbb{Z})))\mathrm{F}$. Let us denote by $\mathbb{P}_{c^{ZX}} \in \mathscr{W}(\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})})$ the joint destribution of (Z, X) which is not fully specified given $c^{ZX} \in \mathbb{D}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)})$. We consider the statistical product experiment

$$((\mathbb{L}_2(\lambda_{[0,1)}))^{2k}, \mathscr{B}^{\otimes 2k}_{\mathbb{L}_2(\lambda_{[0,1)})}, \mathbb{P}^{\otimes k}_{\mathbb{D}_2} = (\mathbb{P}^{\otimes k}_{c^{ZX}})_{c^{ZX}\in\mathbb{D}_2}).$$

The stochastic process $\psi_{\bullet} = (\psi_j(Z,X) := \langle Z, e_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}\langle e_j, X\rangle_{\mathbb{L}_2(\lambda_{[0,1)})})_{j\in\mathbb{Z}} \in \mathcal{M}(\mathscr{B}^2_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes 2^{\mathbb{Z}})$ fulfils Assumption §01|01.04 and $c_{\bullet}^{ZX} = \mathbb{P}_{c^{ZX}}(\psi_{\bullet})$ since for each $j \in \mathbb{Z}$ we have

$$\begin{aligned}\mathbb{P}_{c^{ZX}}(\langle Z, e_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}\langle e_j, X\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) &= \mathbb{Cov}(\langle Z, e_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}, \langle X, e_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) \\ &= \langle \Gamma^{ZX} e_j, e_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} = \langle \circledast_{c^{ZX}} e_j, e_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} = \langle c_{\bullet}^{ZX}\mathrm{F}e_j, \mathrm{F}e_j\rangle_{\ell_2(\mathbb{Z})} = c_j^{ZX}.\end{aligned}$$

Similar to an Empirical mean model §01|02.04 we define $\widehat{c}_{\bullet}^{ZX} = (\widehat{c}_j^{ZX} := \widehat{\mathbb{P}}_k(\psi_j))_{j\in\mathbb{Z}} \in \mathcal{M}(\mathscr{B}^{\otimes 2k}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes 2^{\mathbb{Z}})$. By construction $c_{\bullet}^{ZX} = \mathbb{P}_{c^{ZX}}(\psi_{\bullet}) \in \mathcal{M}(2^{\mathbb{Z}})$ is the mean sequence of $\widehat{c}_{\bullet}^{ZX}$. For each $j \in \mathbb{Z}$ the statistic $\dot{\eta}_j := k^{1/2}(\widehat{\mathbb{P}}_k(\psi_j) - \mathbb{P}_{c^{ZX}}(\psi_j)) \in \mathcal{M}(\mathscr{B}^{\otimes 2k}_{\mathbb{L}_2(\lambda_{[0,1)})})$ is centred, i.e. $\dot{\eta}_j \in \mathbb{L}_1(\mathbb{P}^{\otimes k}_{c^{ZX}})$ with $\mathbb{P}^{\otimes k}_{c^{ZX}}(\dot{\eta}_j) = 0$, and exploiting $\psi_{\bullet} \in \mathcal{M}(\mathscr{B}^2_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes 2^{\mathbb{Z}})$ the stochastic process

$$\dot{\eta}_{\bullet} = (\dot{\eta}_j)_{j\in\mathbb{Z}} = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_{c^{ZX}})(\psi_{\bullet}) = k^{1/2}(\widehat{\mathbb{P}}_k(\psi_{\bullet}) - \mathbb{P}_{c^{ZX}}(\psi_{\bullet})) \in \mathcal{M}(\mathscr{B}^{\otimes 2k}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes 2^{\mathbb{Z}})$$

satisfies Assumption §01|01.04 and by construction $\widehat{c}_{\bullet}^{ZX} = c_{\bullet}^{ZX} + k^{-1/2}\dot{\eta}_{\bullet}$ is a noisy version of $c_{\bullet}^{ZX}$. □

§02|04 Diagonal statistical inverse problem with noisy operator

§02|04.01 **Assumption**. For $\mathbb{J} = \mathbb{L}_2(\nu)$, surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \mathbb{J})$, fixed and presumed to be known in advance, the operator $\mathrm{T} \in \mathbb{L}^{\mathrm{U},\mathrm{V}}(\mathbb{L}^{\mathrm{M}}(\mathbb{J})) \subseteq \mathbb{L}(\mathbb{H}, \mathbb{G})$ and hence $\mathrm{M}_{\mathfrak{s}} = \mathrm{V}\mathrm{T}\mathrm{U}^\star \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$ or in equal $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$ is *not known* in advance where $g_{\bullet} = \mathrm{V}\mathrm{T}\theta = \mathrm{M}_{\mathfrak{s}}\theta_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathbb{J}$ or in equal $g_{\bullet} \in \mathbb{J}\mathfrak{s}_{\bullet}$. □

§02|04.02 **Notation**. Under Assumption §02|04.01 given $g_{\bullet} \in \mathbb{J}\mathfrak{s}_{\bullet}$ for $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$ we consider the reconstruction of $\theta_{\bullet} = \mathrm{U}\theta \in \mathbb{J}$ (or in equal $\theta = \mathrm{U}^\star\theta_{\bullet} \in \mathbb{H}$) from a noisy version of $g_{\bullet} = \mathrm{V}g = \mathrm{M}_{\mathfrak{s}}\theta_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathbb{J}$ and a noisy version of $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$. Keep in mind, that we identify the equivalence class and its representative $g_{\bullet}$. □

§02|04.03 **Diagonal statistical inverse problem with noisy operator.** Consider as in Definition §01|02.03 stochastic processes $\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathcal{J}}$ and $\dot{\eta}_{\bullet} = (\dot{\eta}_j)_{j\in\mathcal{J}}$ satisfying Assumption §01|01.04 with mean zero and sample sizes $n, k \in \mathbb{N}$. Under Assumption §02|04.01 where $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$ is *not known* in advanced, the observable noisy image and operator, respectively, has $\mathbb{J}$-mean $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet}$ and mean-function $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$, and takes the form $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ and $\widehat{\mathfrak{s}}_{\bullet} = \mathfrak{s}_{\bullet} + k^{-1/2}\dot{\eta}_{\bullet}$ or in equal

$$\widehat{g}_j = \mathfrak{s}_j\theta_j + n^{-1/2}\dot{\varepsilon}_j \quad \text{and} \quad \widehat{\mathfrak{s}}_j = \mathfrak{s}_j + k^{-1/2}\dot{\eta}_j, \quad j \in \mathcal{J}. \tag{02.06}$$

We denote by $\mathbb{P}^{n,k}_{\theta|\mathfrak{s}}$ the joint distribution of $(\widehat{g}_{\bullet}, \widehat{\mathfrak{s}}_{\bullet})$. The reconstruction of $\theta_{\bullet} \in \mathbb{J}$ (in equal $\theta = \mathrm{U}^\star\theta_{\bullet} \in \mathbb{H}$) from a noisy version $(\widehat{g}_{\bullet}, \widehat{\mathfrak{s}}_{\bullet}) \sim \mathbb{P}^{n,k}_{\theta|\mathfrak{s}}$ of the image $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathbb{J}$ and $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$ is called a *diagonal statistical inverse problem with noisy operator*. □

§02|04.04 **Diagonal inverse empirical mean model (dieMM) with noisy operator.** Consider the reconstruction of $\theta_{\bullet} \in \mathbb{J}$ (in equal $\theta = \mathrm{U}^\star\theta_{\bullet} \in \mathbb{H}$) in an Empirical mean model as in §01|02.04. Under Assumption §02|04.01, where $\mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$ with $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$ is *not known* in advance, the observable noisy image has $\mathbb{J}$-mean $\mathrm{V}g = g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathbb{J}$ and the observable noisy diagonal operator has mean function $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$, and takes each the form of an Empirical mean model as in §01|02.04. More precisely, for each $\theta \in \Theta \subseteq \mathbb{H}$ and $\mathfrak{s}_{\bullet} \in \mathcal{S} \subseteq \mathbb{L}_\infty(\nu)$ let $\mathbb{P}_{\theta|\mathfrak{s}} \in \mathscr{W}(\mathscr{Z})$ be a probability measure on a measurable space $(\mathcal{Z}, \mathscr{Z})$. Similar to §01|02.04 consider stochastic processes $\psi^{\Theta}_{\bullet}, \psi^{\mathcal{S}}_{\bullet} \in \mathscr{Z} \otimes \mathscr{J}$ which in addition for all $\theta \in \Theta$ and $\mathfrak{s}_{\bullet} \in \mathcal{S}$ satisfy $\psi^{\Theta}_j, \psi^{\mathcal{S}}_j \in \mathcal{L}_1(\mathbb{P}_{\theta|\mathfrak{s}})$ for each $j \in \mathcal{J}$ and $\mathbb{P}_{\theta|\mathfrak{s}}(\psi^{\Theta}_{\bullet}) = g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet}$ and $\mathbb{P}_{\theta|\mathfrak{s}}(\psi^{\mathcal{S}}_{\bullet}) = \mathfrak{s}_{\bullet}$. The observable noisy versions take the form $\widehat{g}_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ and $\widehat{\mathfrak{s}}_{\bullet} = \mathfrak{s}_{\bullet} + k^{-1/2}\dot{\eta}_{\bullet}$, or in equal (02.06) with error processes

$$\dot{\varepsilon}_{\bullet} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_{\theta|\mathfrak{s}})(\psi^{\Theta}_{\bullet}) = n^{1/2}(\widehat{\mathbb{P}}_n(\psi^{\Theta}_{\bullet}) - \mathbb{P}_{\theta|\mathfrak{s}}(\psi^{\Theta}_{\bullet})) \in \mathcal{M}(\mathscr{Z}^{\otimes n} \otimes \mathscr{J}) \quad \text{and}$$
$$\dot{\eta}_{\bullet} = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_{\theta|\mathfrak{s}})(\psi^{\mathcal{S}}_{\bullet}) = k^{1/2}(\widehat{\mathbb{P}}_k(\psi^{\mathcal{S}}_{\bullet}) - \mathbb{P}_{\theta|\mathfrak{s}}(\psi^{\mathcal{S}}_{\bullet})) \in \mathcal{M}(\mathscr{Z}^{\otimes k} \otimes \mathscr{J})$$

satisfying Assumption §01|01.04. □

§02|04.05 **Diagonal inverse sequence model (diSM) with noisy operator.** Consider $\mathbb{J} = \ell_2 = \mathbb{L}_2(\nu_{\mathbb{N}})$ as in §01|01.14. Let $\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j\in\mathbb{N}}$ and $\dot{\eta}_{\bullet} = (\dot{\eta}_j)_{j\in\mathbb{N}}$ be real-valued stochastic processes satisfying Assumption §01|01.04 with mean zero and let $n, k \in \mathbb{N}$ be sample sizes. Under Assumption §02|04.01, where $\mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$ with $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$ is *not known* in advance, the observable noisy image has ℓ_2-mean $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet}$ and the observable noisy operator has mean function $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$, and take both the form of a Sequence model as in §01|02.05, that is $\widehat{g}_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ and $\widehat{\mathfrak{s}}_{\bullet} = \mathfrak{s}_{\bullet} + k^{-1/2}\dot{\eta}_{\bullet}$ or in equal

$$\widehat{g}_j = \mathfrak{s}_j\theta_j + n^{-1/2}\dot{\varepsilon}_j \quad \text{and} \quad \widehat{\mathfrak{s}}_j = \mathfrak{s}_j + k^{-1/2}\dot{\eta}_j, \quad j \in \mathbb{N}. \tag{02.07}$$

We denote by $\mathbb{P}^{n,k}_{\theta|\mathfrak{s}}$ the joint distribution of $(\widehat{g}_{\bullet}, \widehat{\mathfrak{s}}_{\bullet})$. □

§02|04.06 **Gaussian diagonal inverse sequence model (GdiSM) with noisy operator.** Consider Gaussian white noise processes $\dot{\mathrm{B}}_{\bullet} := (\dot{\mathrm{B}}_j)_{j\in\mathbb{N}} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$ and $\dot{\mathrm{W}}_{\bullet} := (\dot{\mathrm{W}}_j)_{j\in\mathbb{N}} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$. The observable noisy versions $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\mathrm{B}}_{\bullet}$ with ℓ_2-mean $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet}$ and $\widehat{\mathfrak{s}}_{\bullet} = \mathfrak{s}_{\bullet} + k^{-1/2}\dot{\mathrm{W}}_{\bullet}$ with mean function $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$ take both the form of a Gaussian sequence model as in §01|02.06, that is

$$\widehat{g}_j = \mathfrak{s}_j\theta_j + n^{-1/2}\dot{\mathrm{B}}_j \quad \text{and} \quad \widehat{\mathfrak{s}}_j = \mathfrak{s}_j + k^{-1/2}\dot{\mathrm{W}}_j, \quad j \in \mathbb{N}$$
$$\text{with} \quad (\dot{\mathrm{B}}_j)_{j\in\mathbb{N}} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)} \quad \text{and} \quad (\dot{\mathrm{W}}_j)_{j\in\mathbb{N}} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}. \tag{02.08}$$

We denote by $\mathrm{N}^{n,k}_{\theta|\mathfrak{s}}$ the joint distribution of the stochastic process $(\widehat{g}_{\bullet}, \widehat{\mathfrak{s}}_{\bullet})$. □

§02|04|01 Examples of diagonal inverse empirical mean models with noisy operator

§02|04.07 **Cicular density deconvolution with unkown error density.** Similar to Model §01|04.12 consider the *complex* Hilbert spaces $\mathbb{L}_2(\lambda_{[0,1)})$ and $\mathbb{J} := \ell_2(\mathbb{Z})$. Let $\mathbb{D}_1 \subseteq \mathbb{L}_1(\lambda_{[0,1)})$ and $\mathbb{D}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)}) \subseteq \mathbb{L}_1(\lambda_{[0,1)})$ be sets of Lebesgue densities on $([0,1), \mathscr{B}_{[0,1)})$ (by the usual embedding of real-valued functions as in Notation §01|02.10). We denote for each density $\mathbb{p} \in \mathbb{L}_1(\lambda_{[0,1)})$ by $\mathbb{P}_{\mathbb{p}} := \mathbb{p}\lambda_{[0,1)} \in \mathscr{W}(\mathscr{B}_{[0,1)})$ the associated probability measure. Given a Lebesgue error density $\mathbb{q} \in \mathbb{D}_1$ which is *not known* anymore for each Lebesgue density $\mathbb{p} \in \mathbb{D}_2$ we consider the Lebesgue density $g = \mathbb{q} \circledast \mathbb{p} \in \mathbb{L}_2(\lambda_{[0,1)})$ (see Notation §01|04.11) and denote by $\mathbb{P}_{\mathbb{p}|\mathbb{q}} := (\mathbb{q} \circledast \mathbb{p})\lambda_{[0,1)} = g\lambda_{[0,1)} \in \mathscr{W}(\mathscr{B}_{[0,1)})$ the associated probability measure. We consider the statistical product experiment

$$\big([0,1)^{n+k}, \mathscr{B}_{[0,1)}^{\otimes(n+k)}, \mathbb{P}^{n\otimes k}_{\mathbb{D}_2\times\mathbb{D}_1} := (\mathbb{P}^{\otimes n}_{\mathbb{p}|\mathbb{q}} \otimes \mathbb{P}^{\otimes k}_{\mathbb{q}})_{\mathbb{p}\in\mathbb{D}_2,\mathbb{q}\in\mathbb{D}_1}\big).$$

Let $\mathrm{F} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}), \ell_2(\mathbb{Z}))$ be the Fourier-series transform (see Notation §01|02.12). Evidently, for $g \in \mathbb{L}_2(\lambda_{[0,1)}) \subseteq \mathbb{L}_1(\lambda_{[0,1)})$ its Fourier-series $g_{\bullet} = (g_j)_{j\in\mathbb{Z}} = \mathrm{F}g$ satisfies $g_j = \lambda_{[0,1)}(g\bar{\mathrm{e}}_j) = \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\bar{\mathrm{e}}_j)$ for each $j \in \mathbb{Z}$. Moreover, considering the Fourier-series $\mathbb{p}_{\bullet} = (\mathbb{p}_j)_{j\in\mathbb{Z}} = \mathrm{F}\mathbb{p}$ of $\mathbb{p} \in \mathbb{D}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)})$ by the *circular convolution theorem* we have $g_{\bullet} = \mathrm{F}(\mathbb{q} \circledast \mathbb{p}) = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet}$ with $\mathbb{q}_{\bullet} = \lambda_{[0,1)}(\mathbb{q}\bar{\mathrm{e}}_{\bullet}) = \mathbb{P}_{\mathbb{q}}(\bar{\mathrm{e}}_j) \in \ell_\infty(\mathbb{Z})$ and $\mathbb{p}_{\bullet} = \mathrm{F}\mathbb{p} \in \ell_2(\mathbb{Z})$ (see Notation §01|04.11). Moreover, the stochastic process $\bar{\mathrm{e}}_{\bullet} = (\bar{\mathrm{e}}_j)_{j\in\mathbb{Z}}$ on $([0,1), \mathscr{B}_{[0,1)})$ is $(\mathscr{B}_{[0,1]} \otimes 2^{\mathbb{Z}})$-$\mathscr{B}$-measurable, i.e. $\bar{\mathrm{e}}_{\bullet} \in \mathcal{M}(\mathscr{B}_{[0,1)} \otimes 2^{\mathbb{Z}})$ for short (compare Model §01|04.12). We define $\widehat{g}_{\bullet} = (\widehat{g}_j := \widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_j))_{j\in\mathbb{Z}} = \widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_{\bullet}) \in \mathcal{M}(\mathscr{B}^{\otimes n}_{[0,1)} \otimes 2^{\mathbb{Z}})$ and $\widehat{\mathbb{q}}_{\bullet} = (\widehat{\mathbb{q}}_j := \widehat{\mathbb{P}}_k(\bar{\mathrm{e}}_j))_{j\in\mathbb{Z}} = \widehat{\mathbb{P}}_k(\bar{\mathrm{e}}_{\bullet}) \in \mathcal{M}(\mathscr{B}^{\otimes k}_{[0,1)} \otimes 2^{\mathbb{Z}})$ similar to an Empirical mean model §01|02.04 where by construction $g_{\bullet} = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet} = \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\bar{\mathrm{e}}_{\bullet})$ is the $\ell_2(\mathbb{Z})$-mean of $\widehat{g}_{\bullet}$ and $\mathbb{q}_{\bullet} = \mathbb{P}_{\mathbb{q}}(\bar{\mathrm{e}}_{\bullet}) \in \ell_\infty(\mathbb{Z})$ is the mean sequence of $\widehat{\mathbb{q}}_{\bullet}$. The observable noisy versions take the form $\widehat{g}_{\bullet} = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ and $\widehat{\mathbb{q}}_{\bullet} = \mathbb{q}_{\bullet} + k^{-1/2}\dot{\eta}_{\bullet}$, or in equal (02.06) with error processes

$$\dot{\varepsilon}_{\bullet} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_{\mathbb{p}|\mathbb{q}})(\bar{\mathrm{e}}_{\bullet}) = n^{1/2}(\widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_{\bullet}) - \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\bar{\mathrm{e}}_{\bullet})) \in \mathcal{M}(\mathscr{B}^{\otimes n}_{[0,1)} \otimes 2^{\mathbb{Z}}) \quad \text{and}$$
$$\dot{\eta}_{\bullet} = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_{\mathbb{q}})(\bar{\mathrm{e}}_{\bullet}) = k^{1/2}(\widehat{\mathbb{P}}_k(\bar{\mathrm{e}}_{\bullet}) - \mathbb{P}_{\mathbb{q}}(\bar{\mathrm{e}}_{\bullet})) \in \mathcal{M}(\mathscr{B}^{\otimes k}_{[0,1)} \otimes 2^{\mathbb{Z}})$$

satisfying Assumption §01|01.04. □

§02|04.08 **Density additive deconvolution on $\mathbb{R}$ with unkown error density.** Similar to Model §01|04.15 consider the *complex* Hilbert space $\mathbb{L}_2 = \mathbb{L}_2(\lambda)$. Let $\mathbb{D}_1 \subseteq \mathbb{L}_1$ and $\mathbb{D}_2 \subseteq \mathbb{L}_2 \cap \mathbb{L}_1$ be sets of Lebesgue densities on $(\mathbb{R}, \mathscr{B})$ (by the usual embedding of real-valued functions as in Notation §01|02.10). We denote for each density $\mathbb{p} \in \mathbb{L}_1$ by $\mathbb{P}_{\mathbb{p}} := \mathbb{p}\lambda \in \mathscr{W}(\mathscr{B})$ the associated probability measure. Given a Lebesgue density $\mathbb{q} \in \mathbb{D}_1$ which is *not known* anymore for each Lebesgue density $\mathbb{p} \in \mathbb{D}_2$ we consider the Lebesgue density $g = *_{\mathbb{q}}\mathbb{p} = \mathbb{q} * \mathbb{p} \in \mathbb{L}_2 \cap \mathbb{L}_1$ (see Notation §01|04.14) and denote by $\mathbb{P}_{\mathbb{p}|\mathbb{q}} := (\mathbb{q} * \mathbb{p})\lambda = g\lambda \in \mathscr{W}(\mathscr{B})$ the associated probability measure. We consider the statistical product experiment

$$\big(\mathbb{R}^{n+k}, \mathscr{B}^{\otimes(n+k)}, \mathbb{P}^{n\otimes k}_{\mathbb{D}_2\times\mathbb{D}_1} := (\mathbb{P}^{\otimes n}_{\mathbb{p}|\mathbb{q}} \otimes \mathbb{P}^{\otimes k}_{\mathbb{q}})_{\mathbb{p}\in\mathbb{D}_2,\mathbb{q}\in\mathbb{D}_1}\big).$$

Let $\mathrm{F} \in \mathbb{L}(\mathbb{L}_2)$ be the Fourier-Plancherel transform (see Notation §01|02.15). Evidently, for $g \in \mathbb{L}_2 \cap \mathbb{L}_1$ its Fourier-Plancherel transform $g_{\bullet} = (g_j)_{j\in\mathbb{R}} = \mathrm{F}g$ satisfies $g_j = \lambda(g\bar{\mathrm{e}}_j) = \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\bar{\mathrm{e}}_j)$ for all $j \in \mathbb{R}$. Moreover, considering the Fourier-Plancherel transform $\mathbb{p}_{\bullet} = (\mathbb{p}_j)_{j\in\mathbb{R}} = \mathrm{F}\mathbb{p}$ of $\mathbb{p} \in \mathbb{D}_2 \subseteq \mathbb{L}_2 \cap \mathbb{L}_1$ by the *additive convolution theorem* we have $g_{\bullet} = \mathrm{F}(*_{\mathbb{q}}\mathbb{p}) = \lambda(\mathbb{q}\bar{\mathrm{e}}_{\bullet})(\mathrm{F}\mathbb{p}) = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet}$ λ-a.s. with $\mathbb{q}_{\bullet} = \lambda(\mathbb{q}\bar{\mathrm{e}}_{\bullet}) = \mathbb{P}_{\mathbb{q}}(\bar{\mathrm{e}}_{\bullet}) \in \mathbb{L}_\infty$ and $\mathbb{p}_{\bullet} = \mathrm{F}\mathbb{p} \in \mathbb{L}_2$ (see Notation §01|04.14). Moreover, the complex-valued stochastic process $\bar{\mathrm{e}}_{\bullet} = (\bar{\mathrm{e}}_j)_{j\in\mathbb{R}}$ on $(\mathbb{R}, \mathscr{B})$ is $\mathscr{B}^2$-$\mathscr{B}$-measurable, i.e. $\bar{\mathrm{e}}_{\bullet} \in \mathcal{M}(\mathscr{B}^2)$ for short (compare Model §01|04.15). We define $\widehat{g}_{\bullet} = (\widehat{g}_j := \widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_j))_{j\in\mathbb{R}} = \widehat{\mathbb{P}}_n(\bar{\mathrm{e}}_{\bullet}) \in \mathcal{M}(\mathscr{B}^{\otimes n} \otimes \mathscr{B})$ and $\widehat{\mathbb{q}}_{\bullet} = (\widehat{\mathbb{q}}_j := \widehat{\mathbb{P}}_k(\bar{\mathrm{e}}_j))_{j\in\mathbb{R}} = \widehat{\mathbb{P}}_k(\bar{\mathrm{e}}_{\bullet}) \in \mathcal{M}(\mathscr{B}^{\otimes k} \otimes \mathscr{B})$ similar to an Empirical mean model §01|02.04

where by construction $g_{\bullet} = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet} = \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\bar{e}_{\bullet})$ is the $\mathbb{L}_2$-mean of $\widehat{g}_{\bullet}$ and $\mathbb{q}_{\bullet} = \mathbb{P}_{\mathbb{q}}(\bar{e}_{\bullet}) \in \mathbb{L}_\infty$ is the mean function of $\widehat{\mathbb{q}}_{\bullet}$. The observable noisy versions take the form $\widehat{g}_{\bullet} = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ and $\widehat{\mathbb{q}}_{\bullet} = \mathbb{q}_{\bullet} + k^{-1/2}\dot{\eta}_{\bullet}$, or in equal (02.06) with error processes

$$\dot{\varepsilon}_{\bullet} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_{\mathbb{p}|\mathbb{q}})(\bar{e}_{\bullet}) = n^{1/2}(\widehat{\mathbb{P}}_n(\bar{e}_{\bullet}) - \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\bar{e}_{\bullet})) \in \mathcal{M}(\mathscr{B}^{\otimes n} \otimes \mathscr{B}) \quad \text{and}$$
$$\dot{\eta}_{\bullet} = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_{\mathbb{q}})(\bar{e}_{\bullet}) = k^{1/2}(\widehat{\mathbb{P}}_k(\bar{e}_{\bullet}) - \mathbb{P}_{\mathbb{q}}(\bar{e}_{\bullet})) \in \mathcal{M}(\mathscr{B}^{\otimes k} \otimes \mathscr{B})$$

satisfying Assumption §01|01.04. □

§02|04.09 **Density multiplicative deconvolution on $\mathbb{R}_{>0}$ with unkown error density**. Consider the *complex* Hilbert spaces $\mathbb{L}_2(\mathrm{x}^{2c-1}) = \mathbb{L}_2(\mathbb{R}_{>0}, \mathscr{B}_{>0}, \mathrm{x}^{2c-1}\lambda_{>0})$ and $\mathbb{L}_2 = \mathbb{L}_2(\lambda)$ as in Model §01|04.17. Let $\mathbb{D}_1 \subseteq \mathbb{L}_1(\mathrm{x}^{c-1})$ and $\mathbb{D}_2 \subseteq \mathbb{L}_1(\mathrm{x}^{c-1}) \cap \mathbb{L}_2(\mathrm{x}^{2c-1})$ be sets of Lebesgue-densities on $(\mathbb{R}_{>0}, \mathscr{B}_{>0})$ (by the usual embedding of real-valued functions as in Notation §01|02.10). We denote for each Lebesgue density $\mathbb{p}$ on $(\mathbb{R}_{>0}, \mathscr{B}_{>0})$ by *colno*$\mathbb{P}_{\mathbb{p}} := \mathbb{p}\lambda_{>0} \in \mathscr{W}(\mathscr{B}_{>0})$ the associated probability measure. Given a Lebesgue density $\mathbb{q} \in \mathbb{L}_1(\mathrm{x}^{c-1})$ which is *not known* anymore for each Lebesgue density $\mathbb{p} \in \mathbb{D}_2$ we consider the Lebesgue density $g = \circledast_{\mathbb{q}}\mathbb{p} = \mathbb{q} \circledast \mathbb{p} \in \mathbb{L}_1(\mathrm{x}^{c-1}) \cap \mathbb{L}_2(\mathrm{x}^{2c-1})$ (see Notation §01|04.16) and denote by $\mathbb{P}_{\mathbb{p}|\mathbb{q}} := (\mathbb{q} \circledast \mathbb{p})\lambda_{>0} = g\lambda_{>0} \in \mathscr{W}(\mathscr{B}_{>0})$ the associated probability measure. We consider the statistical product experiment

$$\big(\mathbb{R}_{>0}^{n+k}, \mathscr{B}_{>0}^{\otimes(n+k)}, \mathbb{P}^{n\otimes k}_{\mathbb{D}_2\times\mathbb{D}_1} := (\mathbb{P}_{\mathbb{p}|\mathbb{q}}^{\otimes n} \otimes \mathbb{P}_{\mathbb{q}}^{\otimes k})_{\mathbb{p}\in\mathbb{D}_2, \mathbb{q}\in\mathbb{D}_1}\big).$$

Let $\mathrm{M}_c \in \mathbb{L}(\mathbb{L}_2(\mathrm{x}^{2c-1}), \mathbb{L}_2)$ be the Mellin transform (see Notation §01|02.17). Evidently, for $g \in \mathbb{L}_1(\mathrm{x}^{c-1}) \cap \mathbb{L}_2(\mathrm{x}^{2c-1})$ its Mellin transform $g_{\bullet} = (g_j)_{j\in\mathbb{R}} = \mathrm{M}_c g$ satisfies $g_j = \mathrm{x}^{c-1}\lambda_{>0}(\bar{\mathrm{x}}_j g) = \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\mathrm{x}^{c-1}\bar{\mathrm{x}}_j)$ for all $j \in \mathbb{R}$. Moreover, considering the Mellin transform $\mathbb{p}_{\bullet} = (\mathbb{p}_j)_{j\in\mathbb{R}} = \mathrm{M}_c\mathbb{p}$ of $\mathbb{p} \in \mathbb{D}_2 \subseteq \mathbb{L}_1(\mathrm{x}^{c-1}) \cap \mathbb{L}_2(\mathrm{x}^{2c-1})$ by the *multiplicative convolution theorem* we have $g_{\bullet} = \mathrm{M}_c(\circledast_{\mathbb{q}}\mathbb{p}) = \mathrm{x}^{c-1}\lambda_{>0}(\bar{\mathrm{x}}_{\bullet}\mathbb{q})(\mathrm{M}_c\mathbb{p}) = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet}$ λ-a.s. with $\mathbb{q}_{\bullet} = \mathrm{x}^{c-1}\lambda_{>0}(\bar{\mathrm{x}}_{\bullet}\mathbb{q}) = \mathbb{P}_{\mathbb{q}}(\mathrm{x}^{c-1}\bar{\mathrm{x}}_{\bullet}) \in \mathbb{L}_\infty$ and $\mathbb{p}_{\bullet} = \mathrm{M}_c\mathbb{p} \in \mathbb{L}_2$ (see Notation §01|04.16). Moreover, the complex-valued stochastic process $\mathrm{x}^{c-1}\bar{\mathrm{x}}_{\bullet} = (\mathrm{x}^{c-1}\bar{\mathrm{x}}_j)_{j\in\mathbb{R}} \in \mathcal{M}(\mathscr{B}_{>0} \otimes \mathscr{B})$ on $(\mathbb{R}_{>0}, \mathscr{B}_{>0})$ is $\mathscr{B}_{>0} \otimes \mathscr{B}$-$\mathscr{B}$-measurable, i.e. $\mathrm{x}^{c-1}\bar{\mathrm{x}}_{\bullet} \in \mathcal{M}(\mathscr{B}_{>0} \otimes \mathscr{B})$ for short. We define $\widehat{g}_{\bullet} = (\widehat{g}_j := \widehat{\mathbb{P}}_n(\mathrm{x}^{c-1}\bar{\mathrm{x}}_j))_{j\in\mathbb{R}} = \widehat{\mathbb{P}}_n(\mathrm{x}^{c-1}\bar{\mathrm{x}}_{\bullet}) \in \mathcal{M}(\mathscr{B}_{>0}^{\otimes n} \otimes \mathscr{B})$ and $\widehat{\mathbb{q}}_{\bullet} = (\widehat{\mathbb{q}}_j := \widehat{\mathbb{P}}_k(\mathrm{x}^{c-1}\bar{\mathrm{x}}_j))_{j\in\mathbb{R}} = \widehat{\mathbb{P}}_k(\mathrm{x}^{c-1}\bar{\mathrm{x}}_{\bullet}) \in \mathcal{M}(\mathscr{B}_{>0}^{\otimes k} \otimes \mathscr{B})$ similar to an Empirical mean model §01|02.04 where by construction $g_{\bullet} = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet} = \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\mathrm{x}^{c-1}\bar{\mathrm{x}}_{\bullet})$ is the $\mathbb{L}_2$-mean of $\widehat{g}_{\bullet}$ and $\mathbb{q}_{\bullet} = \mathbb{P}_{\mathbb{q}}(\mathrm{x}^{c-1}\bar{\mathrm{x}}_{\bullet}) \in \mathbb{L}_\infty$ is the mean function of $\widehat{\mathbb{q}}_{\bullet}$. The observable noisy versions take the form $\widehat{g}_{\bullet} = \mathbb{q}_{\bullet}\mathbb{p}_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ and $\widehat{\mathbb{q}}_{\bullet} = \mathbb{q}_{\bullet} + k^{-1/2}\dot{\eta}_{\bullet}$, or in equal (02.06) with error processes

$$\dot{\varepsilon}_{\bullet} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_{\mathbb{p}|\mathbb{q}})(\mathrm{x}^{c-1}\bar{\mathrm{x}}_{\bullet}) = n^{1/2}(\widehat{\mathbb{P}}_n(\mathrm{x}^{c-1}\bar{\mathrm{x}}_{\bullet}) - \mathbb{P}_{\mathbb{p}|\mathbb{q}}(\mathrm{x}^{c-1}\bar{\mathrm{x}}_{\bullet})) \in \mathcal{M}(\mathscr{B}_{>0}^{\otimes n} \otimes \mathscr{B}) \quad \text{and}$$
$$\dot{\eta}_{\bullet} = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_{\mathbb{q}})(\mathrm{x}^{c-1}\bar{\mathrm{x}}_{\bullet}) = k^{1/2}(\widehat{\mathbb{P}}_k(\mathrm{x}^{c-1}\bar{\mathrm{x}}_{\bullet}) - \mathbb{P}_{\mathbb{q}}(\mathrm{x}^{c-1}\bar{\mathrm{x}}_{\bullet})) \in \mathcal{M}(\mathscr{B}_{>0}^{\otimes k} \otimes \mathscr{B})$$

satisfying Assumption §01|01.04. □

§02|04.10 **Functional linear regression under second order stationarity**. Consider the *complex* Hilbert spaces $\mathbb{L}_2(\lambda_{[0,1)})$ and $\mathbb{J} := \ell_2(\mathbb{Z})$. Let $(\mathbb{L}_2(\lambda_{[0,1)}), \langle\cdot,\cdot\rangle_{\mathbb{L}_2(\lambda_{[0,1)})})$ be equipped with its Borel-σ-algebra $\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})}$ and let (X, Y) be an $\mathbb{L}_2(\lambda_{[0,1)}) \times \mathbb{R}$-valued random vector with joint distribution $\mathbb{P}^{X,Y} \in \mathscr{W}(\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes \mathscr{B})$. We denote by $\mathbb{P}^X \in \mathscr{W}(\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})})$ the marginal distribution of the *real random function* X (by the usual embedding of real-valued functions as in Notation §01|02.10). We tactically identify X and Y with the coordinate map $\Pi_{\mathbb{L}_2(\lambda_{[0,1)})}$ and $\Pi_{\mathbb{R}}$, respectively, and thus (X, Y) with the identity $\mathrm{id}_{\mathbb{L}_2(\lambda_{[0,1)})\times\mathbb{R}}$ such that $\mathbb{P} = \mathbb{P}^{X,Y} \in \mathscr{W}(\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes \mathscr{B})$. Here and subsequently, we assume that $Y, \|X\|^2_{\mathbb{L}_2(\lambda_{[0,1)})} \in \mathbb{L}_2(\mathbb{P}) = \mathbb{L}_2(\mathbb{L}_2(\lambda_{[0,1)}) \times \mathbb{R}, \mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes \mathscr{B}, \mathbb{P})$, $\mathbb{P}(\langle x, X\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = 0$ for all $x \in \mathbb{L}_2(\lambda_{[0,1)})$, and that X is *second order stationary* as in Model §02|03.04. In this situation X admits a *covariance operator* $\Gamma^X \in \mathbb{L}^{\geq}(\mathbb{L}_2(\lambda_{[0,1)}))$ which equals a *circular additive convolution*, that is $\circledast_{c^X} = \Gamma^X \in \mathbb{L}^{\geq}(\mathbb{L}_2(\lambda_{[0,1)}))$ with $c^X \in \mathbb{L}_2(\lambda_{[0,1)})$ (see Model §02|03.04) and there is $g \in \mathbb{L}_2(\lambda_{[0,1)})$ satisfying $\langle g, x\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} =$

$\mathbb{P}(Y\langle X,x\rangle_{\mathbb{L}_2(\lambda_{[0,1)})})$ for all $x \in \mathbb{L}_2(\lambda_{[0,1)})$. In what follows we assume that in addition $g \in \operatorname{ran}(\Gamma^X) \subseteq \mathbb{L}_2(\lambda_{[0,1)})$. In this situation there exists $f \in \mathbb{L}_2(\lambda_{[0,1)})$ such that

$$\langle g,x\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} = \mathbb{P}(Y\langle X,x\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = \mathbb{P}(\langle f,X\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}\langle X,x\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = \langle \Gamma^X f,x\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} \quad \forall x \in \mathbb{L}_2(\lambda_{[0,1)})$$

or in equal $\mathbb{P}$-a.s. we have $Y = \langle f,X\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} + \xi$ with $\mathbb{P}(\xi\langle X,x\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = 0$ for all $x \in \mathbb{L}_2(\lambda_{[0,1)})$. Let us denote by $\mathbb{P}_{c^X} \in \mathscr{W}(\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})})$ the marginal destribution of X which is not fully specified given $c^X \in \mathbb{D}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)})$ (see Model §02|03.04). Moreover, for $c^X \in \mathbb{D}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)})$ and $f \in \mathbb{F}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)})$, and hence $g := \circledast_{c^X} f \in \mathbb{L}_2(\lambda_{[0,1)})$, we denote by $\mathbb{P}_{f|c^X}$ the joint distribution of (X,Y) without fully specifying the distribution which however is assumed to satisfy $\mathbb{P}_{f|c^X}(Y\langle X,x\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = \mathbb{P}_{c^X}(\langle f,X\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}\langle X,x\rangle_{\mathbb{L}_2(\lambda_{[0,1)})})$ for all $x \in \mathbb{L}_2(\lambda_{[0,1)})$. Let $\mathrm{F} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}), \ell_2(\mathbb{Z}))$ be the *Fourier-series transform* with $h \mapsto \mathrm{F}h := h_\bullet = \lambda_{[0,1)}(h\bar{\mathrm{e}}_\bullet)$ and exponential basis $\mathrm{e}_\bullet := (\mathrm{e}_j)_{j\in\mathbb{Z}}$ (see Notations §01|02.10 and §01|02.12). Following Model §02|03.04 $\mathrm{M}_{c^X} = \mathrm{F}\circledast_{c^X}\mathrm{F}^\star = \mathrm{F}\Gamma^X\mathrm{F}^\star \in \mathbb{L}^{\mathrm{M}}(\ell_2(\mathbb{Z}))$ is a multiplication operator with $c^X_\bullet \in \ell_2(\mathbb{Z}) \subseteq \ell_\infty(\mathbb{Z})$ satisfying $c^X_\bullet = (c^X_j = \mathbb{P}_{c^X}(|\langle X,\mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}|^2))_{j\in\mathbb{Z}}$. Therefore the complex-valued stochastic process $|\lambda_{[0,1)}(X\bar{\mathrm{e}}_\bullet)|^2 = (|\lambda_{[0,1)}(X\bar{\mathrm{e}}_j)|^2 = |\langle X,\mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}|^2)_{j\in\mathbb{Z}} \in \mathcal{M}(\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes 2^{\mathbb{Z}})$ fulfils Assumption §01|01.04 and $c^X_\bullet = \mathbb{P}_{c^X}(|\lambda_{[0,1)}(X\bar{\mathrm{e}}_\bullet)|^2)$. Moreover, for each $f \in \mathbb{F}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)})$ the Fourier coefficients $g_\bullet = (g_j)_{j\in\mathbb{Z}} = \mathrm{F}g = \mathrm{F}\Gamma^X\mathrm{F}^\star\mathrm{F}f = \mathrm{M}_{c^X}f_\bullet = c^X_\bullet f_\bullet$, satisfy $g_\bullet = (g_j = \mathbb{P}_{f|c^X}(Y\langle X,\mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}))_{j\in\mathbb{Z}}$. The complex-valued stochastic process $Y\lambda_{[0,1)}(X\bar{\mathrm{e}}_\bullet) = (Y\lambda_{[0,1)}(X\bar{\mathrm{e}}_j) = Y\langle X,\mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})})_{j\in\mathbb{Z}} \in \mathcal{M}((\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes \mathscr{B}) \otimes 2^{\mathbb{Z}})$ fulfils Assumption §01|01.04 and $g_\bullet = c^X_\bullet f_\bullet = \mathbb{P}_{f|c^X}(Y\lambda_{[0,1)}(X\bar{\mathrm{e}}_\bullet))$. The observable noisy versions take the form $\widehat{g}_\bullet = c^X_\bullet f_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ and $\widehat{c}^X_\bullet = c^X_\bullet + k^{-1/2}\dot{\eta}_\bullet$, or in equal (02.06) with error processes

$$\dot{\varepsilon}_\bullet = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_{f|c^X})(|\lambda_{[0,1)}(X\bar{\mathrm{e}}_\bullet)|^2) \in \mathcal{M}((\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes \mathscr{B})^{\otimes n} \otimes 2^{\mathbb{Z}}) \quad \text{and}$$
$$\dot{\eta}_\bullet = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_{c^X})(Y\lambda_{[0,1)}(X\bar{\mathrm{e}}_\bullet)) \in \mathcal{M}(\mathscr{B}^{\otimes k}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes 2^{\mathbb{Z}})$$

satisfying Assumption §01|01.04. □

§02|04.11 **Functional linear instrumental regression under second order stationarity**. Consider the *complex* Hilbert spaces $\mathbb{L}_2(\lambda_{[0,1)})$ and $\mathbb{J} := \ell_2(\mathbb{Z})$. Let $(\mathbb{L}_2(\lambda_{[0,1)}), \langle\cdot,\cdot\rangle_{\mathbb{L}_2(\lambda_{[0,1)})})$ be equipped with its Borel-σ-algebra $\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})}$ and let (Z,X,Y) be an $\mathbb{L}_2(\lambda_{[0,1)})^2 \times \mathbb{R}$-valued random vector with joint distribution $\mathbb{P}^{Z,X,Y} \in \mathscr{W}(\mathscr{B}^2_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes \mathscr{B})$. We denote by $\mathbb{P}^Z, \mathbb{P}^Z \in \mathscr{W}(\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})})$ the marginal distributions of the *real random functions* Z and X (by the usual embedding of real-valued functions as in Notation §01|02.10). Moreover, denote by $\mathbb{P}^{Z,X} \in \mathscr{W}(\mathscr{B}^2_{\mathbb{L}_2(\lambda_{[0,1)})})$, and $\mathbb{P}^{Z,Y} \in \mathscr{W}(\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes \mathscr{B})$ the marginal distribution of (Z,X) and (Z,Y), respectively. We tactically take Z, X and Y as coordinate maps and thus identify (Z,X,Y) with the identity $\mathrm{id}_{\mathbb{L}_2(\lambda_{[0,1)})^2\times\mathbb{R}}$ such that $\mathbb{P} = \mathbb{P}^{Z,X,Y} \in \mathscr{W}(\mathscr{B}^2_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes \mathscr{B})$. Here and subsequently, given $\mathbb{L}_2(\mathbb{P}) := \mathbb{L}_2((\mathbb{L}_2(\lambda_{[0,1)}))^2 \times \mathbb{R}, \mathscr{B}^2_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes \mathscr{B}, \mathbb{P})$ we assume that $Y, \|Z\|^2_{\mathbb{L}_2(\lambda_{[0,1)})}, \|X\|^2_{\mathbb{L}_2(\lambda_{[0,1)})} \in \mathbb{L}_2(\mathbb{P})$, $\mathbb{P}(\langle Z,z\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = 0$, $\mathbb{P}(\langle x,X\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = 0$ for all $z,x \in \mathbb{L}_2(\lambda_{[0,1)})$, and that (Z,X) is *second order stationary* as in Model §02|03.05. In this situation (Z,X) admits a *cross-covariance operator* $\Gamma^{ZX} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}))$ which equals a *circular additive convolution*, that is $\circledast_{c^{ZX}} = \Gamma^{ZX} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}))$ with $c^{ZX} \in \mathbb{L}_2(\lambda_{[0,1)})$ (see Model §02|01.09) and there is $g \in \mathbb{L}_2(\lambda_{[0,1)})$ satisfying $\langle g,z\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} = \mathbb{P}(Y\langle Z,z\rangle_{\mathbb{L}_2(\lambda_{[0,1)})})$ for all $z \in \mathbb{L}_2(\lambda_{[0,1)})$. In what follows we assume that in addition $g \in \operatorname{ran}(\Gamma^{ZX}) \subseteq \mathbb{L}_2(\lambda_{[0,1)})$. In this situation there exists $f \in \mathbb{L}_2(\lambda_{[0,1)})$ such that

$$\langle g,z\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} = \mathbb{P}(Y\langle Z,z\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = \mathbb{P}(\langle f,X\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}\langle Z,z\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = \langle \Gamma^{ZX} f,z\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} \quad \forall z \in \mathbb{L}_2(\lambda_{[0,1)})$$

or in equal $\mathbb{P}$-a.s. we have $Y = \langle f,X\rangle_{\mathbb{L}_2(\lambda_{[0,1)})} + \xi$ with $\mathbb{P}(\xi\langle Z,z\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = 0$ for all $z \in \mathbb{L}_2(\lambda_{[0,1)})$. Let us denote by $\mathbb{P}_{c^{ZX}} \in \mathscr{W}(\mathscr{B}^2_{\mathbb{L}_2(\lambda_{[0,1)})})$ the marginal destribution of (Z,X) which is not fully specified given $c^{ZX} \in \mathbb{D}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)})$ (see Model §02|03.05). Moreover, for $c^{ZX} \in \mathbb{D}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)})$ and $f \in \mathbb{F}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)})$, and hence $g := \circledast_{c^{ZX}} f \in \mathbb{L}_2(\lambda_{[0,1)})$, we denote by $\mathbb{P}_{f|c^{ZX}}$ the joint distribution of (Z,X,Y) without fully specifying the distribution which however is assumed to satisfy

$\mathbb{P}_{f|\mathrm{c}^{ZX}}(Y\langle Z, z\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}) = \mathbb{P}_{\mathrm{c}^{ZX}}(\langle f, X\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}\langle Z, z\rangle_{\mathbb{L}_2(\lambda_{[0,1)})})$ for all $z \in \mathbb{L}_2(\lambda_{[0,1)})$. Let $\mathrm{F} \in \mathbb{L}(\mathbb{L}_2(\lambda_{[0,1)}), \ell_2(\mathbb{Z}))$ be the *Fourier-series transform* with $h \mapsto \mathrm{F}h := h_\bullet = \lambda_{[0,1)}(h\overline{\mathrm{e}}_\bullet)$ and exponential basis $\mathrm{e}_\bullet := (\mathrm{e}_j)_{j\in\mathbb{Z}}$ (see Notations §01|02.10 and §01|02.12). Similar to Model §02|03.05 $\mathrm{M}_{\mathrm{c}^{ZX}} = \mathrm{F}\circledast_{\mathrm{c}^{ZX}}\mathrm{F}^\star = \mathrm{F}\Gamma^{ZX}\mathrm{F}^\star \in \mathbb{L}^{\mathrm{M}}(\ell_2(\mathbb{Z}))$ is a multiplication operator with $\mathrm{c}_\bullet^{ZX} \in \ell_2(\mathbb{Z}) \subseteq \ell_\infty(\mathbb{Z})$ which satisfies $\mathrm{c}_\bullet^{ZX} = (\mathrm{c}_j^{ZX} = \mathbb{P}_{\mathrm{c}^{ZX}}(\langle Z, \mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}\langle \mathrm{e}_j, X\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}))_{j\in\mathbb{Z}}$. Therefore the complex-valued stochastic process $\lambda_{[0,1)}(Z\overline{\mathrm{e}}_\bullet)\lambda_{[0,1)}(\overline{X}\mathrm{e}_\bullet) = (\lambda_{[0,1)}(Z\overline{\mathrm{e}}_j)\lambda_{[0,1)}(\overline{X}\mathrm{e}_j) = \langle Z, \mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}\langle \mathrm{e}_j, X\rangle_{\mathbb{L}_2(\lambda_{[0,1)})})_{j\in\mathbb{Z}} \in \mathcal{M}(\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes 2^{\mathbb{Z}})$ fulfils Assumption §01|01.04 and $\mathrm{c}_\bullet^{ZX} = \mathbb{P}_{\mathrm{c}^{ZX}}(\lambda_{[0,1)}(Z\overline{\mathrm{e}}_\bullet)\lambda_{[0,1)}(\overline{X}\mathrm{e}_\bullet))$. Moreover, for each $f \in \mathbb{F}_2 \subseteq \mathbb{L}_2(\lambda_{[0,1)})$ the Fourier coefficients $g_\bullet = (g_j)_{j\in\mathbb{Z}} = \mathrm{F}g = \mathrm{F}\Gamma^{ZX}\mathrm{F}^\star\mathrm{F}f = \mathrm{M}_{\mathrm{c}^{ZX}}f_\bullet = \mathrm{c}_\bullet^{ZX}f_\bullet$, satisfy $g_\bullet = (g_j = \mathbb{P}_{f|\mathrm{c}^{ZX}}(Y\langle Z, \mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})}))_{j\in\mathbb{Z}}$. The complex-valued stochastic process $Y\lambda_{[0,1)}(Z\overline{\mathrm{e}}_\bullet) = (Y\lambda_{[0,1)}(Z\overline{\mathrm{e}}_j) = Y\langle Z, \mathrm{e}_j\rangle_{\mathbb{L}_2(\lambda_{[0,1)})})_{j\in\mathbb{Z}} \in \mathcal{M}((\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes \mathscr{B}) \otimes 2^{\mathbb{Z}})$ fulfils Assumption §01|01.04 and $g_\bullet = \mathrm{c}_\bullet^{ZX}f_\bullet = \mathbb{P}_{f|\mathrm{c}^{ZX}}(Y\lambda_{[0,1)}(Z\overline{\mathrm{e}}_\bullet))$. The observable noisy versions take the form $\widehat{g}_\bullet = \mathrm{c}_\bullet^{ZX}f_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ and $\widehat{\mathrm{c}}_\bullet^{ZX} = \mathrm{c}_\bullet^{ZX} + k^{-1/2}\dot{\eta}_\bullet$, or in equal (02.06) with error processes

$$\dot{\varepsilon}_\bullet = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_{f|\mathrm{c}^{ZX}})(\lambda_{[0,1)}(Z\overline{\mathrm{e}}_\bullet)\lambda_{[0,1)}(\overline{X}\mathrm{e}_\bullet)) \in \mathcal{M}((\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})} \otimes \mathscr{B})^{\otimes n} \otimes 2^{\mathbb{Z}}) \quad \text{and}$$
$$\dot{\eta}_\bullet = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_{\mathrm{c}^X})(Y\lambda_{[0,1)}(Z\overline{\mathrm{e}}_\bullet)) \in \mathcal{M}(\mathscr{B}_{\mathbb{L}_2(\lambda_{[0,1)})}^{\otimes k} \otimes 2^{\mathbb{Z}})$$

satisfying Assumption §01|01.04. □

Chapter 2

Regularisation of inverse problems

Given $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet}$ the regularised reconstruction of $\theta_{\bullet}$ in a direct problem and an inverse problem with linear operator $\mathrm{T}_{\bullet|\bullet}$ in a diagonal or general case is presented.

Overview

§03 Ill-posed inverse problems

Let $\mathbb{H}$ and $\mathbb{G}$ be separable Hilbert spaces over $\mathbb{K} \in \{\mathbb{R}, \mathbb{C}\}$ endowed with inner product $\langle\cdot,\cdot\rangle_{\mathbb{H}}$ and $\langle\cdot,\cdot\rangle_{\mathbb{G}}$ and induced norm $\|\cdot\|_{\mathbb{H}}$ and $\|\cdot\|_{\mathbb{G}}$, respectively. Consider a linear bounded operator $\mathrm{T} : \mathbb{H} \to \mathbb{G}$, for short $\mathrm{T} \in \mathbb{L}(\mathbb{H}, \mathbb{G})$.

§03|00.01 **Definition**. Given $g \in \mathbb{G}$ the reconstruction of a solution $\theta \in \mathbb{H}$ of the equation $g = \mathrm{T}\theta$ is called *inverse problem*. □

§03|00.02 **Definition** (Hadamard [1932]). An inverse problem $g = \mathrm{T}\theta$ is called *well-posed* if (a) a solution θ *exists*, (b) the solution θ is *unique*, and (c) the solution depends continuously on g. An inverse problem which is not well-posed is called *ill-posed*. □

For a broader overview on inverse problems we refer the reader to the monograph by Kress [1989] or Engl et al. [2000].

§03|00.03 **Property** (Existence and identification). *There exists an unique solution of the equation* $g = \mathrm{T}\theta$ *if and only if the following two conditions are satisfied*

(existence) *g belongs to the range* $\mathrm{ran}(\mathrm{T})$ *of* T,

(identification) *The operator* T *is injective, i.e., its null space* $\ker(\mathrm{T}) = \{0\}$ *is trivial.* □

§03|00.04 **Remark**. If there does not exist a solution typically one might consider a least-square solution which exists if and only if $g \in \mathrm{ran}(\mathrm{T}) \oplus \ker(\mathrm{T}^\star)$. A least-square solution with minimal norm, if it exists, could be recovered, in case the solution is not unique. Nevertheless, the main issue is often the stability of the inverse problem. More precisely, if the solution does not depend continuously on g, i.e., the inverse T^{-1} of T is not continuous, a reconstruction $\widehat{\theta} = \mathrm{T}^{-1}\widehat{g}$ may be far from the solution θ even if the noisy version $\widehat{g}$ is close to g. □

§03|00.05 **Property**. *Denote by* $\Pi_{\overline{\mathrm{ran}}(\mathrm{T})}$ *the orthogonal projection onto the closure* $\overline{\mathrm{ran}}(\mathrm{T})$ *of the range of* T. *For each* $g \in \mathbb{G}$ *the following assertions are equivalent* (i) θ *minimises* $h \mapsto \|g - \mathrm{T}h\|_{\mathbb{G}}$ *over* $\mathbb{H}$ *(least square solution);* (ii) $\Pi_{\overline{\mathrm{ran}}(\mathrm{T})}g = \mathrm{T}\theta$; (iii) $\mathrm{T}^\star g = \mathrm{T}^\star\mathrm{T}\theta$ *(normal equation).* □

§03|00.06 **Remark**. We note that $g \in \mathrm{ran}(\mathrm{T}) \oplus \mathrm{ran}(\mathrm{T})^\perp$ implies $\Pi_{\overline{\mathrm{ran}}(\mathrm{T})}g \in \mathrm{ran}(\mathrm{T})$ and hence the preimage $\mathrm{T}^{-1}(\{\Pi_{\overline{\mathrm{ran}}(\mathrm{T})}g\}) = \{h \in \mathbb{H} : \mathrm{T}h = \Pi_{\overline{\mathrm{ran}}(\mathrm{T})}g\}$ is not empty. More precisely, due to Property §03|00.05 it equals the *set of least square solutions*, i.e. $\mathrm{T}^{-1}(\{\Pi_{\overline{\mathcal{R}}(\mathrm{T})}g\}) = \{\theta \in \mathbb{H} : \mathrm{T}^\star g = \mathrm{T}^\star\mathrm{T}\theta\}$. □

§03|00.07 **Notation**. In the sequel keep in mind that for each $\mathrm{T} \in \mathbb{L}(\mathbb{H},\mathbb{G})$ its restriction $\mathrm{T}_{\mathrm{res}} : \ker(\mathrm{T})^\perp \to \mathrm{ran}(\mathrm{T})$ is bijective and thus has an inverse $\mathrm{T}_{\mathrm{res}}^{-1} : \mathrm{ran}(\mathrm{T}) \to \ker(\mathrm{T})^\perp$. Here and subsequently we identify T and $\mathrm{T}_{\mathrm{res}}$. □

§03|00.08 **Definition**. For $\mathrm{T} \in \mathbb{L}(\mathbb{H},\mathbb{G})$ the *Moore-Penrose inverse* (generalised or pseudo inverse) $\mathrm{T}^\dagger$ is the unique linear extension of $\mathrm{T}^{-1} : \mathrm{ran}(\mathrm{T}) \to \ker(\mathrm{T})^\perp$ to the domain $\mathrm{dom}(\mathrm{T}^\dagger) := \mathrm{ran}(\mathrm{T}) \oplus \mathrm{ran}(\mathrm{T})^\perp$ with $\ker(\mathrm{T}^\dagger) = \mathrm{ran}(\mathrm{T})^\perp$ satisfying $\mathrm{T}^\dagger g := \mathrm{T}^{-1}\Pi_{\overline{\mathrm{ran}}(\mathrm{T})}g$ for any $g \in \mathrm{dom}(\mathrm{T}^\dagger)$. □

§03|00.09 **Remark**. We note that $\mathrm{T}\mathrm{T}^\dagger\mathrm{T} = \mathrm{T}$, $\mathrm{T}^\dagger\mathrm{T}\mathrm{T}^\dagger = \mathrm{T}^\dagger$, $\mathrm{T}^\dagger\mathrm{T} = \Pi_{\ker(\mathrm{T})^\perp}$ and $\mathrm{T}\mathrm{T}^\dagger g = \Pi_{\overline{\mathrm{ran}}(\mathrm{T})}g$ for any $g \in \mathrm{dom}(\mathrm{T}^\dagger)$. If T is injective, and hence $\mathrm{T}^\star\mathrm{T}$, then $\mathrm{T}^\star\mathrm{T} : \mathbb{H} \to \mathrm{ran}(\mathrm{T}^\star\mathrm{T})$ is invertible, which in turn, for any $g \in \mathrm{ran}(\mathrm{T}) \oplus \mathrm{ran}(\mathrm{T})^\perp$, implies that $(\mathrm{T}^\star\mathrm{T})^\dagger\mathrm{T}^\star g$ is the unique solution of the normal equation, and thus $\mathrm{T}^{-1}(\{\Pi_{\overline{\mathrm{ran}}(\mathrm{T})}g\}) = \{\mathrm{T}^\dagger g\} = \{(\mathrm{T}^\star\mathrm{T})^\dagger\mathrm{T}^\star g\}$. If T is invertible then $\mathrm{T}^\dagger = \mathrm{T}^{-1}$. □

§03|00.10 **Property**. *For each* $g \in \mathrm{dom}(\mathrm{T}^\dagger)$, $\mathrm{T}^\dagger g$ *belongs to* $\mathrm{T}_{\mathrm{res}}^{-1}(\{\Pi_{\overline{\mathrm{ran}}(\mathrm{T})}g\})$ *and, hence is a least square solution. Moreover,* $\mathrm{T}^\dagger g$ *is the unique least square solution with minimal* $\|\cdot\|_{\mathbb{H}}$*-norm, that is,* $\|\mathrm{T}^\dagger g\|_{\mathbb{H}} = \inf\{\|h\|_{\mathbb{H}} : h \in \mathrm{T}_{\mathrm{res}}^{-1}(\{\Pi_{\overline{\mathrm{ran}}(\mathrm{T})}g\})\}$. □

We eventually approximate the operator T by sequence $(\mathrm{T}^m)_{m\in\mathbb{N}}$ of operators in $\mathbb{L}(\mathbb{H},\mathbb{G})$, where for each m the operator $\mathrm{T}^m \in \mathbb{L}(\mathbb{H},\mathbb{G})$ has a finite dimensional image. If $\|\mathrm{T}^m - \mathrm{T}\|_{\mathbb{L}(\mathbb{H},\mathbb{G})} = o(1)$ as $m \to \infty$, then T is compact (reference), i.e. $\mathrm{T} \in \mathbb{K}(\mathbb{H},\mathbb{G})$ for short, and the inverse problem is generally ill-posed due to the next property.

§03|00.11 **Property**. *If* $\mathbb{H}$ *and* $\mathbb{G}$ *are infinite dimensional and* $\mathrm{T} \in \mathbb{K}(\mathbb{H},\mathbb{G})$ *is injective, then*

$$\inf\left\{\|\mathrm{T}h\|_{\mathbb{G}} : \|h\|_{\mathbb{H}} = 1, h \in \mathbb{H}\right\} = 0,$$

which in turn implies that $\mathrm{T}_{\mathrm{res}}^{-1} : \mathrm{ran}(\mathrm{T}) \to \mathbb{H}$*, and hence also* $\mathrm{T}^\dagger$*, is not continuous.* □

Coming back to the reconstruction of $\theta \in \mathbb{H}$ from a noisy version of the image $g = \mathrm{T}\theta \in \mathbb{G}$ and eventually a noisy version of the operator $\mathrm{T} \in \mathbb{L}(\mathbb{H}, \mathbb{G})$ as introduced in Chapter 1. Throughout this manuscript introducing the measure space $(\mathcal{J}, \mathscr{J}, \nu)$ with index set $\mathcal{J}$ being contained in $\mathbb{N}$ $\mathbb{Z}$ or $\mathbb{R}$, surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{L}_2(\nu))$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \mathbb{L}_2(\nu))$, which are fixed and presumed to be known in advance, we write $g_{\bullet} = \mathrm{A}\theta$ with $\mathrm{A} := \mathrm{VT} \in \mathbb{L}(\mathbb{H}, \mathbb{L}_2(\nu))$, $\mathrm{VTU}^{\star} \in \mathbb{L}(\mathbb{L}_2(\nu))$, $g_{\bullet} = (g_j)_{j\in\mathcal{J}} := \mathrm{V}g \in \mathbb{L}_2(\nu)$ and $\theta_{\bullet} = (\theta_j)_{j\in\mathcal{J}} := \mathrm{U}\theta \in \mathbb{L}_2(\nu)$. Concerning the operator $\mathrm{VTU}^{\star}$ we distinguish two cases, in Section §04 it behaves like a multiplication, i.e. $\mathrm{VTU}^{\star} = \mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$ for some $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$, while in Section §05 and Section §06 we consider the non-diagonal case $\mathrm{VTU}^{\star} = \mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\bullet}(\ell_2)$.

§04 Regularisation by orthogonal projection

§04|00.01 **Notation** (Reminder). Consider the measure space $(\mathcal{J}, \mathscr{J}, \nu)$ and the Hilbert space $\mathbb{J} = \mathbb{L}_2(\nu)$ as in Notation §01|01.01. For $\mathrm{w}_{\bullet} \in \mathbb{R}^{\mathcal{J}}$ define the multiplication map $\mathrm{M}_{\mathrm{w}} : \mathbb{R}^{\mathcal{J}} \to \mathbb{R}^{\mathcal{J}}$ with $a_{\bullet} \mapsto \mathrm{M}_{\mathrm{w}}a_{\bullet} := \mathrm{w}_{\bullet}a_{\bullet} := (\mathrm{w}_j a_j)_{j\in\mathcal{J}}$. If $\mathrm{w}_{\bullet} \in \mathcal{M}(\mathscr{J})$, i.e. $\mathrm{w}_{\bullet}$ is $\mathscr{J}$-$\mathscr{B}$-measurable, then we have $\mathrm{M}_{\mathrm{w}} : \mathcal{M}(\mathscr{J}) \to \mathcal{M}(\mathscr{J})$ too. If in addition $\mathrm{w}_{\bullet} \in \mathbb{L}_\infty(\nu)$ then we have also $\mathrm{M}_{\mathrm{w}} \in \mathbb{L}(\mathbb{J})$ identifying again equivalence classes and representatives. We set $\mathbb{L}^{\mathrm{M}}(\mathbb{J}) := \big\{\mathrm{M}_{\mathrm{w}} : \mathrm{w}_{\bullet} \in \mathbb{L}_\infty(\nu)\big\} \subseteq \mathbb{L}(\mathbb{J})$ noting that $\|\mathrm{M}_{\mathrm{w}}\|_{\mathbb{L}(\mathbb{J})} = \sup\big\{\|\mathrm{w}_{\bullet}a_{\bullet}\|_{\mathbb{J}} : \|a_{\bullet}\|_{\mathbb{J}} \leqslant 1\big\} \leqslant \|\mathrm{w}_{\bullet}\|_{\mathbb{L}_\infty(\nu)}$ for each $\mathrm{M}_{\mathrm{w}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$ (see Notation §01|04.01). Finally, given surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \mathbb{J})$ we define $\mathbb{L}^{\mathrm{U,V}}(\mathbb{L}^{\mathrm{M}}(\mathbb{J})) := \mathrm{V}^{\star}(\mathbb{L}^{\mathrm{M}}(\mathbb{J}))\mathrm{U} := \big\{\mathrm{V}^{\star}\mathrm{M}_{\mathrm{w}}\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{G}) : \mathrm{M}_{\mathrm{w}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})\big\}$. As a consequence, for each $\mathrm{T} \in \mathbb{L}^{\mathrm{U,V}}(\mathbb{L}^{\mathrm{M}}(\mathbb{J}))$ we have $\mathrm{VTU}^{\star} = \mathrm{M}_{\mathrm{w}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$ for some $\mathrm{w}_{\bullet} \in \mathbb{L}_\infty(\nu)$. □

§04|00.02 **Notation** (Reminder (see §01|04.02)). For $A \in \mathscr{J}$ we denote by $\mathbb{1}^A_{\bullet} = (\mathbb{1}^A_j)_{j\in\mathcal{J}}$ the indicator function where for each $j \in \mathcal{J}$, $\mathbb{1}^A_j = 1$ if $j \in A$ and $\mathbb{1}^A_j = 0$ otherwise. Obviously, $\mathbb{1}^A_{\bullet}$ is $\mathscr{J}$-$\mathscr{B}$-measurable, i.e. $\mathbb{1}^A_{\bullet} \in \mathcal{M}(\mathscr{J})$, and it belongs to $\mathbb{L}_\infty(\nu)$, and to $\mathbb{L}_2(\nu)$ whenever $\nu(A) \in \mathbb{R}_{\geqslant 0}$. Since $\{j\} \in \mathscr{J}$ we have $\mathbb{1}^{\{j\}}_{\bullet} \in \mathscr{J}$ and $\mathbb{1}^{\{j\}}_{\bullet} \in \mathbb{L}_\infty(\nu)$. Obviously, we have $\mathbb{1}_{\bullet} = \mathbb{1}^{\mathcal{J}}_{\bullet} \in \mathbb{L}_\infty(\nu)$ and $\mathrm{M}_{\mathbb{1}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$. For each $\mathrm{w}_{\bullet} \in \mathbb{L}_\infty(\nu)$ set $\mathbb{J}\mathrm{w}_{\bullet} := \{\{a_{\bullet}\mathrm{w}_{\bullet}\}_\nu : a_{\bullet} \in \mathcal{L}_2(\nu)\} = \{a_{\bullet}\mathrm{w}_{\bullet} : a_{\bullet} \in \mathbb{J} = \mathbb{L}_2(\nu)\}$ and hence in particular $\mathbb{J}\mathbb{1}^A_{\bullet} = \{a_{\bullet}\mathbb{1}^A_{\bullet} : a_{\bullet} \in \mathbb{J}\}$. Given $0_{\bullet} = (0)_{j\in\mathcal{J}}$ for $\mathrm{w}_{\bullet} \in \mathcal{M}(\mathscr{J})$ we write further $\mathcal{N}_{\mathrm{w}} := \{\mathrm{w}_{\bullet} = 0_{\bullet}\} := \{j \in \mathcal{J} : \mathrm{w}_j = 0\} \in \mathscr{J}$, and denote by $\mathrm{dom}(\mathrm{M}_{\mathrm{w}}) = \{a_{\bullet} \in \mathbb{J} : a_{\bullet}\mathrm{w}_{\bullet} \in \mathbb{J}\}$, $\mathrm{ran}(\mathrm{M}_{\mathrm{w}}) = \{a_{\bullet}\mathrm{w}_{\bullet} : a_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathrm{w}}) \subseteq \mathbb{J}\}$ and $\ker(\mathrm{M}_{\mathrm{w}}) = \{a_{\bullet} \in \mathbb{J} : \{a_{\bullet}\mathrm{w}_{\bullet}\}_\nu = 0_{\bullet}\}$, respectively, the domain, range and nullspace of $\mathrm{M}_{\mathrm{w}} : \mathbb{J} \supseteq \mathrm{dom}(\mathrm{M}_{\mathrm{w}}) \to \mathbb{J}$. We write $\mathrm{w}_{\bullet} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$, if $\mathrm{w}_{\bullet} \in \mathcal{M}(\mathscr{J})$ and $\nu(\mathcal{N}_{\mathrm{w}}) = 0$. Similarly, for $\mathrm{w}_{\bullet} \in \mathcal{M}(\mathscr{J})$ with $\nu(\{\mathrm{w}_{\bullet} \leqslant 0_{\bullet}\}) = 0$ we write $\mathrm{w}_{\bullet} \in \mathcal{M}_{>0,\nu}(\mathscr{J})$. For $\mathrm{w}_{\bullet} \in \mathcal{M}(\mathscr{J})$ we denote its Moore-Penrose inverse by $\mathrm{w}^{\dagger}_{\bullet} := \mathrm{w}^{-1}_{\bullet}\mathbb{1}^{\mathcal{N}^c_{\mathrm{w}}}_{\bullet} \in \mathcal{M}(\mathscr{J})$ meaning $\mathrm{w}^{\dagger}_j := \mathrm{w}^{-1}_j$ if $j \in \mathcal{N}^c_{\mathrm{w}}$ and $\mathrm{w}^{\dagger}_j := 0$ if $j \in \mathcal{N}_{\mathrm{w}}$. Obviously, we have $\mathrm{w}^{\dagger}_{\bullet}\mathrm{w}_{\bullet}\mathrm{w}^{\dagger}_{\bullet} = \mathrm{w}^{\dagger}_{\bullet}$, $\mathrm{w}_{\bullet}\mathrm{w}^{\dagger}_{\bullet}\mathrm{w}_{\bullet} = \mathrm{w}_{\bullet}$ and $\mathrm{w}_{\bullet}\mathrm{w}^{\dagger}_{\bullet} = \mathrm{w}^{\dagger}_{\bullet}\mathrm{w}_{\bullet} = \mathbb{1}^{\mathcal{N}^c_{\mathrm{w}}}_{\bullet}$. □

§04|00.03 **Property**. *For each $\mathrm{w}_{\bullet} \in \mathcal{L}_\infty(\nu)$ the multiplication $\mathrm{M}_{\mathrm{w}} \in \mathbb{L}(\mathbb{J})$ is a linear bounded operator. Keeping $\mathcal{N}_{\mathrm{w}} = \{\mathrm{w}_{\bullet} = 0\} \in \mathscr{J}$ in mind its range and null space is given by $\mathrm{ran}(\mathrm{M}_{\mathrm{w}}) = \mathbb{J}\mathrm{w}_{\bullet}$ and $\ker(\mathrm{M}_{\mathrm{w}}) = \mathbb{J}\mathbb{1}^{\mathcal{N}_{\mathrm{w}}}_{\bullet} = \mathrm{ran}(\mathrm{M}_{\mathbb{1}^{\mathcal{N}_{\mathrm{w}}}})$, respectively. $\mathrm{M}_{\mathrm{w}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$ is consequently injective if and only if $\mathrm{w}_{\bullet} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$, i.e. $\mathrm{w}_{\bullet} \in \mathcal{M}(\mathscr{J})$ and $\nu(\mathcal{N}_{\mathrm{w}}) = 0$. If in addition $\mathrm{w}_{\bullet} \in \mathcal{M}_{\geqslant 0,\nu}(\mathscr{J}) \cap \mathcal{L}_\infty(\nu)$ then the multiplication $\mathrm{M}_{\mathrm{w}} \in \mathbb{L}^{\geqslant}(\mathbb{J}) \subseteq \mathbb{L}(\mathbb{J})$ is a positive semi-definite operator, which is injective if and only if $\mathrm{w}_{\bullet} \in \mathcal{M}_{>0,\nu}(\mathscr{J})$. For each $A \in \mathscr{J}$ setting $A^c := \mathcal{J}\backslash A \in \mathscr{J}$ the range and null space of the multiplication $\mathrm{M}_{\mathbb{1}^A} \in \mathbb{L}^{\geqslant}(\mathbb{J}) \subseteq \mathbb{L}(\mathbb{J})$ is given by $\mathrm{ran}(\mathrm{M}_{\mathbb{1}^A}) = \mathbb{J}\mathbb{1}^A_{\bullet}$ and $\ker(\mathrm{M}_{\mathbb{1}^A}) = \mathbb{J}\mathbb{1}^{A^c}_{\bullet}$, respectively. Obviously, we have $\mathrm{M}^2_{\mathbb{1}^A} = \mathrm{M}_{\mathbb{1}^A}$ and hence $\mathrm{M}_{\mathbb{1}^A}$ is an orthogonal projection and $\mathbb{J} = \mathbb{J}\mathbb{1}^A_{\bullet} \oplus \mathbb{J}\mathbb{1}^{A^c}_{\bullet}$. Moreover, the map $\mathrm{M}_{\mathbb{1}} = \mathrm{id}_{\mathbb{J}}$ equals the identity on $\mathbb{J}$.* □

§04|00.04 **Assumption**. For $\mathbb{J} = \mathbb{L}_2(\nu)$, surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \mathbb{J})$, fixed and presumed to be *known* in advance, let $\mathrm{T} \in \mathbb{L}^{\mathrm{U,V}}(\mathbb{L}^{\mathrm{M}}(\mathbb{J})) \subseteq \mathbb{L}(\mathbb{H}, \mathbb{G})$ and hence $\mathrm{VTU}^{\star} = \mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$ for some $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$, let $g_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^{\dagger}})$, and hence $\mathfrak{s}^{\dagger}_{\bullet}g_{\bullet} \in \mathbb{J} = \mathbb{L}_2(\nu)$. □

In the sequel we consider first in Subsection §04|02 the direct problem, that is $\mathfrak{s}_\bullet = \mathbb{1}_\bullet$ and then secondly in Subsection §04|03 the diagonal inverse problem, that is $\mathfrak{s}_\bullet \in \mathbb{L}_\infty(\nu)$.

§04|01 Weigthed norms and inner products

§04|01.01 **Notation** (Reminder (see §01|00.01)). Extending the real line by the points $-\infty$ and $+\infty$ we define $\overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ and denote by $\overline{\mathscr{B}}$ the Borel-σ-field over $\overline{\mathbb{R}}$ where the trace of $\overline{\mathscr{B}}_{\mathbb{R}} = \overline{\mathscr{B}} \cap \mathbb{R}$ over $\mathbb{R}$ equals $\mathscr{B}$. Thereby, each $a_\bullet \in \mathcal{M}(\mathscr{J})$ is in a canonical way also $\mathscr{J}$-$\overline{\mathscr{B}}$ measurable, $a_\bullet \in \overline{\mathcal{M}}(\mathscr{J})$ for short. For $\mathrm{w}_\bullet \in \overline{\mathcal{M}}(\mathscr{J})$ and hence $\mathrm{w}_\bullet^2 \in \overline{\mathcal{M}}_{\geqslant 0}(\mathscr{J})$, consider the measure $\mathrm{w}_\bullet^2\nu$ on $(\mathcal{J}, \mathscr{J})$, i.e., $\mathrm{w}_\bullet^2 = \mathrm{d}\mathrm{w}_\bullet^2\nu/\mathrm{d}\nu$ is the Radon-Nikodym density of $\mathrm{w}_\bullet^2\nu$ with respect to ν. We write shortly $\langle\cdot,\cdot\rangle_{\mathrm{w}} := \langle\cdot,\cdot\rangle_{\mathbb{L}_2(\mathrm{w}_\bullet^2\nu)}$ and $\|\cdot\|_{\mathrm{w}} := \|\cdot\|_{\mathbb{L}_2(\mathrm{w}_\bullet^2\nu)}$. For $\mathrm{w}_\bullet \in \mathcal{M}(\mathscr{J})$ with Moore-Penrose inverse $\mathrm{w}_\bullet^\dagger := \mathrm{w}_\bullet^{-1}\mathbb{1}_\bullet^{\mathcal{N}_{\mathrm{w}}^c} \in \mathcal{M}(\mathscr{J})$ we set $\mathbb{J}^{\mathrm{w}} := \mathbb{L}_2^{\mathrm{w}}(\nu) := \mathrm{dom}(\mathrm{M}_{\mathrm{w}^\dagger})$ and write $\mathrm{w}_\bullet^{2|\dagger} := (\mathrm{w}_\bullet^\dagger)^2 = (\mathrm{w}_\bullet^2)^\dagger$ for short. □

§04|01.02 **Property**. *Let $\mathrm{w}_\bullet \in \mathcal{M}(\mathscr{J})$. Then $\mathrm{w}_\bullet^2\nu \in \mathscr{M}_\sigma(\mathscr{J})$ is a σ-finite measure satisfying $\mathrm{w}_\bullet^2\nu(|a_\bullet|^2) = \nu(|\mathrm{w}_\bullet a_\bullet|^2)$ for each $a_\bullet \in \mathcal{L}_2(\mathrm{w}_\bullet^2\nu)$, and $\mathbb{L}_2(\mathrm{w}_\bullet^2\nu)$ endowed with inner product $\langle\cdot,\cdot\rangle_{\mathrm{w}} = \langle\cdot,\cdot\rangle_{\mathbb{L}_2(\mathrm{w}_\bullet^2\nu)} = \langle\mathrm{M}_{\mathrm{w}}\cdot, \mathrm{M}_{\mathrm{w}}\cdot\rangle_{\mathbb{L}_2(\nu)}$ is a separable Hilbert space. If in addition $\mathrm{w}_\bullet \in \mathcal{M}(\mathscr{J}) \cap \mathcal{L}_\infty(\nu)$, then*

$$\mathcal{L}_2(\mathrm{w}_\bullet^{2|\dagger}\nu) = {}_{\mathcal{L}_2(\nu)}\mathrm{w}_\bullet + {}_{\overline{\mathcal{M}}(\mathscr{J})}\mathbb{1}_\bullet^{\mathcal{N}_{\mathrm{w}}} = \{\mathrm{w}_\bullet h_\bullet : h_\bullet \in \mathcal{L}_2(\nu)\} + \{h_\bullet\mathbb{1}_\bullet^{\mathcal{N}_{\mathrm{w}}} : h_\bullet \in \overline{\mathcal{M}}(\mathscr{J})\}. \tag{04.01}$$

Indeed, for each $h_\bullet \in \overline{\mathcal{M}}(\mathscr{J})$ consider the decomposition $h_\bullet = \mathrm{w}_\bullet\mathrm{w}_\bullet^\dagger h_\bullet + h_\bullet\mathbb{1}_\bullet^{\mathcal{N}_{\mathrm{w}}}$. The claim follows immediately from the equivalence of $h_\bullet \in \mathcal{L}_2(\mathrm{w}_\bullet^{2|\dagger}\nu)$ and $\mathrm{w}_\bullet^\dagger h_\bullet \in \mathcal{L}_2(\nu)$. Since $\mathrm{w}_\bullet \in \mathcal{L}_\infty(\nu)$ the map $\mathrm{M}_{\mathrm{w}} : \mathcal{L}_2(\nu) \to \mathcal{L}_2(\nu)$ is well-defined, and (similar to (04.01))

$$\mathrm{dom}(\mathrm{M}_{\mathrm{w}^\dagger}) = \{h_\bullet \in \mathcal{L}_2(\nu) : \mathrm{w}_\bullet^\dagger h_\bullet \in \mathcal{L}_2(\nu)\} = {}_{\mathcal{L}_2(\nu)}\mathrm{w}_\bullet + {}_{\mathcal{L}_2(\nu)}\mathbb{1}_\bullet^{\mathcal{N}_{\mathrm{w}}} \subseteq \mathcal{L}_2(\mathrm{w}_\bullet^{2|\dagger}\nu).$$

Consequently, if in addition $\mathcal{N}_{\mathrm{w}} = \emptyset$, then $\mathrm{dom}(\mathrm{M}_{\mathrm{w}^\dagger}) = \mathcal{L}_2(\mathrm{w}_\bullet^{2|\dagger}\nu)$. If $\mathrm{w}_\bullet \in \mathbb{L}_\infty(\nu)$ then $\mathrm{M}_{\mathrm{w}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$, and (with Moore-Penrose inverse $\mathrm{w}_\bullet^\dagger$ of a representative $\mathrm{w}_\bullet \in \mathcal{M}(\mathscr{J})$) $\mathrm{M}_{\mathrm{w}^\dagger} : \mathbb{J} \supseteq \mathrm{dom}(\mathrm{M}_{\mathrm{w}^\dagger}) \to \mathbb{J}$. Moreover, we have $\mathrm{dom}(\mathrm{M}_{\mathrm{w}}) = \mathbb{J}$, $\mathrm{ran}(\mathrm{M}_{\mathrm{w}}) = {}_{\mathbb{J}}\mathrm{w}_\bullet$ and $\ker(\mathrm{M}_{\mathrm{w}}) = {}_{\mathbb{J}}\mathbb{1}_\bullet^{\mathcal{N}_{\mathrm{w}}}$ (see Property §04|00.03). Therewith, it follows $\mathrm{dom}(\mathrm{M}_{\mathrm{w}^\dagger}) = {}_{\mathbb{J}}\mathrm{w}_\bullet \oplus {}_{\mathbb{J}}\mathbb{1}_\bullet^{\mathcal{N}_{\mathrm{w}}}$. Consequently, if in addition $\nu(\mathcal{N}_{\mathrm{w}}) = 0$, then $\mathbb{J}^{\mathrm{w}} = \mathbb{L}_2^{\mathrm{w}}(\nu) = \mathrm{dom}(\mathrm{M}_{\mathrm{w}^\dagger}) = {}_{\mathbb{J}}\mathrm{w}_\bullet = \mathbb{L}_2(\mathrm{w}_\bullet^{2|\dagger}\nu)$. The last equality follows from (04.01) since both measures $\mathrm{w}_\bullet^{2|\dagger}\nu$ and ν share the same null sets (i.e. they mutually dominate each other). □

§04|02 Direct problem

We assume throughout this subsection that Assumption §04|00.04 is satisfied with $\mathfrak{s}_\bullet = \mathbb{1}_\bullet \in \mathbb{L}_\infty(\nu)$.

§04|02.01 **Notation**. For a non-empty and generally non-finite subset $\mathcal{J}$ of $\mathbb{N}$, $\mathbb{Z}$ or $\mathbb{R}$ and $m \in \mathbb{N}$ we set $[\![m]\!] := [-m, m] \cap \mathcal{J}$ and assuming $[\![m]\!] \in \mathscr{J}$ we write shortly $\mathbb{1}_\bullet^m = (\mathbb{1}_j^m)_{j\in\mathcal{J}} := \mathbb{1}_\bullet^{[\![m]\!]} \in \mathcal{M}(\mathscr{J})$. Furthermore, we define $\mathbb{1}_\bullet^{m|\perp} := \mathbb{1}_\bullet - \mathbb{1}_\bullet^m \in \mathcal{M}(\mathscr{J})$. □

§04|02.02 **Property**. *For each $m \in \mathbb{N}$, $\mathrm{M}_{\mathbb{1}^m} \in \mathbb{L}^{\geqslant}(\mathbb{J})$ and $\mathrm{M}_{\mathbb{1}^{m|\perp}} \in \mathbb{L}^{\geqslant}(\mathbb{J})$ is the orthogonal projection onto the linear subspace ${}_{\mathbb{J}}\mathbb{1}_\bullet^m \subseteq \mathbb{J}$ and its orthogonal complement ${}_{\mathbb{J}}\mathbb{1}_\bullet^{m|\perp} = ({}_{\mathbb{J}}\mathbb{1}_\bullet^m)^\perp \subseteq \mathbb{J}$, respectively, that is $\mathbb{J} = {}_{\mathbb{J}}\mathbb{1}_\bullet^m \oplus {}_{\mathbb{J}}\mathbb{1}_\bullet^{m|\perp}$. We have point-wise $\mathbb{1}_\bullet^m - \mathbb{1}_\bullet = \mathrm{o}(1)$ as $m \to \infty$ meaning that for each $j \in \mathcal{J}$ holds $\mathbb{1}_j^m - \mathbb{1}_j = \mathrm{o}(1)$ as $m \to \infty$. Considering the orthogonal projection $\mathrm{M}_{\mathbb{1}^m} \in \mathbb{L}^{\geqslant}(\mathbb{J})$ and the identity $\mathrm{M}_{\mathbb{1}} = \mathrm{id}_{\mathbb{J}} \in \mathbb{L}^{\geqslant}(\mathbb{J})$ point-wise convergence $\mathrm{M}_{\mathbb{1}^m} - \mathrm{id}_{\mathbb{J}} = \mathrm{o}(1)$ as $m \to \infty$ holds too, that is, $\|(\mathrm{M}_{\mathbb{1}^m} - \mathrm{id}_{\mathbb{J}})a_\bullet\|_{\mathbb{J}} = \|(\mathbb{1}_\bullet^m - \mathbb{1}_\bullet)a_\bullet\|_{\mathbb{J}} = \|\mathbb{1}_\bullet^{m|\perp}a_\bullet\|_{\mathbb{J}} = \mathrm{o}(1)$ as $m \to \infty$ for all $a_\bullet \in \mathbb{J}$.* □

§04|02.03 **Orthogonal projection**. Given $m \in \mathbb{N}$ we define for each $\theta_\bullet = \mathrm{U}\theta \in \mathbb{J}$ its orthogonal projection $\theta_\bullet^m := \theta_\bullet\mathbb{1}_\bullet^m \in {}_{\mathbb{J}}\mathbb{1}_\bullet^m$ (and $\theta^m := \mathrm{U}^\star\theta_\bullet^m \in \mathbb{H}$). □

§04|02|01 Global and maximal global $\mathfrak{v}$-error

We shall measure first globally the accuracy of the orthogonal projection $\theta_{\bullet}^{m} := \theta_{\bullet}\mathbb{1}_{\bullet}^{m}$ of $\theta_{\bullet} \in \mathbb{J}$.

§04|02.04 **Property**. *If $\mathfrak{v}_{\bullet} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ (i.e. $\mathfrak{v}_{\bullet} \in \mathcal{M}(\mathscr{J})$ and $\nu(\mathcal{N}_{\mathfrak{v}}) = 0$) and $\theta_{\bullet} \in \mathbb{L}_2(\mathfrak{v}_{\bullet}^2\nu)$ (i.e. $\|\theta_{\bullet}\|_{\mathfrak{v}}^2 = \mathfrak{v}_{\bullet}^2\nu(\theta_{\bullet}^2) \in \mathbb{R}_{\geqslant 0}$), then for each $m \in \mathbb{N}$ we have $\theta_{\bullet}^{m} \in \mathbb{L}_2(\mathfrak{v}_{\bullet}^2\nu)$ too, since $\|\theta_{\bullet}^{m}\|_{\mathfrak{v}}^2 = \mathfrak{v}_{\bullet}^2\nu(\theta_{\bullet}^2\mathbb{1}_{\bullet}^{m}) \leqslant \mathfrak{v}_{\bullet}^2\nu(\theta_{\bullet}^2)$. Moreover, it holds $\|\theta_{\bullet}^{m} - \theta_{\bullet}\|_{\mathfrak{v}}^2 = \|\theta_{\bullet}\mathbb{1}_{\bullet}^{m|\perp}\|_{\mathfrak{v}}^2 = \mathfrak{v}_{\bullet}^2\nu(\theta_{\bullet}^2\mathbb{1}_{\bullet}^{m|\perp}) \leqslant \mathfrak{v}_{\bullet}^2\nu(\theta_{\bullet}^2) \in \mathbb{R}_{\geqslant 0}$ and $\|\theta_{\bullet}^{m} - \theta_{\bullet}\|_{\mathfrak{v}}^2 = o(1)$ as $m \to \infty$ by dominated convergence.* □

§04|02.05 **Comment**. We assume throughout this chapter that the Hilbert space $\mathbb{J} = \mathbb{L}_2(\mathcal{J}, \mathscr{J}, \nu)$ and the *surjective partial isometry* $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ is fixed and *known* in advance. Considering a $\mathfrak{v}$-error means the weight sequences $\mathfrak{v}_{\bullet} \in \mathcal{M}(\mathscr{J})$ is also fixed and known in advance. Consequently, the condition $\mathfrak{v}_{\bullet} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ does not impose an additional restriction. □

§04|02.06 **Global $\mathfrak{v}$-error**. Given $\mathfrak{v}_{\bullet} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$, $m \in \mathbb{N}$, a solution $\theta_{\bullet} = \mathrm{U}\theta \in \mathbb{L}_2(\mathfrak{v}_{\bullet}^2\nu)$ and its orthogonal projection $\theta_{\bullet}^{m} = \theta_{\bullet}\mathbb{1}_{\bullet}^{m} \in \mathbb{J}\mathbb{1}_{\bullet}^{m}$ we call $\|\theta_{\bullet}^{m} - \theta_{\bullet}\|_{\mathfrak{v}} = \|\theta_{\bullet}\mathbb{1}_{\bullet}^{m|\perp}\|_{\mathfrak{v}} \in \mathbb{R}_{\geqslant 0}$ *global $\mathfrak{v}$-error*. □

§04|02.07 **Assumption**. Consider weights $\mathfrak{a}_{\bullet}, \mathfrak{v}_{\bullet} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ (i.e. $\mathfrak{a}_{\bullet}, \mathfrak{v}_{\bullet} \in \mathcal{M}(\mathscr{J})$ and $\nu(\mathcal{N}_{\mathfrak{a}}) = 0 = \nu(\mathcal{N}_{\mathfrak{v}})$), such that $\mathfrak{a}_{\bullet} \in \mathbb{L}_\infty(\nu)$ and $(\mathfrak{a}\mathfrak{v})_{\bullet} := (\mathfrak{a}_j\mathfrak{v}_j)_{j\in\mathcal{J}} = \mathfrak{a}_{\bullet}\mathfrak{v}_{\bullet} \in \mathbb{L}_\infty(\nu)$. We write $(\mathfrak{a}\mathfrak{v})_{(m)} := \|(\mathfrak{a}\mathfrak{v})_{\bullet}\mathbb{1}_{\bullet}^{m|\perp}\|_{\mathbb{L}_\infty(\nu)} \in \mathbb{R}_{\geqslant 0}$ for each $m \in \mathbb{N}$. □

§04|02.08 **Notation**. For sequences $a_{\bullet}, b_{\bullet} \in (\mathbb{K})^{\mathbb{N}}$ taking its values in $\mathbb{K} \in \{\mathbb{R}, \mathbb{R}_{\geqslant 0}, \mathbb{R}_{>0}, \mathbb{Q}, \mathbb{Z}, \dots\}$ we write $a_{\bullet} \in (\mathbb{K})^{\mathbb{N}}_{\nearrow}$ and $b_{\bullet} \in (\mathbb{K})^{\mathbb{N}}_{\searrow}$ if $a_{\bullet}$ and $b_{\bullet}$, respectively, is monotonically *non-decreasing* and *non-increasing*. If in addition $a_n \to \infty$ and $b_n \to 0$ as $n \to \infty$, then we write $a_{\bullet} \in (\mathbb{K})^{\mathbb{N}}_{\uparrow\infty}$ and $b_{\bullet} \in (\mathbb{K})^{\mathbb{N}}_{\downarrow 0}$ for short. For $\mathrm{w}_{\bullet} \in \mathbb{L}_\infty(\nu)$ we set $\mathrm{w}_{(0)} := \|\mathrm{w}_{\bullet}\|_{\mathbb{L}_\infty(\nu)}$ and $\mathrm{w}_{(\bullet)} = (\mathrm{w}_{(j)} := \|\mathrm{w}_{\bullet}\mathbb{1}_{\bullet}^{j|\perp}\|_{\mathbb{L}_\infty(\nu)})_{j\in\mathbb{N}}$, where by construction $\mathrm{w}_{(\bullet)} \in (\mathbb{R}_{\geqslant 0})^{\mathbb{N}}_{\searrow}$. □

§04|02.09 **Reminder**. Under Assumption §04|02.07 we have $\mathbb{J}^{\mathfrak{a}} = \mathbb{L}_2^{\mathfrak{a}}(\nu) = \mathrm{dom}(\mathrm{M}_{\mathfrak{a}^\dagger}) = \mathbb{J}\mathfrak{a}_{\bullet} = \mathbb{L}_2(\mathfrak{a}_{\bullet}^{2|\dagger}\nu)$ and the three measures ν, $\mathfrak{a}_{\bullet}^{2|\dagger}\nu$ and $\mathfrak{v}_{\bullet}^2\nu$ dominate mutually each other, i.e. they share the same null sets (see Property §04|01.02). Consequently, $\mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{J} = \mathbb{L}_2(\nu)$ and if $h_{\bullet} \in \mathbb{L}_2(\mathfrak{a}_{\bullet}^{2|\dagger}\nu)$ satisfies $\mathfrak{v}_{\bullet}^2\nu(h_{\bullet}^2) \in \mathbb{R}_{\geqslant 0}$, for example, then $h_{\bullet} \in \mathbb{L}_2(\mathfrak{v}_{\bullet}^2\nu)$ too. □

§04|02.10 **Notation**. Under Assumption §04|02.07 and given a constant $\mathrm{r} \in \mathbb{R}_{>0}$ we consider $\mathbb{J}^{\mathfrak{a}} = \mathbb{L}_2^{\mathfrak{a}}(\nu) = \mathbb{L}_2(\mathfrak{a}_{\bullet}^{2|\dagger}\nu)$ endowed with $\|\cdot\|_{\mathfrak{a}^\dagger} := \|\cdot\|_{\mathbb{J}^{\mathfrak{a}}} := \|\cdot\|_{\mathbb{L}_2(\mathfrak{a}_{\bullet}^{2|\dagger}\nu)}$ and the ellipsoid

$$\mathbb{J}^{\mathfrak{a},\mathrm{r}} := \left\{h_{\bullet} \in \mathbb{J}^{\mathfrak{a}}\colon \|h_{\bullet}\|_{\mathfrak{a}^\dagger}^2 = \mathfrak{a}_{\bullet}^{2|\dagger}\nu(h_{\bullet}^2) = \nu(\mathfrak{a}_{\bullet}^{2|\dagger}h_{\bullet}^2) \leqslant \mathrm{r}^2\right\} \subseteq \mathbb{J}^{\mathfrak{a}}.$$ □

§04|02.11 **Property**. *Under Assumption §04|02.07 we have $\mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{L}_2(\mathfrak{v}_{\bullet}^2\nu)$. Indeed, for each $h_{\bullet} \in \mathbb{J}^{\mathfrak{a}}$ (i.e., $\|h_{\bullet}\|_{\mathfrak{a}^\dagger} \in \mathbb{R}_{\geqslant 0}$) follows $\|h_{\bullet}\|_{\mathfrak{v}}^2 = \nu(h_{\bullet}^2\mathfrak{a}_{\bullet}^{2|\dagger}(\mathfrak{a}\mathfrak{v})_{\bullet}^2) \leqslant \|h_{\bullet}\|_{\mathfrak{a}^\dagger}^2\|(\mathfrak{a}\mathfrak{v})_{\bullet}\|_{\mathbb{L}_\infty(\nu)}^2 \in \mathbb{R}_{\geqslant 0}$.* □

§04|02.12 **Abstract smoothness condition**. Under Assumption §04|02.07 a solution $\theta_{\bullet} \in \mathbb{J}$ satisfies an *abstract smoothness condition* if there is $\mathrm{r} \in \mathbb{R}_{>0}$ such that $\theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}} \subseteq \mathbb{J}^{\mathfrak{a}}$. □

§04|02.13 **Lemma**. *Under Assumption §04|02.07 for each $m \in \mathbb{N}$ and solution $\theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}} \subseteq \mathbb{L}_2(\mathfrak{v}_{\bullet}^2\nu)$ its orthogonal projection $\theta_{\bullet}^{m} := \theta_{\bullet}\mathbb{1}_{\bullet}^{m} \in \mathbb{J}\mathbb{1}_{\bullet}^{m}$ satisfies $\|\theta_{\bullet}^{m} - \theta_{\bullet}\|_{\mathfrak{v}} = \|\theta_{\bullet}\mathbb{1}_{\bullet}^{m|\perp}\|_{\mathfrak{v}} \leqslant \mathrm{r}\,(\mathfrak{a}\mathfrak{v})_{(m)}$.*

§04|02.14 **Proof** of Lemma §04|02.13. Given in the lecture. □

§04|02.15 **Maximal global $\mathfrak{v}$-error**. Under Assumption §04|02.07 for $m \in \mathbb{N}$, a solution $\theta_{\bullet} = \mathrm{U}\theta \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ and its orthogonal projection $\theta_{\bullet}^{m} = \theta_{\bullet}\mathbb{1}_{\bullet}^{m} \in \mathbb{J}\mathbb{1}_{\bullet}^{m}$ we call $\sup\left\{\|\theta_{\bullet}^{m} - \theta_{\bullet}\|_{\mathfrak{v}}\colon \theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}\right\}$ *maximal global $\mathfrak{v}$-error* over the class of solutions $\mathbb{J}^{\mathfrak{a},\mathrm{r}}$. □

§04|02|02 Local and maximal local ϕ-error

Secondly, we measure locally the accuracy of the orthogonal projection $\theta^m_{\bullet} := \theta_{\bullet}\mathbb{1}^m_{\bullet}$ of $\theta_{\bullet} \in \mathbb{J}$.

§04|02.16 **Notation**. For $\phi \in \mathcal{M}(\mathscr{J})$ and $\mathrm{dom}(\phi\nu) := \{h_{\bullet} \in \mathbb{J} = \mathbb{L}_2(\nu): \phi h_{\bullet} \in \mathbb{L}_1(\nu)\}$ we consider the linear functional $\phi\nu : \mathbb{J} \supseteq \mathrm{dom}(\phi\nu) \to \mathbb{R}$ given by $h_{\bullet} \mapsto \phi\nu(h_{\bullet}) := \nu(\phi h_{\bullet})$ with a slight abuse of notations. □

§04|02.17 **Comment**. If $\phi \in \mathbb{J} = \mathbb{L}_2(\nu)$, then it follows $\mathrm{dom}(\phi\nu) = \mathbb{J}$ and $\|\phi\nu\|_{\mathbb{L}(\mathbb{J},\mathbb{R})} = \|\phi\|_{\mathbb{J}} \in \mathbb{R}_{\geqslant 0}$. Consequently, we have $\phi\nu \in \mathbb{L}(\mathbb{J},\mathbb{R})$ and $\phi\nu(h_{\bullet}) = \langle h_{\bullet}, \phi\rangle_{\mathbb{J}}$, in other words ϕ is a Fréchet-Riesz representative of the continuous linear functional $\phi\nu$. □

§04|02.18 **Property**. *If $\phi \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ (i.e. $\phi \in \mathcal{M}(\mathscr{J})$ and $\nu(\mathcal{N}_{\phi}) = 0$) and $\theta_{\bullet} \in \mathrm{dom}(\phi\nu)$ (i.e. $\theta_{\bullet}\phi \in \mathbb{L}_1(\nu)$), then for each $m \in \mathbb{N}$ we have $\theta^m_{\bullet} \in \mathrm{dom}(\phi\nu)$ too, since $\|\phi\theta^m_{\bullet}\|_{\mathbb{L}_1(\nu)} = \nu(|\phi\theta_{\bullet}|\mathbb{1}^m_{\bullet}) \leqslant \nu(|\phi\theta_{\bullet}|)$. Moreover, it holds*

$$|\phi\nu(\theta_{\bullet}) - \phi\nu(\theta^m_{\bullet})| \leqslant |\phi|\nu(|\theta^m_{\bullet} - \theta_{\bullet}|) = |\phi|\nu(|\theta_{\bullet}|\mathbb{1}^{m|\perp}_{\bullet}) \leqslant \nu(|\phi\theta_{\bullet}|) \in \mathbb{R}_{\geqslant 0}$$

and $|\phi\nu(\theta_{\bullet}) - \phi\nu(\theta^m_{\bullet})| = \mathrm{o}(1)$ as $m \to \infty$ by dominated convergence. □

§04|02.19 **Comment**. We assume throughout this chapter that the Hilbert space $\mathbb{J} = \mathbb{L}_2(\mathcal{J}, \mathscr{J}, \nu)$ and the *surjective partial isometry* $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ is fixed and known in advance (Assumption §04|00.04). Considering a ϕ-error means the linear function $\phi\nu$ and hence in equal $\phi \in \mathscr{J}$ is also fixed and known in advance. Consequently, the condition $\phi \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ does not impose an additional restriction. □

§04|02.20 **Local ϕ-error**. Given $\phi \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$, $m \in \mathbb{N}$, a solution $\theta_{\bullet} = \mathrm{U}\theta \in \mathrm{dom}(\phi\nu)$ and its orthogonal projection $\theta^m_{\bullet} = \theta_{\bullet}\mathbb{1}^m_{\bullet} \in \mathbb{J}\mathbb{1}^m_{\bullet}$ we call $|\phi\nu(\theta_{\bullet}) - \phi\nu(\theta^m_{\bullet})| = |\phi\nu(\theta_{\bullet}\mathbb{1}^{m|\perp}_{\bullet})| \in \mathbb{R}_{\geqslant 0}$ *local ϕ-error*. □

§04|02.21 **Assumption**. Consider $\phi, \mathfrak{a}_{\bullet} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ (i.e. $\phi, \mathfrak{a}_{\bullet} \in \mathcal{M}(\mathscr{J})$ and $\nu(\mathcal{N}_{\phi}) = 0 = \nu(\mathcal{N}_{\mathfrak{a}})$), such that $\mathfrak{a}_{\bullet} \in \mathbb{L}_{\infty}(\nu)$ and $(\mathfrak{a}\phi)_{\bullet} := (\mathfrak{a}_j\phi_j)_{j\in\mathcal{J}} = \mathfrak{a}_{\bullet}\phi \in \mathbb{L}_2(\nu)$ and hence $\|\mathfrak{a}_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\phi} = \|(\mathfrak{a}\phi)_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\mathbb{L}_2(\nu)} = \mathrm{o}(1)$ as $m \to \infty$. □

§04|02.22 **Reminder**. Under Assumption §04|02.21 we have $\mathbb{J}^{\mathfrak{a}} = \mathbb{L}^{\mathfrak{a}}_2(\nu) = \mathrm{dom}(\mathrm{M}_{\mathfrak{a}^{\dagger}}) = \mathbb{J}\mathfrak{a}_{\bullet} = \mathbb{L}_2(\mathfrak{a}^{2|\dagger}_{\bullet}\nu)$ and the three measures ν, $|\phi|\nu$ and $\mathfrak{a}^{2|\dagger}_{\bullet}\nu$ dominate mutually each other (see Property §04|01.02). Consequently, $\mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{J} = \mathbb{L}_2(\nu)$ and if $h_{\bullet} \in \mathbb{L}_2(\mathfrak{a}^{2|\dagger}_{\bullet}\nu)$ satisfies $\nu(|\phi h_{\bullet}|) \in \mathbb{R}_{\geqslant 0}$, for example, then $h_{\bullet} \in \mathbb{L}_1(|\phi|\nu)$ too. □

§04|02.23 **Property**. *Under Assumption §04|02.21 we have $\mathbb{J}^{\mathfrak{a}} \subseteq \mathrm{dom}(\phi\nu)$. Indeed, for each $h_{\bullet} \in \mathbb{J}^{\mathfrak{a}}$, i.e. $\|h_{\bullet}\|_{\mathfrak{a}^{\dagger}} \in \mathbb{R}_{\geqslant 0}$, we have $\|\phi h_{\bullet}\|_{\mathbb{L}_1(\nu)} = \nu(|h_{\bullet}\mathfrak{a}^{\dagger}_{\bullet}(\mathfrak{a}\phi)_{\bullet}|) \leqslant \|h_{\bullet}\|_{\mathfrak{a}^{\dagger}}\|(\mathfrak{a}\phi)_{\bullet}\|_{\mathbb{L}_2(\nu)} \in \mathbb{R}_{\geqslant 0}$.* □

§04|02.24 **Notation** (Reminder). Under Assumption §04|02.21 a solution $\theta_{\bullet} = \mathrm{U}\theta \in \mathbb{J}$ satisfies an abstract smoothness condition if there is $\mathrm{r} \in \mathbb{R}_{>0}$ such that $\theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}} = \{h_{\bullet} \in \mathbb{J}^{\mathfrak{a}}: \|h_{\bullet}\|^2_{\mathfrak{a}^{\dagger}} \leqslant \mathrm{r}^2\} \subseteq \mathbb{J}^{\mathfrak{a}}$ where $\|\cdot\|_{\mathfrak{a}^{\dagger}} = \|\cdot\|_{\mathbb{J}^{\mathfrak{a}}} := \|\cdot\|_{\mathbb{L}_2(\mathfrak{a}^{2|\dagger}_{\bullet}\nu)}$ (see Definition §04|02.12). Since $(\mathfrak{a}\phi)_{\bullet} \in \mathbb{L}_2(\nu)$ we have $\|\mathfrak{a}_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\phi} = \|(\mathfrak{a}\phi)_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\mathbb{L}_2(\nu)} = \mathrm{o}(1)$ as $m \to \infty$ by dominated convergence. □

§04|02.25 **Lemma**. *Under Assumption §04|02.21 for each $m \in \mathbb{N}$ and $\theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}} \subseteq \mathrm{dom}(\phi\nu)$ its orthogonal projection $\theta^m_{\bullet} := \theta_{\bullet}\mathbb{1}^m_{\bullet} \in \mathbb{J}\mathbb{1}^m_{\bullet}$ of satisfies $|\phi\nu(\theta_{\bullet} - \theta^m_{\bullet})| = |\phi\nu(\theta_{\bullet}\mathbb{1}^{m|\perp}_{\bullet})| \leqslant \nu(|\phi\theta_{\bullet}|\mathbb{1}^{m|\perp}_{\bullet}) \leqslant \mathrm{r}\,\|\mathfrak{a}_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\phi}$.*

§04|02.26 **Proof** of Lemma §04|02.25. Given in the lecture. □

§04|02.27 **Maximal local ϕ-error**. Under Assumption §04|02.21 for $m \in \mathbb{N}$, a solution $\theta_{\bullet} = \mathrm{U}\theta \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ and its orthogonal projection $\theta^m_{\bullet} = \theta_{\bullet}\mathbb{1}^m_{\bullet} \in \mathbb{J}\mathbb{1}^m_{\bullet}$ we call $\sup\{|\phi\nu(\theta_{\bullet}) - \phi\nu(\theta^m_{\bullet})|: \theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}\}$ *maximal local ϕ-error* over the class of solutions $\mathbb{J}^{\mathfrak{a},\mathrm{r}}$. □

§04|03 Diagonal inverse problem

We assume throughout this subsection that Assumption §04|00.04 is satisfied with $\mathfrak{s}_\bullet \in \mathbb{L}_\infty(\nu)$.

§04|03.01 **Reminder**. Under Assumption §04|00.04 we consider $T \in \mathbb{L}^{U,V}(\mathbb{L}^{M}(\mathbb{J})) \subseteq \mathbb{L}(\mathbb{H},\mathbb{G})$, and hence $VTU^\star = M_{\mathfrak{s}} \in \mathbb{L}^{M}(\mathbb{J})$ and $g_\bullet = M_{\mathfrak{s}}\theta_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathbb{J}$ for some $\mathfrak{s}_\bullet \in \mathbb{L}_\infty(\nu)$. Due to Property §04|01.02 the Moore-Penrose inverse of $M_{\mathfrak{s}} \in \mathbb{L}^{M}(\mathbb{J})$ satisfies $M_{\mathfrak{s}}^\dagger = M_{\mathfrak{s}^\dagger} : \mathbb{J} \supseteq \mathrm{dom}(M_{\mathfrak{s}^\dagger}) \to \mathbb{J}$ with $\mathrm{dom}(M_{\mathfrak{s}^\dagger}) = \mathbb{J}\mathfrak{s}_\bullet \oplus \mathbb{J}\mathbb{1}^{\mathcal{N}_{\mathfrak{s}}}_\bullet = \mathbb{J}^{\mathfrak{s}}$. For each $m \in \mathbb{N}$, $M_{\mathbb{1}^m} \in \mathbb{L}^{\geqslant}(\mathbb{J})$ and $M_{\mathbb{1}^{m|\perp}} \in \mathbb{L}^{\geqslant}(\mathbb{J})$ is the *orthogonal projection* onto the linear subspace $\mathbb{J}\mathbb{1}^m_\bullet \subseteq \mathbb{J}$ and its orthogonal complement $\mathbb{J}\mathbb{1}^{m|\perp}_\bullet = (\mathbb{J}\mathbb{1}^m_\bullet)^\perp \subseteq \mathbb{J}$, respectively, that is $\mathbb{J} = \mathbb{J}\mathbb{1}^m_\bullet \oplus \mathbb{J}\mathbb{1}^{m|\perp}_\bullet$ (see Property §04|02.02). Given $g_\bullet \in \mathbb{J}$ we call $\theta_\bullet \in \mathbb{J}$ satisfying $\|g_\bullet - \mathfrak{s}_\bullet\theta_\bullet\|_{\mathbb{J}} = \inf\{\|g_\bullet - \mathfrak{s}_\bullet h_\bullet\|_{\mathbb{J}} : h_\bullet \in \mathbb{J}\}$ a least squares solution, if it exists (see Property §03|00.05). □

§04|03.02 **Property**. *For $\mathfrak{s}_\bullet \in \mathbb{L}_\infty(\nu)$ and each $g_\bullet \in \mathrm{dom}(M_{\mathfrak{s}^\dagger}) = \mathbb{J}\mathfrak{s}_\bullet \oplus \mathbb{J}\mathbb{1}^{\mathcal{N}_{\mathfrak{s}}}_\bullet$ is $\theta_\bullet = M_{\mathfrak{s}^\dagger}g_\bullet = \mathfrak{s}^\dagger_\bullet g_\bullet$ the unique least square solution with minimal $\|\cdot\|_{\mathbb{J}}$-norm in the set $\mathfrak{s}^\dagger_\bullet g_\bullet + \mathbb{J}\mathbb{1}^{\mathcal{N}_{\mathfrak{s}}}_\bullet$ of all least square solutions. If in addition $\nu(\mathcal{N}_{\mathfrak{s}}) = 0$ (i.e. $M_{\mathfrak{s}}$ is injective), then $\theta_\bullet = \mathfrak{s}^\dagger_\bullet g_\bullet$ is the unique least square solution. Given $m \in \mathbb{N}$ for each $g_\bullet \in \mathrm{dom}(M_{\mathfrak{s}^\dagger})$ we have $g_\bullet\mathbb{1}^m_\bullet \in \mathrm{dom}(M_{\mathfrak{s}^\dagger})$ too. In particular, for $\theta_\bullet = \mathfrak{s}^\dagger_\bullet g_\bullet$ it follows $\theta_\bullet\mathbb{1}^m_\bullet = (\mathfrak{s}^\dagger_\bullet g_\bullet)\mathbb{1}^m_\bullet = \mathfrak{s}^\dagger_\bullet(g_\bullet\mathbb{1}^m_\bullet) \in \mathbb{J}\mathbb{1}^m_\bullet$.* □

§04|03.03 **Orthogonal projection**. Given $m \in \mathbb{N}$ we define for each $g_\bullet \in \mathrm{dom}(M_{\mathfrak{s}^\dagger})$ and $\theta_\bullet = \mathfrak{s}^\dagger_\bullet g_\bullet \in \mathbb{J}$ the orthogonal projections $g^m_\bullet = g_\bullet\mathbb{1}^m_\bullet \in \mathbb{J}\mathbb{1}^m_\bullet$ and $\theta^m_\bullet = (\mathfrak{s}^\dagger_\bullet g_\bullet)\mathbb{1}^m_\bullet = \mathfrak{s}^\dagger_\bullet g^m_\bullet \in \mathbb{J}\mathbb{1}^m_\bullet$. □

§04|03.04 **Assumption**. Consider weights $\mathfrak{a}_\bullet, \mathfrak{t}_\bullet \in \mathcal{M}_{>0,\nu}(\mathscr{J}) \cap \mathcal{L}_\infty(\nu)$ and hence $(\mathfrak{t}\mathfrak{a})_\bullet := \mathfrak{t}_\bullet\mathfrak{a}_\bullet \in \mathcal{L}_\infty(\nu)$. □

§04|03.05 **Link condition**. Given weights $\mathfrak{t}_\bullet \in \mathcal{M}_{>0,\nu}(\mathscr{J}) \cap \mathcal{L}_\infty(\nu)$, an operator $M_{\mathfrak{s}} \in \mathbb{L}^{M}(\mathbb{J})$, and hence $\mathfrak{s}_\bullet \in \mathbb{L}_\infty(\nu)$, satisfies a *link condition* if there is $d \in \mathbb{R}_{\geqslant 1}$ such that

$$M_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},d} := \left\{M_{w} \in \mathbb{L}^{M}(\mathbb{J}) : |w_\bullet| \leqslant d\mathfrak{t}_\bullet \wedge \mathfrak{t}_\bullet \leqslant d|w_\bullet| \ \nu\text{-a.e.}\right\}.$$ □

§04|03.06 **Property**. *Given weights $\mathfrak{t}_\bullet \in \mathcal{M}_{>0,\nu}(\mathscr{J}) \cap \mathcal{L}_\infty(\nu)$ and introducing $\|\cdot\|_{\mathfrak{t}} := \|M_{\mathfrak{t}}\cdot\|_{\mathbb{J}}$ we have*

$$\mathbb{M}_{\mathfrak{t},d} = \left\{M \in \mathbb{L}^{M}(\mathbb{J}) : d^{-1}\|h_\bullet\|_{\mathfrak{t}} \leqslant \|Mh_\bullet\|_{\mathbb{J}} \leqslant d\|h_\bullet\|_{\mathfrak{t}},\ \forall h_\bullet \in \mathbb{J}\right\}$$

moreover, for each $M \in \mathbb{M}_{\mathfrak{t},d}$ and for all $s \in \mathbb{R}$ (exploiting $M_{\mathfrak{t}}^s = M_{\mathfrak{t}^s}$) holds

$$d^{-|s|}\|h_\bullet\|_{\mathfrak{t}^s} \leqslant \|M^s h_\bullet\|_{\mathbb{J}} \leqslant d^{|s|}\|h_\bullet\|_{\mathfrak{t}^s}, \qquad \forall h_\bullet \in \mathrm{dom}(M_{\mathfrak{t}^s}).$$ □

§04|03.07 **Comment**. Given $M \in \mathbb{M}_{\mathfrak{t},d}$ there exists $w_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J}) \cap \mathcal{L}_\infty(\nu)$ such that $M = M_{w}$. Consequently, we have $\nu(\mathcal{N}_{w}) = 0$ and hence for each $s \in \mathbb{R}$ the value w_j^s is well-defined for all $j \in \mathcal{N}_{w}^c$, and thus $w_\bullet^s$ is ν-a.e. defined. In particular it follows $w_\bullet^\dagger = w_\bullet^{-1}$ ν-a.e., and hence $M^\dagger = M_{w^{-1}}$. Similarly, we have $M^s = M_{w^s}$ for each $s \in \mathbb{R}$. □

§04|03.08 **Property**. *Under Assumption §04|03.04 let $\theta_\bullet \in \mathbb{J}$ and $M_{\mathfrak{s}} \in \mathbb{L}^{M}(\mathbb{J})$ satisfy, respectively, an abstract smoothness condition $\theta_\bullet \in \mathbb{J}^{\mathfrak{a},r}$ as in Definition §04|02.12 and a link condition $M_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},d}$ as in Definition §04|03.05, then $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathbb{J}$ fulfils an abstract smoothness condition $g_\bullet \in \mathbb{J}^{(\mathfrak{t}\mathfrak{a}),dr}$, since*

$$d^{-2}\|g_\bullet\|^2_{(\mathfrak{t}\mathfrak{a})^\dagger} = d^{-2}(\mathfrak{t}\mathfrak{a})^{2|\dagger}_\bullet\nu(g_\bullet^2) \leqslant \mathfrak{a}^{2|\dagger}_\bullet\mathfrak{s}^{2|\dagger}_\bullet\nu(g_\bullet^2) = \mathfrak{a}^{2|\dagger}_\bullet\nu(\mathfrak{s}^{2|\dagger}_\bullet g_\bullet^2) = \mathfrak{a}^{2|\dagger}_\bullet\nu(\theta_\bullet^2) = \|\theta_\bullet\|^2_{\mathfrak{a}^\dagger} \leqslant r^2.$$ □

§04|03|01 Global and maximal global $\mathfrak{v}$-error

We measure similar to Subsubsection §04|02|01 first globally the accuracy of the orthogonal projection $\theta^m_\bullet = \mathfrak{s}^\dagger_\bullet g^m_\bullet \in \mathbb{J}\mathbb{1}^m_\bullet$ of $\theta_\bullet = \mathfrak{s}^\dagger_\bullet g_\bullet \in \mathbb{J}$.

§04|03.09 **Property** (Global $\mathfrak{v}$-error). *If $\mathfrak{v}_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ and $\theta_\bullet = \mathfrak{s}_\bullet^\dagger g_\bullet \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$, then for each $m \in \mathbb{N}$ we have $\theta_\bullet^m \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$ too, since $\|\theta_\bullet^m\|_{\mathfrak{v}}^2 = \mathfrak{v}_\bullet^2\nu(\theta_\bullet^2 \mathbb{1}_\bullet^m) \leqslant \mathfrak{v}_\bullet^2\nu(\theta_\bullet^2)$. Moreover, it holds*

$$\|\theta_\bullet^m - \theta_\bullet\|_{\mathfrak{v}}^2 = \|\theta_\bullet \mathbb{1}_\bullet^{m|\perp}\|_{\mathfrak{v}}^2 = \mathfrak{v}_\bullet^2\nu(\theta_\bullet^2 \mathbb{1}_\bullet^{m|\perp}) \leqslant \mathfrak{v}_\bullet^2\nu(\theta_\bullet^2) \in \mathbb{R}_{\geqslant 0}$$

and $\|\theta_\bullet^m - \theta_\bullet\|_{\mathfrak{v}}^2 = o(1)$ as $m \to \infty$ by dominated convergence. □

§04|03.10 **Assumption**. Consider weights $\mathfrak{a}_\bullet, \mathfrak{v}_\bullet \in \mathcal{M}_{>0,\nu}(\mathscr{J})$ such that $\mathfrak{a}_\bullet \in \mathbb{L}_\infty(\nu)$ and $(\mathfrak{a}\mathfrak{v})_\bullet := \mathfrak{a}_\bullet\mathfrak{v}_\bullet \in \mathbb{L}_\infty(\nu)$. We write $(\mathfrak{a}\mathfrak{v})_{(\bullet)} = ((\mathfrak{a}\mathfrak{v})_{(m)} := \|(\mathfrak{a}\mathfrak{v})_\bullet \mathbb{1}_\bullet^{m|\perp}\|_{\mathbb{L}_\infty(\nu)})_{m\in\mathbb{N}}$, where by construction $(\mathfrak{a}\mathfrak{v})_{(\bullet)} \in (\mathbb{R}_{>0})_{\searrow}^{\mathbb{N}}$ (compare Notation §04|02.08). □

§04|03.11 **Reminder**. Under Assumption §04|03.10 we have $\mathbb{J}^{\mathfrak{a}} = \mathbb{L}_2^{\mathfrak{a}}(\nu) = \mathrm{dom}(\mathrm{M}_{\mathfrak{a}_\bullet^{!}}) = \mathbb{J}\mathfrak{a}_\bullet = \mathbb{L}_2(\mathfrak{a}_\bullet^{2|\dagger}\nu)$ and the three measures ν, $\mathfrak{a}_\bullet^{2|\dagger}\nu$ and $\mathfrak{v}_\bullet^2\nu$ dominate mutually each other, i.e. they share the same null sets (see Property §04|01.02). Consequently, $\mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{J} = \mathbb{L}_2(\nu)$ and if $h_\bullet \in \mathbb{L}_2(\mathfrak{a}_\bullet^{2|\dagger}\nu)$ satisfies $\mathfrak{v}_\bullet^2\nu(h_\bullet^2) \in \mathbb{R}_{\geqslant 0}$, for example, then $h_\bullet \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$ too. Moreover under Assumption §04|03.10 we have $\mathbb{J}^{\mathfrak{a},\mathrm{r}} \subseteq \mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$ (see Definition §04|02.12 and Property §04|02.11). □

§04|03.12 **Property** (Maximal global $\mathfrak{v}$-error). *Under Assumption §04|03.10 for each $m \in \mathbb{N}$ and for each solution $\theta_\bullet = \mathfrak{s}_\bullet^\dagger g_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}} \subseteq \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$ its orthogonal projection $\theta_\bullet^m := \theta_\bullet\mathbb{1}_\bullet^m = \mathfrak{s}_\bullet^\dagger g_\bullet^m \in \mathbb{J}\mathbb{1}_\bullet^m$ satisfies*

$$\|\theta_\bullet^m - \theta_\bullet\|_{\mathfrak{v}} = \|\theta_\bullet\mathbb{1}_\bullet^{m|\perp}\|_{\mathfrak{v}} \leqslant \mathrm{r}\,(\mathfrak{a}\mathfrak{v})_{(m)}$$

as shown in Proof §04|02.14. □

§04|03|02 Local and maximal local ϕ-error

Secondly, we measure locally the accuracy of the orthogonal projection $\theta_\bullet^m = \mathfrak{s}_\bullet^\dagger g_\bullet^m \in \mathbb{J}\mathbb{1}_\bullet^m$ of $\theta_\bullet = \mathfrak{s}_\bullet^\dagger g_\bullet \in \mathbb{J}$.

§04|03.13 **Property** (Local ϕ-error). *If $\phi_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ and $\theta_\bullet = \mathfrak{s}_\bullet^\dagger g_\bullet \in \mathrm{dom}(\phi\nu)$, then for each $m \in \mathbb{N}$ we have $\theta_\bullet^m \in \mathrm{dom}(\phi\nu)$ too, since $\|\phi_\bullet\theta_\bullet^m\|_{\mathbb{L}_1(\nu)} = \nu(|\phi_\bullet\theta_\bullet|\mathbb{1}_\bullet^m) \leqslant \nu(|\phi_\bullet\theta_\bullet|)$. Moreover, it holds*

$$|\phi\nu(\theta_\bullet) - \phi\nu(\theta_\bullet^m)| \leqslant |\phi_\bullet|\nu(|\theta_\bullet^m - \theta_\bullet|) = |\phi_\bullet|\nu(|\theta_\bullet|\mathbb{1}_\bullet^{m|\perp}) \leqslant \nu(|\phi_\bullet\theta_\bullet|) \in \mathbb{R}_{\geqslant 0}$$

and $|\phi\nu(\theta_\bullet) - \phi\nu(\theta_\bullet^m)| = o(1)$ as $m \to \infty$ by dominated convergence. □

§04|03.14 **Assumption**. Consider $\mathfrak{a}_\bullet, \phi_\bullet \in \mathcal{M}_{>0,\nu}(\mathscr{J})$ such that $\mathfrak{a}_\bullet \in \mathbb{L}_\infty(\nu)$ and $(\mathfrak{a}\phi)_\bullet := (\mathfrak{a}_j\phi_j)_{j\in\mathcal{J}} = \mathfrak{a}_\bullet\phi_\bullet \in \mathbb{L}_2(\nu)$ and hence $\|\mathfrak{a}_\bullet\mathbb{1}_\bullet^{m|\perp}\|_{\phi} = \|(\mathfrak{a}\phi)_\bullet\mathbb{1}_\bullet^{m|\perp}\|_{\mathbb{L}_2(\nu)} = o(1)$ as $m \to \infty$. □

§04|03.15 **Reminder**. Under Assumption §04|03.14 we have $\mathbb{J}^{\mathfrak{a}} = \mathbb{L}_2^{\mathfrak{a}}(\nu) = \mathrm{dom}(\mathrm{M}_{\mathfrak{a}_\bullet^{!}}) = \mathbb{J}\mathfrak{a}_\bullet = \mathbb{L}_2(\mathfrak{a}_\bullet^{2|\dagger}\nu)$ and the three measures ν, $|\phi_\bullet|\nu$ and $\mathfrak{a}_\bullet^{2|\dagger}\nu$ dominate mutually each other (see Property §04|01.02). Consequently, $\mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{J} = \mathbb{L}_2(\nu)$ and if $h_\bullet \in \mathbb{L}_2(\mathfrak{a}_\bullet^{2|\dagger}\nu)$ satisfies $\nu(|\phi_\bullet h_\bullet|) \in \mathbb{R}_{\geqslant 0}$, for example, then $h_\bullet \in \mathbb{L}_1(|\phi_\bullet|\nu)$ too. Moreover, under Assumption §04|03.14 we have $\mathbb{J}^{\mathfrak{a},\mathrm{r}} \subseteq \mathbb{J}^{\mathfrak{a}} \subseteq \mathrm{dom}(\phi\nu)$ (see Definition §04|02.12 and Property §04|02.11). □

§04|03.16 **Property** (Maximal local ϕ-error). *Under Assumption §04|03.14 for each $m \in \mathbb{N}$ and for each solution $\theta_\bullet = \mathfrak{s}_\bullet^\dagger g_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}} \subseteq \mathrm{dom}(\phi\nu)$ its orthogonal projection $\theta_\bullet^m := \theta_\bullet\mathbb{1}_\bullet^m = \mathfrak{s}_\bullet^\dagger g_\bullet^m \in \mathbb{J}\mathbb{1}_\bullet^m$ satisfies*

$$|\phi\nu(\theta_\bullet - \theta_\bullet^m)| = |\phi\nu(\theta_\bullet\mathbb{1}_\bullet^{m|\perp})| \leqslant \nu(|\phi_\bullet\theta_\bullet|\mathbb{1}_\bullet^{m|\perp}) \leqslant \mathrm{r}\,\|\mathfrak{a}_\bullet\mathbb{1}_\bullet^{m|\perp}\|_{\phi}$$

as shown in Proof §04|02.26. □

§05 (Generalised) linear Galerkin approach

§05|00.01 **Notation**. Consider $\mathbb{J} = \ell_2 := \mathbb{L}_2(\nu_{\mathbb{N}}) = \mathbb{L}_2(\mathbb{N}, 2^{\mathbb{N}}, \nu_{\mathbb{N}})$ with counting measure $\nu_{\mathbb{N}} := \sum_{j\in\mathbb{N}} \delta_{\{j\}}$, surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \ell_2)$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \ell_2)$. For each $\mathrm{T} \in \mathbb{L}(\mathbb{H}, \mathbb{G})$ and $\mathrm{T}_{\bullet|\bullet} := \mathrm{V}\mathrm{T}\mathrm{U}^\star \in \mathbb{L}(\ell_2) = \mathbb{L}_{\bullet}(\ell_2)$ (compare Notation §01|04.03) we identify the kernel (infinite dimensional matrix) $\mathrm{T}_{\bullet|\bullet} = (\mathrm{T}_{j,j_o})_{j,j_o\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}^2}$ and the map from ℓ_2 into itself given by

$$a_\bullet \mapsto \mathrm{T}_{\bullet|\bullet} a_\bullet := \Big(\sum_{j_o\in\mathbb{N}} \mathrm{T}_{j|j_o} a_{j_o} = \langle \mathrm{T}_{j|\bullet}, a_\bullet\rangle_{\ell_2} = \nu_{\mathbb{N}}(\mathrm{T}_{j,\bullet} a_\bullet)\Big)_{j\in\mathbb{N}}$$

(compare Notation §01|05.01). Moreover, we denote by $\mathbb{L}_{\bullet}^{\gtrless}(\ell_2)$ the subset of all strictly positive definite operator in $\mathbb{L}_{\bullet}(\ell_2)$. For each $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}_{\bullet}^{\gtrless}(\ell_2)$ we denote its Moore-Penrose inverse by $\mathrm{T}^\dagger_{\bullet|\bullet} : \ell_2 \supseteq \mathrm{dom}(\mathrm{T}^\dagger_{\bullet|\bullet}) \to \ell_2$ (see Definition §03|00.08). □

§05|00.02 **Notation** (Property). For $m \in \mathbb{N}$ set $\mathbb{1}^{m|\perp}_\bullet := \mathbb{1}_\bullet - \mathbb{1}^m_\bullet \in \mathbb{R}^{\mathbb{N}}$ where $0_\bullet = \mathbb{1}^{m|\perp}_\bullet \mathbb{1}^m_\bullet = \mathbb{1}^m_\bullet \mathbb{1}^{m|\perp}_\bullet \in \mathbb{R}^{\mathbb{N}}$.

(a) For $a_\bullet \in \mathbb{R}^{\mathbb{N}}$ introduce its sub-vector $[a_\bullet]_{\underline{m}} := (a_i)_{i\in[\![m]\!]} \in \mathbb{R}^m$ where $[a_\bullet]_{\underline{m}} = [a_\bullet \mathbb{1}^m_\bullet]_{\underline{m}}$.

(b) For $\mathrm{A}_{\bullet|\bullet} = (\mathrm{A}_{j|j_o})_{j,j_o\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}^2}$ introduce its sub-matrix $[\mathrm{A}_{\bullet|\bullet}]_{\underline{m}} := (\mathrm{A}_{j|j_o})_{j,j_o\in[\![m]\!]} \in \mathbb{R}^{(m,m)}$. Clearly, if we restrict $\mathrm{A}^m_{\bullet|\bullet} := \mathrm{M}_{\mathbb{1}^m}\mathrm{A}_{\bullet|\bullet}\mathrm{M}_{\mathbb{1}^m} \in \mathbb{L}(\ell_2)$ with

$$a_\bullet \mapsto \mathrm{A}^m_{\bullet|\bullet} a_\bullet = \Big(\mathbb{1}^m_j \sum_{j_o\in[\![m]\!]} \mathrm{A}_{j,j_o} a_{j_o} = \mathbb{1}^m_j \langle \mathrm{A}_{j|\bullet}\mathbb{1}^m_\bullet, a_\bullet\rangle_{\ell_2} = \mathbb{1}^m_j \nu_{\mathbb{N}}(\mathrm{A}_{j|\bullet} a_\bullet \mathbb{1}^m_\bullet)\Big)_{j\in\mathbb{N}}$$

to an operator from $\mathbb{R}^m$ ($\mathrm{ran}(\mathrm{M}_{\mathbb{1}^m}) = {}_{\ell_2}\mathbb{1}^m_\bullet$) to itself, then it is represented by the matrix $[\mathrm{A}_{\bullet|\bullet}]_{\underline{m}}$. Note that the adjoint $\mathrm{A}^\star_{\bullet|\bullet} \in \mathbb{L}(\ell_2)$ of $\mathrm{A}_{\bullet|\bullet} \in \mathbb{L}(\ell_2)$ and the transposed matrix $[\mathrm{A}_{\bullet|\bullet}]^\star_{\underline{m}} \in \mathbb{R}^{(m,m)}$ of $[\mathrm{A}_{\bullet|\bullet}]_{\underline{m}}$ satisfy $[\mathrm{A}^\star_{\bullet|\bullet}]_{\underline{m}} = [\mathrm{A}_{\bullet|\bullet}]^\star_{\underline{m}}$. If $[\mathrm{A}_{\bullet|\bullet}]^\dagger_{\underline{m}} \in \mathbb{R}^{(m,m)}$ denotes the Moore-Penrose inverse of $[\mathrm{A}_{\bullet|\bullet}]_{\underline{m}}$ (as linear map from $\mathbb{R}^m$ into itself), then the Moore-Penrose inverse $\mathrm{A}^{m|\dagger}_{\bullet|\bullet} \in \mathbb{L}(\ell_2)$ of $\mathrm{A}^m_{\bullet|\bullet}$ (see Definition §03|00.08), restricted to an operator from $\mathbb{R}^m$ to itself can be represented by the matrix $[\mathrm{A}_{\bullet|\bullet}]^\dagger_{\underline{m}}$. In particular, if $[\mathrm{A}_{\bullet|\bullet}]_{\underline{m}}$ is *regular* (invertible), and hence $[\mathrm{A}_{\bullet|\bullet}]^\dagger_{\underline{m}} = [\mathrm{A}_{\bullet|\bullet}]^{-1}_{\underline{m}}$, then we have $\mathrm{A}^m_{\bullet|\bullet}\mathrm{A}^{m|\dagger}_{\bullet|\bullet} = \mathrm{M}_{\mathbb{1}^m} = \mathrm{A}^{m|\dagger}_{\bullet|\bullet}\mathrm{A}^m_{\bullet|\bullet}$.

(c) Given $\mathrm{M}_{\mathrm{w}} \in \mathbb{L}^{\mathrm{M}}(\ell_2)$, the diagonal matrix $[\mathrm{M}_{\mathrm{w}}]_{\underline{m}} \in \mathbb{R}^{(m,m)}$ has $[\mathrm{w}_\bullet]_{\underline{m}}$ as its diagonal entries. Note that $[\mathrm{M}_{\mathrm{w}}]^s_{\underline{m}} = [\mathrm{M}_{\mathrm{w}^s}]_{\underline{m}} = [\mathrm{M}^s_{\mathrm{w}}]_{\underline{m}}$ for all $s \in \mathbb{R}_{>0}$ and $[\mathrm{M}_{\mathrm{w}}]^\dagger_{\underline{m}} = [\mathrm{M}_{\mathrm{w}^\dagger}]_{\underline{m}} = [\mathrm{M}^\dagger_{\mathrm{w}}]_{\underline{m}}$.

(d) Keep in mind the Euclidean norm $\|\cdot\|$ of a vector and the weighted norm $\|\cdot\|_{\mathfrak{t}} := \|\mathrm{M}_{\mathfrak{t}}\cdot\|_{\ell_2}$ with $\mathfrak{t}_\bullet \in \mathbb{R}^{\mathbb{N}}_{>0}$. For all $a_\bullet \in {}_{\ell_2}\mathbb{1}^m_\bullet$ (and $(\mathfrak{t}a)_\bullet := \mathfrak{t}_\bullet a_\bullet \in {}_{\ell_2}\mathbb{1}^m_\bullet$) we have

$$\|a_\bullet\|^2_{\mathfrak{t}} = \|\mathrm{M}_{\mathfrak{t}} a_\bullet\|^2_{\ell_2} = \|(\mathfrak{t}a)_\bullet\|^2_{\ell_2} = \|(\mathfrak{t}a)_\bullet \mathbb{1}^m_\bullet\|^2_{\ell_2} = [a_\bullet]^\star_{\underline{m}}[\mathrm{M}_{\mathfrak{t}^2}]_{\underline{m}}[a_\bullet]_{\underline{m}}$$
$$= [(\mathfrak{t}a)_\bullet]^\star_{\underline{m}}[(\mathfrak{t}a)_\bullet]_{\underline{m}} = \|[\mathrm{M}_{\mathfrak{t}}]_{\underline{m}}[a_\bullet]_{\underline{m}}\|^2 = \|[(\mathfrak{t}a)_\bullet]_{\underline{m}}\|^2.$$

(e) Let $\|A\|_{\mathrm{spec}} := \sup\{\|Ax\| : \|x\| \in [0,1]\}$ denote the spectral norm of a matrix A. Then we have $\|\mathrm{A}^m_{\bullet|\bullet}\|_{\mathbb{L}(\ell_2)} = \|\mathrm{M}_{\mathbb{1}^m}\mathrm{A}_{\bullet|\bullet}\mathrm{M}_{\mathbb{1}^m}\|_{\mathbb{L}(\ell_2)} = \|[\mathrm{A}_{\bullet|\bullet}]_{\underline{m}}\|_{\mathrm{spec}}$ and for $s \in \mathbb{R}_{>0}$ hence

$$\|(\mathrm{M}^m_{\mathfrak{t}})^s\|_{\mathbb{L}(\ell_2)} = \|\mathrm{M}_{\mathbb{1}^m}\mathrm{M}^s_{\mathfrak{t}}\mathrm{M}_{\mathbb{1}^m}\|_{\mathbb{L}(\ell_2)} = \|[\mathrm{M}_{\mathfrak{t}^s}]_{\underline{m}}\|_{\mathrm{spec}} = \|\mathfrak{t}^s_\bullet\mathbb{1}^m_\bullet\|_{\ell_\infty} = \max\{|\mathfrak{t}^s_j| : j \in [\![m]\!]\}. \quad □$$

§05|01 Linear Galerkin approach

§05|01.01 **Assumption**. For $\mathbb{J} = \ell_2$, surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \ell_2)$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \ell_2)$ fixed and presumed to be *known* in advance, the operator $\mathrm{T} \in \mathbb{L}(\mathbb{H}, \mathbb{G})$ satisfies $\mathrm{T}_{\bullet|\bullet} = \mathrm{V}\mathrm{T}\mathrm{U}^\star \in \mathbb{L}_{\bullet}^{\gtrless}(\ell_2) \subseteq \mathbb{L}_{\bullet}(\ell_2) = \mathbb{L}(\ell_2)$. Let $g_\bullet \in \mathrm{dom}(\mathrm{T}^\dagger_{\bullet|\bullet}) = \mathrm{ran}(\mathrm{T}_{\bullet|\bullet})$, and hence $\theta_\bullet = \mathrm{T}^\dagger_{\bullet|\bullet} g_\bullet = \mathrm{T}^{-1}_{\bullet|\bullet} g_\bullet \in \ell_2$. □

§05|01.02 **Linear Galerkin approach**. Let $T_{\bullet|\bullet} \in \mathbb{L}^{\gtrless}_{\bullet}(\ell_2)$ and $g_\bullet \in \ell_2$. For $m \in \mathbb{N}$ any element $\theta^m_\bullet \in {}_{\ell_2}\mathbb{1}^m_\bullet$, i.e. $0_\bullet = \theta^m_\bullet(\mathbb{1}_\bullet - \mathbb{1}^m_\bullet) = \theta^m_\bullet \mathbb{1}^{m|\perp}_\bullet$, satisfying

$$\langle \theta^m_\bullet, T_{\bullet|\bullet}\theta^m_\bullet\rangle_{\ell_2} - 2\langle \theta^m_\bullet, g_\bullet\rangle_{\ell_2} \leqslant \langle a_\bullet, T_{\bullet|\bullet} a_\bullet\rangle_{\ell_2} - 2\langle a_\bullet, g_\bullet\rangle_{\ell_2} \quad \text{for all } a_\bullet \in {}_{\ell_2}\mathbb{1}^m_\bullet$$

is called a *Galerkin solution* in ${}_{\ell_2}\mathbb{1}^m_\bullet$. □

§05|01.03 **Lemma**. *Let $T_{\bullet|\bullet} \in \mathbb{L}^{\gtrless}_{\bullet}(\ell_2)$ and $g_\bullet \in \ell_2$. (i) For all $m \in \mathbb{N}$ the matrix $[T_{\bullet|\bullet}]_{\underline{m}} \in \mathbb{R}^{(m,m)}$ is strictly positive definite. (ii) The Galerkin solution $\theta^m_\bullet \in {}_{\ell_2}\mathbb{1}^m_\bullet$, i.e. $0_\bullet = \theta^m_\bullet(\mathbb{1}_\bullet - \mathbb{1}^m_\bullet) = \theta^m_\bullet \mathbb{1}^{m|\perp}_\bullet$, is uniquely determined by $[\theta^m_\bullet]_{\underline{m}} = [T_{\bullet|\bullet}]^{-1}_{\underline{m}}[g_\bullet]_{\underline{m}}$, and hence $\theta^m_\bullet = T^{m|\dagger}_{\bullet|\bullet} g_\bullet$. (iii) If in addition $g_\bullet \in \mathrm{dom}(T^\dagger_{\bullet|\bullet})$ and $\theta_\bullet := T^\dagger_{\bullet|\bullet} g_\bullet \in \ell_2$, then the Galerkin solution $\theta^m_\bullet$ minimises in ${}_{\ell_2}\mathbb{1}^m_\bullet$ the functional $a_\bullet \to F(a_\bullet) = \|T^{1/2}_{\bullet|\bullet}(a_\bullet - \theta_\bullet)\|^2_{\ell_2}$.*

§05|01.04 **Proof** of Lemma §05|01.03. Given in the lecture. □

§05|01.05 **Remark**. Consider for $\theta_\bullet \in \ell_2$ its orthogonal projection $\mathbb{1}^m_\bullet\theta_\bullet$ and $\mathbb{1}^{m|\perp}_\bullet\theta_\bullet$ onto the subspace ${}_{\ell_2}\mathbb{1}^m_\bullet$ and its orthogonal complement ${}_{\ell_2}\mathbb{1}^{m|\perp}_\bullet := ({}_{\ell_2}\mathbb{1}^m_\bullet)^\perp$, respectively, then the approximation error $\|\mathbb{1}^m_\bullet\theta_\bullet - \theta_\bullet\|_{\ell_2} = \|(\mathbb{1}_\bullet - \mathbb{1}^m_\bullet)\theta_\bullet\|_{\ell_2} = \|\mathbb{1}^{m|\perp}_\bullet\theta_\bullet\|_{\ell_2}$ converges to zero as $m \to \infty$ by Lebesgue's dominated convergence theorem. On the other hand, if $g_\bullet \in \mathrm{dom}(T^\dagger_{\bullet|\bullet})$ and $\theta_\bullet := T^\dagger_{\bullet|\bullet} g_\bullet \in \ell_2$ then the Galerkin solution $\theta^m_\bullet \in {}_{\ell_2}\mathbb{1}^m_\bullet$ satisfies $[\mathbb{1}^m_\bullet\theta_\bullet - \theta^m_\bullet]_{\underline{m}} = -[T_{\bullet|\bullet}]^{-1}_{\underline{m}}[T_{\bullet|\bullet}(\mathbb{1}_\bullet - \mathbb{1}^m_\bullet)\theta_\bullet]_{\underline{m}} = -[T_{\bullet|\bullet}]^{-1}_{\underline{m}}[T_{\bullet|\bullet}\mathbb{1}^{m|\perp}_\bullet\theta_\bullet]_{\underline{m}}$ and, hence it does generally not correspond to the orthogonal projection $\mathbb{1}^{m|\perp}_\bullet\theta_\bullet = (\mathbb{1}_\bullet - \mathbb{1}^m_\bullet)\theta_\bullet$. Moreover, the approximation error $\sup\{\|\theta^m_\bullet - \theta_\bullet\|_{\ell_2} : m \in \mathbb{N}_{\geqslant n}\}$ does generally not converge to zero as $n \to \infty$. However, if

$$\begin{aligned} C_T &:= \sup\{\|T^{m|\dagger}_{\bullet|\bullet} T_{\bullet|\bullet} M_{\mathbb{1}^{m|\perp}_\bullet}\|_{\mathbb{L}(\ell_2)} : m \in \mathbb{N}\} \\ &= \sup\{\|[T_{\bullet|\bullet}]^{-1}_{\underline{m}}[T_{\bullet|\bullet}\mathbb{1}^{m|\perp}_\bullet a_\bullet]_{\underline{m}}\| : \|a_\bullet\|_{\ell_2} = 1, a_\bullet \in \ell_2, m \in \mathbb{N}\} \in \mathbb{R}_{\geqslant 0} \end{aligned}$$

then $\|\theta^m_\bullet - \theta_\bullet\|_{\ell_2} \leqslant (1 + C_T)\|\mathbb{1}^{m|\perp}_\bullet\theta_\bullet\|_{\ell_2}$ which implies $\sup\{\|\theta^m_\bullet - \theta_\bullet\|_{\ell_2} : m \in \mathbb{N}_{\geqslant n}\} = o(1)$ as $m \to \infty$. Here and subsequently, we will restrict ourselves to classes of solutions and operators which ensure the convergence. Obviously, this is a minimal regularity condition for us if we aim to estimate the Galerkin solution. □

§05|01.06 **Notation** (Reminder §04|02.08). For sequences $a_\bullet, b_\bullet \in (\mathbb{K})^{\mathbb{N}}$ taking its values in $\mathbb{K} \in \{\mathbb{R}_{>0}, \mathbb{Q}, \mathbb{Z}, \dots\}$ we write $a_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\nearrow}$ and $b_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\searrow}$ if $a_\bullet$ and $b_\bullet$ is monotonically, respectively, *non-decreasing* and *non-increasing*. If in addition $a_n \to \infty$ and $b_n \to 0$ as $n \to \infty$, then we write $a_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\uparrow\infty}$ and $b_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\downarrow 0}$ for short. □

§05|01.07 **Property**. *If $\mathfrak{t}_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ is monotonically non-increasing, then for all $m \in \mathbb{N}$ we have*

$$\|\mathfrak{t}^{-1}_\bullet \mathbb{1}^m_\bullet\|^{-1}_{\ell_\infty} = \min\{\mathfrak{t}_j : j \in [\![m]\!]\} \geqslant \sup\{\mathfrak{t}_j : j \in \mathbb{N}_{>m}\} = \|\mathfrak{t}_\bullet \mathbb{1}^{m|\perp}_\bullet\|_{\ell_\infty} = \mathfrak{t}_{(m)},$$

and hence $1 \geqslant \mathfrak{t}_{(m)}\|\mathfrak{t}^{-1}_\bullet \mathbb{1}^m_\bullet\|_{\ell_\infty} = \|\mathfrak{t}_\bullet \mathbb{1}^{m|\perp}_\bullet\|_{\ell_\infty}\|\mathfrak{t}^{-1}_\bullet \mathbb{1}^m_\bullet\|_{\ell_\infty}$. □

§05|01.08 **Link condition**. Given weights $\mathfrak{t}_\bullet \in \mathbb{R}^{\mathbb{N}}_{>0} \cap \ell_\infty$ an operator $T_{\bullet|\bullet} \in \mathbb{L}^{\gtrless}_{\bullet}(\ell_2)$ satisfies a *link condition* if there is $d \in \mathbb{R}_{\geqslant 1}$ such that

$$T_{\bullet|\bullet} \in \mathbb{T}^{\gtrless}_{\mathfrak{t},d} := \{A_{\bullet|\bullet} \in \mathbb{L}^{\gtrless}_{\bullet}(\ell_2) : d^{-1}\|a_\bullet\|_{\mathfrak{t}} \leqslant \|T_{\bullet|\bullet} a_\bullet\|_{\ell_2} \leqslant d\|a_\bullet\|_{\mathfrak{t}} \text{ for all } a_\bullet \in \ell_2\}$$

and we set $\mathbb{T}_{\mathfrak{t},d} := \{A_{\bullet|\bullet} \in \mathbb{L}_{\bullet}(\ell_2) : (A^\star_{\bullet|\bullet}A_{\bullet|\bullet})^{1/2} \in \mathbb{T}^{\gtrless}_{\mathfrak{t},d}\}$.

§05|01.09 **Remark**. Note that $\mathfrak{t}_{\bullet} \in \mathbb{R}_{>0}^{\mathbb{N}} \cap \ell_\infty$ for each $A_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},d}^{\gtrless}$ implies $\ker(A_{\bullet|\bullet}) = \{0_{\bullet}\}$, i.e. $A_{\bullet|\bullet}$ is injective and hence strictly positive definite. We shall emphasise that for $T_{\bullet|\bullet} \in \mathbb{L}^{\geqslant}(\ell_2)$ the condition $T_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},d}$ is equivalent to

$$d^{-1}\|a_{\bullet}\|_{\mathfrak{t}} \leqslant \|T_{\bullet|\bullet}a_{\bullet}\|_{\ell_2} \leqslant d\|a_{\bullet}\|_{\mathfrak{t}} \quad \text{for all } a_{\bullet} \in \ell_2. \tag{05.01}$$

Observe further that $M_{\mathfrak{s}} \in \mathbb{L}^{M}(\ell_2)$ satisfies the link condition $M_{\mathfrak{s}} \in \mathbb{T}_{\mathfrak{t},d}$ as in Definition §05|01.08 if and only if $|\mathfrak{s}_{\bullet}| \leqslant d\mathfrak{t}_{\bullet} \wedge \mathfrak{t}_{\bullet} \leqslant d|\mathfrak{s}_{\bullet}|$ ($\nu_{\mathbb{N}}$-a.e.), i.e. $M_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},d}$ as in Definition §04|03.05. Thereby, we have $\mathbb{M}_{\mathfrak{t},d} \subseteq \mathbb{T}_{\mathfrak{t},d}$. We shall emphasise, that there are operators satisfying the link condition $\mathbb{T}_{\mathfrak{t},d}^{\gtrless}$ which do not belong to $\mathbb{M}_{\mathfrak{t},d}$, i.e., they are non-diagonal. Let us briefly give a construction of one of those. We consider a small perturbation of $M_{\mathfrak{t}}$, that is, $T_{\bullet|\bullet} = M_{\mathfrak{t}} + M_{\mathfrak{t}}A_{\bullet|\bullet}M_{\mathfrak{t}}$ where $A_{\bullet|\bullet} \in \mathbb{L}^{\gtrless}(\ell_2)$ is a positive definite operator with spectral norm $c := \|M_{\mathfrak{t}}A_{\bullet|\bullet}\|_{\mathbb{L}(\ell_2)} < 1$. Obviously, $T_{\bullet|\bullet}$ is strictly positive definite, and $\|T_{\bullet|\bullet}a_{\bullet}\|_{\ell_2} \leqslant \|\mathrm{id}_{\ell_2} + M_{\mathfrak{t}}A_{\bullet|\bullet}\|_{\mathbb{L}(\ell_2)}\|M_{\mathfrak{t}}a_{\bullet}\|_{\ell_2} \leqslant (1+c)\|a_{\bullet}\|_{\mathfrak{t}}$. On the other hand, we have $\|(\mathrm{id}_{\ell_2} + M_{\mathfrak{t}}A_{\bullet|\bullet})^{-1}\|_{\mathbb{L}(\ell_2)} = \frac{1}{1-\|M_{\mathfrak{t}}A_{\bullet|\bullet}\|_{\mathbb{L}(\ell_2)}} = \frac{1}{1-c}$ by the Neumann series argument **??**, which in turn implies $\|a_{\bullet}\|_{\mathfrak{t}} = \|M_{\mathfrak{t}}a_{\bullet}\|_{\ell_2} = \|(\mathrm{id}_{\ell_2} + M_{\mathfrak{t}}A_{\bullet|\bullet})^{-1}\|_{\mathbb{L}(\ell_2)}\|T_{\bullet|\bullet}a_{\bullet}\|_{\ell_2} \leqslant \frac{1}{1-c}\|T_{\bullet|\bullet}a_{\bullet}\|_{\ell_2}$. Combining both bounds the operator $T_{\bullet|\bullet}$ satisfies the link condition $T_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},d}^{\gtrless}$ for all $d \geqslant \max(1+c, \frac{1}{1-c})$ and is obviously not a multiplication operator, i.e. diagonal. □

§05|01.10 **Property**. *If* $T_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},d}^{\gtrless}$ *with* $\mathfrak{t}_{\bullet} \in \mathbb{R}_{>0}^{\mathbb{N}} \cap \ell_\infty$ *and* $d \in \mathbb{R}_{\geqslant 1}$ *then for all* $s \in [-1,1]$ *we have*

(inequality of Heinz [1951]) $d^{-|s|}\|a_{\bullet}\|_{\mathfrak{t}^s} \leqslant \|T_{\bullet|\bullet}^{s}a_{\bullet}\|_{\ell_2} \leqslant d^{|s|}\|a_{\bullet}\|_{\mathfrak{t}^s}$ *for all* $a_{\bullet} \in \mathrm{dom}(M_{\mathfrak{t}^s})$. □

§05|01.11 **Comment**. Given $T_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},d}^{\gtrless}$ we have $\ker(T_{\bullet|\bullet}) = \{0_{\bullet}\} = \ker(T_{\bullet|\bullet}^{\star})$ and on $\mathrm{ran}(T_{\bullet|\bullet}) = \mathrm{dom}(T_{\bullet|\bullet}^{\dagger})$ (which is dense in ℓ_2) we have $T_{\bullet|\bullet}^{-1} = T_{\bullet|\bullet}^{\dagger}$. Similarly, for each $s \in \mathbb{R}_{>0}$ on $\mathrm{ran}(T_{\bullet|\bullet}^{s}) = \mathrm{dom}(T_{\bullet|\bullet}^{s|\dagger})$ we have $T_{\bullet|\bullet}^{-s} = T_{\bullet|\bullet}^{s|\dagger} = (T_{\bullet|\bullet}^{s})^{\dagger}$. □

§05|01.12 **Notation**. Given weights $\mathfrak{a}_{\bullet} \in \mathbb{R}_{>0}^{\mathbb{N}} \cap \ell_\infty$ introduce $\ell_2^{\mathfrak{a}} := \ell_2(\mathfrak{a}_{\bullet}^{-2}) := \mathbb{L}_2(\mathfrak{a}_{\bullet}^{-2}\nu_{\mathbb{N}}) = \ell_2\mathfrak{a}_{\bullet} = \mathrm{ran}(M_{\mathfrak{a}}) = \mathbb{J}^{\mathfrak{a}} \subseteq \ell_2 = \mathbb{J}$ endowed with $\|\cdot\|_{\mathfrak{a}^{-1}} := \|M_{\mathfrak{a}^{-1}}\cdot\|_{\ell_2}$ (as in Property §04|01.02). We assume in the following that $\theta_{\bullet} \in \ell_2$ satisfies an abstract smoothness condition (Definition §04|02.12), i.e., there is $r \in \mathbb{R}_{>0}$ such that $\theta_{\bullet} \in \ell_2^{\mathfrak{a},r} := \mathbb{J}^{\mathfrak{a},r} = \{h_{\bullet} \in \ell_2^{\mathfrak{a}} : \|h_{\bullet}\|_{\mathfrak{a}^{-1}} \leqslant r\} \subseteq \ell_2^{\mathfrak{a}} \subseteq \ell_2$. □

§05|01.13 **Source condition**. Given $T_{\bullet|\bullet} \in \mathbb{L}^{\geqslant}(\ell_2)$, the solution $\theta_{\bullet} \in \ell_2$ satisfies a *source condition*, if there is $s \in \mathbb{R}_{>0}$ such that $\theta_{\bullet} \in \mathrm{ran}((T_{\bullet|\bullet}^{\star}T_{\bullet|\bullet})^{s/2})$, that is, $\theta_{\bullet} = (T_{\bullet|\bullet}^{\star}T_{\bullet|\bullet})^{s/2}a_{\bullet}$ for some $a_{\bullet} \in \ell_2$. □

§05|01.14 **Corollary**. *For* $a, t \in \mathbb{R}_{>0}$ *and* $\mathfrak{v}_{\bullet} \in \mathbb{R}_{>0}^{\mathbb{N}} \cap \ell_\infty$ *set* $\mathfrak{t}_{\bullet} := \mathfrak{v}_{\bullet}^{t}$, $\mathfrak{a}_{\bullet} := \mathfrak{v}_{\bullet}^{a} \in \mathbb{R}_{>0}^{\mathbb{N}} \cap \ell_\infty$. *Consider* $\ell_2^{\mathfrak{a}_{\bullet}} = \ell_2\mathfrak{a}_{\bullet}$ *and assume that* $T_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},d}^{\gtrless}$. *If* $a \leqslant t$ *then* (i) *for any* $\theta_{\bullet} \in \ell_2^{\mathfrak{a}}$ *we have* $\theta_{\bullet} = T_{\bullet|\bullet}^{a/t}h_{\bullet}$ *with* $\|h_{\bullet}\|_{\ell_2} \leqslant d^{a/t}\|\theta_{\bullet}\|_{\mathfrak{a}^{-1}}$, *and conversely* (ii) *for any* $\theta_{\bullet} = T_{\bullet|\bullet}^{a/t}h_{\bullet}$ *with* $h_{\bullet} \in \ell_2$ *we obtain* $\theta_{\bullet} \in \ell_2^{\mathfrak{a}}$ *with* $\|\theta_{\bullet}\|_{\mathfrak{a}^{-1}} \leqslant d^{a/t}\|h_{\bullet}\|_{\ell_2}$.

§05|01.15 **Proof** of Corollary §05|01.14. Given in the lecture. □

§05|01.16 **Comment**. Under the assumptions of Corollary §05|01.14 if $T_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},d}$ and $a \leqslant t$ then (i) for any $\theta_{\bullet} \in \ell_2^{\mathfrak{a}}$ we have $\theta_{\bullet} = (T_{\bullet|\bullet}^{\star}T_{\bullet|\bullet})^{a/(2t)}h_{\bullet}$ with $\|h_{\bullet}\|_{\ell_2} \leqslant d^{a/t}\|\theta_{\bullet}\|_{\mathfrak{a}^{-1}}$, and conversely (ii) for any $\theta_{\bullet} = (T_{\bullet|\bullet}^{\star}T_{\bullet|\bullet})^{a/(2t)}h_{\bullet}$ with $h_{\bullet} \in \ell_2$ we obtain $\theta_{\bullet} \in \ell_2^{\mathfrak{a}}$ with $\|\theta_{\bullet}\|_{\mathfrak{a}^{-1}} \leqslant d^{a/t}\|h_{\bullet}\|_{\ell_2}$. □

§05|01.17 **Corollary**. *Given* $d, r \in \mathbb{R}_{>0}$ *and* $\mathfrak{t}_{\bullet}, \mathfrak{a}_{\bullet} \in \mathbb{R}_{>0}^{\mathbb{N}} \cap \ell_\infty$ *set* $(\mathfrak{t}\mathfrak{a})_{\bullet} := \mathfrak{t}_{\bullet}\mathfrak{a}_{\bullet} \in \mathbb{R}_{>0}^{\mathbb{N}} \cap \ell_\infty$. *If* $T_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},d}$ *and* $\theta_{\bullet} \in \ell_2^{\mathfrak{a},r}$, *then we have* $g_{\bullet} = T_{\bullet|\bullet}\theta_{\bullet} \in \ell_2^{(\mathfrak{t}\mathfrak{a}),dr}$.

§05|01.18 **Proof** of Corollary §05|01.17. Given in the lecture. □

§05|01.19 **Remark**. Keeping the orthonormal basis $(\mathbb{1}_{\bullet}^{\{j\}})_{j\in\mathbb{N}}$ in ℓ_2 in mind (Notation §01|04.02) each $M_{\mathfrak{c}} \in \mathbb{L}^{M}(\ell_2)$ with $\mathfrak{c}_{\bullet} \in \ell_\infty$ is *self-adjoint* with eigensystem $((\mathfrak{c}_j, \mathbb{1}_{\bullet}^{\{j\}}))_{j\in\mathbb{N}}$. Indeed, for all $j \in \mathbb{N}$ we have

$\mathrm{M}_{\mathfrak{e}}\mathbb{1}^{\{j\}}_{\bullet} = \mathfrak{e}_j\mathbb{1}^{\{j\}}_{\bullet}$. Recall that $\mathbb{K}(\ell_2)$ denotes the subset of $\mathbb{L}(\ell_2)$ containing all compact operators. If $\mathrm{A}_{\bullet|\bullet} \in \mathbb{K}(\ell_2)$ is *compact* and in addition *self-adjoint*, then $\mathrm{A}_{\bullet|\bullet}$ admits an eigensystem $((\mathfrak{e}_j, \mathrm{e}_j))_{j\in\mathbb{N}}$ where $\mathfrak{e}_{\bullet} \in \ell_\infty$ contains each eigenvalue of $\mathrm{A}_{\bullet|\bullet}$ repeated according to its multiplicity (with zero as only accumulation point) and $\mathrm{e}_{\bullet} = (\mathrm{e}_j)_{j\in\mathbb{N}}$ is the associated eigenbasis. We denote by $\mathbb{K}^{\gtrapprox}(\ell_2)$ the subset of $\mathbb{L}^{\gtrapprox}(\ell_2)$ containing all compact operators. If $\mathrm{A}_{\bullet|\bullet} \in \mathbb{K}^{\gtrapprox}(\ell_2)$ then we have $\mathfrak{e}_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$ (possibly after a reordering). □

§05|01.20 **Lemma**. *Consider as in Definition §05|01.08 a link condition $\mathbb{T}^{\gtrapprox}_{\mathfrak{t},\mathrm{d}}$ with $\mathfrak{t}_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$. Let the operator $\mathrm{T}_{\bullet|\bullet} \in \mathbb{K}^{\gtrapprox}(\ell_2)$ admit $((\mathfrak{e}_j, \mathrm{e}_j))_{j\in\mathbb{N}}$ as eigensystem where $\mathfrak{e}_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ contains each eigenvalue of $\mathrm{T}_{\bullet|\bullet}$ repeated according to its multiplicity and the associated eigenbasis $\mathrm{e}_{\bullet} = (\mathrm{e}_j)_{j\in\mathbb{N}}$ does eventually not correspond to the ONB $(\mathbb{1}^{\{j\}}_{\bullet})_{j\in\mathbb{N}}$. If $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}^{\gtrapprox}_{\mathfrak{t},\mathrm{d}}$, then we have $\mathrm{d}^{-1} \leqslant \mathfrak{e}_j/\mathfrak{t}_j \leqslant \mathrm{d}$ for all $j \in \mathbb{N}$.*

§05|01.21 **Proof** of Lemma §05|01.20. Given in the lecture. □

§05|01.22 **Lemma**. *Consider the link condition $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}^{\gtrapprox}_{\mathfrak{t},\mathrm{d}}$ as in Definition §05|01.08 with $\mathfrak{t}_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$. For all $m \in \mathbb{N}$ and $s \in [0, 1]$ we have (i) $\mathfrak{t}^s_{(m)}\|[\mathrm{T}_{\bullet|\bullet}]^{-s}_{\underline{m}}\|_{\text{spec}} \leqslant \big(\mathrm{d}(\mathrm{d}+2)\big)^s \leqslant \big(3\mathrm{d}^2\big)^s$, (ii) $\|[\mathrm{T}_{\bullet|\bullet}]^{-s}_{\underline{m}}[\mathrm{M}_{\mathfrak{t}}]^{s}_{\underline{m}}\|_{\text{spec}} \leqslant \big(\mathrm{d}(\mathrm{d}+2)\big)^s \leqslant \big(3\mathrm{d}^2\big)^s$ and (iii) $\|[\mathrm{T}_{\bullet|\bullet}]^{s}_{\underline{m}}[\mathrm{M}_{\mathfrak{t}}]^{-s}_{\underline{m}}\|_{\text{spec}} \leqslant \mathrm{d}^s$.*

§05|01.23 **Proof** of Lemma §05|01.22. Given in the lecture. □

§05|01|01 Global and maximal global $\mathfrak{v}$-error

We shall measure first globally the accuracy of the Galerkin solution $\theta^m_{\bullet} \in \ell_2\mathbb{1}^m_{\bullet}$ of $\theta_{\bullet} = \mathrm{T}^{\dagger}_{\bullet|\bullet}g_{\bullet} \in \ell_2$.

§05|01.24 **Property** (Global $\mathfrak{v}$-error). *Consider $\mathfrak{v}_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}$, $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\gtrapprox}_{\bullet}(\ell_2)$ and $g_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) = \mathrm{ran}(\mathrm{T}_{\bullet|\bullet}) \subseteq \ell_2$ and hence $\theta_{\bullet} = \mathrm{T}^{\dagger}_{\bullet|\bullet}g_{\bullet} = \mathrm{T}^{-1}_{\bullet|\bullet}g_{\bullet} \in \ell_2$. Given $m \in \mathbb{N}$ we have $\mathfrak{v}^2_{\bullet}\mathbb{1}^m_{\bullet} \in \ell_\infty$ and hence $\ell_2\mathbb{1}^m_{\bullet} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$. Consequently, denoting by $\theta^m_{\bullet} = \mathrm{T}^{m|\dagger}_{\bullet|\bullet}g_{\bullet} \in \ell_2\mathbb{1}^m_{\bullet}$ a Galerkin solution we have $\theta^m_{\bullet} \in \ell_2(\mathfrak{v}^2_{\bullet})$ with*

$$\|\theta^m_{\bullet}\|_{\mathfrak{v}} \leqslant \|[\mathrm{M}_{\mathfrak{v}}]_{\underline{m}}[\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}}\|_{\text{spec}}\|[g_{\bullet}]_{\underline{m}}\| \in \mathbb{R}_{\geqslant 0}.$$

If $\mathrm{C}_{\mathrm{T}} := \sup\big\{\|\mathrm{M}_{\mathfrak{v}}\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\mathrm{T}_{\bullet|\bullet}\mathrm{M}_{\mathbb{1}^{m|\perp}}\|_{\mathbb{L}(\ell_2)} : m \in \mathbb{N}\big\} \in \mathbb{R}_{\geqslant 0}$ then

$$\|\theta^m_{\bullet} - \theta_{\bullet}\|_{\mathfrak{v}} \leqslant (1 + \mathrm{C}_{\mathrm{T}})\|\mathbb{1}^{m|\perp}_{\bullet}\theta_{\bullet}\|_{\ell_2}$$

which implies $\sup\big\{\|\theta^j_{\bullet} - \theta_{\bullet}\|_{\mathfrak{v}} : j \in [\![m, \infty)\!)\big\} = \mathrm{o}(1)$ as $m \to \infty$. □

§05|01.25 **Notation** (Reminder §04|02.08). For $\mathrm{w}_{\bullet} \in \ell_\infty$ we set $\mathrm{w}^2_{(0)} := \|\mathrm{w}^2_{\bullet}\|_{\ell_\infty}$ and $\mathrm{w}^2_{(\bullet)} = (\mathrm{w}^2_{(j)} := \|\mathrm{w}^2_{\bullet}\mathbb{1}^{j|\perp}_{\bullet}\|_{\ell_\infty})_{j\in\mathbb{N}}$, where by construction $\mathrm{w}^2_{(j)} = \sup\big\{\mathrm{w}^2_i : i \in \mathbb{N}_{>j}\big\}$, $j \in \mathbb{Z}_{\geqslant 0}$ and $\mathrm{w}^2_{(\bullet)} \in (\mathbb{R}_{\geqslant 0})^{\mathbb{N}}_{\searrow}$. Evidently, if in addition $\mathrm{w}^2_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ then we have $\mathrm{w}^2_{(\bullet)} = (\mathrm{w}^2_{(j)} = \mathrm{w}^2_{j+1})_{j\in\mathbb{N}}$. □

§05|01.26 **Assumption**. Consider weights $\mathfrak{t}_{\bullet}, \mathfrak{a}_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ and $\mathfrak{v}_{\bullet} \in (\mathbb{R}_{\setminus 0})^{\mathbb{N}}$ such that $(\mathfrak{a}\mathfrak{v})_{\bullet} := \mathfrak{a}_{\bullet}\mathfrak{v}_{\bullet} \in \ell_\infty$ and $(\mathfrak{t}/\mathfrak{v})_{\bullet} = \mathfrak{t}_{\bullet}\mathfrak{v}^{-1}_{\bullet} \in \ell_\infty$ are satisfied. In addition there exists $\mathrm{C}_{(\mathfrak{t}/\mathfrak{v})} \in (0, 1]$ such that for all $m \in \mathbb{N}$

$$(\mathfrak{t}/\mathfrak{v})^2_{(m-1)} \geqslant \min\big\{(\mathfrak{t}/\mathfrak{v})^2_j : j \in [\![m]\!]\big\} \geqslant \mathrm{C}_{(\mathfrak{t}/\mathfrak{v})}(\mathfrak{t}/\mathfrak{v})^2_{(m)} \tag{05.02}$$

or in equal $\mathrm{C}_{(\mathfrak{t}/\mathfrak{v})}\|(\mathfrak{t}/\mathfrak{v})^{-2}_{\bullet}\mathbb{1}^m_{\bullet}\|_{\ell_\infty} \leqslant (\mathfrak{t}/\mathfrak{v})^{-2}_{(m)}$. □

§05|01.27 **Reminder**. Under Assumption §05|01.26 we have $\mathbb{J}^{\mathfrak{a}} = \ell^{\mathfrak{a}}_2 = \mathrm{dom}(\mathrm{M}_{\mathfrak{a}^{-1}}) = \ell_2\mathfrak{a}_{\bullet} = \ell_2(\mathfrak{a}^{-2}_{\bullet})$ and the three measures $\nu_{\mathbb{N}}$, $\mathfrak{a}^{-2}_{\bullet}\nu_{\mathbb{N}}$ and $\mathfrak{v}^2_{\bullet}\nu_{\mathbb{N}}$ dominate mutually each other, i.e. they share the same null sets (see Property §04|01.02). Consequently, since $(\mathfrak{a}\mathfrak{v})_{\bullet} \in \ell_\infty$ and

$$\|h_{\bullet}\|_{\mathfrak{v}} = \|(\mathfrak{a}\mathfrak{v})_{\bullet}\mathfrak{a}^{-1}_{\bullet}h_{\bullet}\|_{\ell_2} \leqslant \|(\mathfrak{a}\mathfrak{v})_{\bullet}\|_{\ell_\infty}\|h_{\bullet}\|_{\mathfrak{a}^{-1}} \in \mathbb{R}_{\geqslant 0} \quad \text{for each } h_{\bullet} \in \ell^{\mathfrak{a}}_2$$

we have $\ell_2^{\mathfrak{a}} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$. Moreover, since $\mathfrak{t}_{\bullet}, \mathfrak{a}_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ for each $s \in [0,1]$ we have $\mathfrak{t}^{1-s}_{\bullet}, \mathfrak{t}^s_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ and $(\mathfrak{a}\mathfrak{t}^s)_{\bullet} = \mathfrak{a}_{\bullet}\mathfrak{t}^s_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$. We note further if in addition $(\mathfrak{t}/\mathfrak{v})_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ is satisfied, then Assumption §05|01.26 (05.02) is fulfilled with $C_{(\mathfrak{t}/\mathfrak{v})} = 1$ due to Property §05|01.07. □

§05|01.28 **Lemma** (Maximal global $\mathfrak{v}$-error). *Let Assumption §05|01.26, $T_{\bullet|\bullet} \in \mathbb{T}^{\gtrsim}_{\mathfrak{t},d}$, $g_{\bullet} \in \operatorname{dom}(T^{\dagger}_{\bullet|\bullet}) = \operatorname{ran}(T_{\bullet|\bullet}) \subseteq \ell_2$ and $\theta_{\bullet} = T^{\dagger}_{\bullet|\bullet} g_{\bullet} = T^{-1}_{\bullet|\bullet} g_{\bullet} \in \ell_2^{\mathfrak{a}_{\bullet}, r}$ be satisfied. Given $m \in \mathbb{N}$ denoting by $\theta^m_{\bullet} = T^{m|\dagger}_{\bullet|\bullet} g_{\bullet} \in \ell_2 \mathbb{1}^m_{\bullet}$ a Galerkin solution for any $s \in [0,1]$ we obtain*

$$\|\theta_{\bullet} - \theta^m_{\bullet}\|^2_{\mathfrak{v}} \leqslant (9d^6 C^{-2}_{(\mathfrak{t}/\mathfrak{v})} + 1)\,(\mathfrak{a}\mathfrak{v})^2_{(m)} \|\mathbb{1}^{m|\perp}_{\bullet}\theta_{\bullet}\|^2_{\mathfrak{a}^{-1}}, \quad \|\theta^m_{\bullet}\|_{\mathfrak{a}^{-1}} \leqslant 3d^3\,\|\theta_{\bullet}\|_{\mathfrak{a}^{-1}}, \quad \text{and}$$
$$\|T^s_{\bullet|\bullet}(\theta_{\bullet} - \theta^m_{\bullet})\|^2_{\ell_2} \leqslant (9d^6 + 1) d^{2s} (\mathfrak{a}\mathfrak{t}^s)^2_{(m)} \|\mathbb{1}^{m|\perp}_{\bullet}\theta_{\bullet}\|^2_{\mathfrak{a}^{-1}}. \quad (05.03)$$

§05|01.29 **Proof** of Lemma §05|01.28. Given in the lecture. □

§05|01|02 Global and maximal global ϕ-error

Secondly we measure locally the accuracy of the Galerkin solution $\theta^m_{\bullet} \in \ell_2 \mathbb{1}^m_{\bullet}$ of $\theta_{\bullet} = T^{\dagger}_{\bullet|\bullet} g_{\bullet} \in \ell_2$.

§05|01.30 **Reminder**. Given $\phi_{\bullet} \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ for $\operatorname{dom}(\phi\nu_{\mathbb{N}}) := \{h_{\bullet} \in \ell_2 : \phi_{\bullet} h_{\bullet} \in \ell_1\}$ we consider as in Notation §04|02.16 the linear functional $\phi\nu_{\mathbb{N}} : \ell_2 \supseteq \operatorname{dom}(\phi\nu_{\mathbb{N}}) \to \mathbb{R}$ defined by

$$h_{\bullet} \mapsto \phi\nu_{\mathbb{N}}(h_{\bullet}) := \nu_{\mathbb{N}}(\phi_{\bullet} h_{\bullet}) = \sum_{j \in \mathbb{N}} \phi_j h_j.$$

For each $\theta_{\bullet} \in \operatorname{dom}(\phi\nu_{\mathbb{N}})$ and $m \in \mathbb{N}$ by Property §04|02.18 we have $\theta_{\bullet}\mathbb{1}^m_{\bullet} \in \operatorname{dom}(\phi\nu_{\mathbb{N}})$ with

$$|\phi\nu_{\mathbb{N}}(\theta_{\bullet} - \theta_{\bullet}\mathbb{1}^m_{\bullet})| \leqslant |\phi|\nu_{\mathbb{N}}(|\theta_{\bullet}|\mathbb{1}^{m|\perp}_{\bullet}) \leqslant \nu_{\mathbb{N}}(|\phi_{\bullet}\theta_{\bullet}|) \in \mathbb{R}_{\geqslant 0},$$

and $|\phi\nu_{\mathbb{N}}(\theta_{\bullet} - \theta_{\bullet}\mathbb{1}^m_{\bullet})||\phi\nu(\theta_{\bullet}) - \phi\nu(\theta^m_{\bullet})| = o(1)$ as $m \to \infty$ by dominated convergence. □

§05|01.31 **Property** (Local ϕ-error). *Consider $\phi_{\bullet} \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$, $T_{\bullet|\bullet} \in \mathbb{L}^{\gtrsim}(\ell_2)$ and $g_{\bullet} \in \operatorname{dom}(T^{\dagger}_{\bullet|\bullet}) = \operatorname{ran}(T_{\bullet|\bullet}) \subseteq \ell_2$ and hence $\theta_{\bullet} = T^{\dagger}_{\bullet|\bullet} g_{\bullet} = T^{-1}_{\bullet|\bullet} g_{\bullet} \in \ell_2$. Given $m \in \mathbb{N}$ we have $\phi^2_{\bullet}\mathbb{1}^m_{\bullet} \in \ell_2$ and hence $\ell_2\mathbb{1}^m_{\bullet} \subseteq \operatorname{dom}(\phi\nu_{\mathbb{N}})$. Consequently, denoting by $\theta^m_{\bullet} = T^{m|\dagger}_{\bullet|\bullet} g_{\bullet} \in \ell_2\mathbb{1}^m_{\bullet}$ a Galerkin solution we have $\theta^m_{\bullet} \in \operatorname{dom}(\phi\nu_{\mathbb{N}})$ with*

$$\|\phi_{\bullet}\theta^m_{\bullet}\|_{\ell_1} \leqslant \|[T_{\bullet|\bullet}]^{-1}_{\underline{m}}[\phi_{\bullet}]_{\underline{m}}\|\,\|[g_{\bullet}]_{\underline{m}}\| \in \mathbb{R}_{\geqslant 0}.$$

If $C_T := \sup\left\{\|M_{\mathbb{1}^{m|\perp}_{\bullet}} T_{\bullet|\bullet} T^{m|\dagger}_{\bullet|\bullet}\phi_{\bullet}\|_{\ell_2} : m \in \mathbb{N}\right\} \in \mathbb{R}_{\geqslant 0}$ *then*

$$|\phi\nu_{\mathbb{N}}(\theta^m_{\bullet} - \theta_{\bullet})| \leqslant (1 + C_T)\|\mathbb{1}^{m|\perp}_{\bullet}\theta_{\bullet}\|_{\ell_2}$$

which implies $\sup\left\{|\phi\nu_{\mathbb{N}}(\theta^j_{\bullet} - \theta_{\bullet})| : j \in [\![m, \infty[\![\right\} = o(1)$ *as* $m \to \infty$. □

§05|01.32 **Assumption**. Let $\mathfrak{t}_{\bullet}, \mathfrak{a}_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ and $\phi_{\bullet} \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ such that $(\mathfrak{a}\mathfrak{t})_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$ and $(\mathfrak{a}\phi)_{\bullet} \in \ell_2$. □

§05|01.33 **Reminder**. Under Assumption §05|01.32 we have $\mathbb{J}^{\mathfrak{a}} = \ell_2^{\mathfrak{a}} = \operatorname{dom}(M_{\mathfrak{a}^{-1}_{\bullet}}) = \ell_2\mathfrak{a}_{\bullet} = \ell_2(\mathfrak{a}^{-2}_{\bullet})$ and the three measures $\nu_{\mathbb{N}}$, $\mathfrak{a}^{-2}_{\bullet}\nu_{\mathbb{N}}$ and $|\phi|\nu_{\mathbb{N}}$ dominate mutually each other, i.e. they share the same null sets (see Property §04|01.02). Consequently, since $(\mathfrak{a}\phi)_{\bullet} \in \ell_2$ and (Property §04|02.23)

$$\|\phi_{\bullet} h_{\bullet}\|_{\ell_1} = \nu_{\mathbb{N}}(|h_{\bullet}\mathfrak{a}^{-1}_{\bullet}(\mathfrak{a}\phi)_{\bullet}|) \leqslant \|(\mathfrak{a}\phi)_{\bullet}\|_{\ell_2}\|h_{\bullet}\|_{\mathfrak{a}^{-1}} \in \mathbb{R}_{\geqslant 0} \quad \text{for each } h_{\bullet} \in \ell_2^{\mathfrak{a}}$$

we have $\ell_2^{\mathfrak{a}} \subseteq \operatorname{dom}(\phi\nu_{\mathbb{N}})$. Moreover, from $(\mathfrak{a}\phi)_{\bullet} \in \ell_2$ follows $\|\mathfrak{a}_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\phi} = \|(\mathfrak{a}\phi)_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\ell_2} = o(1)$ as $m \to \infty$. For $s \in [0,1]$ from $(\mathfrak{a}\mathfrak{t}^s)_{\bullet} = \mathfrak{a}_{\bullet}\mathfrak{t}^s_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ follows $(\mathfrak{a}\mathfrak{t}^s)_{(\bullet)} = ((\mathfrak{a}\mathfrak{t}^s)_{(m)} := (\mathfrak{a}\mathfrak{t}^s)_{m+1} = \|(\mathfrak{a}\mathfrak{t}^s)_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\ell_\infty})_{m \in \mathbb{N}} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$. □

§05|01.34 **Lemma** (Maximal local ϕ-error). *Let Assumption §05|01.32, $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}}$, $g_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) = \mathrm{ran}(\mathrm{T}_{\bullet|\bullet}) \subseteq \ell_2$ and $\theta_{\bullet} = \mathrm{T}^{\dagger}_{\bullet|\bullet} g_{\bullet} = \mathrm{T}^{-1}_{\bullet|\bullet} g_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}$ be satisfied. Given $m \in \mathbb{N}$ denoting by $\theta^{m}_{\bullet} = \mathrm{T}^{m|\dagger}_{\bullet|\bullet} g_{\bullet} \in {}_{\ell_2}\mathbb{1}^{m}_{\bullet}$ a Galerkin solution for any $s \in [0,1]$ we obtain*

$$|\phi\nu_{\mathbb{N}}(\theta^{m}_{\bullet} - \theta_{\bullet})|^2 \leqslant 3\mathrm{d}^3(3\mathrm{d}^3+1)\|\mathbb{1}^{m|\perp}_{\bullet}\theta_{\bullet}\|^2_{\mathfrak{a}^{-1}}\big(\|\mathfrak{a}_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|^2_{\phi} + (\mathfrak{a}\mathfrak{t}^{s})^2_{(m)}\|\mathfrak{t}^{-s}_{\bullet}\mathbb{1}^{m}_{\bullet}\|^2_{\phi}\big). \tag{05.04}$$

§05|01.35 **Proof** of Lemma §05|01.34. Given in the lecture. □

§05|01.36 **Lemma**. *For each $m \in \mathbb{N}$ denote $\mathfrak{bias}^2_m := \|\mathfrak{a}_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|^2_{\phi} + (\mathfrak{a}\mathfrak{t}^{s})^2_{(m)}\|\mathfrak{t}^{-s}_{\bullet}\mathbb{1}^{m}_{\bullet}\|^2_{\phi}$. If $(\mathfrak{a}\phi)_{\bullet} \in \ell_2$ and $(\mathfrak{a}\mathfrak{t}^{s})_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$ then it follows $\mathfrak{bias}^2_{\bullet} \in (\mathbb{R}_{\geqslant 0})^{\mathbb{N}}_{\downarrow 0}$.*

§05|01.37 **Proof** of Lemma §05|01.36. Given in the lecture. □

§05|02 Generalised linear Galerkin approach

§05|02.01 **Generalised linear Galerkin approach**. Given $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}(\ell_2)$ and $g_{\bullet} \in \ell_2$ any element $\theta^{m}_{\bullet} \in {}_{\ell_2}\mathbb{1}^{m}_{\bullet}$ satisfying $\mathrm{T}^{m}_{\bullet|\bullet}\theta^{m}_{\bullet} = \mathbb{1}^{m}_{\bullet} g_{\bullet}$, i.e., $[\mathrm{T}_{\bullet|\bullet}]_{\underline{m}}[\theta^{m}_{\bullet}]_{\underline{m}} = [g_{\bullet}]_{\underline{m}}$, is called a *generalised Galerkin solution*. □

§05|02.02 **Notation**. We denote by $\mathbb{L}^{\mathbb{R}}(\ell_2)$ the subset of all injective $\mathrm{A}_{\bullet|\bullet} \in \mathbb{L}(\ell_2)$ such that $[\mathrm{A}_{\bullet|\bullet}]_{\underline{m}} \in \mathbb{R}^{(m,m)}$ is regular for all $m \in \mathbb{N}$. For each $m \in \mathbb{N}$ and $\mathrm{A}_{\bullet|\bullet} \in \mathbb{L}^{\mathbb{R}}(\ell_2)$, the inverse $[\mathrm{A}_{\bullet|\bullet}]^{-1}_{\underline{m}} \in \mathbb{R}^{(m,m)}$ of $[\mathrm{A}_{\bullet|\bullet}]_{\underline{m}} \in \mathbb{R}^{(m,m)}$ exists. Note that $\mathbb{L}^{\gtrless}(\ell_2) \subseteq \mathbb{L}^{\mathbb{R}}(\ell_2)$ (Lemma §05|01.22). □

§05|02.03 **Assumption**. For $\mathbb{J} = \ell_2$, surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H},\ell_2)$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G},\ell_2)$ fixed and presumed to be *known* in advance, the operator $\mathrm{T} \in \mathbb{L}(\mathbb{H},\mathbb{G})$ satisfies $\mathrm{T}_{\bullet|\bullet} = \mathrm{V}\mathrm{T}\mathrm{U}^{\star} \in \mathbb{L}^{\mathbb{R}}(\ell_2) \subseteq \mathbb{L}(\ell_2) = \mathbb{L}_{\bullet|\bullet}(\ell_2)$. Let $g_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) = \mathrm{ran}(\mathrm{T}_{\bullet|\bullet})$, and hence $\theta_{\bullet} = \mathrm{T}^{\dagger}_{\bullet|\bullet} g_{\bullet} = \mathrm{T}^{-1}_{\bullet|\bullet} g_{\bullet} \in \ell_2$. □

§05|02.04 **Remark**. We consider a generalised linear Galerkin approach under Assumption §05|02.03, i.e. $[\mathrm{T}_{\bullet|\bullet}]_{\underline{m}}$ is assumed to be regular for each $m \in \mathbb{N}$, so that $[\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}}$ always exists. We shall emphasise that it is a non-trivial problem to determine when such an assumption holds (cf. Efromovich and Koltchinskii [2001] and references therein). However, if $[\mathrm{T}_{\bullet|\bullet}]_{\underline{m}}$ is regular, then for each $g_{\bullet} \in \ell_2$ the *generalised Galerkin solution* $\theta^{m}_{\bullet} = \mathrm{T}^{m|\dagger}_{\bullet|\bullet} g_{\bullet} \in {}_{\ell_2}\mathbb{1}^{m}_{\bullet}$ is by $[\theta^{m}_{\bullet}]_{\underline{m}} = [\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}}[g_{\bullet}]_{\underline{m}}$ uniquely determined. □

§05|02.05 **Generalised link condition**. Given weights $\mathfrak{t}_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}} \cap \ell_{\infty}$ an operator $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\mathbb{R}}(\ell_2)$ satisfies a *generalised link condition* if there exist $\mathrm{D} \in \mathbb{R}_{\geqslant 1}$ and $\mathrm{d} \in [1,\mathrm{D}]$ such that

$$\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d,D}} := \big\{\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d}} \cap \mathbb{L}^{\mathbb{R}}(\ell_2) \colon \|[\mathrm{M}_{\mathrm{t}}]_{\underline{m}}[\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}}\|_{\mathrm{spec}} = \|[\mathrm{T}^{\star}_{\bullet|\bullet}]^{-1}_{\underline{m}}[\mathrm{M}_{\mathrm{t}}]_{\underline{m}}\|_{\mathrm{spec}} \leqslant \mathrm{D} \text{ for all } m \in \mathbb{N}\big\}. \quad \square$$

§05|02.06 **Remark**. We shall emphasise that $\mathbb{T}_{\mathrm{t,d,D}}$ contains the subset $\mathbb{M}_{\mathrm{t,d}}$ of diagonal operator satisfying the link condition, i.e. $\mathbb{M}_{\mathrm{t,d}} \subseteq \mathbb{T}_{\mathrm{t,d}}$ (see Remark §05|01.09). Indeed, any $\mathrm{M}_{\mathrm{w}} \in \mathbb{M}_{\mathrm{t,d}}$ satisfies $\|[\mathrm{M}_{\mathrm{t}}]_{\underline{m}}[\mathrm{M}_{\mathrm{w}}]^{-1}_{\underline{m}}\|_{\mathrm{spec}} = \|\mathfrak{t}_{\bullet}\mathrm{w}^{-1}_{\bullet}\mathbb{1}^{m}_{\bullet}\|_{\ell_{\infty}} \leqslant \mathrm{d} \leqslant \mathrm{D}$. Moreover, we have $\mathbb{T}^{\gtrless}_{\mathrm{t,d}} \subseteq \mathbb{T}_{\mathrm{t,d,D}}$ whenever $\mathrm{D} \geqslant 3\mathrm{d}^2$ due to Lemma §05|01.22 (ii). The link condition $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d}}$ or in equal $(\mathrm{T}^{\star}_{\bullet|\bullet}\mathrm{T}_{\bullet|\bullet})^{1/2} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}}$ does not depend on an unitary V, i.e. $\mathrm{V}^{\star}\mathrm{V} = \mathrm{id}_{\mathbb{G}}$, (or more generally surjective partial isometry with $\mathrm{ran}(\mathrm{T}) \subseteq \overline{\mathrm{ran}}(\mathrm{V}^{\star})$ implying $\mathrm{T}^{\star}\mathrm{V}^{\star}\mathrm{V}\mathrm{T} = \mathrm{T}^{\star}\mathrm{T}$) since for each $\mathrm{T} \in \mathbb{L}(\mathbb{H},\mathbb{G})$ with $\mathrm{V}\mathrm{T}\mathrm{U}^{\star} = \mathrm{T}_{\bullet|\bullet}$ we have $\mathrm{T}^{\star}_{\bullet|\bullet}\mathrm{T}_{\bullet|\bullet} = \mathrm{U}\mathrm{T}^{\star}\mathrm{V}^{\star}\mathrm{V}\mathrm{T}\mathrm{U}^{\star} = \mathrm{U}\mathrm{T}^{\star}\mathrm{T}\mathrm{U}^{\star}$. The general link condition Definition §05|02.05 however involves both surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H},\ell_2)$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G},\ell_2)$. It is worth pointing out, that for each $\mathrm{T} \in \mathbb{L}(\mathbb{H},\mathbb{G})$ and surjective partial isometry $\mathrm{U} \in \mathbb{L}(\mathbb{H},\ell_2)$ satisfying $(\mathrm{U}\mathrm{T}^{\star}\mathrm{T}\mathrm{U}^{\star})^{1/2} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}}$ we can theoretically construct a surjective partial isometry $\mathrm{V} \in \mathbb{L}(\mathbb{G},\ell_2)$ such that $\mathrm{V}\mathrm{T}\mathrm{U}^{\star} = \mathrm{T}_{\bullet|\bullet} \in \mathbb{L}(\ell_2)$ satisfies $\mathrm{T}_{\bullet|\bullet}\mathrm{T}^{\star}_{\bullet|\bullet} = \mathrm{U}\mathrm{T}^{\star}\mathrm{T}\mathrm{U}^{\star}$ and $\mathrm{T}^{\star}_{\bullet|\bullet} = \mathrm{U}\mathrm{T}^{\star}\mathrm{V}^{\star} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}}$. Consequently, from Lemma §05|01.22 (ii) it follows $\|[\mathrm{M}_{\mathrm{t}}]_{\underline{m}}[\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}}\|_{\mathrm{spec}} = \|[\mathrm{T}^{\star}_{\bullet|\bullet}]^{-1}_{\underline{m}}[\mathrm{M}_{\mathrm{t}}]_{\underline{m}}\|_{\mathrm{spec}} \leqslant 3\mathrm{d}^2$ for each $m \in \mathbb{N}$, which implies $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d,D}}$ for all $\mathrm{D} \geqslant 3\mathrm{d}^2$. The fundamental inequality of Heinz [1951] in

Property §05|01.10 implies $\|(\mathrm{UT^{\star}TU^{\star}})^{-1/2}\mathbb{1}^{\{j\}}_{\bullet}\|_{\ell_2} \leqslant \mathrm{d}\mathfrak{t}_j^{-1} \in \mathbb{R}_{\geqslant 0}$ for each $j \in \mathbb{N}$. Thereby, the sequence $(\mathrm{UT^{\star}TU^{\star}})^{-1/2}\mathbb{1}^{\{j\}}_{\bullet}$ is an element of ℓ_2 and, hence $\mathrm{v}_j := \mathrm{TU^{\star}}(\mathrm{UT^{\star}TU^{\star}})^{-1/2}\mathbb{1}^{\{j\}}_{\bullet}$, $j \in \mathbb{N}$ belongs to $\mathbb{G}$. Then it is easily checked that $(\mathrm{v}_j)_{j\in\mathbb{N}}$ is an *orthonormal sequence* in $\mathbb{G}$ which determines a surjective partial isometry $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \ell_2)$ (Notation §01|02.07). By construction we have $\mathrm{T}^{\star}_{\bullet|\bullet} = \mathrm{UT^{\star}V^{\star}} = \mathrm{UT^{\star}TU^{\star}}(\mathrm{UT^{\star}TU^{\star}})^{-1/2} = (\mathrm{UT^{\star}TU^{\star}})^{1/2} \in \mathbb{L}^{\geqslant}(\ell_2)$, hence $\mathrm{T}_{\bullet|\bullet} = (\mathrm{UT^{\star}TU^{\star}})^{1/2} \in \mathbb{L}^{\geqslant}(\ell_2)$, and thus $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}}$ or in equal $(\mathrm{T}_{\bullet|\bullet}\mathrm{T}^{\star}_{\bullet|\bullet})^{1/2} = (\mathrm{T}^{\star}_{\bullet|\bullet}\mathrm{T}_{\bullet|\bullet})^{1/2} = (\mathrm{UT^{\star}TU^{\star}})^{1/2} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}}$ and $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d}}$. □

§05|02.07 **Property**. *If* $\mathrm{T}_{\bullet|\bullet}, \mathrm{T}^{\star}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d}}$ *then also* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d,D}}$ *for each* $\mathrm{D} \geqslant 3\mathrm{d}^2$.(!) □ (!) Wieso ??

§05|02|01 Global and maximal global $\mathfrak{v}$-error

We shall measure first globally the accuracy of the Galerkin solution $\theta^m_{\bullet} \in {}_{\ell_2}\mathbb{1}^m_{\bullet}$ of $\theta_{\bullet} = \mathrm{T}^{\dagger}_{\bullet|\bullet} g_{\bullet} \in \ell_2$.

§05|02.08 **Property** (Global $\mathfrak{v}$-error). *Consider* $\mathfrak{v}_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}$, $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\mathrm{R}}(\ell_2)$ *and* $g_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) = \mathrm{ran}(\mathrm{T}_{\bullet|\bullet}) \subseteq \ell_2$ *and hence* $\theta_{\bullet} = \mathrm{T}^{\dagger}_{\bullet|\bullet} g_{\bullet} = \mathrm{T}^{-1}_{\bullet|\bullet} g_{\bullet} \in \ell_2$. *Given* $m \in \mathbb{N}$ *we have* $\mathfrak{v}^2_{\bullet}\mathbb{1}^m_{\bullet} \in \ell_\infty$ *and hence* ${}_{\ell_2}\mathbb{1}^m_{\bullet} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$. *Consequently, denoting by* $\theta^m_{\bullet} = \mathrm{T}^{m|\dagger}_{\bullet|\bullet} g_{\bullet} \in {}_{\ell_2}\mathbb{1}^m_{\bullet}$ *a generalised Galerkin solution we have* $\theta^m_{\bullet} \in \ell_2(\mathfrak{v}^2_{\bullet})$ *with*

$$\|\theta^m_{\bullet}\|_{\ell_2(\mathfrak{v}_{\bullet})} \leqslant \|[\mathrm{M}_{\mathfrak{v}_{\bullet}}]_{\underline{m}}[\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}}\|_{\mathrm{spec}}\|[g_{\bullet}]_{\underline{m}}\| \in \mathbb{R}_{\geqslant 0}.$$

If $\mathrm{C}_{\mathrm{T}} := \sup\left\{\|\mathrm{M}_{\mathfrak{v}}\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\mathrm{T}_{\bullet|\bullet}\mathrm{M}_{\mathbb{1}^{m|\perp}}\|_{\mathbb{L}(\ell_2)} : m \in \mathbb{N}\right\} \in \mathbb{R}_{\geqslant 0}$ *then*

$$\|\theta^m_{\bullet} - \theta_{\bullet}\|_{\mathfrak{v}} \leqslant (1 + \mathrm{C}_{\mathrm{T}})\|\mathbb{1}^{m|\perp}_{\bullet}\theta_{\bullet}\|_{\ell_2}$$

which implies $\sup\left\{\|\theta^j_{\bullet} - \theta_{\bullet}\|_{\mathfrak{v}} : j \in \mathbb{N}_{\geqslant m}\right\} = \mathrm{o}(1)$ *as* $m \to \infty$. □

§05|02.09 **Lemma** (Maximal global $\mathfrak{v}$-error). *Under Assumption §05|01.26 let* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d,D}}$, $g_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) = \mathrm{ran}(\mathrm{T}_{\bullet|\bullet}) \subseteq \ell_2$ *and* $\theta_{\bullet} = \mathrm{T}^{\dagger}_{\bullet|\bullet} g_{\bullet} = \mathrm{T}^{-1}_{\bullet|\bullet} g_{\bullet} \in \ell_2^{\mathfrak{a}_{\bullet},\mathrm{r}}$. *Given* $m \in \mathbb{N}$ *denoting by* $\theta^m_{\bullet} = \mathrm{T}^{m|\dagger}_{\bullet|\bullet} g_{\bullet} \in {}_{\ell_2}\mathbb{1}^m_{\bullet}$ *a generalised Galerkin solution for any* $s \in [0,1]$ *we obtain*

$$\|\theta_{\bullet} - \theta^m_{\bullet}\|^2_{\mathfrak{v}} \leqslant (\mathrm{D}^2\mathrm{d}^2\mathrm{C}^{-2}_{(\mathfrak{t}/\mathfrak{v})} + 1)\,(\mathfrak{a}\mathfrak{v})^2_{(m)}\|\mathbb{1}^{m|\perp}_{\bullet}\theta_{\bullet}\|^2_{\mathfrak{a}^{-1}}, \quad \|\theta^m_{\bullet}\|_{\mathfrak{a}^{-1}} \leqslant \mathrm{Dd}\,\|\theta_{\bullet}\|_{\mathfrak{a}^{-1}}, \quad \textit{and}$$
$$\|\mathrm{T}^s_{\bullet|\bullet}(\theta_{\bullet} - \theta^m_{\bullet})\|_{\ell_2} \leqslant (\mathrm{Dd}+1)\mathrm{d}^s(\mathfrak{a}\mathfrak{t}^s)_{(m)}\|\mathbb{1}^{m|\perp}_{\bullet}\theta_{\bullet}\|_{\mathfrak{a}^{-1}}. \quad (05.05)$$

§05|02.10 **Proof** of Lemma §05|02.09. Given in the lecture. □

§05|02|02 Global and maximal global ϕ-error

Secondly we measure locally the accuracy of the generalised Galerkin solution $\theta^m_{\bullet} \in {}_{\ell_2}\mathbb{1}^m_{\bullet}$ of $\theta_{\bullet} = \mathrm{T}^{\dagger}_{\bullet|\bullet} g_{\bullet} \in \ell_2$.

§05|02.11 **Reminder**. Given $\phi_{\bullet} \in (\mathbb{R}_{\setminus 0})^{\mathbb{N}}$ for $\mathrm{dom}(\phi\nu_{\mathbb{N}}) := \{h_{\bullet} \in \ell_2 : \phi h_{\bullet} \in \ell_1\}$ we consider as in Notation §04|02.16 the linear functional $\phi\nu_{\mathbb{N}} : \ell_2 \supseteq \mathrm{dom}(\phi\nu_{\mathbb{N}}) \to \mathbb{R}$ defined by

$$h_{\bullet} \mapsto \phi\nu_{\mathbb{N}}(h_{\bullet}) := \nu_{\mathbb{N}}(\phi h_{\bullet}) = \sum_{j\in\mathbb{N}} \phi_j h_j.$$

For each $\theta_{\bullet} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ and $m \in \mathbb{N}$ by Property §04|02.18 we have $\theta_{\bullet}\mathbb{1}^m_{\bullet} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ with

$$|\phi\nu_{\mathbb{N}}(\theta_{\bullet} - \theta_{\bullet}\mathbb{1}^m_{\bullet})| \leqslant |\phi|\nu_{\mathbb{N}}(|\theta_{\bullet}|\mathbb{1}^{m|\perp}_{\bullet}) \leqslant \nu_{\mathbb{N}}(|\phi\theta_{\bullet}|) \in \mathbb{R}_{\geqslant 0},$$

and $|\phi\nu_{\mathbb{N}}(\theta_{\bullet} - \theta_{\bullet}\mathbb{1}^m_{\bullet})||\phi\nu(\theta_{\bullet}) - \phi\nu(\theta^m_{\bullet})| = \mathrm{o}(1)$ as $m \to \infty$ by dominated convergence. □

§05|02.12 **Property** (Local ϕ-error). *Consider* $\phi \in (\mathbb{R}_{\setminus 0})^{\mathbb{N}}$, $T_{\bullet|\bullet} \in \mathbb{L}^{\geq}(\ell_2)$ *and* $g_\bullet \in \mathrm{dom}(T^\dagger_{\bullet|\bullet}) = \mathrm{ran}(T_{\bullet|\bullet}) \subseteq \ell_2$ *and hence* $\theta_\bullet = T^\dagger_{\bullet|\bullet} g_\bullet = T^{-1}_{\bullet|\bullet} g_\bullet \in \ell_2$. *Given* $m \in \mathbb{N}$ *we have* $\phi^2 \mathbb{1}^m_\bullet \in \ell_2$ *and hence* $\ell_2 \mathbb{1}^m_\bullet \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$. *Consequently, denoting by* $\theta^m_\bullet = T^{m|\dagger}_{\bullet|\bullet} g_\bullet \in \ell_2 \mathbb{1}^m_\bullet$ *a Galerkin solution we have* $\theta^m_\bullet \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *with*

$$\|\phi\theta^m_\bullet\|_{\ell_1} \leqslant \|[T_{\bullet|\bullet}]^{-1}_{\underline{m}}[\phi]_{\underline{m}}\|\|[g_\bullet]_{\underline{m}}\| \in \mathbb{R}_{\geqslant 0}.$$

If $C_T := \sup\left\{\|M_{\mathbb{1}^{m|\perp}} T^\star_{\bullet|\bullet}(T^{m|\dagger}_{\bullet|\bullet})^\star \phi\|_{\ell_2} : m \in \mathbb{N}\right\} \in \mathbb{R}_{\geqslant 0}$ *then*

$$|\phi\nu_{\mathbb{N}}(\theta^m_\bullet - \theta_\bullet)| \leqslant (1 + C_T)\|\mathbb{1}^{m|\perp}_\bullet \theta_\bullet\|_{\ell_2}$$

which implies $\sup\left\{|\phi\nu_{\mathbb{N}}(\theta^j_\bullet - \theta_\bullet)| : j \in \mathbb{N}_{\geqslant m}\right\} = \mathrm{o}(1)$ *as* $m \to \infty$. □

§05|02.13 **Reminder**. Under Assumption §05|01.32 we have $\mathbb{J}^{\mathfrak{a}} = \ell_2^{\mathfrak{a}} = \mathrm{dom}(M_{\mathfrak{a}^\dagger}) = \ell_2 \mathfrak{a}_\bullet = \ell_2(\mathfrak{a}_\bullet^{-2})$ and the three measures $\nu_{\mathbb{N}}$, $\mathfrak{a}_\bullet^{-2}\nu_{\mathbb{N}}$ and $|\phi|\nu_{\mathbb{N}}$ dominate mutually each other, i.e. they share the same null sets (see Property §04|01.02). Consequently, since $(\mathfrak{a}\phi)_\bullet \in \ell_2$ and (Property §04|02.23)

$$\|\phi h_\bullet\|_{\ell_1} = \nu_{\mathbb{N}}(|h_\bullet \mathfrak{a}^\dagger_\bullet (\mathfrak{a}\phi)_\bullet|) \leqslant \|(\mathfrak{a}\phi)_\bullet\|_{\ell_2}\|h_\bullet\|_{\mathfrak{a}^{-1}} \in \mathbb{R}_{\geqslant 0} \quad \text{for each } h_\bullet \in \ell_2^{\mathfrak{a}}$$

we have $\ell_2^{\mathfrak{a}} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$. Moreover, from $(\mathfrak{a}\phi)_\bullet \in \ell_2$ follows $\|\mathfrak{a}_\bullet \mathbb{1}^{m|\perp}_\bullet\|_\phi = \|(\mathfrak{a}\phi)_\bullet \mathbb{1}^{m|\perp}\|_{\ell_2} = \mathrm{o}(1)$ as $m \to \infty$. For $s \in [0,1]$ from $(\mathfrak{a}\mathfrak{t}^s)_\bullet = \mathfrak{a}_\bullet \mathfrak{t}^s_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ follows $(\mathfrak{a}\mathfrak{t}^s)_{(\bullet)} = ((\mathfrak{a}\mathfrak{t}^s)_{(m)} := (\mathfrak{a}\mathfrak{t}^s)_{m+1} = \|(\mathfrak{a}\mathfrak{t}^s)_\bullet \mathbb{1}^{m|\perp}_\bullet\|_{\ell_\infty})_{m\in\mathbb{N}} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$. □

§05|02.14 **Lemma** (Maximal local ϕ-error). *Under Assumption §05|01.32 let* $T \in \mathbb{T}_{\mathrm{t,d,D}}$, $g_\bullet \in \mathrm{dom}(T^\dagger_{\bullet|\bullet}) = \mathrm{ran}(T_{\bullet|\bullet}) \subseteq \ell_2$ *and* $\theta_\bullet = T^\dagger_{\bullet|\bullet} g_\bullet = T^{-1}_{\bullet|\bullet} g_\bullet \in \ell_2^{\mathfrak{a},r}$. *Given* $m \in \mathbb{N}$ *denoting by* $\theta^m_\bullet = T^{m|\dagger}_{\bullet|\bullet} g_\bullet \in \ell_2 \mathbb{1}^m_\bullet$ *a generalised Galerkin solution for any* $s \in [0,1]$ *we obtain*

$$|\phi\nu_{\mathbb{N}}(\theta^m_\bullet - \theta_\bullet)|^2 \leqslant (1 + \mathrm{Dd})\mathrm{Dd}\|\mathbb{1}^{m|\perp}_\bullet \theta_\bullet\|^2_{\mathfrak{a}^{-1}}\left(\|\mathfrak{a}_\bullet \mathbb{1}^{m|\perp}_\bullet\|^2_\phi + (\mathfrak{a}\mathfrak{t}^s)^2_{(m)}\|\mathfrak{t}^{-s}_\bullet \mathbb{1}^m_\bullet\|^2_\phi\right). \qquad (05.06)$$

§05|02.15 **Proof** of Lemma §05|02.14. Given in the lecture. □

§06 Spectral regularisation

§06|00.01 **Notation**. Consider the measure space $(\mathcal{J}, \mathscr{J}, \nu)$ and the Hilbert space $\mathbb{J} = \mathbb{L}_2(\nu)$ as in Notation §01|01.01. We suppose that $U \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ and $V \in \mathbb{L}(\mathbb{G}, \mathbb{J})$ are surjective partial isometries, hence $VV^\star = \mathrm{id}_{\mathbb{J}} = UU^\star$. As in Definition §03|00.08 we denote for $A := VTU^\star \in \mathbb{L}(\mathbb{J})$ its Moore-Penrose inverse by $A^\dagger : \mathbb{J} \supseteq \mathrm{dom}(A^\dagger) \to \mathbb{J}$. □

§06|00.02 **Comment**. In case the operator $T \in \mathbb{L}(\mathbb{H}, \mathbb{G})$ is fixed and presumed to be *known* in advance, a spectral regularisation is formally not restricted to the diagonal or non-diagonal case as considered in Subsection §01|04 and Subsection §01|05, respectively. Consequently, we use in this section the symbol $A := VTU^\star \in \mathbb{L}(\mathbb{J})$. However, in case of a noisy operator we will restrict ourselves to the diagonal and non-diagonal case introduced in Definition §02|04.03 and Definition §02|02.03. □

§06|00.03 **Assumption**. For $\mathbb{J} = \mathbb{L}_2(\nu)$ let $U \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ and $V \in \mathbb{L}(\mathbb{G}, \mathbb{J})$ be surjective partial isometries fixed and presumed to be *known* in advance, let $T \in \mathbb{L}(\mathbb{H}, \mathbb{G})$, hence $A = VTU^\star \in \mathbb{L}(\mathbb{J})$ with Moore-Penrose inverse $A^\dagger : \mathbb{J} \supseteq \mathrm{dom}(A^\dagger) \to \mathbb{J}$ and let $g_\bullet \in \mathrm{dom}(A^\dagger)$, and hence $\theta_\bullet = A^\dagger g_\bullet \in \mathbb{J}$. □

§06|00.04 **Definition**. A collection $\left\{R_\alpha \in \mathbb{L}(\mathbb{J}) : \alpha \in (0,1)\right\}$ of operators is called *regularisation* of $A^\dagger$ if for any $g_\bullet \in \mathrm{dom}(A^\dagger)$ holds $\|R_\alpha g_\bullet - A^\dagger g_\bullet\|_{\mathbb{J}} \to 0$ as $\alpha \to 0$. □

§06|00.05 **Remark**. If $A^\dagger$ is not bounded, then we have $\|R_\alpha\|_{\mathbb{L}(\mathbb{J})} \to \infty$ as $\alpha \to 0$. However, for $g_\bullet \in \mathrm{dom}(A^\dagger)$ if $(g^n_\bullet)_{n\in\mathbb{N}}$ is a sequence in $\mathbb{J}$ such that $\|g^n_\bullet - g_\bullet\|_{\mathbb{J}} \leqslant n^{-1}$ for all $n \in \mathbb{N}$, then there exists a sequence $(\alpha_n)_{n\in\mathbb{N}}$ in $(0,1)$ such that $\|R_{\alpha_n} g^n_\bullet - A^\dagger g_\bullet\|_{\mathbb{J}} = \mathrm{o}(1)$ as $n \to \infty$. □

§06|01 (Generalised) Tikhonov regularisation

§06|01.01 **Definition**. The collection $\{\mathrm{A}^{\alpha|-1}\mathrm{A}^\star \in \mathbb{L}(\mathbb{J})\colon \mathrm{A}^{\alpha|-1} := (\mathrm{A}^\star\mathrm{A} + \alpha\mathrm{id}_\mathbb{J})^{-1} \in \mathbb{L}^{\gtrless}(\mathbb{J}), \alpha \in (0,1)\}$ of operators is called *Tikhonov regularisation* of $\mathrm{A}^\dagger : \mathbb{J} \supseteq \mathrm{dom}(\mathrm{A}^\dagger) \to \mathbb{J}$. □

§06|01.02 **Remark**. Given $\mathrm{A} \in \mathbb{L}(\mathbb{J})$ consider for each $\alpha \in (0,1)$ the strictly positive definite operator $\mathrm{A}^\alpha = \mathrm{A}^\star\mathrm{A} + \alpha\mathrm{id}_\mathbb{J} \in \mathbb{L}^{\gtrless}(\mathbb{J})$ where

$$(\|\mathrm{A}\|^2_{\mathbb{L}(\mathbb{J})} + \alpha)\|h_\bullet\|^2_\mathbb{J} \geqslant \|\mathrm{A}^\alpha h_\bullet\|_\mathbb{J}\|h_\bullet\|_\mathbb{J} \geqslant \langle \mathrm{A}^\alpha h_\bullet, h_\bullet\rangle_\mathbb{J} \geqslant \alpha\|h_\bullet\|^2_\mathbb{J} \in \mathbb{R}_{>0} \tag{06.01}$$

for any $h_\bullet \in \mathbb{J}_{\setminus 0} = \mathbb{J} \setminus \{0_\bullet\}$ by applying the Cauchy-Schwarz inequality and, hence

$$\inf\{\|\mathrm{A}^\alpha h_\bullet\|_\mathbb{J}\colon \|h_\bullet\|_\mathbb{J} = 1, h_\bullet \in \mathbb{J}\} \geqslant \alpha \in \mathbb{R}_{>0}. \tag{06.02}$$

Using the last bound $\mathrm{A}^\alpha \in \mathbb{L}^{\gtrless}(\mathbb{J})$ has a closed range $\mathrm{ran}(\mathrm{A}^\alpha)$. Indeed, if $(\mathrm{A}^\alpha a^n_\bullet)_{j\in\mathbb{N}}$ converges, say to $g_\bullet \in \mathbb{J}$, then $(\mathrm{A}^\alpha h^n_\bullet)_{j\in\mathbb{N}}$ is a Cauchy sequence and also $(h^n_\bullet)_{j\in\mathbb{N}}$ by (06.01). Since $\mathbb{J}$ is complete, $(h^n_\bullet)_{j\in\mathbb{N}}$ converges, say to $h_\bullet \in \mathbb{J}$. Since A^α is continuous, $(\mathrm{A}^\alpha h^n_\bullet)_{j\in\mathbb{N}}$ converges to $\mathrm{A}^\alpha h_\bullet = g_\bullet$. In other words the range is closed. Since $\mathrm{A}^\alpha \in \mathbb{L}(\ell_2)$ is injective with closed range it follows $\mathrm{ran}(\mathrm{A}^\alpha) = \ker(\mathrm{A}^\alpha)^\perp = \mathbb{J}$, which in turn implies A^α is invertible, and due to the open mapping theorem with inverse $\mathrm{A}^{\alpha|-1} = (\mathrm{A}^\alpha)^{-1} \in \mathbb{L}(\mathbb{J})$. Moreover, exploiting $\mathrm{ran}(\mathrm{A}^\alpha) = \mathbb{J}$ and (06.02) we have $\|\mathrm{A}^{\alpha|-1}\|_{\mathbb{L}(2)} \leqslant \alpha^{-1}$ since

$$\begin{aligned}\|\mathrm{A}^{\alpha|-1}\|_{\mathbb{L}(2)} &= \sup\{\|\mathrm{A}^{\alpha|-1}g_\bullet\|_\mathbb{J}\colon g_\bullet \in \mathbb{J}, \|g_\bullet\|_\mathbb{J} = 1\} = \sup\{\tfrac{\|\mathrm{A}^{\alpha|-1}g_\bullet\|_\mathbb{J}}{\|g_\bullet\|_\mathbb{J}}\colon g_\bullet \in \mathbb{J}_{\setminus 0} = \mathrm{ran}(\mathrm{A}^\alpha)\setminus\{0_\bullet\}\} \\ &= \sup\{\tfrac{\|h_\bullet\|_\mathbb{J}}{\|\mathrm{A}^\alpha h_\bullet\|_\mathbb{J}}\colon h_\bullet \in \mathbb{J}_{\setminus 0}\} = \sup\{\|\mathrm{A}^\alpha h_\bullet\|^{-1}_\mathbb{J}\colon h_\bullet \in \mathbb{J}, \|h_\bullet\|_\mathbb{J} = 1\} \leqslant \alpha^{-1}.\end{aligned}$$

Consequently, the collection $\{\mathrm{A}^{\alpha|-1}\mathrm{A}^\star = (\mathrm{A}^\star\mathrm{A} + \alpha\mathrm{id}_\mathbb{J})^{-1}\mathrm{A}^\star \in \mathbb{L}(\mathbb{J})\colon \alpha \in (0,1)\}$ is well-defined. □

§06|01.03 **Lemma**. *For each* $h_\bullet \in \ker(\mathrm{A})^\perp$ *holds* $\|\alpha(\mathrm{A}^\star\mathrm{A} + \alpha\mathrm{id}_\mathbb{J})^{-1}h_\bullet\|_\mathbb{J} = \mathrm{o}(1)$ *as* $\alpha \to 0$.

§06|01.04 **Proof** of Lemma §06|01.03. Given in the lecture. □

§06|01.05 **Remark**. Let $g_\bullet \in \mathrm{dom}(\mathrm{A}^\dagger)$, $\theta_\bullet = \mathrm{A}^\dagger g_\bullet \in \mathbb{J}$ and $\theta^\alpha_\bullet := \mathrm{A}^{\alpha|-1}\mathrm{A}^\star g_\bullet \in \mathbb{J}$ we have

$$\mathrm{A}^\alpha(\theta_\bullet - \theta^\alpha_\bullet) = \mathrm{A}^\star\mathrm{A}\mathrm{A}^\dagger g_\bullet + \alpha\theta_\bullet - \mathrm{A}^\alpha\mathrm{A}^{\alpha|-1}\mathrm{A}^\star g_\bullet = \mathrm{A}^\star g_\bullet + \alpha\theta_\bullet - \mathrm{A}^\star g_\bullet = \alpha\theta_\bullet,$$

and rewriting the last identity $\mathrm{A}^{\alpha|-1}\mathrm{A}^\star g_\bullet - \mathrm{A}^\dagger g_\bullet = -\alpha\mathrm{A}^{\alpha|-1}\theta_\bullet$. Consequently, from Lemma §06|01.03 follows $\|\mathrm{A}^{\alpha|-1}\mathrm{A}^\star g_\bullet - \mathrm{A}^\dagger g_\bullet\|_\mathbb{J} = \mathrm{o}(1)$ as $\alpha \to 0$ since $\theta_\bullet = \mathrm{A}^\dagger g_\bullet \in \mathbb{J}$. Thereby, the Tikhonov collection as in Definition §06|01.01 is indeed a regularisation in the sense of Definition §06|00.04. □

§06|01.06 **Lemma**. *For each* $\mathrm{C} \in \mathbb{L}(\mathbb{J})$ *the following statements are equivalent:*

(i) $\theta^\alpha_\bullet$ *minimises the generalised Tikhonov functional* $h_\bullet \mapsto F_\alpha(h_\bullet) := \frac{1}{2}\|g_\bullet - \mathrm{A}h_\bullet\|^2_\mathbb{J} + \frac{\alpha}{2}\|\mathrm{C}h_\bullet\|^2_\mathbb{J}$

(ii) $\theta^\alpha_\bullet$ *is solution of the normal equation:* $\mathrm{A}^\star g_\bullet = (\mathrm{A}^\star\mathrm{A} + \alpha\mathrm{C}^\star\mathrm{C})\theta^\alpha_\bullet$.

§06|01.07 **Proof** of Lemma §06|01.06. Given in the lecture. □

§06|01.08 **Remark**. Observe that $\ker(\mathrm{A})\cap\ker(\mathrm{C}) = \ker(\mathrm{A}^\star\mathrm{A} + \alpha\mathrm{C}^\star\mathrm{C})$ which in turn implies, that the solution of the generalised Tikhonov functional, if it exists, is unique if and only if $\ker(\mathrm{A})\cap\ker(\mathrm{C}) = \{0_\bullet\}$. Recall that there exists a solution, for example, if $(\mathrm{A}^\star\mathrm{A} + \alpha\mathrm{C}^\star\mathrm{C})$ has a continuous inverse. □

§06|01.09 **Corollary**. *Given the Tikhonov regularisation* $\{A^{\alpha|-1}A^\star = (A^\star A + \alpha \mathrm{id}_{\mathbb{J}})^{-1}A^\star \in \mathbb{L}(\mathbb{J}) \colon \alpha \in (0,1)\}$ *as in Definition §06|01.01 for each* $g_\bullet \in \mathbb{J}$, $\theta_\bullet^\alpha := A^{\alpha|-1}A^\star g_\bullet \in \mathbb{J}$ *is the unique minimiser in* $\mathbb{J}$ *of the Tikhonov functional* $h_\bullet \mapsto \frac{1}{2}\|g_\bullet - Ah_\bullet\|_{\mathbb{J}}^2 + \frac{\alpha}{2}\|h_\bullet\|_{\mathbb{J}}^2$. □

§06|01.10 **Proof** of Corollary §06|01.09. Given in the lecture. □

§06|01.11 **Definition**. Given an operator $C \in \mathbb{L}(\mathbb{J})$ satisfying (gTR1) $\mathrm{ran}(C)$ is closed and (gTR2) there exists $c \in \mathbb{R}_{>0}$ such that for any $h_\bullet \in \ker(C)$ it holds $\|Ah_\bullet\|_{\mathbb{J}} \geqslant c\|h_\bullet\|_{\mathbb{J}}$, the collection

$$\left\{\mathrm{gTR}_\alpha := (A^\star A + \alpha C^\star C)^{-1}A^\star \in \mathbb{L}(\mathbb{J}) \colon \alpha \in (0,1)\right\}$$

is called *generalised Tikhonov regularisation* of $A^\dagger$. □

§06|01.12 **Remark**. Assumption (gTR1) and (gTR2) ensure together that the generalised Tikhonov regularisation is well-defined. More precisely, introduce inner products $\langle\cdot,\cdot\rangle_* := \langle A\cdot, A\cdot\rangle_{\mathbb{J}} + \langle C\cdot, C\cdot\rangle_{\mathbb{J}}$ and $\langle\cdot,\cdot\rangle_C := \langle\cdot,\cdot\rangle_{\mathbb{J}} + \langle C\cdot, C\cdot\rangle_{\mathbb{J}}$ on $\mathbb{J}$ with associated norms $\|\cdot\|_*$ and $\|\cdot\|_C$. Since $\mathbb{J}$ is complete with respect to both norms (due to (gTR1) and (gTR2)), it follows from **??** that $\|\cdot\|_*$ and $\|\cdot\|_C$ are equivalent (keeping in mind that $\|h_\bullet\|_*^2 \leqslant \max(\|A\|_{\mathbb{L}(\mathbb{J})}^2, 1)\|h_\bullet\|_C^2$). Consequently, there is $K > 0$ such that $\|h_\bullet\|_* \geqslant K\|a_\bullet\|_C$ and thus $\|Ah_\bullet\|_{\mathbb{J}}^2 + \|Ch_\bullet\|_{\mathbb{J}}^2 \geqslant K^2(\|h_\bullet\|_{\mathbb{J}}^2 + \|Ch_\bullet\|_{\mathbb{J}}^2)$. Exploiting the last inequality we obtain $\|A^\star Ah_\bullet + \alpha C^\star Ch_\bullet\|_{\mathbb{J}} \geqslant K^2\min(1,\alpha)\|h_\bullet\|_{\mathbb{J}}$ for any $h_\bullet \in \mathbb{J}$. In analogy to the arguments exploiting (06.01) in Remark §06|01.02, $A^\star A + \alpha C^\star C$ is injective with closed range and, thus it has a continuous inverse, i.e., $(A^\star A + \alpha C^\star C)^{-1} \in \mathbb{L}(\mathbb{J})$. Consequently, the generalised Tikhonov regularisation $\{\mathrm{gTR}_\alpha := (A^\star A + \alpha C^\star C)^{-1}A^\star \in \mathbb{L}(\mathbb{J}) \colon \alpha \in (0,1)\}$ is well-defined. Moreover, keeping in mind Lemma §06|01.06 $\theta_\bullet^\alpha := \mathrm{gTR}_\alpha g_\bullet \in \mathbb{J}$ is obviously a solution of the normal equation, and thus the unique minimiser of the generalised Tikhonov functional. □

§06|01.13 **Corollary**. *Consider the generalised Tikhonov regularisation as in Definition §06|01.11. For each* $g_\bullet \in \mathbb{J}$, $\theta_\bullet^\alpha := \mathrm{gTR}_\alpha g_\bullet = (A^\star A + \alpha C^\star C)^{-1}A^\star g_\bullet$ *is the unique minimiser in* $\mathbb{J}$ *of the generalised Tikhonov functional* $h_\bullet \mapsto \frac{1}{2}\|g_\bullet - Ah_\bullet\|_{\mathbb{J}}^2 + \frac{\alpha}{2}\|Ch_\bullet\|_{\mathbb{J}}^2$. □

§06|01.14 **Proof** of Corollary §06|01.13. Given in the lecture. □

§06|01.15 **Remark**. Introduce further the adjoint $A_*^\star$ and $C_*^\star$ of A and C, respectively, with respect to the inner product $\langle\cdot,\cdot\rangle_*$ introduced in Remark §06|01.12, i.e., $\langle Ah_\bullet, g_\bullet\rangle_{\mathbb{J}} = \langle h_\bullet, A_*^\star g_\bullet\rangle_*$ and $\langle Ch_\bullet, g_\bullet\rangle_{\mathbb{J}} = \langle h_\bullet, C_*^\star g_\bullet\rangle_*$ for all $h_\bullet, g_\bullet \in \mathbb{J}$. In particular, for each $g_\bullet, h_\bullet \in \mathbb{J}$ we have (a) $A_*^\star g_\bullet = (A^\star A + C^\star C)^{-1}A^\star g_\bullet$, (b) $C_*^\star g_\bullet = (A^\star A + C^\star C)^{-1}C^\star g_\bullet$ and (c) $(A_*^\star A + C_*^\star C)h_\bullet = h_\bullet$ (i.e., $A_*^\star A + C_*^\star C = \mathrm{id}_{\mathbb{J}}$). We note that $\ker(A_*^\star) = \ker(A^\star)$ and $\overline{\mathrm{ran}}(A_*^\star) = \ker(A)^{\perp_*}$ where $\ker(A)^{\perp_*}$ denotes the orthogonal complement of $\ker(A)$ in $(\mathbb{J}, \langle\cdot,\cdot\rangle_*)$. □

Consider the restriction of A as bijective map from $\ker(A)^{\perp_*}$ to $\mathrm{ran}(A)$ and denote its inverse by $A_*^{-1} : \mathrm{ran}(A) \to \ker(A)^{\perp_*}$. Given the orthogonal projection $\Pi_{\overline{\mathrm{ran}}(A)}$ onto $\overline{\mathrm{ran}}(A)$ its associated Moore-Penrose inverse $A_*^\dagger$ (see Definition §03|00.08) defined on $\mathrm{dom}(A_*^\dagger) = \mathrm{ran}(A) \oplus \mathrm{ran}(A)^\perp = \mathrm{dom}(A^\dagger)$ is given by $A_*^\dagger := A_*^{-1}\Pi_{\overline{\mathrm{ran}}(A)}$.

§06|01.16 **Proposition**. *Consider the generalised Tikhonov regularisation* $\{\mathrm{gTR}_\alpha \in \mathbb{L}(\mathbb{J}) \colon \alpha \in (0,1)\}$ *as in Definition §06|01.11. Under the conditions (gTR1) and (gTR2) of Definition §06|01.11 for* $g_\bullet \in \mathbb{J}$ *and* $\theta_\bullet^\alpha = \mathrm{gTR}_\alpha g_\bullet = (A^\star A + \alpha C^\star C)^{-1}A^\star g_\bullet \in \mathbb{J}$ *the following statements are equivalent:*

(i) $g_\bullet \in \mathrm{dom}(A_*^\dagger) = \mathrm{ran}(A) \oplus \mathrm{ran}(A)^\perp = \mathrm{dom}(A^\dagger)$;

(ii) *there is* $\theta_\bullet^* \in \mathbb{J}$ *such that* $\|\theta_\bullet^\alpha - \theta_\bullet^*\|_* = o(1)$ *as* $\alpha \to 0$.

Moreover, under the equivalent conditions we have $\theta_\bullet^* = A_*^\dagger g_\bullet$.

§06|01.17 **Proof** of Proposition §06|01.16. Given in the lecture. □

§06|01.18 **Remark**. Due to Proposition §06|01.16 the generalised Tikhonov regularisation as in Definition §06|01.11 is indeed a regularisation in the sense of Definition §06|00.04. Moreover, we shall emphasise that $\|\theta^\alpha_\bullet - \theta^*_\bullet\|_* = o(1)$ if and only if $\|A\theta^\alpha_\bullet - A\theta^*_\bullet\|_{\mathbb{J}} = o(1)$ and $\|C\theta^\alpha_\bullet - C\theta^*_\bullet\|_{\mathbb{J}} = o(1)$, which in turn implies $\|\theta^\alpha_\bullet - \theta^*_\bullet\|_{\mathbb{J}} = o(1)$. Keep further in mind that $A^\star_* g_\bullet = A^\star_* A\theta_\bullet$ holds if and only if $A^\star g_\bullet = A^\star A\theta_\bullet$ is true, since $A^\star A + C^\star C$ is continuously invertible. Thereby, for each $g_\bullet \in \mathrm{dom}(A^\dagger)$ the set of least squares solution $A^{-1}(\Pi_{\overline{\mathrm{ran}(A)}} g_\bullet)$ satisfies $A^{-1}(\Pi_{\overline{\mathrm{ran}(A)}} g_\bullet) = \{h_\bullet \in \mathbb{J} \colon A^\star A h_\bullet = A^\star g_\bullet\} = \{h_\bullet \in \mathbb{J} \colon A^\star_* A h_\bullet = A^\star_* g_\bullet\} = \{\theta^*_\bullet\} + \ker(A)$ with $\theta^*_\bullet = A^\dagger_* g_\bullet$. Each $\theta_\bullet \in A^{-1}(\Pi_{\overline{\mathrm{ran}(A)}} g_\bullet)$ can thus be written as $\theta_\bullet = \theta^*_\bullet + h_\bullet$ for some $h_\bullet \in \ker(A)$ with $\theta^*_\bullet \in \ker(A)^{\perp_*}$, and hence, $A\theta_\bullet = A\theta^*_\bullet$ and $\|\theta^*_\bullet\|^2_* \leqslant \|\theta^*_\bullet\|^2_* + \|h_\bullet\|^2_* = \|\theta_\bullet\|^2_*$, which together implies $\|C\theta^*_\bullet\|^2_{\mathbb{J}} \leqslant \|C\theta_\bullet\|^2_{\mathbb{J}}$ for any $\theta_\bullet \in A^{-1}(\Pi_{\overline{\mathrm{ran}(A)}} g_\bullet)$. In other words, $\theta^*_\bullet$ is the unique least squares solution with minimal $\|C\cdot\|_{\mathbb{J}}$-norm. □

§06|02 Spectral regularisation

§06|02.01 **Definition**. For $A \in \mathbb{L}(\mathbb{J})$ let $\{r_\alpha, \alpha \in (0,1)\}$ be a collection of real-valued Borel-measurable functions defined on $[0, \|A\|^2_{\mathbb{L}(\mathbb{J})}]$. The collection $\{R_\alpha := r_\alpha(A^\star A)A^\star \in \mathbb{L}(\mathbb{J}) \colon \alpha \in (0,1)\}$ of operators is called *spectral regularisation* of $A^\dagger : \mathbb{J} \supseteq \mathrm{dom}(A^\dagger) \to \mathbb{J}$ if

(sR1) for all $\alpha \in (0,1)$ there exists $C_\alpha \in \mathbb{R}_{\geqslant 0}$ such that $|r_\alpha(x)| \leqslant C_\alpha$ for all $x \in [0, \|A\|^2_{\mathbb{L}(\mathbb{J})}]$,

(sR2) for all $x \in (0, \|A\|^2_{\mathbb{L}(\mathbb{J})}]$ holds $|1 - x r_\alpha(x)| = o(1)$ as $\alpha \to 0$, and

(sR3) there is $K \in \mathbb{R}_{>0}$ such that $|x r_\alpha(x)| \leqslant K$ for all $x \in [0, \|A\|^2_{\mathbb{L}(\mathbb{J})}]$ and $\alpha \in (0,1)$. □

§06|02.02 **Proposition**. *For* $A \in \mathbb{L}(\mathbb{J})$ *a* spectral regularisation $\{R_\alpha := r_\alpha(A^\star A)A^\star \in \mathbb{L}(\mathbb{J}) \colon \alpha \in (0,1)\}$ *as in Definition §06|02.01 is a regularisation in the sense of Definition §06|00.04.*

§06|02.03 **Proof** of Proposition §06|02.02. Given in the lecture. □

§06|02.04 **Remark**. We shall emphasise that under (sR3) for any $g_\bullet \notin \mathrm{dom}(A^\dagger)$ it can be shown that $\|R_\alpha g_\bullet\|_{\mathbb{J}} = \|r_\alpha(A^\star A)A^\star g_\bullet\|_{\mathbb{J}} \to \infty$ as $\alpha \to 0$ (Engl et al. [2000], Theorem 4.1, p. 72). □

§06|02|01 Maximal global $\mathfrak{v}$-error

Given $A \in \mathbb{L}(\mathbb{J})$ and a spectral regularisation $\{R_\alpha := r_\alpha(A^\star A)A^\star \in \mathbb{L}(\mathbb{J}) \colon \alpha \in (0,1)\}$ of $A^\dagger : \mathbb{J} \supseteq \mathrm{dom}(A^\dagger) \to \mathbb{J}$ as in Definition §06|02.01 for $g_\bullet \in \mathrm{dom}(A^\dagger)$ and $\alpha \in (0,1)$ we shall measure globally the accuracy of the approximation $\theta^\alpha_\bullet := R_\alpha g_\bullet = r_\alpha(A^\star A)A^\star g_\bullet \in \mathbb{J}$ of $\theta_\bullet := A^\dagger g_\bullet \in \mathbb{J}$.

§06|02.05 **Source condition**. Given $A \in \mathbb{L}(\mathbb{J})$ and $g_\bullet \in \mathrm{dom}(A^\dagger)$, the solution $\theta_\bullet = A^\dagger g_\bullet \in \mathbb{J}$ satisfies a *source condition*, if there is $s \in \mathbb{R}_{>0}$ such that $\theta_\bullet \in \mathrm{ran}((A^\star A)^{s/2})$, i.e. $\theta_\bullet = (A^\star A)^{s/2} h_\bullet$ for $h_\bullet \in \mathbb{J}$. □

§06|02.06 **Proposition**. *Given* $A \in \mathbb{L}(\mathbb{J})$ *let* $\{R_\alpha := r_\alpha(A^\star A)A^\star \in \mathbb{L}(\mathbb{J}) \colon \alpha \in (0,1)\}$ *be a spectral regularisation of* $A^\dagger : \mathbb{J} \supseteq \mathrm{dom}(A^\dagger) \to \mathbb{J}$ *as in Definition §06|02.01. Assume Definition §06|02.01 (sR1), and (sR3), and in addition replace (sR2) by*

(sR2a) *there is* $s_\circ \in \mathbb{R}_{\geqslant 1}$ *such that for all* $s \in [0, s_\circ]$ *there is a constant* $C_s \in \mathbb{R}_{\geqslant 0}$ *satisfying*

$$\sup\{x^s|1 - x r_\alpha(x)| \colon x \in [0, \|A\|^2_{\mathbb{L}(\mathbb{J})}]\} \leqslant C_s \alpha^s \quad \forall \alpha \in (0,1).$$

For $g_\bullet \in \mathrm{dom}(A^\dagger)$ *and* $\alpha \in (0,1)$ *consider* $\theta^\alpha_\bullet = R_\alpha g_\bullet = r_\alpha(A^\star A)A^\star g_\bullet \in \mathbb{J}$ *and* $\theta_\bullet := A^\dagger g_\bullet \in \mathbb{J}$. *If there are* $s \in [0, 2s_\circ]$ *and* $h_\bullet \in \mathbb{J}$ *such that* $\theta_\bullet = (A^\star A)^{s/2} h_\bullet$ *(i.e.* $\theta_\bullet \in \mathrm{ran}((A^\star A)^{s/2})$ *satisfies a source condition as in Definition §06|02.05), then we have*

$$\|\theta^\alpha_\bullet - \theta_\bullet\|_{\mathbb{J}} \leqslant C_{s/2} \alpha^{s/2} \|h_\bullet\|_{\mathbb{J}} \quad \forall \alpha \in (0,1). \tag{06.03}$$

§06|02.07 **Proof** of Proposition §06|02.06. Given in the lecture. □

§06|02.08 **Link condition**. Given weights $\mathfrak{t}_\bullet \in \mathcal{M}_{>0}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$ an operator $\mathrm{A} \in \mathbb{L}^{\geqslant}(\mathbb{J})$ satisfies a *link condition* if there is $\mathrm{d} \in \mathbb{R}_{\geqslant 1}$ such that

$$\mathrm{A} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}} := \left\{ \mathrm{T} \in \mathbb{L}^{\geqslant}(\mathbb{J}) \colon \mathrm{d}^{-1}\|a_\bullet\|_{\mathfrak{t}} \leqslant \|\mathrm{T}a_\bullet\|_{\mathbb{J}} \leqslant \mathrm{d}\|a_\bullet\|_{\mathfrak{t}} \text{ for all } a_\bullet \in \mathbb{J} \right\}$$

and we set $\mathbb{T}_{\mathrm{t,d}} := \left\{ \mathrm{T} \in \mathbb{L}(\mathbb{J}) \colon (\mathrm{T}^\star\mathrm{T})^{1/2} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}} \right\}$. □

§06|02.09 **Property**. *If* $\mathrm{A} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}}$ *with* $\mathfrak{t}_\bullet \in \mathcal{M}_{>0}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$ *and* $\mathrm{d} \in \mathbb{R}_{\geqslant 1}$ *then for all* $s \in [-1, 1]$ *we have*

(inequality of Heinz [1951]) $\mathrm{d}^{-|s|}\|a_\bullet\|_{\mathfrak{t}^s} \leqslant \|\mathrm{A}^s a_\bullet\|_{\mathbb{J}} \leqslant \mathrm{d}^{|s|}\|a_\bullet\|_{\mathfrak{t}^s}$ *for all* $a_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{t}^s})$. □

§06|02.10 **Comment**. Given $\mathrm{A} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}}$ we have $\ker(\mathrm{A}) = \{0_\bullet\}$ and on $\mathrm{ran}(\mathrm{A})$ (which is dense in $\mathbb{J}$) we have $\mathrm{A}^{-1} = \mathrm{A}^\dagger$. Similarly, for each $s \in \mathbb{R}_{\geqslant 0}$ on $\mathrm{ran}(\mathrm{A}^s)$ we have $\mathrm{A}^{-s} = \mathrm{A}^{s|\dagger} = (\mathrm{A}^s)^\dagger$. □

§06|02.11 **Assumption**. Consider $\mathfrak{v}_\bullet \in \mathcal{M}_{>0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$, and for $\mathrm{t} \in \mathbb{R}_{>0}$, $\mathrm{a} \in (0, \mathrm{t}]$ set $\mathfrak{t}_\bullet := \mathfrak{v}^{\mathrm{t}}_\bullet$ and $\mathfrak{a}_\bullet := \mathfrak{v}^{\mathrm{a}}_\bullet$ where $\mathfrak{t}_\bullet, \mathfrak{a}_\bullet \in \mathcal{M}_{>0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$ and hence $\nu(\mathcal{N}_{\mathfrak{v}}) = \nu(\mathcal{N}_{\mathfrak{a}}) = \nu(\mathcal{N}_{\mathfrak{t}}) = 0$. □

§06|02.12 **Reminder**. Under Assumption §06|02.11 we have $\mathbb{J}^{\mathfrak{a}} = \mathbb{L}^{\mathfrak{a}}_2(\nu) = \mathrm{dom}(\mathrm{M}_{\mathfrak{a}^{-1}}) = \mathbb{J}\mathfrak{a}_\bullet = \mathbb{L}_2(\mathfrak{a}^{-2}_\bullet\nu)$ and the measures ν, $\mathfrak{v}^2_\bullet\nu$, $\mathfrak{t}^2_\bullet\nu$ and $\mathfrak{a}^{-2}_\bullet\nu$ dominate mutually each other (see Property §04|01.02). Consequently, $\mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{J} = \mathbb{L}_2(\nu)$ and $\mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{L}_2(\mathfrak{v}^2_\bullet\nu)$ (Property §04|02.11) since $(\mathfrak{a}\mathfrak{v})_\bullet = \mathfrak{v}^{1+\mathrm{a}}_\bullet \in \mathbb{L}_\infty(\nu)$. We assume in the following that $\theta_\bullet \in \mathbb{J}$ satisfies an abstract smoothness condition (Definition §04|02.12), i.e., there is $\mathrm{r} \in \mathbb{R}_{>0}$ such that $\theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}} = \{h_\bullet \in \mathbb{J}^{\mathfrak{a}} : \|h_\bullet\|_{\mathfrak{a}^{-1}} \leqslant \mathrm{r}\} \subseteq \mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{J}$. Under Assumption §06|02.11 by Corollary §05|01.14 (see Comment §05|01.16) if $\mathrm{A} \in \mathbb{T}_{\mathrm{t,d}}$ (or in equal $(\mathrm{A}^\star\mathrm{A})^{1/2} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}}$) then (i) for any $\theta_\bullet \in \mathbb{J}^{\mathfrak{a}}$ we have $\theta_\bullet = (\mathrm{A}^\star\mathrm{A})^{\mathrm{a}/(2\mathrm{t})}h_\bullet$ with $\|h_\bullet\|_{\mathbb{J}} \leqslant \mathrm{d}^{\mathrm{a}/\mathrm{t}}\|\theta_\bullet\|_{\mathfrak{a}^{-1}}$, and conversely (ii) for any $\theta_\bullet = (\mathrm{A}^\star\mathrm{A})^{\mathrm{a}/(2\mathrm{t})}h_\bullet$ with $h_\bullet \in \mathbb{L}_2(\nu)$ we obtain $\theta_\bullet \in \mathbb{J}^{\mathfrak{a}}$ with $\|\theta_\bullet\|_{\mathfrak{a}^{-1}} \leqslant \mathrm{d}^{\mathrm{a}/\mathrm{t}}\|h_\bullet\|_{\mathbb{J}}$. □

§06|02.13 **Corollary**. *Let Assumption §06|02.11 with* $(\mathfrak{t}\mathfrak{a})_\bullet = \mathfrak{v}^{\mathrm{t+a}}_\bullet \in \mathcal{M}_{>0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$ *and* $\mathrm{d}, \mathrm{r} \in \mathbb{R}_{>0}$ *be satisfied. If* $\mathrm{A} \in \mathbb{T}_{\mathrm{t,d}}$ *and* $\theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$*, then we have* $g_\bullet = \mathrm{A}\theta_\bullet \in \mathbb{J}^{(\mathfrak{t}\mathfrak{a}),\mathrm{dr}}$.

§06|02.14 **Proof** of Corollary §06|02.13. Given in the lecture. □

§06|02.15 **Proposition**. *Given* $\mathrm{A} \in \mathbb{L}(\mathbb{J})$ *let* $\left\{\mathrm{R}_\alpha := \mathrm{r}_\alpha(\mathrm{A}^\star\mathrm{A})\mathrm{A}^\star \in \mathbb{L}(\mathbb{J}) \colon \alpha \in (0,1)\right\}$ *be a spectral regularisation of* $\mathrm{A}^\dagger : \mathbb{J} \supseteq \mathrm{dom}(\mathrm{A}^\dagger) \to \mathbb{J}$ *as in Definition §06|02.01. Assume* (sR1), (sR2), *and* (sR3) *(Definition §06|02.01) and* (sR2a) *(Proposition §06|02.06). For* $g_\bullet \in \mathrm{dom}(\mathrm{A}^\dagger)$ *and* $\alpha \in (0, 1)$ *consider* $\theta^\alpha_\bullet = \mathrm{R}_\alpha g_\bullet = \mathrm{r}_\alpha(\mathrm{A}^\star\mathrm{A})\mathrm{A}^\star g_\bullet \in \mathbb{J}$ *and* $\theta_\bullet := \mathrm{A}^\dagger g_\bullet \in \mathbb{J}$*. Under Assumption §06|02.11 if* $\mathrm{T} \in \mathbb{T}_{\mathrm{t,d}}$ *(link condition as in Definition §06|02.08) and* $\theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ *(abstract smoothness condition as in Definition §04|02.12), then for any* $q \in [-\mathrm{a}, \mathrm{t}]$ *we have*

$$\|\theta^\alpha_\bullet - \theta_\bullet\|_{\mathfrak{v}^q} \leqslant \mathrm{C}_{(q+\mathrm{a})/(2\mathrm{t})}\, \mathrm{d}^{(\mathrm{a}+|q|)/\mathrm{t}}\, \mathrm{r}\, \alpha^{(\mathrm{a}+q)/(2\mathrm{t})}, \quad \forall \alpha \in (0, 1). \tag{06.04}$$

§06|02.16 **Proof** of Proof §06|02.16. Given in the lecture. □

§06|02.17 **Remark**. Let us briefly comment on the Assumption §06|02.11 imposed in Proposition §06|02.15. We set $\theta^0_\bullet := \theta_\bullet$ and write $\left\{\theta^\alpha_\bullet \colon \alpha \in [0,1)\right\} = \{\theta_\bullet\} \cup \left\{\theta^\alpha_\bullet = \mathrm{R}_\alpha\mathrm{A}\theta_\bullet = \mathrm{r}_\alpha(\mathrm{A}^\star\mathrm{A})\mathrm{A}^\star\mathrm{A}\theta_\bullet \colon \alpha \in (0,1)\right\}$, shortly. Note that, under $q \geqslant -\mathrm{a}$ the *global* $\mathfrak{v}^q$*-error* is well-defined on $\mathbb{J}^{\mathfrak{a}}$ since $\left\{\theta^\alpha_\bullet \colon \alpha \in [0,1)\right\} \subseteq \mathbb{L}_2(\mathfrak{v}^{2q}\nu)$ for all $\theta_\bullet \in \mathbb{J}^{\mathfrak{a}}$. Moreover, the additional condition $q \leqslant t$ together with $a \leqslant t$ allows us to apply the inequality of Heinz [1951] Property §06|02.09. We can dismiss those upper bounds, if A and $\mathrm{M}_{\mathfrak{v}}$ commute. However, if A and $\mathrm{M}_{\mathfrak{v}}$ do not commute, then the smallest upper bound of the global approximation bias is up to a constant α since $(a + q)/(2t) \in [0, 1]$. □

§06|02.18 **Example**. Let us discuss certain spectral regularisations satisfying (sR1), (sR2a) and (sR3).

(a) *Tikhonov regularisation* as defined in §06|01.01 is given by $x \mapsto \mathrm{r}_\alpha(x) = (x+\alpha)^{-1}$ and satisfies (sR1) and (sR3) with $\mathrm{C}_\alpha = \alpha^{-1}$ and $\mathrm{K} = 1$, and (sR2a) with $\mathrm{s}_\circ = 1$ and $\mathrm{C}_s = s^s(1-s)^{1-s}$.

(b) *Spectral cut-off* given by the piecewise continuous function $x \mapsto \mathrm{r}_\alpha(x) = \frac{1}{x}\mathbb{1}_{[\alpha,\infty)}(x)$ satisfies (sR1) and (sR3) with $\mathrm{C}_\alpha = \alpha^{-1}$ and $\mathrm{K} = 1$, and (sR2a) with $\mathrm{s}_\circ = \infty$ and $\mathrm{C}_s = 1$.

(c) A special iterative regularisation is the *Landweber iteration*. This method is based on a transformation of the normal equation into an equivalent fixed point equation $\theta_\circ = \theta_\circ + \omega \mathrm{A}^\star(g_\circ - \mathrm{A}\theta_\circ)$ with $\omega \in (0, \|\mathrm{A}\|_{\mathbb{L}(\mathbb{J})}^{-2}]$. Then the corresponding fixed point operator $\mathrm{id}_{\mathbb{J}} - \omega\mathrm{A}^\star\mathrm{A}$ is non-expansive and $\theta_\circ$ may be approximated by $\theta_\circ^m$ determined by $\theta_\circ^j := \theta_\circ^{j-1} + \omega\mathrm{A}^\star(g_\circ - \mathrm{A}\theta_\circ^{j-1})$, $j \in [\![m]\!]$, and $\theta_\circ^0 := 0$. Note, that without loss of generality, we can assume $\|\mathrm{A}\|_{\mathbb{L}(\mathbb{J})} \leqslant 1$ and drop the parameter ω. By induction the iterate $\theta_\circ^m$ can be expressed non-recursively through $\theta_\circ^m = \sum_{j\in[\![m]\!]}(\mathrm{id}_{\mathbb{J}} - \mathrm{A}^\star\mathrm{A})^{j-1}\mathrm{A}^\star g_\circ$ and thus $x \mapsto \mathrm{r}_{1/m}(x) = \sum_{j\in[\![m]\!]}(1-x)^{j-1}$ where $1 - x\mathrm{r}_{1/m}(x) = (1-x)^m$. Under the assumption $\|\mathrm{A}\|_{\mathbb{L}(\mathbb{J})} \leqslant 1$, the Landweber iteration is thus a spectral regularisation with $\alpha = 1/m$ satisfying (sR1) and (sR3) with $\mathrm{C}_\alpha = \alpha^{-1}$ and $\mathrm{K} = 1$. Moreover, (sR2a) holds with $\mathrm{s}_\circ = \infty$ and $\mathrm{C}_s = s^s e^{-s}$. □

§06|02.19 **Notation**. Given $\mathrm{A} \in \mathbb{L}^{>}(\mathbb{J})$, i.e., A is positive definite, we eventually consider a spectral regularisation $\{\mathrm{R}_\alpha := \mathrm{r}_\alpha(\mathrm{A}) \in \mathbb{L}(\mathbb{J}) : \alpha \in (0,1)\}$ of $\mathrm{A}^\dagger$ for a given collection $\{\mathrm{r}_\alpha : \alpha \in (0,1)\}$ of real-valued Borel-measurable functions defined on $[0, \|\mathrm{A}\|_{\mathbb{L}(\mathbb{J})}]$ satisfying

(sR1') for all $\alpha \in (0,1)$ there exists $\mathrm{C}_\alpha \in \mathbb{R}_{\geqslant 0}$ such that $|\mathrm{r}_\alpha(x)| \leqslant \mathrm{C}_\alpha$ for all $x \in [0, \|\mathrm{A}\|_{\mathbb{L}(\mathbb{J})}]$,

(sR2'a) there are $\mathrm{s}_\circ \in [1,\infty)$ and $\mathrm{C}_s \in \mathbb{R}_{\geqslant 0}$ for all $s \in [0, \mathrm{s}_\circ]$ such that $x^s|1 - x\mathrm{r}_\alpha(x)| \leqslant \mathrm{C}_s\alpha^s$ for all $x \in [0, \|\mathrm{A}\|_{\mathbb{L}(\mathbb{J})}]$ and $\alpha \in (0,1)$,

(sR3') there is $\mathrm{K} \in \mathbb{R}_{>0}$ such that $|x\mathrm{r}_\alpha(x)| \leqslant \mathrm{K}$ for all $x \in [0, \|\mathrm{A}\|_{\mathbb{L}(\mathbb{J})}]$ and $\alpha \in (0,1)$.

We shall measure in the sequel the accuracy of the approximation $\theta_\circ^\alpha = \mathrm{R}_\alpha g_\circ = \mathrm{r}_\alpha(\mathrm{A})g_\circ \in \mathbb{J}$ of $\theta_\circ := \mathrm{A}^\dagger g_\circ \in \mathbb{J}$ for $g_\circ \in \mathrm{dom}(\mathrm{A}^\dagger)$, by its global approximation error. For convenient notation we eventually use the notation $\theta_\circ^0 := \theta_\circ$ and write and write $\{\theta_\circ^\alpha : \alpha \in [0,1)\} = \{\theta_\circ\} \cup \{\theta_\circ^\alpha = \mathrm{R}_\alpha\mathrm{A}\theta_\circ = \mathrm{r}_\alpha(\mathrm{A})\mathrm{A}\theta_\circ : \alpha \in (0,1)\}$. □

§06|02.20 **Proposition**. *Given* $\mathrm{A} \in \mathbb{L}^{>}(\mathbb{J})$ *let* $\{\mathrm{R}_\alpha = \mathrm{r}_\alpha(\mathrm{A}) \in \mathbb{L}^{>}(\mathbb{J}) : \alpha \in (0,1)\}$ *be a spectral regularisation of* $\mathrm{A}^\dagger$ *satisfying* (sR1'), (sR2'a) *and* (sR3') *in Notation* §06|02.19. *For* $g_\circ \in \mathrm{dom}(\mathrm{A}^\dagger)$ *and* $\alpha \in (0,1)$ *consider* $\theta_\circ^\alpha = \mathrm{R}_\alpha g_\circ = \mathrm{r}_\alpha(\mathrm{A})g_\circ \in \mathbb{J}$ *and* $\theta_\circ := \mathrm{A}^\dagger g_\circ \in \mathbb{J}$.

(i) *If there are* $s \in [0, \mathrm{s}_\circ]$ *and* $h_\circ \in \mathbb{J}$ *such that* $\theta_\circ = \mathrm{A}^s h_\circ$ *(i.e.* $\theta_\circ \in \mathrm{ran}(\mathrm{A}^s)$ *satisfies a source condition as in Definition* §06|02.05*), then we have*

$$\|\theta_\circ^\alpha - \theta_\circ\|_{\mathbb{J}} \leqslant \mathrm{C}_s\alpha^s\|h_\circ\|_{\mathbb{J}} \quad \forall \alpha \in (0,1). \tag{06.05}$$

(ii) *Under Assumption* §06|02.11 *if* $\mathrm{T} \in \mathbb{T}_{\mathrm{t,d}}^{\gtrless}$ *(link condition as in Definition* §06|02.08*) and* $\theta_\circ \in \mathbb{J}^{\mathrm{a,r}}$ *(abstract smoothness condition as in Definition* §04|02.12*), then for any* $q \in [-\mathrm{a}, \mathrm{t} \wedge (\mathrm{t}\mathrm{s}_\circ - \mathrm{a})]$ *we have*

$$\|\theta_\circ^\alpha - \theta_\circ\|_{\mathfrak{v}^q} \leqslant \mathrm{C}_{(q+\mathrm{a})/\mathrm{t}}\ \mathrm{d}^{(\mathrm{a}+|q|)/\mathrm{t}}\ \mathrm{r}\ \alpha^{(\mathrm{a}+q)/\mathrm{t}} \quad \forall \alpha \in (0,1). \tag{06.06}$$

§06|02.21 **Proof** of Proposition §06|02.20. Given in the lecture. □

§06|02.22 **Remark**. If (sR2'a) is satisfied for some $\mathrm{s}_\circ \geqslant 2$ (excluding the Tikhonov regularisation as discussed in Example §06|02.18 (a)) then (06.06) in Proposition §06|02.20 (ii) holds for any $q \in [-\mathrm{a}, \mathrm{t}]$ as in Proposition §06|02.15. □

§06|02|02 Maximal local ϕ-error

Given $A \in \mathbb{L}(\mathbb{J})$ and a spectral regularisation $\{R_\alpha := r_\alpha(A^\star A)A^\star \in \mathbb{L}(\mathbb{J}) \colon \alpha \in (0,1)\}$ of $A^\dagger : \mathbb{J} \supseteq \mathrm{dom}(A^\dagger) \to \mathbb{J}$ as in Definition §06|02.01 for $g_\bullet \in \mathrm{dom}(A^\dagger)$ and $\alpha \in (0,1)$ we shall measure locally the accuracy of the approximation $\theta_\bullet^\alpha := R_\alpha g_\bullet = r_\alpha(A^\star A)A^\star g_\bullet \in \mathbb{J}$ of $\theta_\bullet := A^\dagger g_\bullet \in \mathbb{J}$.

§06|02.23 **Reminder**. For $\phi \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ and $\mathrm{dom}(\phi\nu) := \{h_\bullet \in \mathbb{J} = \mathbb{L}_2(\nu) \colon \phi h_\bullet \in \mathbb{L}_1(\nu)\}$ we consider the linear functional $\phi\nu : \mathbb{J} \supseteq \mathrm{dom}(\phi\nu) \to \mathbb{R}$ given by $h_\bullet \mapsto \phi\nu(h_\bullet) := \nu(\phi h_\bullet)$ with a slight abuse of notations. Under Assumption §06|02.11 we have $\mathbb{J}^{\mathfrak{a}} = \mathbb{L}_2^{\mathfrak{a}}(\nu) = \mathrm{dom}(M_{\mathfrak{a}^{-1}}) = \mathbb{J}\mathfrak{a}_\bullet = \mathbb{L}_2(\mathfrak{a}_\bullet^{-2}\nu)$ and the measures ν, $\mathfrak{v}_\bullet^2\nu$, $\phi_\bullet^2\nu$, $\mathfrak{t}_\bullet^2\nu$ and $\mathfrak{a}_\bullet^{-2}\nu$ dominate mutually each other (see Property §04|01.02). Consequently, $\mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{J} = \mathbb{L}_2(\nu)$ and $\mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$ (Property §04|02.11) since $(\mathfrak{a}\mathfrak{v})_\bullet = \mathfrak{v}_\bullet^{1+\mathrm{a}} \in \mathbb{L}_\infty(\nu)$. We assume in the following that $\theta_\bullet \in \mathbb{J}^{\mathrm{a,r}}$ and $A \in \mathbb{T}_{\mathrm{t,d}}$ satisfies, respectively, an abstract smoothness condition (Definition §04|02.12) and link condition (Definition §06|02.08). Under Assumption §06|02.11 due to Proposition §06|02.15 we have $\theta_\bullet^\alpha - \theta_\bullet \in \mathbb{L}_2(\mathfrak{v}_\bullet^{2q}\nu)$, and thus if in addition $\mathfrak{v}_\bullet^{-q} \in \mathbb{L}_2(\phi_\bullet^2\nu)$ also $\theta_\bullet^\alpha - \theta_\bullet \in \mathrm{dom}(\phi\nu)$. □

§06|02.24 **Proposition**. *Given $A \in \mathbb{L}(\mathbb{J})$ let $\{R_\alpha := r_\alpha(A^\star A)A^\star \in \mathbb{L}(\mathbb{J}) \colon \alpha \in (0,1)\}$ be a spectral regularisation of $A^\dagger : \mathbb{J} \supseteq \mathrm{dom}(A^\dagger) \to \mathbb{J}$ as in Definition §06|02.01. Assume (sR1), (sR3) (Definition §06|02.01) and (sR2a) (Proposition §06|02.06). For $g_\bullet \in \mathrm{dom}(A^\dagger)$ and $\alpha \in (0,1)$ consider $\theta_\bullet^\alpha = R_\alpha g_\bullet = r_\alpha(A^\star A)A^\star g_\bullet \in \mathbb{J}$ and $\theta_\bullet := A^\dagger g_\bullet \in \mathbb{J}$. Under Assumption §06|02.11 if $A \in \mathbb{T}_{\mathrm{t,d}}$ (link condition) and $\theta_\bullet \in \mathbb{J}^{\mathrm{a,r}}$ (abstract smoothness condition), then for any $q \in [-\mathrm{a}, \mathrm{t}]$ such that $\mathfrak{v}_\bullet^{-q} \in \mathbb{L}_2(\phi_\bullet^2\nu)$ with $\phi \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ we have*

$$|\phi\nu(\theta_\bullet^\alpha - \theta_\bullet)| \leqslant C_{(q+\mathrm{a})/(2\mathrm{t})}\, \mathrm{d}^{(\mathrm{a}+|q|)/\mathrm{t}}\, \mathrm{r}\, \|\mathfrak{v}_\bullet^{-q}\|_\phi\, \alpha^{(\mathrm{a}+q)/(2\mathrm{t})}, \quad \forall \alpha \in (0,1). \tag{06.07}$$

§06|02.25 **Proof** of Proposition §06|02.24. Given in the lecture. □

§06|02.26 **Proposition**. *Given $A \in \mathbb{L}^{\geqslant}(\mathbb{J})$ let $\{R_\alpha = r_\alpha(A) \in \mathbb{L}^{\geqslant}(\mathbb{J}) \colon \alpha \in (0,1)\}$ be a spectral regularisation of $A^\dagger$ satisfying (sR1'), (sR2'a) and (sR3') in Notation §06|02.19. For $g_\bullet \in \mathrm{dom}(A^\dagger)$ and $\alpha \in (0,1)$ consider $\theta_\bullet^\alpha = R_\alpha g_\bullet = r_\alpha(A)g_\bullet \in \mathbb{J}$ and $\theta_\bullet := A^\dagger g_\bullet \in \mathbb{J}$. Under Assumption §06|02.11 if $T \in \mathbb{T}_{\mathrm{t,d}}^{\geqslant}$ (link condition) and $\theta_\bullet \in \mathbb{J}^{\mathrm{a,r}}$ (abstract smoothness condition), then for any $q \in [-\mathrm{a}, \mathrm{t} \wedge (\mathrm{t}s_\circ - \mathrm{a})]$ such that $\mathfrak{v}_\bullet^{-q} \in \mathbb{L}_2(\phi_\bullet^2\nu)$ with $\phi \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ we have*

$$|\phi\nu(\theta_\bullet^\alpha - \theta_\bullet)| \leqslant C_{(q+\mathrm{a})/\mathrm{t}}\, \mathrm{d}^{(\mathrm{a}+|q|)/\mathrm{t}}\, \mathrm{r}\, \|\mathfrak{v}_\bullet^{-q}\|_\phi\, \alpha^{(\mathrm{a}+q)/\mathrm{t}} \quad \forall \alpha \in (0,1). \tag{06.08}$$

§06|02.27 **Proof** of Proposition §06|02.26. Given in the lecture. □

§06|02.28 **Remark**. If (sR2'a) is satisfied for some $s_\circ \geqslant 2$ (excluding the Tikhonov regularisation as discussed in Example §06|02.18 (a)) then Proposition §06|02.26 holds for any $q \in [-\mathrm{a}, \mathrm{t}]$ as in Proposition §06|02.24. □

Chapter 3

Regularised estimation

Making use of the regularisation approaches presented in Chapter 2 we introduce estimators of the solution $\theta \in \mathbb{H}$ based on a noisy observation of the image $g = \mathrm{T}\theta$ and eventually in addition of the operator T.

Overview

§07 Orthogonal projection estimator

§07|00.01 **Notation** (§04|00.01 continued). Consider the measure space $(\mathcal{J}, \mathscr{J}, \nu)$ and the Hilbert space $\mathbb{J} = \mathbb{L}_2(\nu)$ as in Notation §01|01.01. For $\mathrm{w}_\bullet \in \mathbb{R}^{\mathcal{J}}$ define the multiplication map $\mathrm{M}_{\mathrm{w}} : \mathbb{R}^{\mathcal{J}} \to \mathbb{R}^{\mathcal{J}}$ with $a_\bullet \mapsto \mathrm{M}_{\mathrm{w}} a_\bullet := \mathrm{w}_\bullet a_\bullet := (\mathrm{w}_j a_j)_{j \in \mathcal{J}}$. If $\mathrm{w}_\bullet \in \mathcal{M}(\mathscr{J})$, i.e. $\mathrm{w}_\bullet$ is $\mathscr{J}$-$\mathscr{B}$-measurable, then we have $\mathrm{M}_{\mathrm{w}} : \mathcal{M}(\mathscr{J}) \to \mathcal{M}(\mathscr{J})$ too. If in addition $\mathrm{w}_\bullet \in \mathbb{L}_\infty(\nu)$ then we have also $\mathrm{M}_{\mathrm{w}} \in \mathbb{L}(\mathbb{J})$ identifying again equivalence classes and representatives. We set $\mathbb{L}^{\mathrm{M}}(\mathbb{J}) := \{\mathrm{M}_{\mathrm{w}} \colon \mathrm{w}_\bullet \in \mathbb{L}_\infty(\nu)\} \subseteq \mathbb{L}(\mathbb{J})$ noting that $\|\mathrm{M}_{\mathrm{w}}\|_{\mathbb{L}(\mathbb{J})} = \sup\{\|\mathrm{w}_\bullet a_\bullet\|_{\mathbb{J}} \colon \|a_\bullet\|_{\mathbb{J}} \leqslant 1\} \leqslant \|\mathrm{w}_\bullet\|_{\mathbb{L}_\infty(\nu)}$ for each $\mathrm{M}_{\mathrm{w}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$ (see Notation §01|04.01). Finally, given surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \mathbb{J})$ we define $\mathbb{L}^{\mathrm{U,V}}(\mathbb{L}^{\mathrm{M}}(\mathbb{J})) := \mathrm{V}^\star(\mathbb{L}^{\mathrm{M}}(\mathbb{J}))\mathrm{U} := \{\mathrm{V}^\star \mathrm{M}_{\mathrm{w}} \mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{G}) \colon \mathrm{M}_{\mathrm{w}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})\}$. As a consequence, for each $\mathrm{T} \in \mathbb{L}^{\mathrm{U,V}}(\mathbb{L}^{\mathrm{M}}(\mathbb{J}))$ we have $\mathrm{V}\mathrm{T}\mathrm{U}^\star = \mathrm{M}_{\mathrm{w}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$ for some $\mathrm{w}_\bullet \in \mathbb{L}_\infty(\nu)$. □

§07|00.02 **Assumption**. The separable Hilbert space $\mathbb{J} = \mathbb{L}_2(\mathcal{J}, \mathscr{J}, \nu)$ with σ-algebra $\mathscr{J}$ over $\mathcal{J}$ containing all elementary events $\{j\}$, $j \in \mathcal{J}$, and all events $[\![m]\!] := [-m, m] \cap \mathcal{J}$, $m \in \mathbb{N}$, and with σ-finite measure $\nu \in \mathscr{M}_\sigma(\mathscr{J})$ such that $\nu([\![m]\!]) \in \mathbb{R}_{\geqslant 0}$, for all $m \in \mathbb{N}$, and the surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \mathbb{J})$ are fixed and presumed to be *known* in advance. The operator satisfies $\mathrm{T} \in \mathbb{L}^{\mathrm{U,V}}(\mathbb{L}^{\mathrm{M}}(\mathbb{J})) \subseteq \mathbb{L}(\mathbb{H}, \mathbb{G})$ and hence $\mathrm{VTU}^\star = \mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$ for some $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$ and the image fulfils $g_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger})$, and hence $\mathfrak{s}^\dagger_{\bullet} g_{\bullet} \in \mathbb{J} = \mathbb{L}_2(\nu)$. □

§07|00.03 **Reminder**. Under Assumption §07|00.02 we consider $\mathrm{T} \in \mathbb{L}^{\mathrm{U,V}}(\mathbb{L}^{\mathrm{M}}(\mathbb{J})) \subseteq \mathbb{L}(\mathbb{H}, \mathbb{G})$, and hence $\mathrm{VTU}^\star = \mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$ and $g_{\bullet} = \mathrm{M}_{\mathfrak{s}}\theta_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathbb{J}$ for some $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$. Due to Property §04|01.02 the Moore-Penrose inverse of $\mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$ satisfies $\mathrm{M}_{\mathfrak{s}}^\dagger = \mathrm{M}_{\mathfrak{s}^\dagger} : \mathbb{J} \supseteq \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \to \mathbb{J}$ with $\mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) = \mathbb{J}\mathfrak{s}_{\bullet} \oplus \mathbb{J}\mathbb{1}^{\mathcal{N}_{\mathfrak{s}}}_{\bullet} = \mathbb{J}^{\mathfrak{s}}$. For each $m \in \mathbb{N}$, $\mathrm{M}_{\mathbb{1}^m} \in \mathbb{L}^{\mathrm{P}}(\mathbb{J})$ and $\mathrm{M}_{\mathbb{1}^{m|\perp}} \in \mathbb{L}^{\mathrm{P}}(\mathbb{J})$ is the *orthogonal projection* onto the linear subspace $\mathbb{J}\mathbb{1}^m_{\bullet} \subseteq \mathbb{J}$ and its orthogonal complement $\mathbb{J}\mathbb{1}^{m|\perp}_{\bullet} = (\mathbb{J}\mathbb{1}^m_{\bullet})^\perp \subseteq \mathbb{J}$, respectively, that is $\mathbb{J} = \mathbb{J}\mathbb{1}^m_{\bullet} \oplus \mathbb{J}\mathbb{1}^{m|\perp}_{\bullet}$ (see Property §04|02.02). Given $g_{\bullet} \in \mathbb{J}$ we call $\theta_{\bullet} \in \mathbb{J}$ satisfying $\|g_{\bullet} - \mathfrak{s}_{\bullet}\theta_{\bullet}\|_{\mathbb{J}} = \inf\{\|g_{\bullet} - \mathfrak{s}_{\bullet}h_{\bullet}\|_{\mathbb{J}} : h_{\bullet} \in \mathbb{J}\}$ a least squares solution, if it exists (see Property §03|00.05). Writing $\mathfrak{s}^\dagger_{\bullet} = \mathfrak{s}^{-1}_{\bullet}\mathbb{1}^{\mathcal{N}^c_{\mathfrak{s}}}_{\bullet}$ and $\mathcal{N}_{\mathfrak{s}} = \{j \in \mathbb{N} : \mathfrak{s}_j \in \mathbb{R}_{\setminus 0}\}$ for each $g_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) = \mathbb{J}\mathfrak{s}_{\bullet} \oplus \mathbb{J}\mathbb{1}^{\mathcal{N}_{\mathfrak{s}}}_{\bullet}$ is $\theta_{\bullet} = \mathrm{M}_{\mathfrak{s}^\dagger} g_{\bullet} = \mathfrak{s}^\dagger_{\bullet} g_{\bullet}$ the unique least square solution with minimal $\|\cdot\|_{\mathbb{J}}$-norm in the set $\mathfrak{s}^\dagger_{\bullet} g_{\bullet} + \mathbb{J}\mathbb{1}^{\mathcal{N}_{\mathfrak{s}}}_{\bullet}$ of all least square solutions (Property §04|03.02). If in addition $\nu(\mathcal{N}_{\mathfrak{s}}) = 0$, i.e. $\mathrm{M}_{\mathfrak{s}}$ is injective, then $\theta_{\bullet} = \mathfrak{s}^\dagger_{\bullet} g_{\bullet}$ is the *unique* least square solution. Given $m \in \mathbb{N}$ for each $g_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger})$ we have $g_{\bullet}\mathbb{1}^m_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger})$ too. In particular, for $\theta_{\bullet} = \mathfrak{s}^\dagger_{\bullet} g_{\bullet}$ follows $\theta_{\bullet}\mathbb{1}^m_{\bullet} = (\mathfrak{s}^\dagger_{\bullet} g_{\bullet})\mathbb{1}^m_{\bullet} = \mathfrak{s}^\dagger_{\bullet}(g_{\bullet}\mathbb{1}^m_{\bullet}) \in \mathbb{J}\mathbb{1}^m_{\bullet}$. □

§07|01 Diagonal statistical inverse problem

§07|01.01 **Assumption**. Consider a stochastic process $\dot{\varepsilon}_{\bullet} = (\dot{\varepsilon}_j)_{j \in \mathcal{J}}$ on a probability space $(\Omega, \mathscr{A}, \mathbb{P})$ satisfying Assumption §01|01.04 (i.e. $\dot{\varepsilon}_{\bullet} \in \mathcal{M}(\mathscr{A} \otimes \mathscr{J})$) with *mean zero* (i.e. $\mathbb{P}(\dot{\varepsilon}_{\bullet}) = (\mathbb{P}(\dot{\varepsilon}_j))_{j \in \mathcal{J}} = 0_{\bullet}$), a sample size $n \in \mathbb{N}$ and let Assumption §07|00.02 be satisfied where $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$ is *known* in advance. For $\theta_{\bullet} \in \mathbb{J}$ the observable noisy image with mean $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathbb{J} = \mathbb{L}_2(\nu)$ takes the form $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$. We denote by $\mathbb{P}^n_{\theta|\mathfrak{s}}$ the distribution of $\widehat{g}_{\bullet}$. □

§07|01.02 **Definition**. Under Assumption §07|01.01 for $\theta_{\bullet} \in \mathbb{J}$ and $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$ consider a noisy version $\widehat{g}_{\bullet} \sim \mathbb{P}^n_{\theta|\mathfrak{s}}$ of $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger})$. For each $m \in \mathbb{N}$ we call $\widehat{g}^m_{\bullet} := \widehat{g}_{\bullet}\mathbb{1}^m_{\bullet}$ and $\widehat{\theta}^m_{\bullet} := \mathfrak{s}^\dagger_{\bullet}\widehat{g}^m_{\bullet} = \mathfrak{s}^\dagger_{\bullet}\widehat{g}_{\bullet}\mathbb{1}^m_{\bullet}$ *orthogonal projection estimator (OPE)* of $g_{\bullet}$ and $\theta_{\bullet} = \mathfrak{s}^\dagger_{\bullet} g_{\bullet} \in \mathbb{J}$, respectively. □

§07|01|01 Examples

§07|01.03 **GdiSM (§01|04.09 continued)**. Considering $\ell_2 = \mathbb{L}_2(\mathbb{N}, 2^{\mathbb{N}}, \nu_{\mathbb{N}})$ for $\mathbb{J} = \ell_2$ let Assumption §07|00.02 be satisfied where $\mathrm{VTU}^\star = \mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathrm{M}}(\ell_2)$ for some $\mathfrak{s}_{\bullet} \in \ell_\infty = \mathbb{L}_\infty(\nu_{\mathbb{N}})$ is *known* in advance. We illustrate the OPE in a Gaussian diagonal inverse sequence model (GdiSM) as in §01|04.09. Here the observable stochastic process $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}^n_{\theta|\mathfrak{s}}$ is a noisy version of $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \ell_2$ with $\theta_{\bullet} = \mathfrak{s}^\dagger_{\bullet} g_{\bullet} \in \Theta \subseteq \ell_2$ and $\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$. Consequently, $\widehat{g}_{\bullet}$ admits a $\mathrm{N}^n_{\theta|\mathfrak{s}}$-distribution belonging to the family $\mathrm{N}^n_{\Theta\times\{\mathfrak{s}_{\bullet}\}} := (\mathrm{N}^n_{\theta|\mathfrak{s}})_{\theta_{\bullet}\in\Theta}$. Summarising the observations satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}^{\otimes\mathbb{N}}, \mathrm{N}^n_{\Theta\times\{\mathfrak{s}_{\bullet}\}})$ where $\Theta \subseteq \ell_2$. □

§07|01.04 **Property** (GdiSM §07|01.03 continued). *The error process $\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$ as in Model §07|01.03 admits a covariance operator $\mathrm{id}_{\ell_2} \in \mathbb{L}^{\mathrm{P}}(\ell_2)$ which is evidently invertible with inverse $\mathrm{id}_{\ell_2} \in \mathbb{L}(\ell_2)$ where $\|\mathrm{id}_{\ell_2}\|_{\mathbb{L}(\ell_2)} = 1$ and $\mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}(\dot{\mathrm{B}}^2_{\bullet}) = \mathbb{1}_{\bullet}$. For all $h_{\bullet} \in \ell_2$ we have $\|h_{\bullet}\|^2_{\ell_2} = \|h_{\bullet}\|^2_{\mathrm{id}_{\ell_2}} = \langle \mathrm{id}_{\ell_2} h_{\bullet}, h_{\bullet}\rangle_{\ell_2}$.* □

§07|01.05 **Property**. *For $\sigma^2_{\bullet} \in \mathbb{R}^{\mathbb{N}}_{>0} \cap \ell_\infty$ and $\mathrm{P}_{(0,\sigma^2)} \in \mathscr{W}_2(\mathscr{B})$, $j \in \mathbb{N}$, a stochastic process $\mathrm{Y}_{\bullet} \sim \otimes_{j\in\mathbb{N}}\mathrm{P}_{(\mu_j,\sigma^2_j)}$ of independent random variables admits $\mathrm{M}_{\sigma^2} \in \mathbb{L}^{\mathrm{M}}(\ell_2) \cap \mathbb{L}^{\mathrm{P}}(\ell_2)$ as covariance operator with $\|\mathrm{M}_{\sigma^2}\|_{\mathbb{L}(\ell_2)} =$*

$\|\sigma_\bullet^2\|_{\ell_\infty}$, *since*

$$\langle M_{\sigma^2} a_\bullet, b_\bullet \rangle_{\ell_2} = \sum_{j\in\mathbb{N}} \sigma_j^2 a_j b_j = \sum_{j\in\mathbb{N}} a_j \sum_{j_\circ\in\mathbb{N}} \mathbb{C}\mathrm{ov}(Y_j, Y_{j_\circ}) b_{j_\circ} \quad \forall a_\bullet, b_\bullet \in \ell_2.$$

If $\sigma_\bullet^2, \sigma_\bullet^{-2} \in \mathbb{R}_{>0}^{\mathbb{N}} \cap \ell_\infty$ *then* $M_{\sigma^2} \in \mathbb{L}^{\mathbb{M}}(\ell_2)$ *is invertible with inverse* $M_{\sigma^2}^{-1} = M_{\sigma^{-2}} \in \mathbb{L}^{\mathbb{M}}(\ell_2)$ *and* $\|M_{\sigma^2}^{-1}\|_{\mathbb{L}(\ell_2)} = \|\sigma_\bullet^{-2}\|_{\ell_\infty}$. □

§07|01.06 **diSM (§01|04.08 continued).** For $\mathbb{J} = \ell_2$ let Assumption §07|00.02 be satisfied where $\mathfrak{s}_\bullet \in \ell_\infty = \mathbb{L}_\infty(\nu_{\mathbb{N}})$ is *known* in advance. We illustrate the OPE in a Diagonal inverse sequence model (diSM) as in §01|04.08. Here the observable stochastic process $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ is a noisy version of $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \ell_2$ with $\theta_\bullet = \mathfrak{s}_\bullet^\dagger g_\bullet \in \Theta \subseteq \ell_2$ and $\dot{\varepsilon}_\bullet \sim \otimes_{j\in\mathbb{N}}\mathbb{P}^{\dot{\varepsilon}_j}$, where

(iSM1) for $\sigma_\bullet \in \Sigma \subseteq \mathbb{R}_{\geq 0}^{\mathbb{N}} \cap \ell_\infty$ and $\mathbb{P}^{\dot{\varepsilon}_j} = \mathrm{P}_{(0,\sigma_j^2)} \in \mathscr{W}_2(\mathscr{B})$ for all $j \in \mathbb{N}$,

(iSM2) $\Sigma \subseteq \mathbb{R}_{>0}^{\mathbb{N}} \cap \ell_\infty$ and for each $\sigma_\bullet \in \Sigma$ we have $\sigma_\bullet^{-1} \in \mathbb{R}_{>0}^{\mathbb{N}} \cap \ell_\infty$ too.

Under (iSM1) $\widehat{g}_\bullet$ admits a $\mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}$-distribution belonging to the family $\mathrm{P}^n_{\Theta\times\{\mathfrak{s}_\bullet\}\times\Sigma} := (\mathrm{P}^n_{\theta|\mathfrak{s}|\sigma})_{\theta_\bullet\in\Theta,\sigma_\bullet\in\Sigma}$. Summarising the observations satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}^{\otimes\mathbb{N}}, \mathrm{P}^n_{\Theta\times\{\mathfrak{s}_\bullet\}\times\Sigma})$ where $\Theta \subseteq \ell_2$ and $\Sigma \subseteq \mathbb{R}_{>0}^{\mathbb{N}} \cap \ell_\infty$. □

§07|01.07 **Property** (diSM §07|01.06 continued).

(i) *Under* (iSM1) *the error process* $\dot{\varepsilon}_\bullet \sim \otimes_{j\in\mathbb{N}}\mathrm{P}_{(0,\sigma_j^2)}$ *admits a covariance operator* $M_{\sigma^2} \in \mathbb{L}^{\mathbb{M}}(\ell_2) \cap \mathbb{L}^{\geq}(\ell_2)$, *i.e.* $\dot{\varepsilon}_\bullet \sim \mathrm{P}_{(0,M_{\sigma^2})}$, *satisfying* $\|M_{\sigma^2}\|_{\mathbb{L}(\ell_2)} = \|\sigma_\bullet^2\|_{\ell_\infty}$ *(Property §07|01.05) and* $\mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}(\dot{\varepsilon}_\bullet^2) = \sigma_\bullet^2$.

(ii) *Under* (iSM1) *and* (iSM2) *the covariance operator* $M_{\sigma^2} \in \mathbb{L}^{\mathbb{M}}(\ell_2)\cap\mathbb{L}^{\geq}(\ell_2)$ *is invertible with inverse* $M_{\sigma^{-2}} \in \mathbb{L}^{\mathbb{M}}(\ell_2) \cap \mathbb{L}(\ell_2)$ *satisfying* $\|M_{\sigma^{-2}}\|_{\mathbb{L}(\ell_2)} = \|\sigma_\bullet^{-2}\|_{\ell_\infty}$.

Under (iSM1) *and* (iSM2) *setting* $\mathrm{v}_\sigma := \max(\|\sigma_\bullet^{-2}\|_{\ell_\infty}, \|\sigma_\bullet^2\|_{\ell_\infty})$ *we evidently have* $\|M_{\sigma^2}\|_{\mathbb{L}(\ell_2)} \leqslant \mathrm{v}_\sigma$ *and* $\|M_{\sigma^{-2}}\|_{\mathbb{L}(\ell_2)} \leqslant \mathrm{v}_\sigma$. *Consequently, from Lemma §01|01.08* (01.03) *we obtain*

$$\mathrm{v}_\sigma^{-1}\|h_\bullet\|_{\ell_2}^2 \leqslant \|h_\bullet\|_{M_{\sigma^2}}^2 = \langle M_{\sigma^2} h_\bullet, h_\bullet\rangle_{\ell_2} \leqslant \mathrm{v}_\sigma\|h_\bullet\|_{\ell_2}^2 \quad \forall h_\bullet \in \ell_2.$$ □

§07|01.08 **dieMM (§01|04.07 continued).** For $\mathbb{J} = \mathbb{L}_2(\nu)$ let Assumption §07|00.02 be satisfied where $\mathfrak{s}_\bullet \in \mathbb{L}_\infty(\nu)$ is *known* in advance. We illustrate the OPE in a Diagonal inverse empirical mean model (dieMM) as in §01|04.07. Here the observable stochastic process $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ is a noisy version of $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathbb{J}$ with $\theta_\bullet = \mathfrak{s}_\bullet^\dagger g_\bullet \in \Theta \subseteq \mathbb{J}$, and error process $\dot{\varepsilon}_\bullet = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_\bullet) - \mathbb{P}_{\theta|\mathfrak{s}}(\psi_\bullet)) \in \mathcal{M}(\mathscr{Z}^{\otimes n} \otimes \mathscr{J})$ satisfying Assumption §01|01.04. More precisely, on a measurable space $(\mathcal{Z}, \mathscr{Z})$ for each $\theta_\bullet \in \Theta \subseteq \mathbb{J}$ there is a probability measure $\mathbb{P}_{\theta|\mathfrak{s}} \in \mathscr{W}(\mathscr{Z})$. Consider a stochastic process $\psi_\bullet = (\psi_j)_{j\in\mathcal{J}} \in \mathcal{M}(\mathscr{Z}\otimes\mathscr{J})$ which in addition for $\mathfrak{s}_\bullet \in \mathbb{L}_\infty(\nu)$ and for each $\theta_\bullet \in \Theta \subseteq \mathbb{J}$ satisfies

(dieMM1) $\psi_j \in \mathcal{L}_1(\mathbb{P}_{\theta|\mathfrak{s}}) := \mathcal{L}_1(\mathcal{Z}, \mathscr{Z}, \mathbb{P}_{\theta|\mathfrak{s}})$ ν-a.e. $j \in \mathcal{J}$ and $\mathbb{P}_{\theta|\mathfrak{s}}(\psi_\bullet) = \mathfrak{s}_\bullet\theta_\bullet = g_\bullet$ ν-a.s.,

(dieMM2) $(\psi_\bullet - \mathbb{P}_{\theta|\mathfrak{s}}(\psi_\bullet))\mathbb{1}^m_\bullet \in \mathbb{L}_\infty(\nu)$ $\mathbb{P}_{\theta|\mathfrak{s}}$-a.s. for each $m \in \mathbb{N}$,

(dieMM3) there is $\mathrm{v}_{\theta|\mathfrak{s}|\psi} \in \mathbb{R}_{\geqslant 1}$ such that $\|\mathbb{P}_{\theta|\mathfrak{s}}(\psi_\bullet^2)\|_{\mathbb{L}_\infty(\nu)} \leqslant \mathrm{v}_{\theta|\mathfrak{s}|\psi}$ and

$$\mathbb{P}_{\theta|\mathfrak{s}}\big(|\nu(h_\bullet\psi_\bullet)|^2\big) \leqslant \mathrm{v}_{\theta|\mathfrak{s}|\psi}\|h_\bullet\|_{\mathbb{J}}^2, \quad \forall h_\bullet \in \mathbb{J},$$

(dieMM4) $\mathrm{v}_\bullet^{\theta|\mathfrak{s}} := \mathbb{P}_{\theta|\mathfrak{s}}(\psi_\bullet^2) - |\mathbb{P}_{\theta|\mathfrak{s}}(\psi_\bullet)|^2 \in \mathcal{M}_{>0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$, $\|(\mathrm{v}_\bullet^{\theta|\mathfrak{s}})^{-1}\|_{\mathbb{L}_\infty(\nu)} \leqslant \mathrm{v}_{\theta|\mathfrak{s}|\psi}$ and

$$\mathbb{P}_{\theta|\mathfrak{s}}\big(|\nu(h_\bullet\psi_\bullet)|^2\big) \geqslant \mathbb{P}_{\theta|\mathfrak{s}}\big(|\nu(h_\bullet\psi_\bullet)|^2\big) - \big|\mathbb{P}_{\theta|\mathfrak{s}}\big(\nu(h_\bullet\psi_\bullet)\big)\big|^2 \geqslant \mathrm{v}_{\theta|\mathfrak{s}|\psi}^{-1}\|h_\bullet\|_{\mathbb{J}}^2, \quad \forall h_\bullet \in \mathbb{J}.$$

We consider a statistical product experiment $(\mathcal{Z}^n, \mathscr{Z}^{\otimes n}, \mathbb{P}^{\otimes n}_{\Theta\times\{\mathfrak{s}_\bullet\}} = (\mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}})_{\theta_\bullet\in\Theta})$ as in an Empirical mean function §01|01.10 where $\Theta \subseteq \mathbb{J}$. □

§07|01.09 **Property** (dieMM §07|01.08 continued).

(i) *Under* (dieMM1)–(dieMM3) *due to Lemma §01|01.08* (i) *the stochastic process* $\psi_{\bullet} \in \mathcal{M}(\mathscr{X} \otimes \mathscr{J})$ *and hence the error process* $\dot{\varepsilon}_{\bullet} = n^{1/2}(\widehat{\mathbb{P}}_n - \mathbb{P}_{\theta|\mathfrak{s}})(\psi_{\bullet}) \in \mathcal{M}(\mathscr{X}^{\otimes n} \otimes \mathscr{J})$ *admits a covariance operator* $\Gamma_{\theta|\mathfrak{s}} \in \mathbb{L}^{\geqslant}(\mathbb{J})$ *satisfying* $\|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})} \leqslant \mathbb{v}_{\theta|\mathfrak{s}|\psi}$.

(ii) *Under* (dieMM1)–(dieMM4) *due to Lemma §01|01.08* (ii) *the covariance operator* $\Gamma_{\theta|\mathfrak{s}} \in \mathbb{L}^{\geqslant}(\mathbb{J})$ *is invertible with inverse* $\Gamma_{\theta|\mathfrak{s}}^{-1} \in \mathbb{L}(\mathbb{J})$ *satisfying* $\|\Gamma_{\theta|\mathfrak{s}}^{-1}\|_{\mathbb{L}(\mathbb{J})} \leqslant \mathbb{v}_{\theta|\mathfrak{s}|\psi}$.

Consequently, from Lemma §01|01.08 (01.03) *we obtain*

$$\mathbb{v}_{\theta|\mathfrak{s}|\psi}^{-1}\|h_{\bullet}\|_{\mathbb{J}}^2 \leqslant \|h_{\bullet}\|_{\Gamma_{\theta|\mathfrak{s}}}^2 = \langle \Gamma_{\theta|\mathfrak{s}} h_{\bullet}, h_{\bullet}\rangle_{\mathbb{J}} \leqslant \mathbb{v}_{\theta|\mathfrak{s}|\psi}\|h_{\bullet}\|_{\mathbb{J}}^2 \quad \forall h_{\bullet} \in \mathbb{J}. \qquad \square$$

§07|01|02 Global and maximal global $\mathfrak{v}$-risk

We measure first the accuracy of the OPE $\widehat{\theta}_{\bullet}^m := \mathfrak{s}_{\bullet}^{\dagger}\widehat{g}_{\bullet}^m$ of the projection $\theta_{\bullet}^m = \mathfrak{s}_{\bullet}^{\dagger} g_{\bullet}^m \in \mathbb{J}\mathbb{1}_{\bullet}^m$ with $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}})$ by the mean of its global $\mathfrak{v}$-error introduced in §04|03|01, i.e. its $\mathfrak{v}$-risk.

§07|01.10 **Reminder**. If $\mathfrak{v}_{\bullet} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ and $\theta_{\bullet} \in \mathbb{L}_2(\mathfrak{v}_{\bullet}^2\nu)$ then we have $\theta_{\bullet}^m \in \mathbb{L}_2(\mathfrak{v}_{\bullet}^2\nu)$ too and $\|\theta_{\bullet}^m - \theta_{\bullet}\|_{\mathfrak{v}}^2 = \mathrm{o}(1)$ as $m \to \infty$ (Property §04|03.09). □

§07|01.11 **Assumption**. Consider a noisy version $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} \sim \mathbb{P}_{\theta|\mathfrak{s}}^n$ satisfying Assumption §07|01.01,

(dSIPg1) $\mathbb{v}_{\bullet}^{\theta|\mathfrak{s}} := \mathbb{P}_{\theta|\mathfrak{s}_{\bullet}}^n(\dot{\varepsilon}_{\bullet}^2) := (\mathbb{v}_j^{\theta|\mathfrak{s}} := \mathbb{P}_{\theta|\mathfrak{s}}^n(|\dot{\varepsilon}_j|^2))_{j\in\mathscr{J}} \in \mathbb{L}_\infty(\nu)$ and

(dSIPg2) $\dot{\varepsilon}_{\bullet}\mathbb{1}_{\bullet}^m \in \mathbb{L}_\infty(\nu)$ $\mathbb{P}_{\theta|\mathfrak{s}}^n$-a.s., for each $m \in \mathbb{N}$. □

§07|01.12 **Comment**. Under Assumption §07|01.11 and $\mathfrak{v}_{\bullet} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ set $(\mathfrak{s}^{\dagger}\mathfrak{v})_{\bullet} := \mathfrak{s}_{\bullet}^{\dagger}\mathfrak{v}_{\bullet} \in \mathcal{M}(\mathscr{J})$. If $\mathfrak{s}_{\bullet}^{\dagger}\mathbb{1}_{\bullet}^m \in \mathbb{L}_2(\mathfrak{v}_{\bullet}^2\nu)$ then we have $(\mathfrak{s}^{\dagger}\mathfrak{v})_{\bullet}\dot{\varepsilon}_{\bullet}\mathbb{1}_{\bullet}^m \in \mathbb{J}$ $\mathbb{P}_{\theta|\mathfrak{s}}^n$-a.s.. If in addition $\theta_{\bullet} \in \mathbb{L}_2(\mathfrak{v}_{\bullet}^2\nu)$, and hence $\theta_{\bullet}^m \in \mathbb{L}_2(\mathfrak{v}_{\bullet}^2\nu)$ (Property §04|03.09), then it follows

$$\mathfrak{v}_{\bullet}\widehat{\theta}_{\bullet}^m = (\mathfrak{s}^{\dagger}\mathfrak{v})_{\bullet}\widehat{g}_{\bullet}\mathbb{1}_{\bullet}^m = n^{-1/2}(\mathfrak{s}^{\dagger}\mathfrak{v})_{\bullet}\dot{\varepsilon}_{\bullet}\mathbb{1}_{\bullet}^m + \mathfrak{v}_{\bullet}\theta_{\bullet}^m \in \mathbb{J} = \mathbb{L}_2(\nu) \quad \mathbb{P}_{\theta|\mathfrak{s}}^n\text{-a.s..} \qquad (07.01)$$

If $\mathscr{J} \subseteq \mathbb{Z}$ (at most countable) and $\nu_{\mathscr{J}}$ is the counting measure over the index set $\mathscr{J}$ then Assumption §01|01.04 and **(dSIPg1)** $\mathbb{v}_{\bullet}^{\theta|\mathfrak{s}} = \mathbb{P}_{\theta|\mathfrak{s}}^n(\dot{\varepsilon}_{\bullet}^2) \in \mathbb{L}_\infty(\nu_{\mathscr{J}})$ implies the additional assumption **(dSIPg2)** $\dot{\varepsilon}_{\bullet}\mathbb{1}_{\bullet}^m \in \mathbb{L}_\infty(\nu_{\mathscr{J}})$ $\mathbb{P}_{\theta|\mathfrak{s}}^n$-a.s.. However, the last implication does generally not hold, if $\mathscr{J} \in \{\mathbb{R}, \mathbb{R}_{\geqslant 0}\}$ for example. □

§07|01|02|01 Global $\mathfrak{v}$-risk

§07|01.13 **Assumption**. Let $\mathfrak{v}_{\bullet} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$, $\theta_{\bullet} \in \mathbb{L}_2(\mathfrak{v}_{\bullet}^2\nu)$, and $\mathfrak{s}_{\bullet}^{\dagger}\mathbb{1}_{\bullet}^m \in \mathbb{L}_2(\mathfrak{v}_{\bullet}^2\nu)$ for all $m \in \mathbb{N}$ be satisfied. □

§07|01.14 **Definition**. Under Assumptions §07|01.11 and §07|01.13 the *global* $\mathfrak{v}$*-risk* of an OPE $\widehat{\theta}_{\bullet}^m = \mathfrak{s}_{\bullet}^{\dagger}\widehat{g}_{\bullet}^m = \mathfrak{s}_{\bullet}^{\dagger}\widehat{g}_{\bullet}\mathbb{1}_{\bullet}^m \in \mathbb{L}_2(\mathfrak{v}_{\bullet}^2\nu)$ $\mathbb{P}_{\theta|\mathfrak{s}}^n$-a.s. satisfies

$$\mathbb{P}_{\theta|\mathfrak{s}}^n(\|\widehat{\theta}_{\bullet}^m - \theta_{\bullet}\|_{\mathfrak{v}}^2) = \mathbb{P}_{\theta|\mathfrak{s}}^n(\|\mathfrak{s}_{\bullet}^{\dagger}(\widehat{g}_{\bullet} - g_{\bullet})\mathbb{1}_{\bullet}^m\|_{\mathfrak{v}}^2) + \|\theta_{\bullet}\mathbb{1}_{\bullet}^{m|\perp}\|_{\mathfrak{v}}^2 \qquad (07.02)$$

with *variance* term $\mathbb{P}_{\theta|\mathfrak{s}}^n(\|\mathfrak{s}_{\bullet}^{\dagger}(\widehat{g}_{\bullet} - g_{\bullet})\mathbb{1}_{\bullet}^m\|_{\mathfrak{v}}^2) = n^{-1}\mathbb{P}_{\theta|\mathfrak{s}}^n(\|(\mathfrak{s}^{\dagger}\mathfrak{v})_{\bullet}\dot{\varepsilon}_{\bullet}\mathbb{1}_{\bullet}^m\|_{\mathbb{J}}^2)$ and *bias* term $\|\theta_{\bullet}\mathbb{1}_{\bullet}^{m|\perp}\|_{\mathfrak{v}}$. □

§07|01.15 **Property**. *Under Assumptions §07|01.11 and §07|01.13 we have*

$$\mathbb{P}_{\theta|\mathfrak{s}}^n\big(\|(\mathfrak{s}^{\dagger}\mathfrak{v})_{\bullet}\dot{\varepsilon}_{\bullet}\mathbb{1}_{\bullet}^m\|_{\mathbb{J}}^2\big) = \int_{\mathscr{J}} \mathbb{P}_{\theta|\mathfrak{s}}^n(|\dot{\varepsilon}_j|^2)(\mathfrak{s}^{\dagger}\mathfrak{v})_j^2\mathbb{1}_j^m\,\nu(dj) = \nu(\mathbb{v}_{\bullet}^{\theta|\mathfrak{s}}(\mathfrak{s}^{\dagger}\mathfrak{v})_{\bullet}^2\mathbb{1}_{\bullet}^m) \qquad (07.03)$$

and consequently $\mathbb{P}_{\theta|\mathfrak{s}}^n(\|\mathfrak{s}_{\bullet}^{\dagger}(\widehat{g}_{\bullet} - g_{\bullet})\mathbb{1}_{\bullet}^m\|_{\mathfrak{v}}^2) \leqslant n^{-1}\|\mathbb{v}_{\bullet}^{\theta|\mathfrak{s}}\|_{\mathbb{L}_\infty(\nu)}\|\mathfrak{s}_{\bullet}^{\dagger}\mathbb{1}_{\bullet}^m\|_{\mathfrak{v}}^2 \in \mathbb{R}_{\geqslant 0}$. □

§07|01.16 **Notation.** For $a_\bullet \in \mathbb{R}^{\mathbb{N}}$ with minimal value in $B \subseteq \mathbb{N}$ we define

$$\arg\min\{a_m : m \in B\} := \min\{m \in B : a_m \leqslant a_j,\ \forall j \in B\}.$$ □

§07|01.17 **Proposition** (Upper bound). *Under Assumptions §07|01.11 and §07|01.13 for all $n, m \in \mathbb{N}$ setting*

$$\mathrm{R}_n^m(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) := \|\theta_\bullet \mathbb{1}_\bullet^{m|\perp}\|_{\mathfrak{v}}^2 + n^{-1}\|\mathfrak{s}_\bullet^\dagger \mathbb{1}_\bullet^m\|_{\mathfrak{v}}^2, \quad m_n^\circ := \arg\min\{\mathrm{R}_n^m(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) : m \in \mathbb{N}\}$$
$$\text{and} \quad \mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) := \mathrm{R}_n^{m_n^\circ}(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) = \min\{\mathrm{R}_n^m(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) : m \in \mathbb{N}\} \quad (07.04)$$

we have $\mathbb{P}_{\theta|\mathfrak{s}}^n(\|\widehat{\theta}_\bullet^{m_n^\circ} - \theta_\bullet\|_{\mathfrak{v}}^2) \leqslant (1 \vee \|\mathbb{v}_\bullet^{\theta|\mathfrak{s}}\|_{\mathbb{L}_\infty(\nu)})\, \mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet)$ *for all $n \in \mathbb{N}$.*

§07|01.18 **Proof** of Proposition §07|01.17. Given in the lecture. □

§07|01.19 **Definition.** Let $\theta_\bullet \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$ and $\widehat{\theta}_\bullet^m \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$ $\mathbb{P}_{\theta|\mathfrak{s}}^n$-a.s. for all $m \in \mathbb{N}$. If there exist $\mathrm{C} \in \mathbb{R}_{>0}$ and for each $n \in \mathbb{N}$, $\mathrm{R}_n^\circ \in \mathbb{R}_{>0}$ and $m_n^\circ \in \mathbb{N}$ satisfying

$$\mathrm{C}^{-1}\,\mathrm{R}_n^\circ \leqslant \inf_{m\in\mathbb{N}} \mathbb{P}_{\theta|\mathfrak{s}}^n\|\widehat{\theta}_\bullet^m - \theta_\bullet\|_{\mathfrak{v}}^2 \leqslant \mathbb{P}_{\theta|\mathfrak{s}}^n\|\widehat{\theta}_\bullet^{m_n^\circ} - \theta_\bullet\|_{\mathfrak{v}}^2 \leqslant \mathrm{C}\ \mathrm{R}_n^\circ \quad \forall n \in \mathbb{N},$$

then we call R_n° *oracle bound*, m_n° *oracle dimension* and $\widehat{\theta}_\bullet^{m_n^\circ}$ *oracle optimal* (up to the constant C). As a consequence, up to the constant C^2 the statistic $\widehat{\theta}_\bullet^{m_n^\circ}$ attains the lower global $\mathfrak{v}$-risk bound within the family of OPE's, that is, $\mathbb{P}_{\theta|\mathfrak{s}}^n\|\widehat{\theta}_\bullet^{m_n^\circ} - \theta_\bullet\|_{\mathfrak{v}}^2 \leqslant \mathrm{C}^2 \inf_{m\in\mathbb{N}} \mathbb{P}_{\theta|\mathfrak{s}}^n\|\widehat{\theta}_\bullet^m - \theta_\bullet\|_{\mathfrak{v}}^2$. □

§07|01.20 **Oracle inequality.** *Under Assumptions §07|01.11 and §07|01.13 if in addition*

$$1 \leqslant \max(\|\mathbb{v}_\bullet^{\theta|\mathfrak{s}}\|_{\mathbb{L}_\infty(\nu)}, \|(\mathbb{v}_\bullet^{\theta|\mathfrak{s}})^{-1}\|_{\mathbb{L}_\infty(\nu)}) \leqslant \mathbb{v}_{\theta|\mathfrak{s}} \in \mathbb{R}_{\geqslant 1}$$

is satisfied then (07.04) *implies*

$$\mathbb{v}_{\theta|\mathfrak{s}}^{-1}\,\mathrm{R}_n^m(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) \leqslant \mathbb{P}_{\theta|\mathfrak{s}}^n(\|\widehat{\theta}_\bullet^m - \theta_\bullet\|_{\mathfrak{v}}^2) = n^{-1}\nu(\mathbb{v}_\bullet^{\theta|\mathfrak{s}}(\mathfrak{s}_\bullet^\dagger\mathfrak{v}_\bullet)^2\mathbb{1}_\bullet^m) + \|\theta_\bullet\mathbb{1}_\bullet^{m|\perp}\|_{\mathfrak{v}}^2$$
$$\leqslant \mathbb{v}_{\theta|\mathfrak{s}}\mathrm{R}_n^m(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) \quad \forall m, n \in \mathbb{N}.$$

As a consequence we immediately obtain the following oracle inequality

$$\mathbb{v}_{\theta|\mathfrak{s}}^{-1}\,\mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) \leqslant \inf_{m\in\mathbb{N}} \mathbb{P}_{\theta|\mathfrak{s}}^n(\|\widehat{\theta}_\bullet^m - \theta_\bullet\|_{\mathfrak{v}}^2) \leqslant \mathbb{P}_{\theta|\mathfrak{s}}^n(\|\widehat{\theta}_\bullet^{m_n^\circ} - \theta_\bullet\|_{\mathfrak{v}}^2)$$
$$\leqslant \mathbb{v}_{\theta|\mathfrak{s}}\,\mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) \leqslant \mathbb{v}_{\theta|\mathfrak{s}}^2 \inf_{m\in\mathbb{N}} \mathbb{P}_{\theta|\mathfrak{s}}^n(\|\widehat{\theta}_\bullet^m - \theta_\bullet\|_{\mathfrak{v}}^2) \quad \forall n \in \mathbb{N}, \quad (07.05)$$

and, hence $\mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet)$, m_n° and the statistic $\widehat{\theta}_\bullet^{m_n^\circ}$, respectively, is an oracle bound, an oracle dimension and oracle optimal (up to the constant $\mathbb{v}_{\theta|\mathfrak{s}}^2$). □

§07|01.21 **Remark.** For each fixed $m \in \mathbb{N}$ with $\|\mathfrak{s}_\bullet^\dagger\mathbb{1}_\bullet^m\|_{\mathfrak{v}} \in \mathbb{R}_{\geqslant 0}$ we have $n^{-1}\|\mathfrak{s}_\bullet^\dagger\mathbb{1}_\bullet^m\|_{\mathfrak{v}} = \mathrm{o}(1)$ as $n \to \infty$. As a consequence, if $\|\mathfrak{s}_\bullet^\dagger\mathbb{1}_\bullet^m\|_{\mathfrak{v}} \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$ and $\|\theta_\bullet\mathbb{1}_\bullet^{m|\perp}\|_{\mathfrak{v}} = \mathrm{o}(1)$ as $m \to \infty$ then we obtain $\mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) = \mathrm{o}(1)$ as $n \to \infty$, and thus, $\mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet)$ is also called an *oracle rate*. Indeed, for all $\delta \in \mathbb{R}_{>0}$ there exists $m_\delta \in \mathbb{N}$ and $n_\delta \in \mathbb{N}$ such that we have both $\|\theta_\bullet\mathbb{1}_\bullet^{m_\delta|\perp}\|_{\mathfrak{v}}^2 \leqslant \delta/2$ and $n^{-1}\|\mathfrak{s}_\bullet^\dagger\mathbb{1}_\bullet^{m_\delta}\|_{\mathfrak{v}}^2 \leqslant \delta/2$ for all $n \in \mathbb{N}_{\geqslant n_\delta}$, and whence $\mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) \leqslant \mathrm{R}_n^{m_\delta}(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) \leqslant \delta$. However, note that the oracle dimension $m_n^\circ = m_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet)$ as defined in Proposition §07|01.17 depends on the unknown parameter of interest $\theta_\bullet$, and thus also the oracle optimal statistic $\widehat{\theta}_\bullet^{m_n^\circ}$. In other words $\widehat{\theta}_\bullet^{m_n^\circ}$ is not a feasible estimator. □

§07|01.22 **Corollary** (GdiSM §07|01.03 continued). *Consider* $\widehat{g} = g + n^{-1/2}\dot{B} \sim N^n_{\theta|\mathfrak{s}}$ *as in Model §07|01.03, where* $\dot{B} \sim N^{\otimes\mathbb{N}}_{(0,1)}$, $\mathfrak{s} \in \ell_\infty$, $\theta \in \ell_2$ *and hence* $g = \mathfrak{s}\theta \in \mathrm{dom}(M_{\mathfrak{s}^\dagger}) \subseteq \ell_2$. *Given* $\mathfrak{v} \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ *and* $\theta \in \ell_2(\mathfrak{v}^2)$ *the (infeasible) OPE* $\widehat{\theta}^{m^\circ_n} = \mathfrak{s}^\dagger \widehat{g}^{m^\circ_n} \in \ell_2(\mathfrak{v}^2)$ *with oracle dimension* m°_n *as in* (07.04) *satisfies*

$$N^n_{\theta|\mathfrak{s}}\big(\|\widehat{\theta}^{m^\circ_n} - \theta\|^2_{\mathfrak{v}}\big) = R^\circ_n(\theta,\mathfrak{s},\mathfrak{v}) = \inf_{m\in\mathbb{N}} N^n_{\theta|\mathfrak{s}}\big(\|\widehat{\theta}^m - \theta\|^2_{\mathfrak{v}}\big) \quad \forall n \in \mathbb{N},$$

and hence it is oracle optimal (with constant 1*).*

§07|01.23 **Proof** of Corollary §07|01.22. Given in the lecture. □

§07|01.24 **Corollary** (diSM §07|01.06 continued). *Consider* $\widehat{g} = g + n^{-1/2}\dot{\varepsilon} \sim P^n_{\theta|\mathfrak{s}|\sigma}$ *as in Model §07|01.06, where* $\dot{\varepsilon} \sim \otimes_{j\in\mathbb{N}} P_{(0,\sigma_j^2)}$ *satisfies* **(iSM1)** *and* **(iSM2)** *with* $\max(\|\sigma^{-2}\|_{\ell_\infty}, \|\sigma^2\|_{\ell_\infty}) =: \mathbb{v}_\sigma \in \mathbb{R}_{\geqslant 1}$, $\mathfrak{s} \in \ell_\infty$, $\theta \in \ell_2$ *and hence* $g = \mathfrak{s}\theta \in \mathrm{dom}(M_{\mathfrak{s}^\dagger}) \subseteq \ell_2$. *Given* $\mathfrak{v} \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ *and* $\theta \in \ell_2(\mathfrak{v}^2)$ *the (infeasible) OPE* $\widehat{\theta}^{m^\circ_n} = \mathfrak{s}^\dagger \widehat{g}^{m^\circ_n} \in \ell_2(\mathfrak{v}^2)$ *with oracle dimension* m°_n *as in* (07.04) *satisfies*

$$P^n_{\theta|\mathfrak{s}|\sigma}\big(\|\widehat{\theta}^{m^\circ_n} - \theta\|^2_{\mathfrak{v}}\big) \leqslant \mathbb{v}_\sigma R^\circ_n(\theta,\mathfrak{s},\mathfrak{v}) \leqslant \mathbb{v}_\sigma^2 \inf_{m\in\mathbb{N}} P^n_{\theta|\mathfrak{s}|\sigma}\big(\|\widehat{\theta}^m - \theta\|^2_{\mathfrak{v}}\big) \quad \forall n \in \mathbb{N},$$

and hence it is oracle optimal (with constant $\mathbb{v}_\sigma$*).*

§07|01.25 **Proof** of Corollary §07|01.24. Given in the lecture. □

§07|01.26 **Corollary** (dieMM §07|01.08 continued). *Let* $\widehat{g} = g + n^{-1/2}\dot{\boldsymbol{\varepsilon}}$ *be defined on* $(\mathcal{Z}^n, \mathscr{Z}^{\otimes n}, \mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}})$ *as in Model §07|01.08, where* $\psi \in \mathcal{M}(\mathscr{Z}\otimes\mathscr{J})$ *satisfies* **(dieMM1)**–**(dieMM4)** *for some* $\mathbb{v}_{\theta|\mathfrak{s}|\psi} \in \mathbb{R}_{\geqslant 1}$, $\mathfrak{s} \in \mathbb{L}_\infty(\nu)$, $\theta \in \mathbb{J}$ *and hence* $g = \mathfrak{s}\theta \in \mathrm{dom}(M_{\mathfrak{s}^\dagger}) \subseteq \mathbb{J}$. *Under Assumption §07|01.13 the (infeasible) OPE* $\widehat{\theta}^{m^\circ_n} = \mathfrak{s}^\dagger \widehat{g}^{m^\circ_n} \in \mathbb{L}_2(\mathfrak{v}^2\nu)$ $\mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}}$-a.s. *with oracle dimension* m°_n *as in* (07.04) *satisfies*

$$\mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}}\big(\|\widehat{\theta}^{m^\circ_n} - \theta\|^2_{\mathfrak{v}}\big) \leqslant \mathbb{v}_{\theta|\mathfrak{s}|\psi} R^\circ_n(\theta,\mathfrak{s},\mathfrak{v}) \leqslant \mathbb{v}_{\theta|\mathfrak{s}|\psi}^2 \inf_{m\in\mathbb{N}} \mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}}\big(\|\widehat{\theta}^m - \theta\|^2_{\mathfrak{v}}\big) \quad \forall n \in \mathbb{N},$$

and hence it is oracle optimal (with constant $\mathbb{v}_{\theta|\mathfrak{s}|\psi}$*).*

§07|01.27 **Proof** of Corollary §07|01.26. Given in the lecture. □

§07|01.28 **Illustration**. We illustrate the last results considering usual behaviour for $\theta, \mathfrak{s}, \mathfrak{v} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$. We distinguish the following two cases

(p) $\mathfrak{s}^\dagger \in \mathbb{L}_2(\mathfrak{v}^2\nu)$ or there is $m \in \mathbb{N}$ with $\|\theta^m - \theta\|^2_{\mathfrak{v}} = 0$,

(np) $\mathfrak{s}^\dagger \notin \mathbb{L}_2(\mathfrak{v}^2\nu)$ and for all $m \in \mathbb{N}$ holds $\|\theta^m - \theta\|^2_{\mathfrak{v}} \in \mathbb{R}_{>0}$.

Interestingly, in case **(p)** the oracle bound is parametric, that is, $nR^\circ_n(\theta,\mathfrak{s},\mathfrak{v}) = O(1)$, in case **(np)** the oracle bound is nonparametric, i.e. $\lim_{n\to\infty} nR^\circ_n(\theta,\mathfrak{s},\mathfrak{v}) = \infty$. In case **(np)** consider the following three specifications:

Table 01 [§07]

Order of the oracle rate $R^\circ_n(\theta,\mathfrak{s},\mathfrak{v})$ as $n \to \infty$

	$(j \in \mathcal{J})$ $\mathfrak{v}_j^2 = j^{2\mathrm{v}}$	$(\mathrm{a} \in \mathbb{R}_{>0})$ θ_j^2	$(\mathrm{t} \in \mathbb{R}_{>0})$ $\mathfrak{s}_j^2$	(squared bias) $\|\theta\mathbb{1}^{m\|\perp}\|^2_{\mathfrak{v}}$	(variance) $\|\mathfrak{s}^\dagger\mathbb{1}^m\|^2_{\mathfrak{v}}$	m°_n	$R^\circ_n(\theta,\mathfrak{s},\mathfrak{v})$
(o-m)	$\mathrm{v} \in (-1/2-\mathrm{t}, \mathrm{a})$	$j^{-2\mathrm{a}-1}$	$j^{-2\mathrm{t}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2\mathrm{v}+2\mathrm{t}+1}$	$n^{\frac{1}{2\mathrm{a}+2\mathrm{t}+1}}$	$n^{-\frac{2(\mathrm{a}-\mathrm{v})}{2\mathrm{a}+2\mathrm{t}+1}}$
	$\mathrm{v}+\mathrm{t} = -1/2$	$j^{-2\mathrm{a}-1}$	$j^{-2\mathrm{t}}$	$m^{-2\mathrm{a}-2\mathrm{t}-1}$	$\log m$	$\left(\frac{n}{\log n}\right)^{\frac{1}{2\mathrm{a}+2\mathrm{t}+1}}$	$\frac{\log n}{n}$
(o-s)	$\mathrm{a}-\mathrm{v} \in \mathbb{R}_{>0}$	$j^{-2\mathrm{a}-1}$	$e^{-j^{2\mathrm{t}}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{(1-2(\mathrm{t}-\mathrm{v}))_+}e^{m^{2\mathrm{t}}}$	$(\log n)^{\frac{1}{2\mathrm{t}}}$	$(\log n)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v}+\mathrm{t}+1/2 \in \mathbb{R}_{>0}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$m^{(1-2(\mathrm{a}-\mathrm{v}))_+}e^{-m^{2\mathrm{a}}}$	$m^{2\mathrm{v}+2\mathrm{t}+1}$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{(\log n)^{\frac{2\mathrm{t}+2\mathrm{v}+1}{2\mathrm{a}}}}{n}$
	$\mathrm{v}+\mathrm{t} = -1/2$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$e^{-m^{2\mathrm{a}}}$	$\log m$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{\log\log n}{n}$

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v}+\mathrm{t} < -1/2$ the oracle rate $\mathrm{R}^{\circ}_{n}(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet})$ is parametric. □

§07|01|02|02 Maximal global $\mathfrak{v}$-risk

§07|01.29 **Notation** (Reminder). For sequences $a_{\bullet}, b_{\bullet} \in (\mathbb{K})^{\mathbb{N}}$ taking its values in $\mathbb{K} \in \{\mathbb{R}, \mathbb{R}_{>0}, \mathbb{Q}, \mathbb{Z}, \dots\}$ we write $a_{\bullet} \in (\mathbb{K})^{\mathbb{N}}_{\nearrow}$ and $b_{\bullet} \in (\mathbb{K})^{\mathbb{N}}_{\searrow}$ if $a_{\bullet}$ and $b_{\bullet}$, respectively, is monotonically *non-decreasing* and *non-increasing*. If in addition $a_n \to \infty$ and $b_n \to 0$ as $n \to \infty$, then we write $a_{\bullet} \in (\mathbb{K})^{\mathbb{N}}_{\uparrow\infty}$ and $b_{\bullet} \in (\mathbb{K})^{\mathbb{N}}_{\downarrow 0}$ for short. For $\mathrm{w}_{\bullet} \in \mathbb{L}_{\infty}(\nu)$ we set $\mathrm{w}_{(0)} := \|\mathrm{w}_{\bullet}\|_{\mathbb{L}_{\infty}(\nu)}$ and $\mathrm{w}_{(\bullet)} = (\mathrm{w}_{(j)} := \|\mathrm{w}_{\bullet}\mathbb{1}^{j|\perp}_{\bullet}\|_{\mathbb{L}_{\infty}(\nu)})_{j\in\mathbb{N}}$, where by construction $\mathrm{w}_{(\bullet)} \in (\mathbb{R}_{\geqslant 0})^{\mathbb{N}}_{\searrow}$. □

§07|01.30 **Assumption**. Consider weights $\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet} \in \mathcal{M}_{>0,\nu}(\mathscr{J})$ (i.e. $\nu(\mathcal{N}_{\mathfrak{a}}) = \nu(\mathcal{N}_{\mathfrak{t}}) = 0 = \nu(\mathcal{N}_{\mathfrak{v}})$), such that $\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet} \in \mathbb{L}_{\infty}(\nu)$, $(\mathfrak{a}\mathfrak{v})_{\bullet} = \mathfrak{a}_{\bullet}\mathfrak{v}_{\bullet} \in \mathbb{L}_{\infty}(\nu)$, $(\mathfrak{a}\mathfrak{v})_{(\bullet)} \in (\mathbb{R}_{\geqslant 0})^{\mathbb{N}}_{\downarrow 0}$, and $\mathfrak{t}^{\dagger}_{\bullet}\mathbb{1}^{m}_{\bullet} \in \mathbb{L}_2(\mathfrak{v}^2_{\bullet}\nu)$ for all $m \in \mathbb{N}$. □

§07|01.31 **Reminder**. Under Assumption §07|01.30 we have $\mathbb{J}^{\mathfrak{a}} = \mathbb{L}^{\mathfrak{a}}_{2}(\nu) = \mathrm{dom}(\mathrm{M}_{\mathfrak{a}^{\dagger}}) = \mathbb{J}\mathfrak{a}_{\bullet} \subseteq \mathbb{J}$ and the three measures ν, $\mathfrak{a}^{2|\dagger}_{\bullet}\nu$ and $\mathfrak{v}^2_{\bullet}\nu$ dominate mutually each other, i.e. they share the same null sets (see Property §04|01.02). We consider $\mathbb{J}^{\mathfrak{a}}$ endowed with $\|\cdot\|_{\mathfrak{a}^{\dagger}} = \|\mathrm{M}_{\mathfrak{a}^{\dagger}}\cdot\|_{\mathbb{J}}$ and given a constant $\mathrm{r} \in \mathbb{R}_{>0}$ the ellipsoid $\mathbb{J}^{\mathfrak{a},\mathrm{r}} := \{h_{\bullet} \in \mathbb{J}^{\mathfrak{a}} : \|h_{\bullet}\|_{\mathfrak{a}^{\dagger}} \leqslant \mathrm{r}\} \subseteq \mathbb{J}^{\mathfrak{a}}$. Since $(\mathfrak{a}\mathfrak{v})_{\bullet} \in \mathbb{L}_{\infty}(\nu)$, and hence $(\mathfrak{a}\mathfrak{v})_{(m)} := \|(\mathfrak{a}\mathfrak{v})_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\mathbb{L}_{\infty}(\nu)} \in \mathbb{R}_{\geqslant 0}$ for each $m \in \mathbb{N}$ we have $\mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{L}_2(\mathfrak{v}^2_{\bullet}\nu)$ (Property §04|02.11), and $\|\theta_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\mathfrak{v}} \leqslant \mathrm{r}\,(\mathfrak{a}\mathfrak{v})_{(m)}$ for all $\theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ (Lemma §04|02.13). Consequently, if Assumption §07|01.30, $\theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ and $\mathfrak{s}^{\dagger}_{\bullet}\mathbb{1}^{m}_{\bullet} \in \mathbb{L}_2(\mathfrak{v}^2_{\bullet}\nu)$ for all $m \in \mathbb{N}$ are satisfied, then Assumption §07|01.13 is fulfilled. Moreover, under Assumption §07|01.30 for each $\mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ we have $\|\mathfrak{s}^{\dagger}_{\bullet}\mathbb{1}^{m}_{\bullet}\|_{\mathfrak{v}} \leqslant \mathrm{d}\|\mathfrak{t}^{\dagger}_{\bullet}\mathbb{1}^{m}_{\bullet}\|_{\mathfrak{v}} \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$ (Definition §04|03.05). Therefore, if Assumption §07|01.30, $\theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ and $\mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ are satisfied, then Assumption §07|01.13 is also fulfilled. □

§07|01.32 **Proposition** (Upper bound). *Under Assumptions §07|01.11 and §07|01.30 let* $\mathfrak{s}^{\dagger}_{\bullet}\mathbb{1}^{m}_{\bullet} \in \mathbb{L}_2(\mathfrak{v}^2_{\bullet}\nu)$ *for all* $m \in \mathbb{N}$. *Setting for* $n, m \in \mathbb{N}$

$$\mathrm{R}^{m}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet}) := [(\mathfrak{a}\mathfrak{v})^2_{(m)} \vee n^{-1}\|\mathfrak{s}^{\dagger}_{\bullet}\mathbb{1}^{m}_{\bullet}\|^2_{\mathfrak{v}}], \quad m^{\star}_{n} := \arg\min\left\{\mathrm{R}^{m}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet}) : m \in \mathbb{N}\right\}$$
$$\text{and} \quad \mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet}) := \mathrm{R}^{m^{\star}_{n}}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet}) = \min\left\{\mathrm{R}^{m}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet}) : m \in \mathbb{N}\right\} \quad (07.06)$$

and $\|\mathrm{v}^{\theta|\mathfrak{s}}_{\bullet}\|_{\mathbb{L}_{\infty}(\nu)} =: \mathrm{v}_{\theta|\mathfrak{s}} \in \mathbb{R}_{>0}$, *for all* $\theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$, *hence* $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^{\dagger}}) \subseteq \mathbb{J}$, *we have*

$$\mathbb{E}^{n}_{\theta|\mathfrak{s}}(\|\widehat{\theta}^{m^{\star}_{n}}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) \leqslant (\mathrm{v}_{\theta|\mathfrak{s}} + \mathrm{r}^2)\,\mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet}) \quad \forall n \in \mathbb{N}.$$

§07|01.33 **Proof** of Proposition §07|01.32. Given in the lecture. □

§07|01.34 **Remark**. Under the assumptions of Proposition §07|01.32 if there exists in addition $\mathrm{v}_{\mathfrak{s}} \in \mathbb{R}_{\geqslant 0}$ satisfying $\|\mathrm{v}^{\theta|\mathfrak{s}}_{\bullet}\|_{\mathbb{L}_{\infty}(\nu)} \leqslant \mathrm{v}_{\mathfrak{s}}$ for all $\theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ then

$$\sup\left\{\mathbb{E}^{n}_{\theta|\mathfrak{s}}(\|\widehat{\theta}^{m^{\star}_{n}}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) \colon \theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}\right\} \leqslant (\mathrm{v}_{\mathfrak{s}} + \mathrm{r}^2)\mathrm{R}^{\star}_{n}(\mathfrak{a}, \mathfrak{s}, \mathfrak{v}_{\bullet}) \quad \forall n \in \mathbb{N}.$$

Arguing similarly as in Remark §07|01.21 we note that $\mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet}) = \mathrm{o}(1)$ as $n \to \infty$, whenever $\|\mathfrak{s}^{\dagger}_{\bullet}\mathbb{1}^{m}_{\bullet}\|_{\mathfrak{v}_{\bullet}} \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$ (note that $(\mathfrak{a}\mathfrak{v})_{(m)} = \mathrm{o}(1)$ as $m \to \infty$ by Assumption §07|01.30 which is satisfied, for example, if $(\mathfrak{a}\mathfrak{v})_{\bullet} = \mathfrak{a}_{\bullet}\mathfrak{v}_{\bullet} \in \mathbb{J}$ or in equal $\mathfrak{a}_{\bullet} \in \mathbb{L}_2(\mathfrak{v}^2_{\bullet}\nu)$). Note that the dimension $m^{\star}_{n} := m^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet})$ as defined in (07.06) does not depend on the unknown parameter of interest $\theta_{\bullet}$ but on the class $\mathbb{J}^{\mathfrak{a},\mathrm{r}}$ only, and thus also the statistic $\widehat{\theta}^{m^{\star}_{n}}_{\bullet}$. In other words, if the regularity of $\theta_{\bullet}$ is known in advance, then the OPE $\widehat{\theta}^{m^{\star}_{n}}_{\bullet}$ is a feasible estimator. □

§07|01.35 **Corollary** (Upper bound). *Under Assumptions §07|01.11 and §07|01.30 setting for* $n, m \in \mathbb{N}$

$$\mathrm{R}^{m}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) := [(\mathfrak{a}\mathfrak{v})^2_{(m)} \vee n^{-1}\|\mathfrak{t}^{\dagger}_{\bullet}\mathbb{1}^{m}_{\bullet}\|^2_{\mathfrak{v}}], \quad m^{\star}_{n} := \arg\min\left\{\mathrm{R}^{m}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) : m \in \mathbb{N}\right\}$$
$$\text{and} \quad \mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) := \mathrm{R}^{m^{\star}_{n}}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) = \min\left\{\mathrm{R}^{m}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) : m \in \mathbb{N}\right\} \quad (07.07)$$

and $\|\mathbb{v}^{\theta|\mathfrak{s}}_{\bullet}\|_{\mathbb{L}_\infty(\nu)} =: \mathbb{v}_{\theta|\mathfrak{s}} \in \mathbb{R}_{\geqslant 0}$*, for each* $\mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ known *in advance, for all* $\theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$*, hence* $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \subseteq \mathbb{J}$*, we have*

$$\mathbb{P}^n_{\theta|\mathfrak{s}}(\|\widehat{\theta}^{m_n^\star}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) \leqslant (\mathrm{d}^2\mathbb{v}_{\theta|\mathfrak{s}} + \mathrm{r}^2)\, \mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) \quad \forall n \in \mathbb{N}.$$

§07|01.36 **Proof** of Corollary §07|01.35. Given in the lecture. □

§07|01.37 **Remark.** Under the assumptions of Corollary §07|01.35 if there exists in addition $\mathbb{v} \in \mathbb{R}_{\geqslant 0}$ satisfying $\|\mathbb{v}^{\theta|\mathfrak{s}}_{\bullet}\|_{\mathbb{L}_\infty(\nu)} \leqslant \mathbb{v}$ for all $\theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ and $\mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ then

$$\sup\left\{\mathbb{P}^n_{\theta|\mathfrak{s}}(\|\widehat{\theta}^{m_n^\star}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}})\colon \theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}\right\} \leqslant (\mathbb{v}\mathrm{d}^2 + \mathrm{r}^2)\, \mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) \quad \forall n \in \mathbb{N}.$$

Arguing similarly as in Remark §07|01.21 we note that $\mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) = \mathrm{o}(1)$ as $n \to \infty$ since $\|\mathfrak{t}^\dagger_{\bullet}\mathbb{1}^m_{\bullet}\|_{\mathfrak{v}} \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$ and $(\mathfrak{a}\mathfrak{v})_{(m)} = \mathrm{o}(1)$ as $m \to \infty$ by Assumption §07|01.30. Note that the dimension $m_n^\star := m_n^\star(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet})$ as defined in (07.07) does neither depend on the unknown parameter of interest $\theta_{\bullet}$ nor on the known multiplication operator $\mathrm{M}_{\mathfrak{s}}$ but on the classes $\mathbb{J}^{\mathfrak{a},\mathrm{r}}$ and $\mathbb{M}_{\mathfrak{t},\mathrm{d}}$ only, and thus also the statistic $\widehat{\theta}^{m_n^\star}_{\bullet}$. In other words, if the regularity of $\theta_{\bullet}$ is known in advance, then the OPE $\widehat{\theta}^{m_n^\star}_{\bullet}$ is a feasible estimator. □

§07|01.38 **Corollary** (GdiSM §07|01.03 continued). *Consider* $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}^n_{\theta|\mathfrak{s}}$ *as in Model §07|01.03, where* $\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$, $\mathfrak{s}_{\bullet} \in \ell_\infty$, $\theta_{\bullet} \in \ell_2$ *and hence* $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \subseteq \ell_2$*. Under Assumption §07|01.30 the OPE* $\widehat{\theta}^{m_n^\star}_{\bullet} = \mathfrak{s}^\dagger_{\bullet}\widehat{g}_{\bullet}\mathbb{1}^{m_n^\star}_{\bullet} \in \ell_2(\mathfrak{v}^2_{\bullet})$ *satisfies*

(i) *with dimension* $m_n^\star = m_n^\star(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet})$ *as in* (07.06) *and constant* $\mathrm{C} = 1 + \mathrm{r}^2$

$$\sup\left\{\mathrm{N}^n_{\theta|\mathfrak{s}}\left(\|\widehat{\theta}^{m_n^\star}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}\right)\colon \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}\right\} \leqslant \mathrm{C}\,\mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet}) \quad \forall n \in \mathbb{N}; \tag{07.08}$$

(ii) *with dimension* $m_n^\star = m_n^\star(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet})$ *as in* (07.07) *and constant* $\mathrm{C} = \mathrm{d}^2 + \mathrm{r}^2$

$$\sup\left\{\mathrm{N}^n_{\theta|\mathfrak{s}}\left(\|\widehat{\theta}^{m_n^\star}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}\right)\colon \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}\right\} \leqslant \mathrm{C}\,\mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) \quad \forall n \in \mathbb{N}. \tag{07.09}$$

§07|01.39 **Proof** of Corollary §07|01.38. Given in the lecture. □

§07|01.40 **Corollary** (diSM §07|01.06 continued). *Consider* $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} \sim \mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}$ *as in Model §07|01.06, where* $\dot{\varepsilon}_{\bullet} \sim \otimes_{j\in\mathbb{N}}\mathrm{P}_{(0,\sigma_j^2)}$ *satisfies* **(iSM1)** *with* $\|\sigma^2_{\bullet}\|_{\ell_\infty} =: \mathbb{v}_\sigma \in \mathbb{R}_{>0}$, $\mathfrak{s}_{\bullet} \in \ell_\infty$, $\theta_{\bullet} \in \ell_2$ *and hence* $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \subseteq \ell_2$*. Under Assumption §07|01.30 the OPE* $\widehat{\theta}^{m_n^\star}_{\bullet} = \mathfrak{s}^\dagger_{\bullet}\widehat{g}_{\bullet}\mathbb{1}^{m_n^\star}_{\bullet} \in \ell_2(\mathfrak{v}^2_{\bullet})$ *satisfies*

(i) *with dimension* $m_n^\star = m_n^\star(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet})$ *as in* (07.06) *and constant* $\mathrm{C} = \mathbb{v}_\sigma + \mathrm{r}^2$

$$\sup\left\{\mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}\left(\|\widehat{\theta}^{m_n^\star}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}\right)\colon \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}\right\} \leqslant \mathrm{C}\,\mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet}) \quad \forall n \in \mathbb{N}; \tag{07.10}$$

(ii) *with dimension* $m_n^\star = m_n^\star(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet})$ *as in* (07.07) *and constant* $\mathrm{C} = \mathbb{v}_\sigma\mathrm{d}^2 + \mathrm{r}^2$

$$\sup\left\{\mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}\left(\|\widehat{\theta}^{m_n^\star}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}\right)\colon \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}\right\} \leqslant \mathrm{C}\,\mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) \quad \forall n \in \mathbb{N}. \tag{07.11}$$

§07|01.41 **Proof** of Corollary §07|01.40. Given in the lecture. □

§07|01.42 **Corollary** (dieMM §07|01.08 continued). *Let* $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ *be defined on* $(\mathcal{Z}^n, \mathscr{Z}^{\otimes n}, \mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}})$ *as in Model §07|01.08, where* $\psi_{\bullet} \in \mathcal{M}(\mathscr{X}\otimes\mathscr{J})$ *satisfies* **(dieMM1)–(dieMM3)** *for some* $\mathbb{v}_{\theta|\mathfrak{s}|\psi} \in \mathbb{R}_{\geqslant 1}$, $\mathfrak{s}_{\bullet} \in \mathbb{L}_\infty(\nu)$, $\theta_{\bullet} \in \mathbb{J}$ *and hence* $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \subseteq \mathbb{J}$*. Under Assumption §07|01.30 the OPE* $\widehat{\theta}^{m_n^\star}_{\bullet} = \mathfrak{s}^\dagger_{\bullet}\widehat{g}_{\bullet}\mathbb{1}^{m_n^\star}_{\bullet} \in \mathbb{L}_2(\mathfrak{v}^2_{\bullet}\nu)$ $\mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}}$-a.s. *satisfies*

(i) *with constant*

$$\mathrm{C}_{\mathfrak{a},\mathrm{r},\mathfrak{s}} := \sup\left\{\mathbb{v}_{\theta|\mathfrak{s}|\psi} \colon \theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}\right\} + \mathrm{r}^2$$

and dimension $m_n^\star = m_n^\star(\mathfrak{a}_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet)$ *as in* (07.06)

$$\sup\left\{\mathbb{E}_{\theta|\mathfrak{s}}^{\otimes n}\big(\|\widehat{\theta}_\bullet^{m_n^\star} - \theta_\bullet\|_{\mathfrak{v}}^2\big) \colon \theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}\right\} \leqslant \mathrm{C}_{\mathfrak{a},\mathrm{r},\mathfrak{s}}\, \mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) \quad \forall n \in \mathbb{N} \tag{07.12}$$

provided $\mathfrak{s}_\bullet^\dagger \mathbb{1}_\bullet^m \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$ *for all* $m \in \mathbb{N}$*;*

(ii) *with constant*

$$\mathrm{C}_{\mathfrak{a},\mathrm{r},\mathfrak{t},\mathrm{d}} := \mathrm{d}^2 \sup\left\{\mathbb{v}_{\theta|\mathfrak{s}|\psi} \colon \theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}\right\} + \mathrm{r}^2$$

and dimension $m_n^\star = m_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet)$ *as in* (07.07)

$$\sup\left\{\mathbb{E}_{\theta|\mathfrak{s}}^{\otimes n}\big(\|\widehat{\theta}_\bullet^{m_n^\star} - \theta_\bullet\|_{\mathfrak{v}}^2\big) \colon \theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}\right\} \leqslant \mathrm{C}_{\mathfrak{a},\mathrm{r},\mathfrak{t},\mathrm{d}}\, \mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) \quad \forall n \in \mathbb{N}. \tag{07.13}$$

§07|01.43 **Proof** of Corollary §07|01.42. Given in the lecture. □

§07|01.44 **Illustration**. We illustrate the last results considering usual behaviour for $\mathfrak{a}_\bullet, \mathfrak{s}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ and $\mathrm{w}_\bullet \in \{\mathfrak{s}_\bullet, \mathfrak{t}_\bullet\}$. We distinguish similar to Illustration §07|01.28 the following two cases **(p)** $\mathrm{w}_\bullet^\dagger \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$, and **(np)** $\mathrm{w}_\bullet^\dagger \notin \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$. Interestingly, in case **(p)** the bounds in Proposition §07|01.32 and Corollary §07|01.35 are parametric, that is, $n\mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathrm{w}_\bullet, \mathfrak{v}_\bullet) = \mathrm{O}(1)$, in case **(np)** the bounds are nonparametric, i.e. $\lim_{n\to\infty} n\mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathrm{w}_\bullet, \mathfrak{v}_\bullet) = \infty$. In case **(np)** consider the following three specifications:

Table 02 [§07]

Order of the oracle rate $\mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathrm{w}_\bullet, \mathfrak{v}_\bullet)$ as $n \to \infty$

	$(j \in \mathscr{J})$ $\mathfrak{v}_j^2 = j^{2v}$	$(\mathrm{a} \in \mathbb{R}_{>0})$ $\mathfrak{a}_j^2$	$(\mathrm{t} \in \mathbb{R}_{>0})$ w_j^2	(squared bias) $(\mathfrak{a}\mathfrak{v})_{(m)}^2$	(variance) $\|\mathrm{w}_\bullet^\dagger \mathbb{1}_\bullet^m\|_{\mathfrak{v}}^2$	$m_n^\star$	$\mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathrm{w}_\bullet, \mathfrak{v}_\bullet)$
(o-m)	$\mathrm{v} \in (-1/2 - \mathrm{t}, \mathrm{a})$	$j^{-2\mathrm{a}}$	$j^{-2\mathrm{t}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2\mathrm{v}+2\mathrm{t}+1}$	$n^{\frac{1}{2\mathrm{a}+2\mathrm{t}+1}}$	$n^{-\frac{2(\mathrm{a}-\mathrm{v})}{2\mathrm{a}+2\mathrm{t}+1}}$
	$\mathrm{v} + \mathrm{t} = -1/2$	$j^{-2\mathrm{a}}$	$j^{-2\mathrm{t}}$	$m^{-2\mathrm{a}-2\mathrm{t}-1}$	$\log m$	$\left(\frac{n}{\log n}\right)^{\frac{1}{2\mathrm{a}+2\mathrm{t}+1}}$	$\frac{\log n}{n}$
(o-s)	$\mathrm{a} - \mathrm{v} \in \mathbb{R}_{>0}$	$j^{-2\mathrm{a}}$	$e^{-j^{2\mathrm{t}}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{(1-2(\mathrm{t}-\mathrm{v}))_+} e^{m^{2\mathrm{t}}}$	$(\log n)^{\frac{1}{2\mathrm{t}}}$	$(\log n)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v} + \mathrm{t} + 1/2 \in \mathbb{R}_{>0}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$m^{2\mathrm{v}} e^{-m^{2\mathrm{a}}}$	$m^{2\mathrm{v}+2\mathrm{t}+1}$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{(\log n)^{\frac{2\mathrm{t}+2\mathrm{v}+1}{2\mathrm{a}}}}{n}$
	$\mathrm{v} + \mathrm{t} = -1/2$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$m^{2\mathrm{v}} e^{-m^{2\mathrm{a}}}$	$\log m$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{\log\log n}{n}$

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v} + \mathrm{t} < -1/2$ the rate $\mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathrm{w}_\bullet, \mathfrak{v}_\bullet)$ is parametric. □

§07|01|03 Local and maximal local ϕ-risk

Secondly, we measure the accuracy of the OPE $\widehat{\theta}_\bullet^m := \mathfrak{s}_\bullet^\dagger \widehat{g}_\bullet^m$ of $\theta_\bullet^m = \mathfrak{s}_\bullet^\dagger g_\bullet^m \in \mathbb{J}\mathbb{1}_\bullet^m$ with $g_\bullet = \mathfrak{s}_\bullet \theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}'})$ by the mean of its local ϕ-error introduced in §04|03|02, i.e. its ϕ-risk.

§07|01.45 **Reminder**. If $\phi_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ and $\theta_\bullet \in \mathrm{dom}(\phi\nu)$, then for each $m \in \mathbb{N}$ we have $\theta_\bullet^m \in \mathrm{dom}(\phi\nu)$ too and $|\phi\nu(\theta_\bullet) - \phi\nu(\theta_\bullet^m)| = \mathrm{o}(1)$ as $m \to \infty$ (Property §04|03.13). □

§07|01.46 **Assumption**. Consider a noisy version $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet \sim \mathbb{P}_{\theta|\mathfrak{s}}^n$ satisfying Assumption §07|01.01. In addition

(dSIPl1) $\dot{\varepsilon}_\bullet$ admits a covariance operator, say $\Gamma_{\theta|\mathfrak{s}} \in \mathbb{L}^2(\mathbb{J})$, i.e. $\dot{\varepsilon}_\bullet \sim \mathrm{P}_{(0,\Gamma_{\theta|\mathfrak{s}})}$, and

(dSIPl2) $\dot{\varepsilon}\mathbb{1}^m_\bullet \in \mathbb{J}$ $\mathbb{P}^n_{\theta|\mathfrak{s}}$-a.s., for each $m \in \mathbb{N}$. □

§07|01.47 **Comment**. Under Assumption §07|01.46 and $\phi \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ set $(\mathfrak{s}^\dagger\phi)_\bullet := \mathfrak{s}^\dagger_\bullet\phi \in \mathcal{M}(\mathscr{J})$. If $\mathfrak{s}^\dagger_\bullet\mathbb{1}^m_\bullet \in \mathbb{L}_2(\phi^2\nu)$ then we have $\mathfrak{s}^\dagger_\bullet\dot{\varepsilon}_\bullet\mathbb{1}^m_\bullet \in \mathrm{dom}(\phi\nu)$ $\mathbb{P}^n_{\theta|\mathfrak{s}}$-a.s. since $\nu(|(\mathfrak{s}^\dagger\phi)_\bullet\dot{\varepsilon}_\bullet\mathbb{1}^m_\bullet|) \leqslant \|\mathfrak{s}^\dagger_\bullet\mathbb{1}^m_\bullet\|_\phi\|\dot{\varepsilon}_\bullet\mathbb{1}^m_\bullet\|_{\mathbb{J}} \in \mathbb{R}_{\geqslant 0}$ $\mathbb{P}^n_{\theta|\mathfrak{s}}$-a.s.. If in addition $\theta_\bullet \in \mathrm{dom}(\phi\nu)$, and hence $\theta^m_\bullet \in \mathrm{dom}(\phi\nu)$ (Property §04|03.13), then it follows

$$\widehat{\theta}^m_\bullet = \mathfrak{s}^\dagger_\bullet\,\widehat{g}_\bullet\,\mathbb{1}^m_\bullet = n^{-1/2}\mathfrak{s}^\dagger_\bullet\dot{\varepsilon}_\bullet\mathbb{1}^m_\bullet + \theta^m_\bullet \in \mathrm{dom}(\phi\nu) \quad \mathbb{P}^n_{\theta|\mathfrak{s}}\text{-a.s..} \tag{07.14}$$

If $\mathscr{J} \subseteq \mathbb{Z}$ (at most countable) and $\nu_{\mathscr{J}}$ is the counting measure over the index set $\mathscr{J}$ then Assumption §01|01.04 and **(dSIPl1)** $\dot{\varepsilon}_\bullet \sim \mathrm{P}_{(0,\Gamma_{\theta|\mathfrak{s}})}$ implies $\mathbb{v}^{\theta|\mathfrak{s}}_\bullet = \mathbb{P}^n_{\theta|\mathfrak{s}}(|\dot{\varepsilon}_\bullet|^2) \in \mathbb{L}_\infty(\nu_{\mathscr{J}})$ and hence the additional assumption **(dSIPl2)** $\dot{\varepsilon}_\bullet\mathbb{1}^m_\bullet \in \mathbb{J} = \mathbb{L}_2(\nu_{\mathscr{J}})$ $\mathbb{P}^n_{\theta|\mathfrak{s}}$-a.s.. However, the last implication does generally not hold, if $\mathscr{J} \in \{\mathbb{R}, \mathbb{R}_{>0}\}$ for example. □

§07|01|03|01 Local ϕ-risk

§07|01.48 **Assumption**. Let $\phi \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$, $\theta_\bullet \in \mathrm{dom}(\phi\nu)$, and $\mathfrak{s}^\dagger_\bullet\mathbb{1}^m_\bullet \in \mathbb{L}_2(\phi^2\nu)$ for all $m \in \mathbb{N}$ be satisfied. □

§07|01.49 **Definition**. Under Assumptions §07|01.46 and §07|01.48 the *local ϕ-risk* of an OPE $\widehat{\theta}^m_\bullet = \mathfrak{s}^\dagger_\bullet\widehat{g}^m_\bullet = \mathfrak{s}^\dagger_\bullet\widehat{g}_\bullet\mathbb{1}^m_\bullet \in \mathrm{dom}(\phi\nu)$ $\mathbb{P}^n_{\theta|\mathfrak{s}}$-a.s. satisfies

$$\mathbb{P}^n_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^m_\bullet - \theta_\bullet)|^2) = \mathbb{P}^n_{\theta|\mathfrak{s}}(|\phi\nu(\mathfrak{s}^\dagger_\bullet(\widehat{g}_\bullet - g_\bullet)\mathbb{1}^m_\bullet)|^2) + |\phi\nu(\theta_\bullet\mathbb{1}^{m|\perp}_\bullet)|^2. \tag{07.15}$$

with *variance* $\mathbb{P}^n_{\theta|\mathfrak{s}}(|\phi\nu(\mathfrak{s}^\dagger_\bullet(\widehat{g}_\bullet - g_\bullet)\mathbb{1}^m_\bullet)|^2) = n^{-1}\mathrm{P}_{(0,\Gamma_{\theta|\mathfrak{s}})}(|\phi\nu(\mathfrak{s}^\dagger_\bullet\dot{\varepsilon}_\bullet\mathbb{1}^m_\bullet)|^2)$ and *bias* $|\phi\nu(\theta_\bullet\mathbb{1}^{m|\perp}_\bullet)|$. □

§07|01.50 **Property**. *Under Assumptions §07|01.46 and §07|01.48 we have*

$$\begin{aligned}\mathrm{P}_{(0,\Gamma_{\theta|\mathfrak{s}})}(|\phi\nu(\mathfrak{s}^\dagger_\bullet\dot{\varepsilon}_\bullet\mathbb{1}^m_\bullet)|^2) &= \mathrm{P}_{(0,\Gamma_{\theta|\mathfrak{s}})}(|\nu(\dot{\varepsilon}_\bullet(\mathfrak{s}^\dagger\phi)_\bullet\mathbb{1}^m_\bullet)|^2)\\ &= \langle\Gamma_{\theta|\mathfrak{s}}((\mathfrak{s}^\dagger\phi)_\bullet\mathbb{1}^m_\bullet), (\mathfrak{s}^\dagger\phi)_\bullet\mathbb{1}^m_\bullet\rangle_{\mathbb{J}} =: \|(\mathfrak{s}^\dagger\phi)_\bullet\mathbb{1}^m_\bullet\|^2_{\Gamma_{\theta|\mathfrak{s}}}\end{aligned} \tag{07.16}$$

and consequently $\mathbb{P}^n_{\theta|\mathfrak{s}}(|\phi\nu(\mathfrak{s}^\dagger_\bullet(\widehat{g}_\bullet - g_\bullet)\mathbb{1}^m_\bullet)|^2) \leqslant n^{-1}\|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})}\|\mathfrak{s}^\dagger_\bullet\mathbb{1}^m_\bullet\|^2_\phi \in \mathbb{R}_{\geqslant 0}$. □

§07|01.51 **Proposition** (Upper bound). *Under Assumptions §07|01.46 and §07|01.48 for all $m, n \in \mathbb{N}$ setting*

$$\begin{aligned}\mathrm{R}^m_n(\theta_\bullet, \mathfrak{s}_\bullet, \phi) &:= |\phi\nu(\theta_\bullet\mathbb{1}^{m|\perp}_\bullet)|^2 + n^{-1}\|\mathfrak{s}^\dagger_\bullet\mathbb{1}^m_\bullet\|^2_\phi, \quad m^\circ_n := \arg\min\left\{\mathrm{R}^m_n(\theta_\bullet, \mathfrak{s}_\bullet, \phi) : m \in \mathbb{N}\right\}\\ \textit{and}\quad \mathrm{R}^\circ_n(\theta_\bullet, \mathfrak{s}_\bullet, \phi) &:= \mathrm{R}^{m^\circ_n}_n(\theta_\bullet, \mathfrak{s}_\bullet, \phi) := \min\left\{\mathrm{R}^m_n(\theta_\bullet, \mathfrak{s}_\bullet, \phi) : m \in \mathbb{N}\right\}\end{aligned} \tag{07.17}$$

we have $\mathbb{P}^n_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^{m^\circ_n}_\bullet - \theta_\bullet)|^2) \leqslant (1 \vee \|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})})\mathrm{R}^\circ_n(\theta_\bullet, \mathfrak{s}_\bullet, \phi)$ *for all $n \in \mathbb{N}$.*

§07|01.52 **Proof** of Proposition §07|01.51. Given in the lecture. □

§07|01.53 **Definition**. Let $\theta_\bullet \in \mathrm{dom}(\phi\nu)$ and $\widehat{\theta}^m_\bullet \in \mathrm{dom}(\phi\nu)$ $\mathbb{P}^n_{\theta|\mathfrak{s}}$-a.s. for all $m \in \mathbb{N}$. If there exist $\mathrm{C} \in \mathbb{R}_{>0}$ and for each $n \in \mathbb{N}$, $\mathrm{R}^\circ_n \in \mathbb{R}_{>0}$ and $m^\circ_n \in \mathbb{N}$ satisfying

$$\mathrm{C}^{-1}\,\mathrm{R}^\circ_n \leqslant \inf_{m\in\mathbb{N}}\mathbb{P}^n_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^m_\bullet - \theta_\bullet)|^2) \leqslant \mathbb{P}^n_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^{m^\circ_n}_\bullet - \theta_\bullet)|^2) \leqslant \mathrm{C}\;\mathrm{R}^\circ_n \quad \forall n \in \mathbb{N},$$

then we call R°_n *oracle bound*, m°_n *oracle dimension* and $\widehat{\theta}^{m^\circ_n}_\bullet$ *oracle optimal* (up to the constant C). As a consequence, up to the constant C^2 the statistic $\widehat{\theta}^{m^\circ_n}_\bullet$ attains the lower local ϕ-risk bound within the family of OPE's, that is, $\mathbb{P}^n_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^{m^\circ_n}_\bullet - \theta_\bullet)|^2) \leqslant \mathrm{C}^2\inf_{m\in\mathbb{N}}\mathbb{P}^n_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^m_\bullet - \theta_\bullet)|^2)$. □

§07|01.54 **Comment**. If $\Gamma_{\theta|\mathfrak{s}} \in \mathbb{L}^2(\mathbb{J})$ is invertible with inverse $\Gamma_{\theta|\mathfrak{s}}^{-1} \in \mathbb{L}(\mathbb{J})$, i.e. $\Gamma_{\mathfrak{a}|\phi}\Gamma_{\theta|\mathfrak{s}}^{-1} = \mathrm{id}_{\mathbb{J}} = \Gamma_{\mathfrak{a}|\phi}^{-1}\Gamma_{\theta|\mathfrak{s}}$, then we write shortly $1 \leqslant \max(\|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})}, \|\Gamma_{\theta|\mathfrak{s}}^{-1}\|_{\mathbb{L}(\mathbb{J})}) \leqslant \mathbb{v}_{\theta|\mathfrak{s}} \in \mathbb{R}_{\geqslant 1}$. In this situation for all $h_{\bullet} \in \mathbb{J}$ we have $\mathbb{v}_{\theta|\mathfrak{s}}^{-1}\|h_{\bullet}\|_{\mathbb{J}}^2 \leqslant \|h_{\bullet}\|_{\Gamma_{\theta|\mathfrak{s}}}^2 = \langle \Gamma_{\theta|\mathfrak{s}} h_{\bullet}, h_{\bullet}\rangle_{\mathbb{J}} \leqslant \mathbb{v}_{\theta|\mathfrak{s}}\|h_{\bullet}\|_{\mathbb{J}}^2$. □

§07|01.55 **Oracle inequality**. *Under Assumptions §07|01.46 and §07|01.48 if in addition*

$$1 \leqslant \max(\|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})}, \|\Gamma_{\theta|\mathfrak{s}}^{-1}\|_{\mathbb{L}(\mathbb{J})}) \leqslant \mathbb{v}_{\theta|\mathfrak{s}} \in \mathbb{R}_{\geqslant 1}$$

is satisfied then (07.17) *(and Comment §07|01.54) implies*

$$\mathbb{v}_{\theta|\mathfrak{s}}^{-1}\mathrm{R}_n^m(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi) \leqslant \mathbb{P}_{\theta|\mathfrak{s}}^n(|\phi\nu(\widehat{\theta}_{\bullet}^m - \theta_{\bullet})|^2) = n^{-1}\|(\mathfrak{s}^{\dagger}\phi)_{\bullet}\mathbb{1}_{\bullet}^m\|_{\Gamma_{\theta|\mathfrak{s}}}^2 + |\phi\nu(\theta_{\bullet}\mathbb{1}_{\bullet}^{m|\perp})|^2$$
$$\leqslant \mathbb{v}_{\theta|\mathfrak{s}}\mathrm{R}_n^m(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi) \quad \forall m, n \in \mathbb{N}.$$

As a consequence we immediately obtain the following oracle inequality

$$\mathbb{v}_{\theta|\mathfrak{s}}^{-1}\mathrm{R}_n^{\circ}(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi) \leqslant \inf_{m\in\mathbb{N}} \mathbb{P}_{\theta|\mathfrak{s}}^n(|\phi\nu(\widehat{\theta}_{\bullet}^m - \theta_{\bullet})|^2) \leqslant \mathbb{P}_{\theta|\mathfrak{s}}^n(|\phi\nu(\widehat{\theta}_{\bullet}^{m_n^{\circ}} - \theta_{\bullet})|^2)$$
$$\leqslant \mathbb{v}_{\theta|\mathfrak{s}}\mathrm{R}_n^{\circ}(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi) \leqslant \mathbb{v}_{\theta|\mathfrak{s}}^2 \inf_{m\in\mathbb{N}} \mathbb{P}_{\theta|\mathfrak{s}}^n(|\phi\nu(\widehat{\theta}_{\bullet}^m - \theta_{\bullet})|^2) \quad \forall n \in \mathbb{N}, \quad (07.18)$$

and hence $\mathrm{R}_n^{\circ}(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi)$, m_n° and the statistic $\widehat{\theta}_{\bullet}^{m_n^{\circ}}$, respectively, is an oracle bound, an oracle dimension and oracle optimal (up to the constant $\mathbb{v}_{\theta|\mathfrak{s}}^2$). □

§07|01.56 **Remark**. Arguing similarly as in Remark §07|01.21 we note that $\mathrm{R}_n^{\circ}(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi) = \mathrm{o}(1)$ as $n \to \infty$, whenever $\|\mathfrak{s}_{\bullet}^{\dagger}\mathbb{1}_{\bullet}^m\|_{\phi}^2 \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$ and $|\phi\nu(\theta_{\bullet}\mathbb{1}_{\bullet}^{m|\perp})| = \mathrm{o}(1)$ as $m \to \infty$. The latter is satisfied, for example, if $\theta_{\bullet} = \mathfrak{s}_{\bullet}^{\dagger}g_{\bullet} \in \mathrm{dom}(\phi\nu)$. The oracle dimension $m_n^{\circ} = m_n^{\circ}(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi)$ as defined in (§07|01.51) depends again on the unknown parameter of interest $\theta_{\bullet}$ and thus also the oracle optimal statistic $\widehat{\theta}_{\bullet}^{m_n^{\circ}}$. In other words $\widehat{\theta}_{\bullet}^{m_n^{\circ}}$ is not a feasible estimator. □

§07|01.57 **Corollary** (GdiSM §07|01.03 continued). *Consider $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}_{\theta|\mathfrak{s}}^n$ as in Model §07|01.03, where $\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}$, $\mathfrak{s}_{\bullet} \in \ell_{\infty}$, $\theta_{\bullet} \in \ell_2$ and hence $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^{\dagger}}) \subseteq \ell_2$. Given $\phi \in \mathbb{R}_{\setminus 0}^{\mathbb{N}}$ and $\theta_{\bullet} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ the (infeasible) OPE $\widehat{\theta}_{\bullet}^{m_n^{\circ}} = \mathfrak{s}_{\bullet}^{\dagger}\widehat{g}_{\bullet}\mathbb{1}_{\bullet}^{m_n^{\circ}} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ with oracle dimension m_n° as in* (07.17) *satisfies*

$$\mathrm{N}_{\theta|\mathfrak{s}}^n(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_{\bullet}^{m_n^{\circ}} - \theta_{\bullet})|^2) = \mathrm{R}_n^{\circ}(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi) = \inf_{m\in\mathbb{N}} \mathrm{N}_{\theta|\mathfrak{s}}^n(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_{\bullet}^m - \theta_{\bullet})|^2),$$

and hence it is oracle optimal (with constant 1).

§07|01.58 **Proof** of Corollary §07|01.57. Given in the lecture. □

§07|01.59 **Corollary** (diSM §07|01.06 continued). *Consider $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} \sim \mathrm{P}_{\theta|\mathfrak{s}|\sigma}^n$ as in Model §07|01.06, where $\dot{\varepsilon}_{\bullet} \sim \otimes_{j\in\mathbb{N}}\mathrm{P}_{(0,\sigma_j^2)}$ satisfies* **(iSM1)** *and* **(iSM2)** *with $\max(\|\sigma_{\bullet}^{-2}\|_{\ell_{\infty}}, \|\sigma_{\bullet}^2\|_{\ell_{\infty}}) =: \mathbb{v}_{\sigma} \in \mathbb{R}_{\geqslant 1}$, $\mathfrak{s}_{\bullet} \in \ell_{\infty}$, $\theta_{\bullet} \in \ell_2$ and hence $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^{\dagger}}) \subseteq \ell_2$. Given $\phi \in \mathbb{R}_{\setminus 0}^{\mathbb{N}}$ and $\theta_{\bullet} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ the (infeasible) OPE $\widehat{\theta}_{\bullet}^{m_n^{\circ}} = \mathfrak{s}_{\bullet}^{\dagger}\widehat{g}_{\bullet}\mathbb{1}_{\bullet}^{m_n^{\circ}} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ with oracle dimension m_n° as in* (07.17) *satisfies*

$$\mathrm{P}_{\theta|\mathfrak{s}|\sigma}^n(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_{\bullet}^{m_n^{\circ}} - \theta_{\bullet})|^2) \leqslant \mathbb{v}_{\sigma}\mathrm{R}_n^{\circ}(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi) \leqslant \mathbb{v}_{\sigma}^2 \inf_{m\in\mathbb{N}} \mathrm{P}_{\theta|\mathfrak{s}|\sigma}^n(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_{\bullet}^m - \theta_{\bullet})|^2),$$

and hence it is oracle optimal (with constant $\mathbb{v}_{\sigma}$).

§07|01.60 **Proof** of Corollary §07|01.59. Given in the lecture. □

§07|01.61 **Corollary** (dieMM §07|01.08 continued). *Let* $\widehat{g} = g + n^{-1/2}\dot{\varepsilon}$ *be defined on* $(\mathcal{Z}^n, \mathscr{Z}^{\otimes n}, \mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}})$ *as in Model §07|01.08, where* $\psi \in \mathcal{M}(\mathscr{X}\otimes\mathscr{J})$ *satisfies* **(dieMM1)–(dieMM4)** *for some* $\mathrm{v}_{\theta|\mathfrak{s}|\psi} \in \mathbb{R}_{\geqslant 1}$, $\mathfrak{s} \in \mathbb{L}_\infty(\nu)$, $\theta \in \mathbb{J}$ *and hence* $g = \mathfrak{s}\theta \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}}) \subseteq \mathbb{J}$. *Under Assumption §07|01.48 the (infeasible) OPE* $\widehat{\theta}^{m_n^\circ} = \mathfrak{s}^\dagger\widehat{g}\mathbb{1}^{m_n^\circ} \in \mathrm{dom}(\phi\nu)$ *with oracle dimension* m_n° *as in* (07.17) *satisfies*

$$\mathbb{E}^n_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^{m_n^\circ} - \theta)|^2) \leqslant \mathrm{v}_{\theta|\mathfrak{s}|\psi}\mathrm{R}^\circ_n(\theta,\mathfrak{s},\phi) \leqslant \mathrm{v}^2_{\theta|\mathfrak{s}|\psi} \inf_{m\in\mathbb{N}} \mathbb{E}^n_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^m - \theta)|^2),$$

and hence it is oracle optimal (with constant $\mathrm{v}_{\theta|\mathfrak{s}|\psi}$*).*

§07|01.62 **Proof** of Corollary §07|01.61. Given in the lecture. □

§07|01.63 **Illustration**. We illustrate the last results considering usual behaviour for $\theta, \mathfrak{s}, \phi \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$. Similar to the two cases **(p)** and **(np)** in Illustration §07|01.28 we distinguish here the following two cases

(p) $\mathfrak{s}^\dagger \in \mathbb{L}_2(\phi^2\nu)$ or there is $K \in \mathbb{N}$ with $\sup\{|\phi\nu(\theta\mathbb{1}^{m|\perp})|^2 : m \in \mathbb{N}_{\geqslant K}\} = 0$,

(np) $\mathfrak{s}^\dagger \notin \mathbb{L}_2(\phi^2\nu)$ and for all $m \in \mathbb{N}$ holds $\sup\{|\phi\nu(\theta\mathbb{1}^{m|\perp})|^2 : m \in \mathbb{N}_{\geqslant K}\} \in \mathbb{R}_{>0}$.

In case **(p)** the oracle bound is again parametric, i.e. $n\mathrm{R}^\circ_n(\theta,\mathfrak{s},\phi) = \mathrm{O}(1)$, while in case **(np)** the oracle bound is nonparametric, i.e. $\lim_{n\to\infty} n\mathrm{R}^\circ_n(\theta,\mathfrak{s},\phi) = \infty$. In case **(np)** consider the following three specifications:

Table 03 [§07]

Order of the oracle rate $\mathrm{R}^\circ_n(\theta,\mathfrak{s},\phi)$ as $n\to\infty$

	$(j\in\mathcal{J})$ $\phi_j = j^{\mathrm{v}-1/2}$	$(\mathrm{a}\in\mathbb{R}_{>0})$ θ_j	$(\mathrm{t}\in\mathbb{R}_{>0})$ $\mathfrak{s}^2_j$	(squared bias) $\lvert\phi\nu(\theta\mathbb{1}^{m\lvert\perp})\rvert^2$	(variance) $\lVert\mathfrak{s}^\dagger\mathbb{1}^m\rVert^2_\phi$	m°_n	$\mathrm{R}^\circ_n(\theta,\mathfrak{s},\phi)$
(o-m)	$\mathrm{v}\in(-\mathrm{t},\mathrm{a})$	$j^{-\mathrm{a}-1/2}$	$j^{-2\mathrm{t}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2\mathrm{v}+2\mathrm{t}}$	$n^{\frac{1}{2\mathrm{a}+2\mathrm{t}}}$	$n^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{a}+\mathrm{t}}}$
	$\mathrm{v}=-\mathrm{t}$	$j^{-\mathrm{a}-1/2}$	$j^{-2\mathrm{t}}$	$m^{-2(\mathrm{a}+\mathrm{t})}$	$\log m$	$\left(\frac{n}{\log n}\right)^{\frac{1}{2(\mathrm{a}+\mathrm{t})}}$	$\frac{\log n}{n}$
(o-s)	$\mathrm{a}-\mathrm{v}\in\mathbb{R}_{>0}$	$j^{-\mathrm{a}-1/2}$	$e^{-j^{2\mathrm{t}}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2(\mathrm{v}-\mathrm{t})_+}e^{m^{2\mathrm{t}}}$	$(\log n)^{\frac{1}{2\mathrm{t}}}$	$(\log n)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v}+\mathrm{t}\in\mathbb{R}_{>0}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$m^{(1-4\mathrm{a}+2\mathrm{v})_+}e^{-2m^{2\mathrm{a}}}$	$m^{2\mathrm{v}+2\mathrm{t}}$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{(\log n)^{\frac{\mathrm{t}+\mathrm{v}}{\mathrm{a}}}}{n}$
	$\mathrm{v}=-\mathrm{t}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$m^{(1-4\mathrm{a}-2\mathrm{t}))_+}e^{-2m^{2\mathrm{a}}}$	$\log m$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{\log\log n}{n}$

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v} < -\mathrm{t}$ the oracle rate $\mathrm{R}^\circ_n(\theta,\mathfrak{s},\phi)$ is parametric. □

§07|01|03|02 Maximal local ϕ-risk

§07|01.64 **Assumption**. Consider weights $\mathfrak{a}, \mathfrak{t} \in \mathcal{M}_{>0,\nu}(\mathscr{J})$ and $\phi \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ (i.e. $\nu(\mathcal{N}_{\mathfrak{a}}) = \nu(\mathcal{N}_{\mathfrak{t}}) = 0 = \nu(\mathcal{N}_\phi)$), such that $\mathfrak{a}, \mathfrak{t} \in \mathbb{L}_\infty(\nu)$, $\mathfrak{a} \in \mathbb{L}_2(\phi^2\nu)$, and $\mathfrak{t}^\dagger\mathbb{1}^m \in \mathbb{L}_2(\phi^2\nu)$ for all $m \in \mathbb{N}$. □

§07|01.65 **Reminder**. Under Assumption §07|01.64 we have $\mathbb{J}^{\mathfrak{a}} = \mathbb{L}^{\mathfrak{a}}_2(\nu) = \mathrm{dom}(\mathrm{M}_{\mathfrak{a}^\dagger}) = \mathbb{J}\mathfrak{a} \subseteq \mathbb{J}$ and the three measures ν, $\mathfrak{a}^{2|\dagger}\nu$ and $|\phi|\nu$ dominate mutually each other, i.e. they share the same null sets (see Property §04|01.02). We consider $\mathbb{J}^{\mathfrak{a}}$ endowed with $\|\cdot\|_{\mathfrak{a}^\dagger} = \|\mathrm{M}_{\mathfrak{a}^\dagger}\cdot\|_{\mathbb{J}}$ and given a constant $\mathrm{r} \in \mathbb{R}_{>0}$ the ellipsoid $\mathbb{J}^{\mathfrak{a},\mathrm{r}} := \{h \in \mathbb{J}^{\mathfrak{a}} : \|h\|_{\mathfrak{a}^\dagger} \leqslant \mathrm{r}\} \subseteq \mathbb{J}^{\mathfrak{a}}$. Since $\mathfrak{a} \in \mathbb{L}_2(\phi^2\nu)$, and hence $\|\mathfrak{a}\mathbb{1}^{m|\perp}\|_\phi = \|(\mathfrak{a}\phi)\mathbb{1}^{m|\perp}\|_{\mathbb{J}} \in \mathbb{R}_{\geqslant 0}$ for each $m \in \mathbb{N}$ ($\|\mathfrak{a}\mathbb{1}^{m|\perp}\|_\phi = \mathrm{o}(1)$ as $m\to\infty$ by dominated convergence) we have $\mathbb{J}^{\mathfrak{a}} \subseteq \mathrm{dom}(\phi\nu)$ (Property §04|02.23), and $|\phi\nu(\theta\mathbb{1}^{m|\perp})| \leqslant \mathrm{r}\,\|\mathfrak{a}\mathbb{1}^{m|\perp}\|_\phi$ for all $\theta \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ (Lemma §04|02.25). Consequently, if Assumption §07|01.64, $\theta \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ and $\mathfrak{s}^\dagger\mathbb{1}^m \in \mathbb{L}_2(\phi^2\nu)$ for all $m \in \mathbb{N}$ are satisfied, then Assumption §07|01.48 is fulfilled. Moreover, under Assumption §07|01.64 for each $\mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ we have $\|\mathfrak{s}^\dagger\mathbb{1}^m\|_\phi \leqslant \mathrm{d}\|\mathfrak{t}^\dagger\mathbb{1}^m\|_\phi \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$ (Definition §04|03.05). Therefore, if Assumption §07|01.64, $\theta \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ and $\mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ are satisfied, then Assumption §07|01.48 is also fulfilled. □

§07|01.66 **Proposition** (Upper bound). *Under Assumptions §07|01.46 and §07|01.64 let* $\mathfrak{s}^{\dagger}_{\bullet}\mathbb{1}^{m}_{\bullet} \in \mathbb{L}_2(\phi^2\nu)$ *for all* $m \in \mathbb{N}$. *Setting for* $n, m \in \mathbb{N}$

$$\mathrm{R}^{m}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \phi) := \|\mathfrak{a}_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|^2_{\phi} + n^{-1}\|\mathfrak{s}^{\dagger}_{\bullet}\mathbb{1}^{m}_{\bullet}\|^2_{\phi}, \quad m^{\star}_{n} := \arg\min\left\{\mathrm{R}^{m}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \phi) : m \in \mathbb{N}\right\}$$
$$\text{and} \quad \mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \phi) := \mathrm{R}^{m^{\star}_{n}}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \phi) = \min\left\{\mathrm{R}^{m}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \phi) : m \in \mathbb{N}\right\} \quad (07.19)$$

and $\|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})} =: \mathrm{v}_{\theta|\mathfrak{s}} \in \mathbb{R}_{\geqslant 0}$, *for all* $\theta_{\bullet} = \mathfrak{s}^{\dagger}_{\bullet} g_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ *we have*

$$\mathbb{E}^{n}_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^{m^{\star}_{n}}_{\bullet} - \theta_{\bullet})|^2) \leqslant (\mathrm{v}_{\theta|\mathfrak{s}} \vee \mathrm{r}^2)\, \mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \phi) \quad \forall n \in \mathbb{N}.$$

§07|01.67 **Proof** of Proposition §07|01.66. Given in the lecture. □

§07|01.68 **Remark**. Under the assumptions of Proposition §07|01.66 if there exists in addition $\mathrm{v}_{\mathfrak{s}} \in \mathbb{R}_{\geqslant 0}$ satisfying $\|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})} \leqslant \mathrm{v}_{\mathfrak{s}}$ for all $\theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ then

$$\sup\left\{\mathbb{E}^{n}_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^{m^{\star}_{n}}_{\bullet} - \theta_{\bullet})|^2) \colon \theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}\right\} \leqslant (\mathrm{v}_{\mathfrak{s}} \vee \mathrm{r}^2)\, \mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \phi) \quad \forall n \in \mathbb{N}.$$

Arguing similarly as in Remark §07|01.21 we note that $\mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \phi) = \mathrm{o}(1)$ as $n \to \infty$, whenever $\|\mathfrak{s}^{\dagger}_{\bullet}\mathbb{1}^{m}_{\bullet}\|^2_{\phi} \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$ and $\|\mathfrak{a}_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\phi} = \mathrm{o}(1)$ as $m \to \infty$. The latter is satisfied since $\mathfrak{a}_{\bullet} \in \mathbb{L}_2(\phi^2\nu)$ by Assumption §07|01.64. Note that the dimension $m^{\star}_{n} := m^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{s}_{\bullet}, \phi)$ as defined in (07.19) does not depend on the unknown parameter of interest $\theta_{\bullet}$ but on the class $\mathbb{J}^{\mathfrak{a},\mathrm{r}}$ only, and thus also the statistic $\widehat{\theta}^{m^{\star}_{n}}_{\bullet}$. In other words, if the regularity of $\theta_{\bullet}$ is known in advance, then the OPE $\widehat{\theta}^{m^{\star}_{n}}_{\bullet}$ is a feasible estimator. □

§07|01.69 **Corollary** (Upper bound). *Under Assumptions §07|01.46 and §07|01.64 setting for* $n, m \in \mathbb{N}$

$$\mathrm{R}^{m}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi) := \|\mathfrak{a}_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|^2_{\phi} + n^{-1}\|\mathfrak{t}^{\dagger}_{\bullet}\mathbb{1}^{m}_{\bullet}\|^2_{\phi}, \quad m^{\star}_{n} := \arg\min\left\{\mathrm{R}^{m}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi) : m \in \mathbb{N}\right\}$$
$$\text{and} \quad \mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi) := \mathrm{R}^{m^{\star}_{n}}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi) = \min\left\{\mathrm{R}^{m}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi) : m \in \mathbb{N}\right\} \quad (07.20)$$

and $\|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})} =: \mathrm{v}_{\theta|\mathfrak{s}} \in \mathbb{R}_{\geqslant 0}$, *for each* $\mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ known *in advance, for all* $\theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$, *hence* $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^{\dagger}}) \subseteq \mathbb{J}$, *we have*

$$\mathbb{E}^{n}_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^{m^{\star}_{n}}_{\bullet} - \theta_{\bullet})|^2) \leqslant (\mathrm{d}^2\mathrm{v}_{\theta|\mathfrak{s}} \vee \mathrm{r}^2)\, \mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi) \quad \forall n \in \mathbb{N}.$$

§07|01.70 **Proof** of Corollary §07|01.69. Given in the lecture. □

§07|01.71 **Remark**. Under the assumptions of Corollary §07|01.69 if there exists in addition $\mathrm{v} \in \mathbb{R}_{\geqslant 0}$ satisfying $\|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})} \leqslant \mathrm{v}$ for all $\theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ and $\mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ then

$$\sup\left\{\mathbb{E}^{n}_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^{m^{\star}_{n}}_{\bullet} - \theta_{\bullet})|^2) \colon \theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}\right\} \leqslant (\mathrm{v}\mathrm{d}^2 \vee \mathrm{r}^2)\, \mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi) \quad \forall n \in \mathbb{N}.$$

Arguing similarly as in Remark §07|01.21 we note that $\mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi) = \mathrm{o}(1)$ as $n \to \infty$ since $\|\mathfrak{t}^{\dagger}_{\bullet}\mathbb{1}^{m}_{\bullet}\|_{\phi} \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$ and $\|\mathfrak{a}_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\phi} = \mathrm{o}(1)$ as $m \to \infty$ by Assumption §07|01.64. Note that the dimension $m^{\star}_{n} := m^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi)$ as defined in (07.20) does neither depend on the unknown parameter of interest $\theta_{\bullet}$ nor on the known multiplication operator $\mathrm{M}_{\mathfrak{s}}$ but on the classes $\mathbb{J}^{\mathfrak{a},\mathrm{r}}$ and $\mathbb{M}_{\mathfrak{t},\mathrm{d}}$ only, and thus also the statistic $\widehat{\theta}^{m^{\star}_{n}}_{\bullet}$. In other words, if the regularity of $\theta_{\bullet}$ is known in advance, then the OPE $\widehat{\theta}^{m^{\star}_{n}}_{\bullet}$ is a feasible estimator. □

§07|01.72 **Corollary** (GdiSM §07|01.03 continued). *Consider* $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}^{n}_{\theta|\mathfrak{s}}$ *as in Model §07|01.03, where* $\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$, $\mathfrak{s}_{\bullet} \in \ell_{\infty}$, $\theta_{\bullet} \in \ell_2$ *and hence* $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^{\dagger}}) \subseteq \ell_2$. *Under Assumption §07|01.64 the OPE* $\widehat{\theta}^{m^{\star}_{n}}_{\bullet} = \mathfrak{s}^{\dagger}_{\bullet}\widehat{g}_{\bullet}\mathbb{1}^{m^{\star}_{n}}_{\bullet} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *satisfies*

(i) *with dimension* $m_n^\star = m_n^\star(\mathfrak{a}_\bullet, \mathfrak{s}_\bullet, \phi)$ *as in* (07.19) *and constant* $\mathrm{C}_{\mathrm{r}} = 1 \vee \mathrm{r}^2$

$$\sup\left\{\mathrm{N}^n_{\theta|\mathfrak{s}}\big(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_\bullet^{m_n^\star} - \theta_\bullet)|^2\big) \colon \theta_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}\right\} \leqslant \mathrm{C}_{\mathrm{r}}\, \mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{s}_\bullet, \phi) \quad \forall n \in \mathbb{N} \tag{07.21}$$

(ii) *with dimension* $m_n^\star = m_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi)$ *as in* (07.20) *and constant* $\mathrm{C}_{\mathrm{r,d}} = \mathrm{d}^2 \vee \mathrm{r}^2$

$$\sup\left\{\mathrm{N}^n_{\theta|\mathfrak{s}}\big(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_\bullet^{m_n^\star} - \theta_\bullet)|^2\big) \colon \theta_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}\right\} \leqslant \mathrm{C}_{\mathrm{r,d}}\, \mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi) \quad \forall n \in \mathbb{N}. \tag{07.22}$$

§07|01.73 **Proof** of Corollary §07|01.72. Given in the lecture. □

§07|01.74 **Corollary** (diSM §07|01.06 continued). *Consider* $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet \sim \mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}$ *as in Model §07|01.06, where* $\dot{\varepsilon}_\bullet \sim \mathrm{P}_{(0,\mathrm{M}_{\sigma^2})}$ *satisfies* **(iSM1)** *with* $\|\sigma^2_\bullet\|_{\ell_\infty} =: \mathrm{v}_\sigma \in \mathbb{R}_{>0}$, $\mathfrak{s}_\bullet \in \ell_\infty$, $\theta_\bullet \in \ell_2$ *and hence* $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \subseteq \ell_2$. *Under Assumption §07|01.64 the OPE* $\widehat{\theta}_\bullet^{m_n^\star} = \mathfrak{s}^\dagger_\bullet \widehat{g}_\bullet \mathbb{1}_\bullet^{m_n^\star} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *satisfies*

(i) *with dimension* $m_n^\star = m_n^\star(\mathfrak{a}_\bullet, \mathfrak{s}_\bullet, \phi)$ *as in* (07.19) *and constant* $\mathrm{C}_{\mathrm{r},\sigma} = \mathrm{v}_\sigma \vee \mathrm{r}^2$

$$\sup\left\{\mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}\big(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_\bullet^{m_n^\star} - \theta_\bullet)|^2\big) \colon \theta_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}\right\} \leqslant \mathrm{C}_{\mathrm{r},\sigma}\, \mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{s}_\bullet, \phi) \quad \forall n \in \mathbb{N} \tag{07.23}$$

(ii) *with dimension* $m_n^\star = m_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi)$ *as in* (07.20) *and constant* $\mathrm{C}_{\mathrm{r,d},\sigma} = \mathrm{v}_\sigma \mathrm{d}^2 \vee \mathrm{r}^2$

$$\sup\left\{\mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}\big(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_\bullet^{m_n^\star} - \theta_\bullet)|^2\big) \colon \theta_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}\right\} \leqslant \mathrm{C}\, \mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi) \quad \forall n \in \mathbb{N}. \tag{07.24}$$

§07|01.75 **Proof** of Corollary §07|01.74. Given in the lecture. □

§07|01.76 **Corollary** (dieMM §07|01.08 continued). *Let* $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ *be defined on* $(\mathcal{Z}^n, \mathscr{Z}^{\otimes n}, \mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}})$ *as in Model §07|01.08, where* $\psi_\bullet \in \mathcal{M}(\mathscr{X} \otimes \mathscr{J})$ *satisfies* **(dieMM1)–(dieMM3)** *for some* $\mathrm{v}_{\theta|\mathfrak{s}|\psi} \in \mathbb{R}_{\geqslant 1}$, $\mathfrak{s}_\bullet \in \mathbb{L}_\infty(\nu)$, $\theta_\bullet \in \mathbb{J}$ *and hence* $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \subseteq \mathbb{J}$. *Under Assumption §07|01.64 the OPE* $\widehat{\theta}_\bullet^{m_n^\star} = \mathfrak{s}^\dagger_\bullet \widehat{g}_\bullet \mathbb{1}_\bullet^{m_n^\star} \in \mathrm{dom}(\phi\nu)$ $\mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}}$-a.s. *satisfies*

(i) *with constant*

$$\mathrm{C}_{\mathfrak{a},\mathrm{r},\mathfrak{s}} := \sup\left\{\mathrm{v}_{\theta|\mathfrak{s}|\psi} \colon \theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}\right\} \vee \mathrm{r}^2$$

and with dimension $m_n^\star = m_n^\star(\mathfrak{a}_\bullet, \mathfrak{s}_\bullet, \phi)$ *as in* (07.19)

$$\sup\left\{\mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}}\big(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_\bullet^{m_n^\star} - \theta_\bullet)|^2\big) \colon \theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}\right\} \leqslant \mathrm{C}_{\mathfrak{a},\mathrm{r},\mathfrak{s}}\, \mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{s}_\bullet, \phi) \quad \forall n \in \mathbb{N} \tag{07.25}$$

provided $\mathfrak{s}^\dagger_\bullet \mathbb{1}^m_\bullet \in \mathbb{L}_2(\phi^2\nu)$ *for all* $m \in \mathbb{N}$;

(ii) *with constant*

$$\mathrm{C}_{\mathfrak{a},\mathrm{r},\mathfrak{t},\mathrm{d}} := \mathrm{d}^2 \sup\left\{\mathrm{v}_{\theta|\mathfrak{s}|\psi} \colon \theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}\right\} \vee \mathrm{r}^2$$

and dimension $m_n^\star = m_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi)$ *as in* (07.20)

$$\sup\left\{\mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}}\big(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_\bullet^{m_n^\star} - \theta_\bullet)|^2\big) \colon \theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}\right\} \leqslant \mathrm{C}_{\mathfrak{a},\mathrm{r},\mathfrak{t},\mathrm{d}}\, \mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi) \quad \forall n \in \mathbb{N}. \tag{07.26}$$

§07|01.77 **Proof** of Corollary §07|01.76. Given in the lecture. □

§07|01.78 **Illustration**. We illustrate the last results considering usual behaviour for $\mathfrak{a}_\bullet, \mathfrak{s}_\bullet, \mathfrak{t}_\bullet, \phi \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ and $\mathrm{w}_\bullet \in \{\mathfrak{s}_\bullet, \mathfrak{t}_\bullet\}$. We distinguish the following two cases **(p)** $\mathrm{w}^\dagger_\bullet \in \mathbb{L}_2(\phi^2\nu)$, and **(np)** $\mathrm{w}^\dagger_\bullet \notin \mathbb{L}_2(\phi^2\nu)$. Interestingly, in case **(p)** the bound in Proposition §07|01.66 is parametric, that is, $n\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathrm{w}_\bullet, \phi) = \mathrm{O}(1)$, in case **(np)** the bound is nonparametric, i.e. $\lim_{n\to\infty} n\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathrm{w}_\bullet, \phi) = \infty$. In case **(np)** consider the following three specifications:

Table 04 [§07]

Order of the rate $\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathrm{w}_\bullet, \phi_\bullet)$ as $n \to \infty$

	$(j \in \mathcal{J})$ $\phi_j^2 = j^{2\mathrm{v}-1}$	$(\mathrm{a} \in \mathbb{R}_{>0})$ $\mathfrak{a}_j^2$	$(\mathrm{t} \in \mathbb{R}_{>0})$ w_j^2	(squared bias) $\|\mathfrak{a}_\bullet \mathbb{1}_\bullet^{m\mid\perp}\|_\phi^2$	(variance) $\|\mathrm{w}_\bullet^\dagger \mathbb{1}_\bullet^m\|_\phi^2$	$m_n^\star$	$\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathrm{w}_\bullet, \phi_\bullet)$
(o-m)	$\mathrm{v} \in (-\mathrm{t}, \mathrm{a})$	$j^{-2\mathrm{a}}$	$j^{-2\mathrm{t}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2\mathrm{v}+2\mathrm{t}}$	$n^{\frac{1}{2\mathrm{a}+2\mathrm{t}}}$	$n^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{a}+\mathrm{t}}}$
	$\mathrm{v} = -\mathrm{t}$	$j^{-2\mathrm{a}}$	$j^{-2\mathrm{t}}$	$m^{-2(\mathrm{a}+\mathrm{t})}$	$\log m$	$\left(\frac{n}{\log n}\right)^{\frac{1}{2(\mathrm{a}+\mathrm{t})}}$	$\frac{\log n}{n}$
(o-s)	$\mathrm{a} - \mathrm{v} \in \mathbb{R}_{>0}$	$j^{-2\mathrm{a}}$	$e^{-j^{2\mathrm{t}}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2(\mathrm{v}-\mathrm{t})_+} e^{m^{2\mathrm{t}}}$	$(\log n)^{\frac{1}{2\mathrm{t}}}$	$(\log n)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v} + \mathrm{t} \in \mathbb{R}_{>0}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$e^{-m^{2\mathrm{a}}}$	$m^{2\mathrm{v}+2\mathrm{t}}$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{(\log n)^{\frac{\mathrm{t}+\mathrm{v}}{\mathrm{a}}}}{n}$
	$\mathrm{v} = -\mathrm{t}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$e^{-m^{2\mathrm{a}}}$	$\log m$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{\log\log n}{n}$

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v} < -\mathrm{t}$ the rate $\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathrm{w}_\bullet, \phi_\bullet)$ is parametric. □

§07|02 Diagonal statistical inverse problem with noisy operator

§07|02.01 **Assumption**. Consider stochastic processes $\dot{\varepsilon}_\bullet = (\dot{\varepsilon}_j)_{j\in\mathcal{J}}$ and $\dot{\eta}_\bullet = (\dot{\eta}_j)_{j\in\mathcal{J}}$ on a probability space $(\Omega, \mathscr{A}, \mathbb{P})$ satisfying Assumption §01|01.04 (i.e. $\dot{\varepsilon}_\bullet, \dot{\eta}_\bullet \in \mathcal{M}(\mathscr{A} \otimes \mathscr{J})$) with *mean zero* (i.e. $\mathbb{P}(\dot{\varepsilon}_\bullet) = 0_\bullet = \mathbb{P}(\dot{\eta}_\bullet)$), sample sizes $n, k \in \mathbb{N}$ and let Assumption §07|00.02 and in addition $\mathfrak{s}_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$ be satisfied where $\mathfrak{s}_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$ is *not known* anymore. The observable noisy image and operator, respectively, has mean $g_\bullet = \mathfrak{s}_\bullet \theta_\bullet \in \mathbb{J} = \mathbb{L}_2(\nu)$ and mean-function $\mathfrak{s}_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$, and takes the form $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ and $\widehat{\mathfrak{s}}_\bullet = \mathfrak{s}_\bullet + k^{-1/2}\dot{\eta}_\bullet$. We denote by $\mathbb{P}^{n,k}_{\theta|\mathfrak{s}}$ the joint distribution of $(\widehat{g}_\bullet, \widehat{\mathfrak{s}}_\bullet)$. Denoting by $\mathbb{P}^n_{\theta|\mathfrak{s}}$ and $\mathbb{P}^k_{\mathfrak{s}}$ the marginal distribution of $\widehat{g}_\bullet$ and $\widehat{\mathfrak{s}}_\bullet$, respectively, if $\dot{\varepsilon}_\bullet$ and $\dot{\eta}_\bullet$ are *independent* then we write $\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}} = \mathbb{P}^n_{\theta|\mathfrak{s}} \otimes \mathbb{P}^k_{\mathfrak{s}}$ for the joint product distribution of $(\widehat{g}_\bullet, \widehat{\mathfrak{s}}_\bullet)$. □

§07|02.02 **Comment**. We restrict ourselves in this section to the case $\mathfrak{s}_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ only, which ensure identification of the solution $\theta_\bullet$ of the equation $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet$. □

§07|02.03 **Notation**. Introduce the random index set $\{\widehat{\mathfrak{s}}_\bullet^2 \geqslant k^{-1}\} := \{j \in \mathcal{J} : k\widehat{\mathfrak{s}}_j^2 \geqslant 1\} \in \mathscr{J}$, for each $j \in \mathcal{J}$ the elementary random variable $\mathbb{1}_j^{\{\widehat{\mathfrak{s}}^2 \geqslant k^{-1}\}}$ taking the value one on the event $\{\widehat{\mathfrak{s}}_j^2 \geqslant k^{-1}\}$ and zero otherwise, and the stochastic process $\mathbb{1}_\bullet^{\{\widehat{\mathfrak{s}}^2 \geqslant k^{-1}\}} := (\mathbb{1}_j^{\{\widehat{\mathfrak{s}}^2 \geqslant k^{-1}\}})_{j\in\mathcal{J}} \in \mathcal{M}(\mathscr{A} \otimes \mathscr{J})$ satisfying hence Assumption §01|01.04. Furthermore, we define $\widehat{\mathfrak{s}}_\bullet^{(k)} := \widehat{\mathfrak{s}}_\bullet \mathbb{1}_\bullet^{\{\widehat{\mathfrak{s}}^2 \geqslant k^{-1}\}}$ and denote its Moore-Penrose inverse by $\widehat{\mathfrak{s}}_\bullet^{(k)|\dagger} = \widehat{\mathfrak{s}}_\bullet^{-1} \mathbb{1}_\bullet^{\{\widehat{\mathfrak{s}}^2 \geqslant k^{-1}\}}$. We eventually use the elementary identity $\widehat{\mathfrak{s}}_\bullet^{(k)} \widehat{\mathfrak{s}}_\bullet^{(k)|\dagger} = \mathbb{1}_\bullet^{\{\widehat{\mathfrak{s}}^2 \geqslant k^{-1}\}} = \widehat{\mathfrak{s}}_\bullet^{(k)|\dagger} \widehat{\mathfrak{s}}_\bullet^{(k)}$ and the upper bound $\|\widehat{\mathfrak{s}}_\bullet^{(k)|\dagger}\|_{\mathbb{L}_\infty(\nu)} \leqslant k^{1/2}$. □

§07|02.04 **Definition**. Under Assumption §07|02.01 for $\theta_\bullet \in \mathbb{J}$ let $(\widehat{g}_\bullet, \widehat{\mathfrak{s}}_\bullet) \sim \mathbb{P}^{n,k}_{\theta|\mathfrak{s}}$ be noisy versions of $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger})$ and $\mathfrak{s}_\bullet \in \mathbb{L}_\infty(\nu)$. For each $m \in \mathbb{N}$ we call $\widehat{\theta}_\bullet^m := \widehat{\mathfrak{s}}_\bullet^{(k)|\dagger} \widehat{g}_\bullet^m = \widehat{\mathfrak{s}}_\bullet^\dagger \mathbb{1}_\bullet^{\{\widehat{\mathfrak{s}}^2 \geqslant k^{-1}\}} \widehat{g}_\bullet \mathbb{1}_\bullet^m$ *thresholded orthogonal projection estimator (tOPE)* of $\theta_\bullet = \mathfrak{s}_\bullet^\dagger g_\bullet \in \mathbb{J}$ where $\widehat{g}_\bullet^m = \widehat{g}_\bullet \mathbb{1}_\bullet^m$ is an orthogonal projection estimator (OPE) of $g_\bullet$. □

§07|02|01 Examples

§07|02.05 **GdiSM with noisy operator** (§02|04.06 continued). Considering $\mathbb{J} = \ell_2 = \mathbb{L}_2(\mathbb{N}, 2^{\mathbb{N}}, \nu_{\mathbb{N}})$ let Assumption §07|00.02 be satisfied where $\mathfrak{s}_\bullet \in \mathcal{S} \subseteq \mathbb{R}^{\mathbb{N}}_{\setminus 0} \cap \ell_\infty$ is *not known* anymore. We illustrate the tOPE in a Gaussian diagonal inverse sequence model (GdiSM) with noisy operator as in §02|04.06. Here the observable process $\widehat{\mathfrak{s}}_\bullet = \mathfrak{s}_\bullet + k^{-1/2}\dot{\mathrm{W}}_\bullet \sim \mathrm{N}^k_{\mathfrak{s}}$ and $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\mathrm{B}}_\bullet \sim \mathrm{N}^n_{\theta|\mathfrak{s}}$ is a noisy version of $\mathfrak{s}_\bullet \in \mathcal{S} \subseteq \mathbb{R}^{\mathbb{N}}_{\setminus 0} \cap \ell_\infty$ and $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \subseteq \ell_2$ with $\theta_\bullet \in \Theta \subseteq \ell_2$, respectively, where $\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}$ and $\dot{\mathrm{W}}_{\bullet} \sim \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}$ are *independent*. Consequently, $(\widehat{g}_{\bullet}, \widehat{\mathfrak{s}}_{\bullet})$ admits a joint $\mathrm{N}_{\theta|\mathfrak{s}}^{n\otimes k} = \mathrm{N}_{\theta|\mathfrak{s}}^{n} \otimes \mathrm{N}_{\mathfrak{s}}^{k}$ distribution belonging to the family $\mathrm{N}_{\Theta\times\mathcal{S}}^{n\otimes k} := (\mathrm{N}_{\theta|\mathfrak{s}}^{n} \otimes \mathrm{N}_{\mathfrak{s}}^{k})_{\theta\in\Theta,\mathfrak{s}\in\mathcal{S}}$. Summarising the observations satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}^2}, \mathscr{B}^{\otimes\mathbb{N}^2}, \mathrm{N}_{\Theta\times\mathcal{S}}^{n\otimes k})$ where $\Theta \subseteq \ell_2$ and $\mathcal{S} \subseteq \mathbb{R}_{\setminus 0}^{\mathbb{N}} \cap \ell_\infty$. □

§07|02.06 **Property** (GdiSM with noisy operator §07|02.05 continued). *For* $\dot{\mathrm{W}}_{\bullet} := (\dot{\mathrm{W}}_j)_{j\in\mathbb{N}} \sim \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}$ *we have* $\mathrm{N}_{(0,1)} \in \mathscr{W}_4(\mathscr{B})$ *with* $3\mathbb{1}_{\bullet} = \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}(\dot{\mathrm{W}}_{\bullet}^4)$, $\mathbb{1}_{\bullet} = \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}(\dot{\mathrm{W}}_{\bullet}^2)$, *and* $0_{\bullet} = \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}(\dot{\mathrm{W}}_{\bullet})$. □

§07|02.07 **diSM with noisy operator (§02|04.05 continued).** For $\mathbb{J} = \ell_2$ let Assumption §07|00.02 be satisfied where $\mathfrak{s}_{\bullet} \in \mathcal{S} \subseteq \mathbb{R}_{\setminus 0}^{\mathbb{N}} \cap \ell_\infty$ is *not known* anymore. We illustrate the tOPE in a Diagonal inverse sequence model (diSM) with noisy operator as in §02|04.05. Here the observable stochastic process $\widehat{\mathfrak{s}}_{\bullet} = \mathfrak{s}_{\bullet} + k^{-1/2}\dot{\eta}_{\bullet}$ and $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ is a noisy version of $\mathfrak{s}_{\bullet} \in \mathcal{S} \subseteq \mathbb{R}_{\setminus 0}^{\mathbb{N}} \cap \ell_\infty$ and $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}_{\bullet}}) \subseteq \ell_2$ with $\theta_{\bullet} \in \Theta \subseteq \ell_2$, respectively, where $\dot{\varepsilon}_{\bullet} \sim \otimes_{j\in\mathbb{N}}\mathrm{P}^{\dot{\varepsilon}_j}$ and $\dot{\eta}_{\bullet} \sim \otimes_{j\in\mathbb{N}}\mathrm{P}^{\dot{\eta}_j}$ are *independent*. In addition, let $\dot{\varepsilon}_{\bullet}$ satisfy **(iSM1)** of Model §07|01.06 for $\sigma_{\bullet} \in \Sigma \subseteq \mathbb{R}_{\geqslant 0}^{\mathbb{N}} \cap \ell_\infty$ and

(diSMnO1) for $\xi_{\bullet} \in \Xi \subseteq \mathbb{R}_{\geqslant 0}^{\mathbb{N}} \cap \ell_\infty$ we have $\mathrm{P}^{\dot{\eta}_j} \in \mathscr{W}_4(\mathscr{B})$ with $\xi_j^4 = \mathbb{P}(\dot{\eta}_j^4)$ and $0 = \mathbb{P}(\dot{\eta}_j)$ for all $j \in \mathbb{N}$.

Under **(iSM1)** $\widehat{g}_{\bullet}$ admits a $\mathrm{P}_{\theta|\mathfrak{s}|\sigma}^{n}$-distribution belonging to the family $\mathrm{P}_{\Theta\times\mathcal{S}\times\Sigma}^{n} := (\mathrm{P}_{\theta|\mathfrak{s}|\sigma}^{n})_{\theta\in\Theta,\mathfrak{s}\in\mathcal{S},\sigma\in\Sigma}$ and under **(diSMnO1)** $\widehat{\mathfrak{s}}_{\bullet}$ admits a $\mathrm{P}_{\mathfrak{s}|\xi}^{k}$-distribution belonging to the family $\mathrm{P}_{\mathcal{S}\times\Xi}^{k} := (\mathrm{P}_{\mathfrak{s}|\xi}^{k})_{\mathfrak{s}\in\mathcal{S},\xi\in\Xi}$. Consequently, $(\widehat{g}_{\bullet}, \widehat{\mathfrak{s}}_{\bullet})$ admits a joint $\mathrm{P}_{\theta|\mathfrak{s}|\sigma|\xi}^{n\otimes k} = \mathrm{P}_{\theta|\mathfrak{s}|\sigma}^{n} \otimes \mathrm{P}_{\mathfrak{s}|\xi}^{k}$ distribution belonging to the family $\mathrm{P}_{\Theta\times\mathcal{S}\times\Sigma\times\Xi}^{n\otimes k} := (\mathrm{P}_{\theta|\mathfrak{s}|\sigma}^{n} \otimes \mathrm{P}_{\mathfrak{s}|\xi}^{k})_{\theta\in\Theta,\mathfrak{s}\in\mathcal{S},\sigma\in\Sigma,\xi\in\Xi}$. Summarising the observations satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}^2}, \mathscr{B}^{\otimes\mathbb{N}^2}, \mathrm{P}_{\Theta\times\mathcal{S}\times\Sigma\times\Xi}^{n\otimes k})$ where $\Sigma, \Xi \subseteq \mathbb{R}_{>0}^{\mathbb{N}} \cap \ell_\infty$. $\mathcal{S} \subseteq \mathbb{R}_{\setminus 0}^{\mathbb{N}} \cap \ell_\infty$ and $\Theta \subseteq \ell_2$. □

§07|02.08 **Property** (diSM with noisy operator §07|02.07 continued). *Under* **(diSMnO1)** *the process* $\dot{\eta}_{\bullet} \sim \otimes_{j\in\mathbb{N}}\mathrm{P}^{\dot{\eta}_j}$ *satisfies* $\mathrm{P}^{\dot{\eta}_j} \in \mathscr{W}_4(\mathscr{B})$ *with* $\xi_j^4 = \mathrm{P}^{\dot{\eta}_j}(\dot{\eta}_j^4)$, $\xi_j^2 \geqslant \mathrm{P}^{\dot{\eta}_j}(\dot{\eta}_j^2)$, *and* $0 = \mathrm{P}^{\dot{\eta}_j}(\dot{\eta}_j)$ *for all* $j \in \mathbb{N}$. □

§07|02.09 **dieMM with noisy operator (§02|04.04 continued).** For $\mathbb{J} = \mathbb{L}_2(\nu)$ let Assumption §07|00.02 be satisfied where $\mathfrak{s}_{\bullet} \in \mathcal{S} \subseteq \mathcal{M}_{\neq 0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$ is *not known* anymore. We illustrate the tOPE in a Diagonal inverse empirical mean model (dieMM) with noisy operator as in §02|04.04. Here the observable stochastic processes $\widehat{\mathfrak{s}}_{\bullet} = \mathfrak{s}_{\bullet} + k^{-1/2}\dot{\eta}_{\bullet}$ and $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ are noisy version of $\mathfrak{s}_{\bullet} \in \mathcal{S}$ and $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathbb{J}$ with $\theta_{\bullet} \in \Theta \subseteq \mathbb{J}$, respectively, and *independent* error processes $\dot{\varepsilon}_{\bullet} = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_{\bullet}) - \mathbb{P}_{\theta|\mathfrak{s}}(\psi_{\bullet})) \in \mathcal{M}(\mathscr{Z}^{\otimes n} \otimes \mathscr{J})$ and $\dot{\eta}_{\bullet} = k^{1/2}(\widehat{\mathbb{P}}_k(\varphi_{\bullet}) - \mathbb{P}_{\mathfrak{s}}(\varphi_{\bullet})) \in \mathcal{M}(\mathscr{Z}^{\otimes k} \otimes \mathscr{J})$ satisfying Assumption §01|01.04. More precisely, on a measurable space $(\mathcal{Z}, \mathscr{Z})$ for each $\theta_{\bullet} \in \Theta$ and $\mathfrak{s}_{\bullet} \in \mathcal{S}$ there are probability measures $\mathbb{P}_{\theta|\mathfrak{s}}, \mathbb{P}_{\mathfrak{s}} \in \mathscr{W}(\mathscr{Z})$. Similar to Model §02|04.04 consider stochastic processes $\psi_{\bullet}, \varphi_{\bullet} \in \mathcal{M}(\mathscr{Z} \otimes \mathscr{J})$. In addition for all $\theta_{\bullet} \in \Theta$ and $\mathfrak{s}_{\bullet} \in \mathcal{S}$ the process $\psi_{\bullet} \in \mathcal{M}(\mathscr{Z} \otimes \mathscr{J})$ satisfies **(dieMM1)-(dieMM3)** of Model §07|01.08 for $\mathrm{v}_{\theta|\mathfrak{s}|\psi} \in \mathbb{R}_{\geqslant 1}$ and the process $\varphi_{\bullet} \in \mathcal{M}(\mathscr{Z} \otimes \mathscr{J})$ fulfils

(dieMMnO1) $\varphi_j \in \mathcal{L}_1(\mathbb{P}_{\mathfrak{s}}) := \mathcal{L}_1(\mathcal{Z}, \mathscr{Z}, \mathbb{P}_{\mathfrak{s}})$ ν-a.e. $j \in \mathcal{J}$ and $\mathbb{P}_{\mathfrak{s}}(\varphi_{\bullet}) = \mathfrak{s}_{\bullet}$ ν-a.s.,

(dieMMnO2) there is $\mathrm{v}_{\mathfrak{s}|\varphi}^2 \in \mathbb{R}_{\geqslant 1}$ such that $\|\mathbb{P}_{\mathfrak{s}}(\varphi_{\bullet}^4)\|_{\mathbb{L}_\infty(\nu)} \leqslant \mathrm{v}_{\mathfrak{s}|\varphi}^2$ and hence $\|\mathbb{P}_{\mathfrak{s}}(\varphi_{\bullet}^2)\|_{\mathbb{L}_\infty(\nu)} \leqslant \mathrm{v}_{\mathfrak{s}|\varphi}$.

We consider a statistical product experiment $(\mathcal{Z}^{n+k}, \mathscr{Z}^{\otimes(n+k)}, \mathbb{P}_{\Theta\times\mathcal{S}}^{n\otimes k} = (\mathbb{P}_{\theta|\mathfrak{s}}^{\otimes n} \otimes \mathbb{P}_{\mathfrak{s}}^{\otimes k})_{\theta\in\Theta,\mathfrak{s}\in\mathcal{S}})$ as in an Empirical mean function §01|01.10 where $\mathcal{S} \subseteq \mathcal{M}_{\neq 0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$ and $\Theta \subseteq \mathbb{J}$. □

§07|02.10 **Property** (dieMM with noisy operator §07|02.09 continued). *Under* **(dieMMnO1)** *and* **(dieMMnO2)** *the process* $\dot{\eta}_{\bullet} = k^{1/2}(\widehat{\mathbb{P}}_k - \mathbb{P}_{\mathfrak{s}})(\varphi_{\bullet}) \in \mathcal{M}(\mathscr{Z}^{\otimes k} \otimes \mathscr{J})$ *satisfies* $\mathrm{v}_{\mathfrak{s}|\varphi}^2 \geqslant \|\mathbb{P}_{\mathfrak{s}}^{\otimes k}(\dot{\eta}_{\bullet}^4)\|_{\mathbb{L}_\infty(\nu)}$, $\mathrm{v}_{\mathfrak{s}|\varphi} \geqslant \|\mathbb{P}_{\mathfrak{s}}^{\otimes k}(\dot{\eta}_{\bullet}^2)\|_{\mathbb{L}_\infty(\nu)}$, *and* $0 = \mathbb{P}_{\mathfrak{s}}^{\otimes k}(\dot{\eta}_j)$ *for* ν-a.e. $j \in \mathcal{J}$. □

§07|02|02 Global and maximal global $\mathfrak{v}$-risk

We measure first the accuracy of the tOPE $\widehat{\theta}_{\bullet}^m := \widehat{\mathfrak{s}}_{\bullet}^{(k)|\dagger}\widehat{g}_{\bullet}^m$ of the projection $\theta_{\bullet}^m = \mathfrak{s}_{\bullet}^{\dagger}g_{\bullet}^m \in \mathbb{J}\mathbb{1}_{\bullet}^m$ with

$g_\bullet = \mathfrak{s}_\bullet \theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}})$ and $\mathfrak{s}_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$ by the mean of its global $\mathfrak{v}$-error introduced in §04|03|01, i.e. its $\mathfrak{v}$-risk.

§07|02.11 **Reminder**. If $\mathfrak{v}_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ and $\theta_\bullet \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$ then for each $m \in \mathbb{N}$ we have $\theta_\bullet^m \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$ too and $\|\theta_\bullet^m - \theta_\bullet\|_{\mathfrak{v}}^2 = \mathrm{o}(1)$ as $m \to \infty$ (Property §04|03.09). □

§07|02.12 **Assumption**. Let $(\widehat{g}_\bullet, \widehat{\mathfrak{s}}_\bullet) = (g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet, \mathfrak{s}_\bullet + k^{-1/2}\dot{\eta}_\bullet) \sim \mathbb{P}_{\theta|\mathfrak{s}}^{n\otimes k} := \mathbb{P}_{\theta|\mathfrak{s}}^n \otimes \mathbb{P}_{\mathfrak{s}}^k$ be *independent* noisy versions satisfying Assumption §07|02.01. In addition

(dSIPg1) $\mathbb{v}_\bullet^{\theta|\mathfrak{s}} := \mathbb{P}_{\theta|\mathfrak{s}}^n(\dot{\varepsilon}_\bullet^2) := (\mathbb{v}_j^{\theta|\mathfrak{s}} := \mathbb{P}_{\theta|\mathfrak{s}}^n(\dot{\varepsilon}_j^2))_{j\in\mathcal{J}} \in \mathbb{L}_\infty(\nu)$, $\mathrm{K}_{\theta|\mathfrak{s}}^2 := 1 \vee \|\mathbb{v}_\bullet^{\theta|\mathfrak{s}}\|_{\mathbb{L}_\infty(\nu)}$,

(dSIPg2) $\dot{\varepsilon}_\bullet \mathbb{1}_\bullet^m \in \mathbb{L}_\infty(\nu)$ $\mathbb{P}_{\theta|\mathfrak{s}}^n$-a.s. for each $m \in \mathbb{N}$, and

(dSIPnO) $\mathbb{v}_\bullet^{\mathfrak{s}|(2)} := \mathbb{P}_{\mathfrak{s}}^k(\dot{\eta}_\bullet^4) := (\mathbb{v}_j^{\mathfrak{s}|(2)} := \mathbb{P}_{\mathfrak{s}}^k(\dot{\eta}_j^4))_{j\in\mathcal{J}} \in \mathbb{L}_\infty(\nu)$, $\mathrm{K}_{\mathfrak{s}}^4 := 1 \vee \|\mathbb{v}_\bullet^{\mathfrak{s}|(2)}\|_{\mathbb{L}_\infty(\nu)}$.

Moreover, from **(dSIPnO)** (i.e. $\mathbb{v}_\bullet^{\mathfrak{s}|(2)} = \mathbb{P}_{\mathfrak{s}}^k(\dot{\eta}_\bullet^4) \in \mathbb{L}_\infty(\nu)$) follows $\mathbb{P}_{\mathfrak{s}}^k(\dot{\eta}_j^2) =: \mathbb{v}_j^{\mathfrak{s}} \leqslant (\mathbb{v}_j^{\mathfrak{s}|(2)})^{1/2}$ for ν-a.e. $j \in \mathcal{J}$, and hence $\|\mathbb{v}_\bullet^{\mathfrak{s}} \vee \mathbb{1}_\bullet\|_{\mathbb{L}_\infty(\nu)} \leqslant \mathrm{K}_{\mathfrak{s}}^2$. □

§07|02.13 **Notation**. Since $\|\widehat{\mathfrak{s}}_\bullet^{(k)|\dagger}\|_{\mathbb{L}_\infty(\nu)} \leqslant k^{1/2}$ (Notation §07|02.03), $\mathfrak{s}_\bullet \in \mathbb{L}_\infty(\nu)$ and $\mathbb{1}_\bullet^m \in \mathbb{L}_\infty(\nu)$ for all $m \in \mathbb{N}$, for $(\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{s})_\bullet := \widehat{\mathfrak{s}}_\bullet^{(k)|\dagger}\mathfrak{s}_\bullet \in \mathcal{M}(\mathscr{A} \otimes \mathscr{J})$ we have $(\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{s})_\bullet \mathbb{1}_\bullet^m \in \mathbb{L}_\infty(\nu)$ for all $m \in \mathbb{N}$ too. If in addition $\mathbb{1}_\bullet^m \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$ for all $m \in \mathbb{N}$ then for $(\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{v})_\bullet := \widehat{\mathfrak{s}}_\bullet^{(k)|\dagger}\mathfrak{v}_\bullet \in \mathcal{M}(\mathscr{A} \otimes \mathscr{J})$ we also have $(\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{v})_\bullet \mathbb{1}_\bullet^m \in \mathbb{J}$ for all $m \in \mathbb{N}$. □

§07|02.14 **Comment**. Under Assumption §07|02.12 and $\mathfrak{v}_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ if $\mathbb{1}_\bullet^m \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$ for all $m \in \mathbb{N}$ then we have $(\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{v})_\bullet \dot{\varepsilon}_\bullet \mathbb{1}_\bullet^m \in \mathbb{J}$ $\mathbb{P}_{\theta|\mathfrak{s}}^{n\otimes k}$-a.s.. If in addition $\theta_\bullet \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$, and hence $\theta_\bullet^m \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$ (Property §04|03.09), then it follows

$$\mathfrak{v}_\bullet \widehat{\theta}_\bullet^m = (\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{v})_\bullet \widehat{g}_\bullet \mathbb{1}_\bullet^m = n^{-1/2}(\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{v})_\bullet \dot{\varepsilon}_\bullet \mathbb{1}_\bullet^m + (\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{s})_\bullet \mathfrak{v}_\bullet \theta_\bullet^m \in \mathbb{J} \quad \mathbb{P}_{\theta|\mathfrak{s}}^{n\otimes k}\text{-a.s.}. \tag{07.27}$$

If $\mathcal{J} \subseteq \mathbb{Z}$ (at most countable) and $\nu_{\mathcal{J}}$ is the counting measure over the index set $\mathcal{J}$ then Assumption §01|01.04 and **(dSIPg1)** (i.e. $\mathbb{v}_\bullet^{\theta|\mathfrak{s}} = \mathbb{P}_{\theta|\mathfrak{s}}^n(\dot{\varepsilon}_\bullet^2) \in \mathbb{L}_\infty(\nu_{\mathcal{J}})$) imply the additional assumption **(dSIPg2)** $\dot{\varepsilon}_\bullet \mathbb{1}_\bullet^m \in \mathbb{L}_\infty(\nu_{\mathcal{J}})$ $\mathbb{P}_{\theta|\mathfrak{s}}^n$-a.s.. However, the last implication does generally not hold, if $\mathcal{J} \in \{\mathbb{R}, \mathbb{R}_{\geqslant 0}\}$ for example. □

§07|02|02|01 Global $\mathfrak{v}$-risk

§07|02.15 **Assumption**. Let $\mathfrak{v}_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$, $\theta_\bullet \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$, and $\mathfrak{s}_\bullet^\dagger \mathbb{1}_\bullet^m, \mathbb{1}_\bullet^m \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$ for $m \in \mathbb{N}$ be satisfied. □

§07|02.16 **Definition**. Under Assumptions §07|02.12 and §07|02.15 for $m \in \mathbb{N}$ the *global $\mathfrak{v}$-risk* of a thresholded OPE $\widehat{\theta}_\bullet^m = \widehat{\mathfrak{s}}_\bullet^{(k)|\dagger}\widehat{g}_\bullet^m = \widehat{\mathfrak{s}}_\bullet^\dagger \mathbb{1}_\bullet^{\{\widehat{\mathfrak{s}}_\bullet^2 \geqslant k^{-1}\}}\widehat{g}_\bullet \mathbb{1}_\bullet^m \in \mathbb{L}_2(\mathfrak{v}_\bullet^2\nu)$ $\mathbb{P}_{\theta|\mathfrak{s}}^{n\otimes k}$-a.s. satisfies

$$\mathbb{P}_{\theta|\mathfrak{s}}^{n\otimes k}(\|\widehat{\theta}_\bullet^m - \theta_\bullet\|_{\mathfrak{v}}^2) = \mathbb{P}_{\theta|\mathfrak{s}}^{n\otimes k}(\|\widehat{\mathfrak{s}}_\bullet^{(k)|\dagger}(\widehat{g}_\bullet - \widehat{\mathfrak{s}}_\bullet\theta_\bullet)\mathbb{1}_\bullet^m\|_{\mathfrak{v}}^2) + \mathbb{P}_{\mathfrak{s}}^k(\|\mathbb{1}_\bullet^{\{\widehat{\mathfrak{s}}_\bullet^2 < k^{-1}\}}\theta_\bullet \mathbb{1}_\bullet^m\|_{\mathfrak{v}}^2) + \|\theta_\bullet \mathbb{1}_\bullet^{m|\perp}\|_{\mathfrak{v}}^2 \tag{07.28}$$

with *variance* terms $\mathbb{P}_{\theta|\mathfrak{s}}^{n\otimes k}(\|\widehat{\mathfrak{s}}_\bullet^{(k)|\dagger}(\widehat{g}_\bullet - \widehat{\mathfrak{s}}_\bullet\theta_\bullet)\mathbb{1}_\bullet^m\|_{\mathfrak{v}}^2)$, $\mathbb{P}_{\mathfrak{s}}^k(\|\mathbb{1}_\bullet^{\{\widehat{\mathfrak{s}}_\bullet^2 < k^{-1}\}}\theta_\bullet \mathbb{1}_\bullet^m\|_{\mathfrak{v}}^2)$ and *bias* term $\|\theta_\bullet \mathbb{1}_\bullet^{m|\perp}\|_{\mathfrak{v}}$.

§07|02.17 **Property**. *Under Assumptions §07|02.12 and §07|02.15 for each $m \in \mathbb{N}$ we have*

$$\begin{aligned}\mathbb{P}_{\theta|\mathfrak{s}}^{n\otimes k}\|\widehat{\mathfrak{s}}_\bullet^{(k)|\dagger}(\widehat{g}_\bullet - \widehat{\mathfrak{s}}_\bullet\theta_\bullet)\mathbb{1}_\bullet^m\|_{\mathfrak{v}}^2 &= \mathbb{P}_{\theta|\mathfrak{s}}^n \otimes \mathbb{P}_{\mathfrak{s}}^k\|\widehat{\mathfrak{s}}_\bullet^{(k)|\dagger}(n^{-1/2}\dot{\varepsilon}_\bullet + (\mathfrak{s}_\bullet - \widehat{\mathfrak{s}}_\bullet)\theta_\bullet)\mathbb{1}_\bullet^m\|_{\mathfrak{v}}^2 \\ &= n^{-1}\nu\big(\mathbb{P}_{\mathfrak{s}}^k((\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{s})_\bullet^2)\,\mathbb{v}_\bullet^{\theta|\mathfrak{s}}\,(\mathfrak{s}^\dagger\mathfrak{v})_\bullet^2\mathbb{1}_\bullet^m\big) + \nu\big(\mathbb{P}_{\mathfrak{s}}^k(|\widehat{\mathfrak{s}}_\bullet^{(k)|\dagger}|^2|\mathfrak{s}_\bullet - \widehat{\mathfrak{s}}_\bullet|^2)\mathfrak{v}_\bullet^2\theta_\bullet^2\mathbb{1}_\bullet^m\big)\end{aligned}$$

($\mathfrak{s}_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ by Assumption §07|02.01) and $\mathbb{P}_{\mathfrak{s}}^k\|\mathbb{1}_\bullet^{\{\widehat{\mathfrak{s}}_\bullet^2 < k^{-1}\}}\theta_\bullet \mathbb{1}_\bullet^m\|_{\mathfrak{v}}^2 = \nu\big(\mathbb{P}_{\mathfrak{s}}^k(\widehat{\mathfrak{s}}_\bullet^2 < k^{-1})\mathfrak{v}_\bullet^2\theta_\bullet^2\mathbb{1}_\bullet^m\big)$. □

§07|02.18 **Lemma**. *Under Assumption §07|02.12 **(dSIPnO)** for all $j \in \mathcal{J}$ we have*

(i) $\mathbb{P}_{\mathfrak{s}}^k((\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{s})_j^2) \leqslant 2(\mathbb{v}_j^{\mathfrak{s}} + 1) \leqslant 4(1 \vee \mathbb{v}_j^{\mathfrak{s}|(2)})^{1/2}$,

(ii) $\mathbb{P}^k_{\mathfrak{s}}(\widehat{\mathfrak{s}}_j^2 < k^{-1}) \leqslant 4(1 \vee \mathfrak{v}_j^{\mathfrak{s}})(1 \vee k\mathfrak{s}_j^2)^{-1} \leqslant 4(1 \vee \mathfrak{v}_j^{\mathfrak{s}|(2)})^{1/2}(1 \vee k\mathfrak{s}_j^2)^{-1}$, *and*

(iii) $\mathbb{P}^k_{\mathfrak{s}}(|\mathfrak{s}_j - \widehat{\mathfrak{s}}_j|^2|\widehat{\mathfrak{s}}_j^{(k)|\dagger}|^2) \leqslant 2(\mathfrak{v}_j^{\mathfrak{s}|(2)} + \mathfrak{v}_j^{\mathfrak{s}})(1 \vee k\mathfrak{s}_j^2)^{-1} \leqslant 4(1 \vee \mathfrak{v}_j^{\mathfrak{s}|(2)})(1 \vee k\mathfrak{s}_j^2)^{-1}$.

§07|02.19 **Proof** of Lemma §07|02.18. Given in the lecture. □

§07|02.20 **Reminder**. If Assumptions §07|01.11 and §07|01.13 are satisfied, then for all $n, m \in \mathbb{N}$ setting

$$\mathrm{R}^m_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) := \|\theta_\bullet \mathbb{1}^{m|\perp}_\bullet\|^2_{\mathfrak{v}} + n^{-1}\|\mathfrak{s}^\dagger_\bullet \mathbb{1}^m_\bullet\|^2_{\mathfrak{v}}, \quad m^\circ_n := \arg\min\left\{\mathrm{R}^m_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) : m \in \mathbb{N}\right\}$$
$$\text{and} \quad \mathrm{R}^\circ_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) := \mathrm{R}^{m^\circ_n}_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) = \min\left\{\mathrm{R}^m_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) : m \in \mathbb{N}\right\} \quad (07.29)$$

the OPE $\widehat{\theta}^{m^\circ_n}_\bullet := \mathfrak{s}^\dagger_\bullet \widehat{g}^m_\bullet$ with *known* $\mathfrak{s}_\bullet \in \mathbb{L}_\infty(\nu)$ fulfils $\mathbb{P}^n_{\theta|\mathfrak{s}}(\|\widehat{\theta}^{m^\circ_n}_\bullet - \theta_\bullet\|^2_{\mathfrak{v}}) \leqslant (1 \vee \|\mathfrak{v}^{\theta|\mathfrak{s}}_\bullet\|_{\mathbb{L}_\infty(\nu)})\,\mathrm{R}^\circ_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet)$ due to Proposition §07|01.17. Keep in mind that Assumption §07|01.11 is part of Assumption §07|02.12 and that Assumption §07|01.13 is part of Assumption §07|02.15. □

§07|02.21 **Proposition** (Upper bound). *Let Assumptions §07|02.12 and §07|02.15 be satisfied. The thresholded OPE* $\widehat{\theta}^m_\bullet = \widehat{\mathfrak{s}}^{(k)|\dagger}_\bullet \widehat{g}^m_\bullet \in \mathbb{L}_2(\mathfrak{v}^2_\bullet\nu)$ $\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}$-a.s. *for all* $n, k, m \in \mathbb{N}$ *fulfils*

$$\begin{aligned}\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}(\|\widehat{\theta}^m_\bullet - \theta_\bullet\|^2_{\mathfrak{v}}) &\leqslant 4\|\mathfrak{v}^{\mathfrak{s}}_\bullet \vee \mathbb{1}_\bullet\|_{\mathbb{L}_\infty(\nu)}\|\mathfrak{v}^{\theta|\mathfrak{s}}_\bullet \vee \mathbb{1}_\bullet\|_{\mathbb{L}_\infty(\nu)}\,\mathrm{R}^m_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet)\\ &\quad + 2(\|\mathfrak{v}^{\mathfrak{s}|(2)}_\bullet\|_{\mathbb{L}_\infty(\nu)} + 3\|\mathfrak{v}^{\mathfrak{s}}_\bullet \vee \mathbb{1}_\bullet\|_{\mathbb{L}_\infty(\nu)})\|(1 \vee k\mathfrak{s}^2_\bullet)^{-1/2}\theta_\bullet\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}} \quad (07.30)\\ &\leqslant 4\mathrm{K}^2_{\mathfrak{s}}\mathrm{K}^2_{\theta|\mathfrak{s}}\,\mathrm{R}^m_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) + 8\mathrm{K}^4_{\mathfrak{s}}\|(1 \vee k\mathfrak{s}^2_\bullet)^{-1/2}\theta_\bullet\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}}. \quad (07.31)\end{aligned}$$

§07|02.22 **Proof** of Proposition §07|02.21. Given in the lecture. □

§07|02.23 **Comment**. For each $m \in \mathbb{N}$ we have

$$\begin{aligned}\|(1 \vee k\mathfrak{s}^2_\bullet)^{-1/2}\theta_\bullet\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}} \leqslant \|(1 \vee k\mathfrak{s}^2_\bullet)^{-1/2}\theta_\bullet\|^2_{\mathfrak{v}} &= \|(1 \vee k\mathfrak{s}^2_\bullet)^{-1/2}\theta_\bullet\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}} + \|(1 \vee k\mathfrak{s}^2_\bullet)^{-1/2}\theta_\bullet\mathbb{1}^{m|\perp}_\bullet\|^2_{\mathfrak{v}}\\ &\leqslant \|(1 \vee k\mathfrak{s}^2_\bullet)^{-1/2}\theta_\bullet\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}} + \|\theta_\bullet\mathbb{1}^{m|\perp}_\bullet\|^2_{\mathfrak{v}}. \quad (07.32)\end{aligned}$$

Consequently, under the assumptions of Proposition §07|02.21 from (??) (Proof §07|02.22) follows

$$\begin{aligned}\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}\left(\|\widehat{\theta}^m_\bullet - \theta_\bullet\|^2_{\mathfrak{v}}\right) &\leqslant 2\mathrm{K}^2_{\mathfrak{s}}\mathrm{K}^2_{\theta|\mathfrak{s}}\,\mathrm{R}^m_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) + 8\mathrm{K}^4_{\mathfrak{s}}\|\theta_\bullet(1 \vee k\mathfrak{s}^2_\bullet)^{-1/2}\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}}\\ &\leqslant 2\mathrm{K}^2_{\mathfrak{s}}\mathrm{K}^2_{\theta|\mathfrak{s}}\,\mathrm{R}^m_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) + 8\mathrm{K}^4_{\mathfrak{s}}\|\theta_\bullet(1 \vee k\mathfrak{s}^2_\bullet)^{-1/2}\|^2_{\mathfrak{v}}\\ &\leqslant 10\mathrm{K}^4_{\mathfrak{s}}\mathrm{K}^2_{\theta|\mathfrak{s}}\,\mathrm{R}^m_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) + 8\mathrm{K}^4_{\mathfrak{s}}\|\theta_\bullet(1 \vee k\mathfrak{s}^2_\bullet)^{-1/2}\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}}.\end{aligned}$$

Selecting $m^\circ_n := \arg\min\left\{\mathrm{R}^m_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) : m \in \mathbb{N}\right\}$ and $\mathrm{R}^\circ_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) = \mathrm{R}^{m^\circ_n}_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet)$ as in (07.29) (Reminder §07|02.20) we obtain

$$\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}\left(\|\widehat{\theta}^{m^\circ_n}_\bullet - \theta_\bullet\|^2_{\mathfrak{v}}\right) \leqslant 2\mathrm{K}^2_{\mathfrak{s}}\mathrm{K}^2_{\theta|\mathfrak{s}}\,\mathrm{R}^\circ_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) + 8\mathrm{K}^4_{\mathfrak{s}}\|\theta_\bullet(1 \vee k\mathfrak{s}^2_\bullet)^{-1/2}\|^2_{\mathfrak{v}}. \quad (07.33)$$

We shall emphasise, that the upper bound consists (up to the constants) of the sum of the two terms $\mathrm{R}^\circ_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet)$ and $\|\theta_\bullet(1 \vee k\mathfrak{s}^2_\bullet)^{-1/2}\|^2_{\mathfrak{v}}$ depending each on one of the sample sizes n and k only. Moreover, $\mathrm{R}^\circ_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet)$ is the oracle rate (Property §07|01.20) in case of an in advanced known $\mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathrm{M}}(\mathbb{J})$. □

§07|02.24 **Corollary** (GdiSM with noisy operator §07|02.05 continued). *Consider* independent *noisy versions* $(\widehat{g}_\bullet, \widehat{\mathfrak{s}}_\bullet) = (g_\bullet + n^{-1/2}\dot{\mathrm{B}}_\bullet, \mathfrak{s}_\bullet + k^{-1/2}\dot{\mathrm{W}}_\bullet) \sim \mathrm{N}^{n\otimes k}_{\theta|\mathfrak{s}} = \mathrm{N}^n_{\theta|\mathfrak{s}} \otimes \mathrm{N}^k_{\mathfrak{s}}$ *as in Model §07|02.05, where* $\dot{\mathrm{B}}_\bullet \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$ *and* $\dot{\mathrm{W}}_\bullet \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$ *are* independent, $\mathfrak{s}_\bullet \in \mathbb{R}^{\mathbb{N}}_{\setminus 0} \cap \ell_\infty$ *and* $\theta_\bullet \in \ell_2$, *and hence* $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \subseteq \ell_2$. *Given* $\mathfrak{v}_\bullet \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ *and* $\theta_\bullet \in \ell_2(\mathfrak{v}^2_\bullet)$ *the (infeasible) thresholded OPE* $\widehat{\theta}^{m^\circ_n}_\bullet = \widehat{\mathfrak{s}}^{(k)|\dagger}_\bullet \widehat{g}^{m^\circ_n}_\bullet \in \ell_2(\mathfrak{v}^2_\bullet)$ *with oracle dimension* m°_n *as in* (07.29) *satisfies*

$$\mathrm{N}^{n\otimes k}_{\theta|\mathfrak{s}}(\|\widehat{\theta}^{m^\circ_n}_\bullet - \theta_\bullet\|^2_{\mathfrak{v}}) \leqslant 4\mathrm{R}^\circ_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) + 12\|(1 \vee k\mathfrak{s}^2_\bullet)^{-1/2}\theta_\bullet\|^2_{\mathfrak{v}} \quad \forall n, k \in \mathbb{N} \quad (07.34)$$

where $\mathrm{R}^\circ_n(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet)$ *is the oracle rate in a* GdiSM *§07|01.03 (see* Corollary *§07|01.22).*

§07|02.25 **Proof** of Corollary §07|02.24. Given in the lecture. □

§07|02.26 **Corollary** (diSM with noisy operator §07|02.07 continued). *Consider independent noisy versions* $(\widehat{g}_{\bullet}, \widehat{\mathfrak{s}}_{\bullet}) = (g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}, \mathfrak{s}_{\bullet} + k^{-1/2}\dot{\eta}_{\bullet}) \sim \mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}|\sigma|\xi} = \mathbb{P}^{n}_{\theta|\mathfrak{s}|\sigma} \otimes \mathbb{P}^{k}_{\mathfrak{s}|\xi}$ *as in Model §07|02.07, where* $\dot{\varepsilon}_{\bullet}$ *and* $\dot{\eta}_{\bullet}$ *satisfy* **(iSM1)** *and* **(diSMnO1)** *with* $\mathrm{K}_{\sigma} := \|\sigma_{\bullet}\|_{\ell_\infty} \vee 1 \in \mathbb{R}_{\geqslant 1}$ *and* $\mathrm{K}_{\xi} := \|\xi_{\bullet}\|_{\ell_\infty} \vee 1 \in \mathbb{R}_{\geqslant 1}$, *respectively,* $\mathfrak{s}_{\bullet} \in \mathbb{R}^{\mathbb{N}}_{\setminus 0} \cap \ell_\infty$ *and* $\theta_{\bullet} \in \ell_2$, *and hence* $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}}) \subseteq \ell_2$. *Given* $\mathfrak{v}_{\bullet} \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ *and* $\theta_{\bullet} \in \ell_2(\mathfrak{v}^2_{\bullet})$ *the (infeasible) thresholded OPE* $\widehat{\theta}^{m_n^\circ}_{\bullet} = \widehat{\mathfrak{s}}^{(k)|\dagger}_{\bullet}\widehat{g}^{m_n^\circ}_{\bullet} \in \ell_2(\mathfrak{v}^2_{\bullet})$ *with oracle dimension* m_n° *as in* (07.29) *satisfies*

$$\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}|\sigma|\xi}(\|\widehat{\theta}^{m_n^\circ}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) \leqslant 4\mathrm{K}^2_{\sigma}\mathrm{K}^2_{\xi}\,\mathrm{R}^\circ_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet}) + 8\mathrm{K}^4_{\xi}\,\|(1 \vee k\mathfrak{s}^2_{\bullet})^{-1/2}\theta_{\bullet}\|^2_{\mathfrak{v}} \quad \forall n, k \in \mathbb{N} \qquad (07.35)$$

where $\mathrm{R}^\circ_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet})$ *is the oracle rate in a* diSM *§07|01.06 (see Corollary §07|01.24).*

§07|02.27 **Proof** of Corollary §07|02.26. Given in the lecture. □

§07|02.28 **Corollary** (dieMM with noisy operator §07|02.09 continued). *Consider independent noisy versions* $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ *and* $\widehat{\mathfrak{s}}_{\bullet} = \mathfrak{s}_{\bullet} + k^{-1/2}\dot{\eta}_{\bullet}$ *defined on* $(\mathcal{Z}^{n+k}, \mathscr{Z}^{\otimes(n+k)}, \mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}} = \mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}} \otimes \mathbb{P}^{\otimes k}_{\mathfrak{s}})$ *as in Model §07|02.09, where* $\psi_{\bullet}, \varphi_{\bullet} \in \mathcal{M}(\mathscr{X} \otimes \mathscr{J})$ *satisfy* **(dieMM1)–(dieMM3)** *(Model §07|01.08) and* **(dieMMnO1)–(dieMMnO2)** *(Model §07|02.09) with* $\mathbb{v}_{\theta|\mathfrak{s}|\psi}, \mathbb{v}_{\mathfrak{s}|\varphi} \in \mathbb{R}_{\geqslant 1}$, *respectively,* $\mathfrak{s}_{\bullet} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$, $\theta_{\bullet} \in \mathbb{J}$ *and hence* $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}}) \subseteq \mathbb{J}$. *Given Assumption §07|02.15 the (infeasible) thresholded OPE* $\widehat{\theta}^{m_n^\circ}_{\bullet} = \widehat{\mathfrak{s}}^{(k)|\dagger}_{\bullet}\widehat{g}^{m_n^\circ}_{\bullet} \in \mathbb{L}_2(\mathfrak{v}^2_{\bullet}\nu)$ *with oracle dimension* m_n° *as in* (07.29) *satisfies*

$$\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}(\|\widehat{\theta}^{m_n^\circ}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) \leqslant 4\mathbb{v}_{\theta|\mathfrak{s}|\psi}\mathbb{v}_{\mathfrak{s}|\varphi}\,\mathrm{R}^\circ_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet}) + 8\mathbb{v}^2_{\mathfrak{s}|\varphi}\,\|(1 \vee k\mathfrak{s}^2_{\bullet})^{-1/2}\theta_{\bullet}\|^2_{\mathfrak{v}} \quad \forall n, k \in \mathbb{N} \qquad (07.36)$$

where $\mathrm{R}^\circ_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet})$ *is the oracle rate in a* dieMM *§07|01.08 (see Corollary §07|01.26).*

§07|02.29 **Proof** of Corollary §07|02.28. Given in the lecture. □

§07|02.30 **Illustration**. We illustrate the last results considering usual behaviour for $\theta_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$. We distinguish again the two cases **(p)** and **(np)** in Illustration §07|01.28, where in case **(p)** the term $\mathrm{R}^\circ_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet})$ is parametric, that is, $n\mathrm{R}^\circ_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet}) = \mathrm{O}(1)$, in case **(np)** it is nonparametric, i.e. $\lim_{n\to\infty} n\mathrm{R}^\circ_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet}) = \infty$. In case **(np)** we consider again the three specifications **(o-m)**, **(o-s)** and **(s-m)** introduced in Illustration §07|01.28 where also in Table 01 [§07] the order of the oracle dimension m_n° and the oracle rate $\mathrm{R}^\circ_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet})$ as $n \to \infty$ are given. The next table depict the oracle rate $\mathrm{R}^\circ_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet})$ and the rate of the additional term $\|(1 \vee k\mathfrak{s}^2_{\bullet})^{-1/2}\theta_{\bullet}\|^2_{\mathfrak{v}}$ as $n, k \to \infty$:

Table 05 [§07]

Order of $\mathrm{R}^\circ_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet})$ and $\|(1 \vee k\mathfrak{s}^2_{\bullet})^{-1/2}\theta_{\bullet}\|^2_{\mathfrak{v}}$ as $n, k \to \infty$

	$(j \in \mathcal{J})$ $\mathfrak{v}^2_j = j^{2\mathrm{v}}$	$(\mathrm{a} \in \mathbb{R}_{>0})$ θ^2_j	$(\mathrm{t} \in \mathbb{R}_{>0})$ $\mathfrak{s}^2_j$	$\mathrm{R}^\circ_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet})$	$\theta^2_j\mathfrak{v}^2_j$	$\mathfrak{s}^{-2}_j\theta^2_j\mathfrak{v}^2_j$	$\|(1 \vee k\mathfrak{s}^2_{\bullet})^{-1/2}\theta_{\bullet}\|^2_{\mathfrak{v}}$
(o-m)	$\mathrm{v} \in (-1/2 - \mathrm{t}, \mathrm{a})$	$j^{-2\mathrm{a}-1}$	$j^{-2\mathrm{t}}$	$n^{-\frac{2(\mathrm{a}-\mathrm{v})}{2\mathrm{a}+2\mathrm{t}+1}}$	$j^{-2(\mathrm{a}-\mathrm{v})-1}$	$j^{2(\mathrm{t}+\mathrm{v}-\mathrm{a})-1}$	
	$\mathrm{a} - \mathrm{v} < \mathrm{t}$						$k^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
	$\mathrm{a} - \mathrm{v} = \mathrm{t}$						$(k/\log k)^{-1}$
	$\mathrm{a} - \mathrm{v} > \mathrm{t}$						k^{-1}
(o-s)	$\mathrm{a} - \mathrm{v} \in \mathbb{R}_{>0}$	$j^{-2\mathrm{a}-1}$	$e^{-j^{2\mathrm{t}}}$	$(\log n)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$	$j^{-2(\mathrm{a}-\mathrm{v})-1}$	$j^{-2(\mathrm{a}-\mathrm{v})-1}e^{j^{2\mathrm{t}}}$	$(\log k)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v} + \mathrm{t} + 1/2 \in \mathbb{R}_{>0}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$n^{-1}(\log n)^{\frac{2(\mathrm{t}+\mathrm{v})+1}{2\mathrm{a}}}$	$j^{2\mathrm{v}}e^{-j^{2\mathrm{a}}}$	$j^{2(\mathrm{t}+\mathrm{v})}e^{-j^{2\mathrm{a}}}$	k^{-1}

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v} + \mathrm{t} < -1/2$ the oracle rate $\mathrm{R}^\circ_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \mathfrak{v}_{\bullet})$ is parametric. □

§07|02|02|02 Maximal global $\mathfrak{v}$-risk

§07|02.31 **Notation** (Reminder). For sequences $a_\bullet, b_\bullet \in (\mathbb{K})^{\mathbb{N}}$ taking its values in $\mathbb{K} \in \{\mathbb{R}, \mathbb{R}_{>0}, \mathbb{Q}, \mathbb{Z}, \dots\}$ we write $a_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\nearrow}$ and $b_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\searrow}$ if $a_\bullet$ and $b_\bullet$, respectively, is monotonically *non-decreasing* and *non-increasing*. If in addition $a_n \to \infty$ and $b_n \to 0$ as $n \to \infty$, then we write $a_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\uparrow\infty}$ and $b_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\downarrow 0}$ for short. For $\mathrm{w}_\bullet \in \mathbb{L}_\infty(\nu)$ we set $\mathrm{w}_{(0)} := \|\mathrm{w}_\bullet\|_{\mathbb{L}_\infty(\nu)}$ and $\mathrm{w}_{(\bullet)} = (\mathrm{w}_{(j)} := \|\mathrm{w}_\bullet \mathbb{1}^{j|\perp}_\bullet\|_{\mathbb{L}_\infty(\nu)})_{j\in\mathbb{N}}$, where by construction $\mathrm{w}_{(\bullet)} \in (\mathbb{R}_{\geqslant 0})^{\mathbb{N}}_{\searrow}$. □

§07|02.32 **Assumption**. Consider weights $\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet \in \mathcal{M}_{>0,\nu}(\mathscr{J})$ (i.e. $\nu(\mathcal{N}_{\mathfrak{a}}) = \nu(\mathcal{N}_{\mathfrak{t}}) = 0 = \nu(\mathcal{N}_{\mathfrak{v}})$), such that $\mathfrak{a}_\bullet, \mathfrak{t}_\bullet \in \mathbb{L}_\infty(\nu)$, $(\mathfrak{a}\mathfrak{v})_\bullet = \mathfrak{a}_\bullet\mathfrak{v}_\bullet \in \mathbb{L}_\infty(\nu)$, $(\mathfrak{a}\mathfrak{v})_{(\bullet)} \in (\mathbb{R}_{\geqslant 0})^{\mathbb{N}}_{\downarrow 0}$, and $\mathfrak{t}^\dagger_\bullet \mathbb{1}^m_\bullet, \mathbb{1}^m_\bullet \in \mathbb{L}_2(\mathfrak{v}^2_\bullet\nu)$ for all $m \in \mathbb{N}$. □

§07|02.33 **Reminder**. Under Assumption §07|02.32 we have $\mathbb{J}^{\mathfrak{a}} = \mathbb{L}^{\mathfrak{a}}_2(\nu) = \mathrm{dom}(\mathrm{M}_{\mathfrak{a}^\dagger_\bullet}) = \mathbb{J}\mathfrak{a}_\bullet \subseteq \mathbb{J}$ and the three measures ν, $\mathfrak{a}^{2|\dagger}_\bullet\nu$ and $\mathfrak{v}^2_\bullet\nu$ dominate mutually each other, i.e. they share the same null sets (see Property §04|01.02). We consider $\mathbb{J}^{\mathfrak{a}}$ endowed with $\|\cdot\|_{\mathfrak{a}^\dagger} = \|\mathrm{M}_{\mathfrak{a}^\dagger_\bullet}\cdot\|_{\mathbb{J}}$ and given a constant $\mathrm{r} \in \mathbb{R}_{>0}$ the ellipsoid $\mathbb{J}^{\mathfrak{a},\mathrm{r}} := \{h_\bullet \in \mathbb{J}^{\mathfrak{a}} : \|h_\bullet\|_{\mathfrak{a}^\dagger} \leqslant \mathrm{r}\} \subseteq \mathbb{J}^{\mathfrak{a}}$. Since $(\mathfrak{a}\mathfrak{v})_\bullet \in \mathbb{L}_\infty(\nu)$, and hence $(\mathfrak{a}\mathfrak{v})_{(m)} := \|(\mathfrak{a}\mathfrak{v})_\bullet \mathbb{1}^{m|\perp}_\bullet\|_{\mathbb{L}_\infty(\nu)} \in \mathbb{R}_{\geqslant 0}$ for each $m \in \mathbb{N}$ we have $\mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{L}_2(\mathfrak{v}^2_\bullet\nu)$ (Property §04|02.11), and $\|\theta_\bullet \mathbb{1}^{m|\perp}_\bullet\|_{\mathfrak{v}} \leqslant \mathrm{r}\,(\mathfrak{a}\mathfrak{v})_{(m)}$ for all $\theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ (Lemma §04|02.13). Let in addition $\mathrm{M}_{\mathfrak{s}_\bullet} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ satisfy a link condition as in Definition §04|03.05 with weights $\mathfrak{t}_\bullet \in \mathcal{M}_{>0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$, and radius $\mathrm{d} \in \mathbb{R}_{>0}$. We set $(\mathfrak{t}^\dagger\mathfrak{v})_\bullet := (\mathfrak{t}^\dagger_j\mathfrak{v}_j)_{j\in\mathscr{J}} = \mathfrak{t}^\dagger_\bullet\mathfrak{v}_\bullet \in \mathcal{M}(\mathscr{J})$. Obviously, for each $m \in \mathbb{N}$ the condition $\mathfrak{t}^\dagger_\bullet\mathbb{1}^m_\bullet \in \mathbb{L}_2(\mathfrak{v}^2_\bullet\nu)$ (due to Assumption §07|02.32) implies $\mathfrak{s}^\dagger_\bullet\mathbb{1}^m_\bullet \in \mathbb{L}_2(\mathfrak{v}^2_\bullet\nu)$ too. Consequently, if Assumption §07|02.32, $\theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ and $\mathrm{M}_{\mathfrak{s}_\bullet} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ are satisfied, then Assumption §07|02.15 is also fulfilled. Keep in mind if Assumptions §07|01.11 and §07|01.30 are satisfied, then for $n, m \in \mathbb{N}$ setting

$$\mathrm{R}^m_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) := [(\mathfrak{a}\mathfrak{v})^2_{(m)} \vee n^{-1}\|\mathfrak{t}^\dagger_\bullet\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}}], \quad m^\star_n := \arg\min\left\{\mathrm{R}^m_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) : m \in \mathbb{N}\right\}$$
$$\text{and} \quad \mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) := \mathrm{R}^{m^\star_n}_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) = \min\left\{\mathrm{R}^m_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) : m \in \mathbb{N}\right\} \quad (07.37)$$

for each $\mathrm{M}_{\mathfrak{s}_\bullet} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ *known* in advance, for all $\theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$, and hence $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger_\bullet}) \subseteq \mathbb{J}$, and for each $n \in \mathbb{N}$ the OPE $\widehat{\theta}^{m^\star_n}_\bullet := \mathfrak{s}^\dagger_\bullet\widehat{g}^{m^\star_n}_\bullet \in \mathbb{L}_2(\mathfrak{v}^2_\bullet\nu)$ fulfils

$$\mathbb{P}^n_{\theta|\mathfrak{s}}(\|\widehat{\theta}^{m^\star_n}_\bullet - \theta_\bullet\|^2_{\mathfrak{v}}) \leqslant (\|\mathrm{v}^{\theta|\mathfrak{s}}_\bullet\|_{\mathbb{L}_\infty(\nu)}\mathrm{d}^2 + \mathrm{r}^2)\,\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) \quad \forall n \in \mathbb{N}.$$

due to Corollary §07|01.35. We shall emphasise that Assumptions §07|02.12 and §07|02.32 contains Assumptions §07|01.11 and §07|01.30, respectively. □

§07|02.34 **Proposition** (Upper bound). *Let Assumptions §07|02.12 and §07|02.32 be satisfied. If* $\mathrm{M}_{\mathfrak{s}_\bullet} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ *with* $\mathrm{d} \in \mathbb{R}_{>0}$, *and* $\theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ *with* $\mathrm{r} \in \mathbb{R}_{>0}$, *then the thresholded OPE* $\widehat{\theta}^m_\bullet = \widehat{\mathfrak{s}}^{(k)|\dagger}_\bullet\widehat{g}^m_\bullet \in \mathbb{L}_2(\mathfrak{v}^2_\bullet\nu)$ $\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}$-a.s. *for all* $n, k, m \in \mathbb{N}$ *fulfils*

$$\begin{aligned}\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}(\|\widehat{\theta}^m_\bullet - \theta_\bullet\|^2_{\mathfrak{v}}) &\leqslant (4\|\mathrm{v}^{\mathfrak{s}}_\bullet \vee \mathbb{1}_\bullet\|_{\mathbb{L}_\infty(\nu)}\|\mathrm{v}^{\theta|\mathfrak{s}}_\bullet \vee \mathbb{1}_\bullet\|_{\mathbb{L}_\infty(\nu)}\mathrm{d}^2 + \mathrm{r}^2)\,\mathrm{R}^m_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet)\\ &\quad + 2(\|\mathrm{v}^{\mathfrak{s}|(2)}_\bullet\|_{\mathbb{L}_\infty(\nu)} + 3\|\mathrm{v}^{\mathfrak{s}}_\bullet \vee \mathbb{1}_\bullet\|_{\mathbb{L}_\infty(\nu)})\mathrm{d}^2\mathrm{r}^2\|(1 \vee k\mathfrak{t}^2_\bullet)^{-1}(\mathfrak{a}\mathfrak{v})^2_\bullet\|_{\mathbb{L}_\infty(\nu)} \quad (07.38)\\ &\leqslant (4\mathrm{K}^2_{\mathfrak{s}}\mathrm{K}^2_{\theta|\mathfrak{s}}\mathrm{d}^2 + \mathrm{r}^2)\,\mathrm{R}^m_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) + 8\mathrm{K}^4_{\mathfrak{s}}\mathrm{d}^2\mathrm{r}^2\|(1 \vee k\mathfrak{t}^2_\bullet)^{-1}(\mathfrak{a}\mathfrak{v})^2_\bullet\|_{\mathbb{L}_\infty(\nu)}. \quad (07.39)\end{aligned}$$

§07|02.35 **Proof** of Proposition §07|02.34. Given in the lecture. □

§07|02.36 **Remark**. If in addition there exists $\mathrm{v} \in \mathbb{R}_{>0}$ satisfying $\mathrm{v} \geqslant (\mathrm{K}_{\mathfrak{s}} \vee \mathrm{K}_{\theta|\mathfrak{s}})$ for all $\theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ and $\mathrm{M}_{\mathfrak{s}_\bullet} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ then the maximal global $\mathfrak{v}$-risk of the thresholded OPE $\widehat{\theta}^{m^\star_n}_\bullet$ with optimally chosen dimension $m^\star_n := \arg\min\left\{\mathrm{R}^m_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) : m \in \mathbb{N}\right\}$ and $\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) = \min\left\{\mathrm{R}^m_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) : m \in \mathbb{N}\right\}$ as in (07.37) fulfils

$$\begin{aligned}&\sup\left\{\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}\|\widehat{\theta}^{m^\star_n}_\bullet - \theta_\bullet\|^2_{\mathfrak{v}} : \theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}_\bullet} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}\right\}\\ &\qquad \leqslant (4\mathrm{v}^4\mathrm{d}^2 + \mathrm{r}^2 + 8\mathrm{v}^4\mathrm{r}^2\mathrm{d}^2)\,\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) \vee \|(\mathfrak{a}\mathfrak{v})^2_\bullet(1 \vee k\mathfrak{t}^2_\bullet)^{-1}\|_{\mathbb{L}_\infty(\nu)} \text{ for all } n, k \in \mathbb{N}.\end{aligned}$$

Arguing similarly as in Remark §07|01.21 we note that $R^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) = o(1)$ as $n \to \infty$ since $(\mathfrak{t}^\dagger\mathfrak{v})_\bullet \mathbb{1}^m_\bullet \in \mathbb{L}_\infty(\nu)$ for all $m \in \mathbb{N}$ and $(\mathfrak{a}\mathfrak{v})_{(m)} = o(1)$ as $m \to \infty$ by Assumption §07|02.32. Moreover, we have $\|(\mathfrak{a}\mathfrak{v})^2_\bullet(1 \vee k\mathfrak{t}^2_\bullet)^{-1}\|_{\mathbb{L}_\infty(\nu)} = o(1)$ as $m \to \infty$ by dominated convergence. Note that the dimension $m^\star_n := m^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet)$ as defined in (07.37) does depend neither on the unknown parameter of interest $\theta_\bullet$ nor on the unknown operator $M_{\mathfrak{s}}$ but on the classes $\mathbb{J}^{\mathfrak{a},r}$ and $\mathbb{M}_{\mathfrak{t},d}$ only, and thus also the statistic $\widehat{\theta}^{m^\star_n}_\bullet$. In other words, if the regularity of $\theta_\bullet$ and $M_{\mathfrak{s}_\bullet}$ is known in advance, then the thresholded OPE $\widehat{\theta}^{m^\star_n}_\bullet$ is a feasible estimator. □

§07|02.37 **Corollary** (GdiSM with noisy operator §07|02.05 continued). *Consider independent noisy versions* $(\widehat{g}_\bullet, \widehat{\mathfrak{s}}_\bullet) = (g_\bullet + n^{-1/2}\dot{B}_\bullet, \mathfrak{s}_\bullet + k^{-1/2}\dot{W}_\bullet) \sim N^{n\otimes k}_{\theta|\mathfrak{s}} = N^n_{\theta|\mathfrak{s}} \otimes N^k_{\mathfrak{s}}$ *as in Model §07|02.05, where* $\dot{B}_\bullet \sim N^{\otimes\mathbb{N}}_{(0,1)}$ *and* $\dot{W}_\bullet \sim N^{\otimes\mathbb{N}}_{(0,1)}$ *are independent,* $\mathfrak{s}_\bullet \in \mathbb{R}^{\mathbb{N}}_{\setminus 0} \cap \ell_\infty$, $\theta_\bullet \in \ell_2$, *and hence* $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(M_{\mathfrak{s}^\dagger}) \subseteq \ell_2$. *Under Assumption §07|02.32 the thresholded OPE* $\widehat{\theta}^{m^\star_n}_\bullet = \widehat{\mathfrak{s}}^{(k)|\dagger}_\bullet \widehat{g}^{m^\star_n}_\bullet \in \ell_2(\mathfrak{v}^2_\bullet)$ *with dimension* $m^\star_n$ *as in* (07.37) *satisfies*

$$\sup\left\{N^{n\otimes k}_{\theta|\mathfrak{s}}(\|\widehat{\theta}^{m^\star_n}_\bullet - \theta_\bullet\|^2_{\mathfrak{v}}) \colon \theta_\bullet \in \ell_2^{\mathfrak{a},r}, M_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},d}\right\} \leqslant C_{r,d}\, R^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) \vee \|(\mathfrak{a}\mathfrak{v})^2_\bullet(1 \vee k\mathfrak{t}^2_\bullet)^{-1}\|_{\ell_\infty} \quad (07.40)$$

with constant $C_{r,d} = r^2 + 4d^2 + 12r^2d^2$.

§07|02.38 **Proof** of Corollary §07|02.37. Given in the lecture. □

§07|02.39 **Corollary** (diSM with noisy operator §07|02.07 continued). *Consider independent noisy versions* $(\widehat{g}_\bullet, \widehat{\mathfrak{s}}_\bullet) = (g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet, \mathfrak{s}_\bullet + k^{-1/2}\dot{\eta}_\bullet) \sim P^{n\otimes k}_{\theta|\mathfrak{s}|\sigma|\xi} = P^n_{\theta|\mathfrak{s}|\sigma} \otimes P^k_{\mathfrak{s}|\xi}$ *as in Model §07|02.07, where* $\dot{\varepsilon}_\bullet$ *and* $\dot{\eta}_\bullet$ *satisfy* **(iSM1)** *and* **(diSMnO1)** *with* $K_\sigma := \|\sigma_\bullet\|_{\ell_\infty} \vee 1 \in \mathbb{R}_{\geqslant 1}$ *and* $K_\xi := \|\xi_\bullet\|_{\ell_\infty} \vee 1 \in \mathbb{R}_{\geqslant 1}$, *respectively,* $\mathfrak{s}_\bullet \in \mathbb{R}^{\mathbb{N}}_{\setminus 0} \cap \ell_\infty$ *and* $\theta_\bullet \in \ell_2$, *and hence* $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(M_{\mathfrak{s}^\dagger}) \subseteq \ell_2$. *Under Assumption §07|02.32 the thresholded OPE* $\widehat{\theta}^{m^\star_n}_\bullet = \widehat{\mathfrak{s}}^{(k)|\dagger}_\bullet \widehat{g}^{m^\star_n}_\bullet \in \ell_2(\mathfrak{v}^2_\bullet)$ *with dimension* $m^\star_n$ *as in* (07.37) *satisfies*

$$\sup\left\{P^{n\otimes k}_{\theta|\mathfrak{s}|\sigma|\xi}(\|\widehat{\theta}^{m^\star_n}_\bullet - \theta_\bullet\|^2_{\mathfrak{v}}) \colon \theta_\bullet \in \ell_2^{\mathfrak{a},r}, M_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},d}\right\} \leqslant C_{r,d,\sigma,\xi}\, R^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) \vee \|(\mathfrak{a}\mathfrak{v})^2_\bullet(1 \vee k\mathfrak{t}^2_\bullet)^{-1}\|_{\ell_\infty} \quad (07.41)$$

with constant $C_{r,d,\sigma,\xi} = r^2 + 4K^2_\xi K^2_\sigma d^2 + 8K^4_\xi r^2 d^2$.

§07|02.40 **Proof** of Corollary §07|02.39. Given in the lecture. □

§07|02.41 **Corollary** (dieMM with noisy operator §07|02.09 continued). *Consider independent noisy versions* $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ *and* $\widehat{\mathfrak{s}}_\bullet = \mathfrak{s}_\bullet + k^{-1/2}\dot{\eta}_\bullet$ *defined on* $(\mathcal{Z}^{n+k}, \mathscr{Z}^{\otimes(n+k)}, \mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}} = \mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}} \otimes \mathbb{P}^{\otimes k}_{\mathfrak{s}})$ *as in Model §07|02.09, where* $\psi_\bullet, \varphi_\bullet \in \mathcal{M}(\mathscr{X} \otimes \mathscr{J})$ *satisfies* **(dieMM1)**–**(dieMM3)** *(Model §07|01.08) and* **(dieMMnO1)**–**(dieMMnO2)** *(Model §07|02.09) with* $\mathbb{v}_{\theta|\mathfrak{s}|\psi}, \mathbb{v}_{\mathfrak{s}|\varphi} \in \mathbb{R}_{\geqslant 1}$, *respectively,* $\mathfrak{s}_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$, $\theta_\bullet \in \mathbb{J}$ *and hence* $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(M_{\mathfrak{s}^\dagger}) \subseteq \mathbb{J}$. *Under Assumption §07|02.32 the thresholded OPE* $\widehat{\theta}^{m^\star_n}_\bullet = \widehat{\mathfrak{s}}^{(k)|\dagger}_\bullet \widehat{g}^{m^\star_n}_\bullet \in \mathbb{L}_2(\mathfrak{v}^2_\bullet\nu)$ *with dimension* $m^\star_n$ *as in* (07.37) *satisfies*

$$\sup\left\{P^{n\otimes k}_{\theta|\mathfrak{s}|\sigma|\xi}(\|\widehat{\theta}^{m^\star_n}_\bullet - \theta_\bullet\|^2_{\mathfrak{v}}) \colon \theta_\bullet \in \mathbb{J}^{\mathfrak{a},r}, M_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},d}\right\} \leqslant C_{\mathfrak{a},r,\mathfrak{t},d}\, R^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) \vee \|(\mathfrak{a}\mathfrak{v})^2_\bullet(1 \vee k\mathfrak{t}^2_\bullet)^{-1}\|_{\ell_\infty} \quad (07.42)$$

with constant $C_{\mathfrak{a},r,\mathfrak{t},d} = r^2 + 4d^2 \sup\left\{\mathbb{v}_{\theta|\mathfrak{s}|\psi}\mathbb{v}_{\mathfrak{s}|\varphi} \colon \theta_\bullet \in \mathbb{J}^{\mathfrak{a},r}, M_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},d}\right\} + 8r^2d^2 \sup\left\{\mathbb{v}^2_{\mathfrak{s}|\varphi} \colon M_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},d}\right\}$.

§07|02.42 **Proof** of Corollary §07|02.41. Given in the lecture. □

§07|02.43 **Illustration**. We illustrate the last results considering usual behaviour for $\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$. We distinguish again the two cases **(p)** and **(np)** in Illustration §07|01.44, where in case **(p)** the term $\mathrm{R}^\star_n(\mathfrak{a}_\bullet,\mathfrak{t}_\bullet,\mathfrak{v}_\bullet)$ is parametric, that is, $n\mathrm{R}^\star_n(\mathfrak{a}_\bullet,\mathfrak{t}_\bullet,\mathfrak{v}_\bullet) = \mathrm{O}(1)$, in case **(np)** it is nonparametric, i.e. $\lim_{n\to\infty} n\mathrm{R}^\star_n(\mathfrak{a}_\bullet,\mathfrak{t}_\bullet,\mathfrak{v}_\bullet) = \infty$. In case **(np)** we consider again three specifications similar to **(o-m)**, **(o-s)** and **(s-m)** introduced in Illustration §07|01.44 where also in Table 02 [§07] the order of the dimension $m^\star_n$ and the rate $\mathrm{R}^\star_n(\mathfrak{a}_\bullet,\mathfrak{t}_\bullet,\mathfrak{v}_\bullet)$ as $n\to\infty$ are given. The next table depict the rate $\mathrm{R}^\star_n(\mathfrak{a}_\bullet,\mathfrak{t}_\bullet,\mathfrak{v}_\bullet)$ and the additional term $\|(1\vee k\mathfrak{t}^2_\bullet)^{-1}(\mathfrak{a}\mathfrak{v})^2_\bullet\|_{\mathbb{L}_\infty(\nu)}$ as $n,k\to\infty$:

Table 06 [§07]

Order of $\mathrm{R}^\star_n(\mathfrak{a}_\bullet,\mathfrak{t}_\bullet,\mathfrak{v}_\bullet)$ and $\|(1\vee k\mathfrak{t}^2_\bullet)^{-1}(\mathfrak{a}\mathfrak{v})^2_\bullet\|_{\mathbb{L}_\infty(\nu)}$ as $n,k\to\infty$

	$(j\in\mathcal{J})$ $\mathfrak{v}^2_j = j^{2\mathrm{v}}$	$(\mathrm{a}\in\mathbb{R}_{>0})$ $\mathfrak{a}^2_j$	$(\mathrm{t}\in\mathbb{R}_{>0})$ $\mathfrak{t}^2_j$	$\mathrm{R}^\star_n(\mathfrak{a}_\bullet,\mathfrak{t}_\bullet,\mathfrak{v}_\bullet)$	$(\mathfrak{a}\mathfrak{v})^2_j$	$\mathfrak{t}^{-2}_j(\mathfrak{a}\mathfrak{v})^2_j$	$\|(1\vee k\mathfrak{t}^2_\bullet)^{-1}(\mathfrak{a}\mathfrak{v})^2_\bullet\|_{\mathbb{L}_\infty(\nu)}$
(o-m)	$\mathrm{v}\in(-1/2-\mathrm{t},\mathrm{a})$	$j^{-2\mathrm{a}}$	$j^{-2\mathrm{t}}$	$n^{-\frac{2(\mathrm{a}-\mathrm{v})}{2\mathrm{a}+2\mathrm{t}+1}}$	$j^{-2(\mathrm{a}-\mathrm{v})}$	$j^{2(\mathrm{t}+\mathrm{v}-\mathrm{a})}$	
	$\mathrm{a}-\mathrm{v}\leqslant\mathrm{t}$						$k^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
	$\mathrm{a}-\mathrm{v}\geqslant\mathrm{t}$						k^{-1}
(o-s)	$\mathrm{a}-\mathrm{v}\in\mathbb{R}_{>0}$	$j^{-2\mathrm{a}}$	$e^{-j^{2\mathrm{t}}}$	$(\log n)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$	$j^{-2(\mathrm{a}-\mathrm{v})}$	$j^{2(\mathrm{v}-\mathrm{a})}e^{j^{2\mathrm{t}}}$	$(\log k)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v}+\mathrm{t}+1/2\in\mathbb{R}_{>0}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$n^{-1}(\log n)^{\frac{2(\mathrm{t}+\mathrm{v})+1}{2\mathrm{a}}}$	$j^{2\mathrm{v}}e^{-j^{2\mathrm{a}}}$	$j^{2(\mathrm{t}+\mathrm{v})}e^{-j^{2\mathrm{a}}}$	k^{-1}

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v}+\mathrm{t}<-1/2$ the rate $\mathrm{R}^\star_n(\mathfrak{a}_\bullet,\mathfrak{t}_\bullet,\mathfrak{v}_\bullet)$ is parametric. □

§07|02|03 Local and maximal local ϕ-risk

Secondly, we measure the accuracy of the tOPE $\widehat{\theta}^m_\bullet := \widehat{\mathfrak{s}}^{(k)|\dagger}_\bullet \widehat{g}^m_\bullet$ of $\theta^m_\bullet = \mathfrak{s}^\dagger_\bullet g^m_\bullet \in \mathbb{J}\mathbb{1}^m_\bullet$ with $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}})$ and $\mathfrak{s}_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})\cap\mathbb{L}_\infty(\nu)$ by the mean of its local ϕ-error introduced in §04|03|02, i.e. its ϕ-risk.

§07|02.44 **Reminder**. If $\phi\in\mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ and $\theta_\bullet\in\mathrm{dom}(\phi\nu)$, then for each $m\in\mathbb{N}$ we have $\theta^m_\bullet\in\mathrm{dom}(\phi\nu)$ too and $|\phi\nu(\theta_\bullet)-\phi\nu(\theta^m_\bullet)| = \mathrm{o}(1)$ as $m\to\infty$ (Property §04|03.13). □

§07|02.45 **Assumption**. Let $(\widehat{g}_\bullet,\widehat{\mathfrak{s}}_\bullet) = (g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet, \mathfrak{s}_\bullet + k^{-1/2}\dot{\eta}_\bullet) \sim \mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}} := \mathbb{P}^n_{\theta|\mathfrak{s}}\otimes\mathbb{P}^k_{\mathfrak{s}}$ be *independent* noisy versions satisfying Assumption §07|02.01. In addition

(dSIPl1) $\dot{\varepsilon}_\bullet$ admit a covariance operator, say $\Gamma_{\theta|\mathfrak{s}}\in\mathbb{L}^{\geqslant}(\mathbb{J})$, i.e. $\dot{\varepsilon}_\bullet\sim\mathrm{P}_{(0,\Gamma_{\theta|\mathfrak{s}})}$, $\mathrm{K}^2_{\theta|\mathfrak{s}} := 1\vee\|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})}$,

(dSIPl2) $\dot{\varepsilon}_\bullet\mathbb{1}^m_\bullet\in\mathbb{J}=\mathbb{L}_2(\nu)$ $\mathbb{P}^n_{\theta|\mathfrak{s}}$-a.s. for each $m\in\mathbb{N}$, and

(dSIPnO) $\mathrm{v}^{\mathfrak{s}|(2)}_\bullet := \mathbb{P}^k_{\mathfrak{s}}(\dot{\eta}^4_\bullet) := (\mathrm{v}^{\mathfrak{s}|(2)}_j := \mathbb{P}^k_{\mathfrak{s}}(\dot{\eta}^4_j))_{j\in\mathcal{J}}\in\mathbb{L}_\infty(\nu)$, $\mathrm{K}^4_{\mathfrak{s}} := 1\vee\|\mathrm{v}^{\mathfrak{s}|(2)}_\bullet\|_{\mathbb{L}_\infty(\nu)}$.

Moreover, from **(dSIPnO)** (i.e. $\mathrm{v}^{\mathfrak{s}|(2)}_\bullet = \mathbb{P}^k_{\mathfrak{s}}(\dot{\eta}^4_\bullet)\in\mathbb{L}_\infty(\nu)$) follows $\mathbb{P}^k_{\mathfrak{s}}(\dot{\eta}^2_j) =: \mathrm{v}^{\mathfrak{s}}_j \leqslant (\mathrm{v}^{\mathfrak{s}|(2)}_j)^{1/2}$ for ν-a.e. $j\in\mathcal{J}$, and hence $\|\mathrm{v}^{\mathfrak{s}}_\bullet\vee\mathbb{1}_\bullet\|_{\mathbb{L}_\infty(\nu)}\leqslant\mathrm{K}^2_{\mathfrak{s}}$. □

§07|02.46 **Notation**. Since $\|\widehat{\mathfrak{s}}^{(k)|\dagger}_\bullet\|_{\mathbb{L}_\infty(\nu)}\leqslant k^{1/2}$ (Notation §07|02.03), $\mathfrak{s}_\bullet\in\mathbb{L}_\infty(\nu)$ and $\mathbb{1}^m_\bullet\in\mathbb{L}_2(\nu)=\mathbb{J}$ (using $\nu([\![m]\!])\in\mathbb{R}_{>0}$ by Assumption §07|00.02) for all $m\in\mathbb{N}$, for $(\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{s})_\bullet := \widehat{\mathfrak{s}}^{(k)|\dagger}_\bullet\mathfrak{s}_\bullet\in\mathcal{M}(\mathscr{A}\otimes\mathscr{J})$ we have $(\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{s})_\bullet\mathbb{1}^m_\bullet\in\mathbb{L}_\infty(\nu)$ for all $m\in\mathbb{N}$ too. If in addition $\mathbb{1}^m_\bullet\in\mathbb{L}_2(\phi^2\nu)$ for all $m\in\mathbb{N}$ then for $(\widehat{\mathfrak{s}}^{(k)|\dagger}\phi)_\bullet := \widehat{\mathfrak{s}}^{(k)|\dagger}_\bullet\phi\in\mathcal{M}(\mathscr{A}\otimes\mathscr{J})$ we also have $(\widehat{\mathfrak{s}}^{(k)|\dagger}\phi)_\bullet\mathbb{1}^m_\bullet\in\mathbb{J}$ for all $m\in\mathbb{N}$. □

§07|02.47 **Comment**. Under Assumption §07|02.45 and $\phi\in\mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ if $\mathbb{1}^m_\bullet\in\mathbb{L}_2(\phi^2\nu)$ for all $m\in\mathbb{N}$ then we have $\widehat{\mathfrak{s}}^{(k)|\dagger}_\bullet\dot{\varepsilon}_\bullet\mathbb{1}^m_\bullet\in\mathrm{dom}(\phi\nu)$ $\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}$-a.s. (since $\nu(|(\widehat{\mathfrak{s}}^{(k)|\dagger}\phi)_\bullet\dot{\varepsilon}_\bullet\mathbb{1}^m_\bullet|)\leqslant\|(\widehat{\mathfrak{s}}^{(k)|\dagger}\phi)_\bullet\mathbb{1}^m_\bullet\|_{\mathbb{J}}\|\dot{\varepsilon}_\bullet\mathbb{1}^m_\bullet\|_{\mathbb{J}}\in\mathbb{R}_{\geqslant 0}$ $\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}$-a.s.). If in addition $\theta_\bullet\in\mathrm{dom}(\phi\nu)$, and hence $\theta^m_\bullet\in\mathrm{dom}(\phi\nu)$ (Property §04|03.13), then it follows

$$\widehat{\theta}^m_\bullet = \widehat{\mathfrak{s}}^{(k)|\dagger}_\bullet\,\widehat{g}_\bullet\,\mathbb{1}^m_\bullet = n^{-1/2}\widehat{\mathfrak{s}}^{(k)|\dagger}_\bullet\dot{\varepsilon}_\bullet\mathbb{1}^m_\bullet + (\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{s})_\bullet\theta^m_\bullet\in\mathrm{dom}(\phi\nu)\quad\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}\text{-a.s.}. \tag{07.43}$$

If $\mathcal{J}\subseteq\mathbb{Z}$ (at most countable) and $\nu_{\mathcal{J}}$ is the counting measure over the index set $\mathcal{J}$ then Assumption §01|01.04 and **(dSIPl1)** (i.e. $\dot{\varepsilon}_\bullet\sim\mathrm{P}_{(0,\Gamma_{\theta|\mathfrak{s}})}$) implies $\mathrm{v}^{\theta|\mathfrak{s}}_\bullet = \mathbb{P}^n_{\theta|\mathfrak{s}}(|\dot{\varepsilon}_\bullet|^2)\in\mathbb{L}_\infty(\nu_{\mathcal{J}})$ and hence the additional assumption **(dSIPl2)** $\dot{\varepsilon}_{\cdot}\mathbb{1}^m_{\cdot} \in \mathbb{J} = \mathbb{L}_2(\nu_{\mathscr{J}})$ $\mathbb{P}^n_{\theta|\mathfrak{s}}$-a.s.. However, the last implication does generally not hold, if $\mathscr{J} \in \{\mathbb{R}, \mathbb{R}_{\geqslant 0}\}$ for example. □

§07|02|03|01 Local ϕ-risk

§07|02.48 **Assumption**. Let $\phi \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$, $\theta_{\cdot} \in \mathrm{dom}(\phi\nu)$, $\mathfrak{s}^{\dagger}_{\cdot}\mathbb{1}^m_{\cdot}, \mathbb{1}^m_{\cdot} \in \mathbb{L}_2(\phi^2\nu)$ for $m \in \mathbb{N}$ be satisfied. □

§07|02.49 **Definition**. Under Assumptions §07|02.45 and §07|02.48for $m \in \mathbb{N}$ the *local ϕ-risk* of a thresholded OPE $\widehat{\theta}^m_{\cdot} = \widehat{\mathfrak{s}}^{(k)|\dagger}_{\cdot}\widehat{g}^m_{\cdot} = \widehat{\mathfrak{s}}^{\dagger}_{\cdot}\mathbb{1}^{\{\widehat{\mathfrak{s}}^2 \geqslant k^{-1}\}}_{\cdot}\widehat{g}_{\cdot}\mathbb{1}^m_{\cdot} \in \mathrm{dom}(\phi\nu)$ $\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}$-a.s. satisfies

$$\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^m_{\cdot} - \theta_{\cdot})|^2) = \mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}(|\nu((\widehat{\mathfrak{s}}^{(k)|\dagger}\phi)_{\cdot}(\widehat{g}_{\cdot} - g_{\cdot})\mathbb{1}^m_{\cdot})|^2) + \mathbb{P}^k_{\mathfrak{s}}(|\nu(((\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{s})_{\cdot}\mathbb{1}^m_{\cdot} - \mathbb{1}_{\cdot})\phi\theta_{\cdot})|^2). \quad (07.44)$$

with *variance* $\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}(|\nu((\widehat{\mathfrak{s}}^{(k)|\dagger}\phi)_{\cdot}(\widehat{g}_{\cdot} - g_{\cdot})\mathbb{1}^m_{\cdot})|^2)$ and *bias* $\mathbb{P}^k_{\mathfrak{s}}(|\nu(((\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{s})_{\cdot}\mathbb{1}^m_{\cdot} - \mathbb{1}_{\cdot})\phi\theta_{\cdot})|^2)$. □

§07|02.50 **Property**. *Under Assumptions §07|02.45 and §07|02.48 (exploiting the independence of $(\widehat{\mathfrak{s}}^{(k)|\dagger}\phi)_{\cdot}$ and $\dot{\varepsilon}_{\cdot}$, $\dot{\varepsilon}_{\cdot} \sim \mathrm{P}_{(0,\Gamma_{\theta|\mathfrak{s}})}$ with $\Gamma_{\theta|\mathfrak{s}} \in \mathbb{L}^{\geqslant}(\mathbb{J})$, and $(\widehat{\mathfrak{s}}^{(k)|\dagger}\phi)_{\cdot}\mathbb{1}^m_{\cdot} \in \mathbb{L}_2(\nu)$) we have*

$$n\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}(|\nu((\widehat{\mathfrak{s}}^{(k)|\dagger}\phi)_{\cdot}(\widehat{g}_{\cdot} - g_{\cdot})\mathbb{1}^m_{\cdot})|^2) = \mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}(|\nu((\widehat{\mathfrak{s}}^{(k)|\dagger}\phi)_{\cdot}\dot{\varepsilon}_{\cdot}\mathbb{1}^m_{\cdot})|^2)$$
$$= \mathbb{P}^k_{\mathfrak{s}}\langle\Gamma_{\theta|\mathfrak{s}}((\widehat{\mathfrak{s}}^{(k)|\dagger}\phi)_{\cdot}\mathbb{1}^m_{\cdot}), (\widehat{\mathfrak{s}}^{(k)|\dagger}\phi)_{\cdot}\mathbb{1}^m_{\cdot}\rangle_{\mathbb{J}} \leqslant \|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})}\nu(\mathbb{P}^k_{\mathfrak{s}}(|\widehat{\mathfrak{s}}^{(k)|\dagger}_{\cdot}|^2)\phi^2_{\cdot}\mathbb{1}^m_{\cdot}).$$

Since $\mathfrak{s}_{\cdot} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ and $\mathfrak{s}^{\dagger}_{\cdot}\mathbb{1}^m_{\cdot} \in \mathbb{L}_2(\phi^2\nu)$ the last bound together with Lemma §07|02.18 (i) implies

$$n\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}(|\nu((\widehat{\mathfrak{s}}^{(k)|\dagger}\phi)_{\cdot}(\widehat{g}_{\cdot} - g_{\cdot})\mathbb{1}^m_{\cdot})|^2) \leqslant \|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})}\nu(\mathbb{P}^k_{\mathfrak{s}}((\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{s})^2_{\cdot})(\mathfrak{s}^{\dagger}\phi)^2_{\cdot}\mathbb{1}^m_{\cdot})$$
$$\leqslant \|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})}2(\|\mathrm{v}^{\mathfrak{s}}_{\cdot}\|_{\mathbb{L}_\infty(\nu)} + 1)\|\mathfrak{s}^{\dagger}_{\cdot}\mathbb{1}^m_{\cdot}\|^2_{\phi} \leqslant \|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})}4\|\mathrm{v}^{\mathfrak{s}}_{\cdot} \vee \mathbb{1}_{\cdot}\|_{\mathbb{L}_\infty(\nu)}\|\mathfrak{s}^{\dagger}_{\cdot}\mathbb{1}^m_{\cdot}\|^2_{\phi}. \quad (07.45)$$

Moreover, assuming $\theta_{\cdot} \in \mathrm{dom}(\phi\nu)$ and hence $\theta_{\cdot}\mathbb{1}^m_{\cdot} \in \mathrm{dom}(\phi\nu)$, we obtain (using $\widehat{\mathfrak{s}}^{(k)|\dagger}_{\cdot}\widehat{\mathfrak{s}}_{\cdot} = \mathbb{1}^{\{\widehat{\mathfrak{s}}^2 < k^{-1}\}}_{\cdot}$ and applying the generalised Minkowski inequality)

$$\tfrac{1}{3}\mathbb{P}^k_{\mathfrak{s}}\left(|\nu(((\widehat{\mathfrak{s}}^{(k)|\dagger}\mathfrak{s})_{\cdot}\mathbb{1}^m_{\cdot} - \mathbb{1}_{\cdot})\phi\theta_{\cdot})|^2\right) - |\phi\nu(\theta_{\cdot}\mathbb{1}^{m|\perp}_{\cdot})|^2$$
$$\leqslant \mathbb{P}^k_{\mathfrak{s}}(|\nu(\widehat{\mathfrak{s}}^{(k)|\dagger}_{\cdot}(\mathfrak{s}_{\cdot} - \widehat{\mathfrak{s}}_{\cdot})\phi\theta_{\cdot}\mathbb{1}^m_{\cdot})|^2) + \mathbb{P}^k_{\mathfrak{s}}(|\nu(\mathbb{1}^{\{\widehat{\mathfrak{s}}^2 < k^{-1}\}}_{\cdot}\phi\theta_{\cdot}\mathbb{1}^m_{\cdot})|^2)$$
$$\leqslant \left|\nu\left(|\phi\theta_{\cdot}\mathbb{1}^m_{\cdot}|(\mathbb{P}^k_{\mathfrak{s}}(|\mathfrak{s}_{\cdot} - \widehat{\mathfrak{s}}_{\cdot}|^2|\widehat{\mathfrak{s}}^{(k)|\dagger}_{\cdot}|^2))^{1/2}\right)\right|^2 + \left|\nu\left(|\phi\theta_{\cdot}\mathbb{1}^m_{\cdot}|(\mathbb{P}^k_{\mathfrak{s}}(\widehat{\mathfrak{s}}^2_{\cdot} < k^{-1}))^{1/2}\right)\right|^2$$
$$\leqslant 2(\|\mathrm{v}^{\mathfrak{s}|(2)}_{\cdot}\|_{\mathbb{L}_\infty(\nu)} + \|\mathrm{v}^{\mathfrak{s}}_{\cdot}\|_{\mathbb{L}_\infty(\nu)} + 2\|\mathrm{v}^{\mathfrak{s}}_{\cdot} \vee \mathbb{1}_{\cdot}\|_{\mathbb{L}_\infty(\nu)})\|\theta_{\cdot}\mathbb{1}^m_{\cdot}(1 \vee k\mathfrak{s}^2_{\cdot})^{-1/2}\|^2_{\mathbb{L}_1(|\phi|\nu)} \quad (07.46)$$

where the last inequality follows from Lemma §07|02.18 (ii) and (iii). □

§07|02.51 **Reminder**. If Assumptions §07|01.46 and §07|01.48 are satisfied then for all $m, n \in \mathbb{N}$ setting

$$\mathrm{R}^m_n(\theta_{\cdot}, \mathfrak{s}_{\cdot}, \phi) := |\phi\nu(\theta_{\cdot}\mathbb{1}^{m|\perp}_{\cdot})|^2 + n^{-1}\|\mathfrak{s}^{\dagger}_{\cdot}\mathbb{1}^m_{\cdot}\|^2_{\phi}, \quad m^{\circ}_n := \arg\min\left\{\mathrm{R}^m_n(\theta_{\cdot}, \mathfrak{s}_{\cdot}, \phi) : m \in \mathbb{N}\right\}$$
$$\text{and} \quad \mathrm{R}^{\circ}_n(\theta_{\cdot}, \mathfrak{s}_{\cdot}, \phi) := \mathrm{R}^{m^{\circ}_n}_n(\theta_{\cdot}, \mathfrak{s}_{\cdot}, \phi) := \min\left\{\mathrm{R}^m_n(\theta_{\cdot}, \mathfrak{s}_{\cdot}, \phi) : m \in \mathbb{N}\right\} \quad (07.47)$$

the OPE $\widehat{\theta}^{m^{\circ}_n}_{\cdot} := \mathfrak{s}^{\dagger}_{\cdot}\widehat{g}^m_{\cdot}$ with *known* $\mathfrak{s}_{\cdot} \in \mathbb{L}_\infty(\nu)$ fulfills $\mathbb{P}^n_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^{m^{\circ}_n}_{\cdot} - \theta_{\cdot})|^2) \leqslant (1 \vee \|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})})\mathrm{R}^{\circ}_n(\theta_{\cdot}, \mathfrak{s}_{\cdot}, \phi)$ due to Proposition §07|01.51. Keep in mind that Assumption §07|01.46 is part of Assumption §07|02.45 and that Assumption §07|01.48 is part of Assumption §07|02.48. □

§07|02.52 **Proposition** (Upper bound). *Let Assumptions §07|02.45 and §07|02.48 be satisfied. The thresholded OPE $\widehat{\theta}^m_{\cdot} = \widehat{\mathfrak{s}}^{(k)|\dagger}_{\cdot}\widehat{g}^m_{\cdot} \in \mathrm{dom}(\phi\nu)$ $\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}$-a.s. for all $n, k, m \in \mathbb{N}$ fulfills*

$$\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^m_{\cdot} - \theta_{\cdot})|^2) \leqslant 4(\|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})} \vee 1)\|\mathrm{v}^{\mathfrak{s}}_{\cdot} \vee \mathbb{1}_{\cdot}\|_{\mathbb{L}_\infty(\nu)}\mathrm{R}^m_n(\theta_{\cdot}, \mathfrak{s}_{\cdot}, \phi)$$
$$+ 6(\|\mathrm{v}^{\mathfrak{s}|(2)}_{\cdot}\|_{\mathbb{L}_\infty(\nu)} + 3\|\mathrm{v}^{\mathfrak{s}}_{\cdot} \vee \mathbb{1}_{\cdot}\|_{\mathbb{L}_\infty(\nu)})\|\theta_{\cdot}\mathbb{1}^m_{\cdot}(1 \vee k\mathfrak{s}^2_{\cdot})^{-1/2}\|^2_{\mathbb{L}_1(|\phi|\nu)} \quad (07.48)$$
$$\leqslant 4\mathrm{K}^2_{\theta|\mathfrak{s}}\mathrm{K}^2_{\mathfrak{s}}\,\mathrm{R}^m_n(\theta_{\cdot}, \mathfrak{s}_{\cdot}, \phi) + 24\mathrm{K}^4_{\mathfrak{s}}\,\|\theta_{\cdot}\mathbb{1}^m_{\cdot}(1 \vee k\mathfrak{s}^2_{\cdot})^{-1/2}\|^2_{\mathbb{L}_1(|\phi|\nu)}. \quad (07.49)$$

§07|02.53 **Proof** of Proposition §07|02.52. Given in the lecture. □

§07|02.54 **Comment**. Selecting $m_n^\circ := \arg\min\{\mathrm{R}_n^m(\theta_\bullet, \mathfrak{s}_\bullet, \phi) : m \in \mathbb{N}\}$ and $\mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \mathfrak{v}_\bullet) = \mathrm{R}_n^{m_n^\circ}(\theta_\bullet, \mathfrak{s}_\bullet, \phi)$ as in (07.47) (Reminder §07|02.51) from Proposition §07|02.52 we obtain immediately

$$\mathbb{E}_{\theta|\mathfrak{s}}^{n\otimes k}(|\phi\nu(\widehat{\theta}_\bullet^{m_n^\circ} - \theta_\bullet)|^2) \leqslant 4\mathrm{K}_{\mathfrak{s}}^2 \mathrm{K}_{\theta|\mathfrak{s}}^2 \,\mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \phi) + 24\mathrm{K}_{\mathfrak{s}}^4 \|\theta_\bullet \mathbb{1}_\bullet^{m_n^\circ}(1 \vee k\mathfrak{s}_\bullet^2)^{-1/2}\|_{\mathbb{L}_1(|\phi|\nu)}^2. \tag{07.50}$$

We shall emphasise, that $\mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \phi)$ is the oracle rate (Property §07|01.55) if $\mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathbb{M}}(\mathbb{J})$ is *known* in advance. Furthermore, for each $m \in \mathbb{N}$ we have

$$\begin{aligned}\|\theta_\bullet \mathbb{1}_\bullet^m (1 \vee k\mathfrak{s}_\bullet^2)^{-1/2}\|_{\mathbb{L}_1(|\phi|\nu)} &\leqslant \|\theta_\bullet (1 \vee k\mathfrak{s}_\bullet^2)^{-1/2}\|_{\mathbb{L}_1(|\phi|\nu)} \\ &= \|\theta_\bullet (1 \vee k\mathfrak{s}_\bullet^2)^{-1/2}\mathbb{1}_\bullet^m\|_{\mathbb{L}_1(|\phi|\nu)} + \|\theta_\bullet (1 \vee k\mathfrak{s}_\bullet^2)^{-1/2}\mathbb{1}_\bullet^{m|\perp}\|_{\mathbb{L}_1(|\phi|\nu)} \\ &\leqslant \|\theta_\bullet (1 \vee k\mathfrak{s}_\bullet^2)^{-1/2}\mathbb{1}_\bullet^m\|_{\mathbb{L}_1(|\phi|\nu)} + \|\theta_\bullet \mathbb{1}_\bullet^{m|\perp}\|_{\mathbb{L}_1(|\phi|\nu)}.\end{aligned} \tag{07.51}$$

Consequently, under the assumptions of Proposition §07|02.52 from (??) (Proof §07|02.53) follows

$$\begin{aligned}\mathbb{E}_{\theta|\mathfrak{s}}^{n\otimes k}(|\phi\nu(\widehat{\theta}_\bullet^m - \theta_\bullet)|^2) &\leqslant 4\mathrm{K}_{\theta|\mathfrak{s}}^2 \mathrm{K}_{\mathfrak{s}}^2 \,\mathrm{R}_n^m(\theta_\bullet, \mathfrak{s}_\bullet, \phi) + 24\mathrm{K}_{\mathfrak{s}}^4 \,\|\theta_\bullet \mathbb{1}_\bullet^m (1 \vee k\mathfrak{s}_\bullet^2)^{-1/2}\|_{\mathbb{L}_1(|\phi|\nu)}^2 \\ &\leqslant 4\mathrm{K}_{\theta|\mathfrak{s}}^2 \mathrm{K}_{\mathfrak{s}}^2 \,\mathrm{R}_n^m(\theta_\bullet, \mathfrak{s}_\bullet, \phi) + 24\mathrm{K}_{\mathfrak{s}}^4 \,\|\theta_\bullet (1 \vee k\mathfrak{s}_\bullet^2)^{-1/2}\|_{\mathbb{L}_1(|\phi|\nu)}^2 \\ &\leqslant 28\mathrm{K}_{\mathfrak{s}}^4 \mathrm{K}_{\theta|\mathfrak{s}}^2 \left(\|\theta_\bullet \mathbb{1}_\bullet^{m|\perp}\|_{\mathbb{L}_1(|\phi|\nu)}^2 + n^{-1}\|\mathfrak{s}_\bullet^\dagger \mathbb{1}_\bullet^m\|_\phi^2\right) + 24\mathrm{K}_{\mathfrak{s}} \|\theta_\bullet \mathbb{1}_\bullet^m (1 \vee k\mathfrak{s}_\bullet^2)^{-1/2}\|_{\mathbb{L}_1(|\phi|\nu)}^2.\end{aligned}$$

Selecting $m_n^\diamond := \arg\min\left\{\|\theta_\bullet \mathbb{1}_\bullet^{m|\perp}\|_{\mathbb{L}_1(|\phi|\nu)}^2 + n^{-1}\|\mathfrak{s}_\bullet^\dagger \mathbb{1}_\bullet^m\|_\phi^2 : m \in \mathbb{N}\right\}$ we obtain

$$\begin{aligned}\mathbb{E}_{\theta|\mathfrak{s}}^{n\otimes k}(|\phi\nu(\widehat{\theta}_\bullet^{m_n^\diamond} - \theta_\bullet)|^2) \leqslant 3\mathrm{K}_{\mathfrak{s}}^{1/2}\mathrm{K}_{\theta|\mathfrak{s}} \min\left\{\|\theta_\bullet \mathbb{1}_\bullet^{m|\perp}\|_{\mathbb{L}_1(|\phi|\nu)}^2 + n^{-1}\|\mathfrak{s}_\bullet^\dagger \mathbb{1}_\bullet^m\|_\phi^2 : m \in \mathbb{N}\right\} \\ + 24\mathrm{K}_{\mathfrak{s}} \|\theta_\bullet (1 \vee k\mathfrak{s}_\bullet^2)^{-1/2}\|_{\mathbb{L}_1(|\phi|\nu)}^2.\end{aligned}$$

We shall emphasise, that the last upper bound consists (up to the constants) of the sum of the two terms depending each on one of the sample sizes n and k only. However, the first term $\min\left\{\|\theta_\bullet \mathbb{1}_\bullet^{m|\perp}\|_{\mathbb{L}_1(|\phi|\nu)}^2 + n^{-1}\|\mathfrak{s}_\bullet^\dagger \mathbb{1}_\bullet^m\|_\phi^2 : m \in \mathbb{N}\right\}$ is generally larger than the oracle rate $\mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \phi)$. □

§07|02.55 **Corollary** (GdiSM with noisy operator §07|02.05 continued). *Consider* independent *noisy versions* $(\widehat{g}_\bullet, \widehat{\mathfrak{s}}_\bullet) = (g_\bullet + n^{-1/2}\dot{\mathrm{B}}_\bullet, \mathfrak{s}_\bullet + k^{-1/2}\dot{\mathrm{W}}_\bullet) \sim \mathrm{N}_{\theta|\mathfrak{s}}^{n\otimes k} = \mathrm{N}_{\theta|\mathfrak{s}}^n \otimes \mathrm{N}_{\mathfrak{s}}^k$ *as in Model §07|02.05, where* $\dot{\mathrm{B}}_\bullet \sim \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}$ *and* $\dot{\mathrm{W}}_\bullet \sim \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}$ *are* independent, $\mathfrak{s}_\bullet \in \mathbb{R}_{\setminus 0}^{\mathbb{N}} \cap \ell_\infty$ *and* $\theta_\bullet \in \ell_2$, *and hence* $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \subseteq \ell_2$. *Given* $\phi \in \mathbb{R}_{\setminus 0}^{\mathbb{N}}$ *and* $\theta_\bullet \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *the (infeasible) thresholded OPE* $\widehat{\theta}_\bullet^{m_n^\circ} = \widehat{\mathfrak{s}}_\bullet^{(k)|\dagger}\widehat{g}_\bullet^{m_n^\circ} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *with oracle dimension* m_n° *as in* (07.47) *satisfies*

$$\mathrm{N}_{\theta|\mathfrak{s}}^{n\otimes k}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_\bullet^{m_n^\circ} - \theta_\bullet)|^2) \leqslant 4\mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \phi) + 36\|(1 \vee k\mathfrak{s}_\bullet^2)^{-1/2}\theta_\bullet\|_{\ell_1(|\phi|)}^2 \tag{07.52}$$

where $\mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \phi)$ *is the oracle rate in a* GdiSM *§07|01.03 (see Corollary §07|01.57).*

§07|02.56 **Proof** of Corollary §07|02.55. Given in the lecture. □

§07|02.57 **Corollary** (diSM with noisy operator §07|02.07 continued). *Consider* independent *noisy versions* $(\widehat{g}_\bullet, \widehat{\mathfrak{s}}_\bullet) = (g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet, \mathfrak{s}_\bullet + k^{-1/2}\dot{\eta}_\bullet) \sim \mathrm{P}_{\theta|\mathfrak{s}|\sigma|\xi}^{n\otimes k} = \mathrm{P}_{\theta|\mathfrak{s}|\sigma}^n \otimes \mathrm{P}_{\mathfrak{s}|\xi}^k$ *as in Model §07|02.07, where* $\dot{\varepsilon}_\bullet$ *and* $\dot{\eta}_\bullet$ *satisfy* **(iSM1)** *and* **(diSMnO1)** *with* $\mathrm{K}_\sigma := \|\sigma\|_{\ell_\infty} \vee 1 \in \mathbb{R}_{\geqslant 1}$ *and* $\mathrm{K}_\xi := \|\xi\|_{\ell_\infty} \vee 1 \in \mathbb{R}_{\geqslant 1}$, *respectively,* $\mathfrak{s}_\bullet \in \mathbb{R}_{\setminus 0}^{\mathbb{N}} \cap \ell_\infty$ *and* $\theta_\bullet \in \ell_2$, *and hence* $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \subseteq \ell_2$. *Given* $\phi \in \mathbb{R}_{\setminus 0}^{\mathbb{N}}$ *and* $\theta_\bullet \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *the (infeasible) thresholded OPE* $\widehat{\theta}_\bullet^{m_n^\circ} = \widehat{\mathfrak{s}}_\bullet^{(k)|\dagger}\widehat{g}_\bullet^{m_n^\circ} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *with oracle dimension* m_n° *as in* (07.47) *satisfies*

$$\mathrm{P}_{\theta|\mathfrak{s}|\sigma|\xi}^{n\otimes k}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_\bullet^{m_n^\circ} - \theta_\bullet)|^2) \leqslant 4\mathrm{K}_\sigma^2\mathrm{K}_\xi^2 \,\mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \phi) + 24\mathrm{K}_\xi^4\|(1 \vee k\mathfrak{s}_\bullet^2)^{-1/2}\theta_\bullet\|_{\ell_1(|\phi|)}^2 \quad \forall n, k \in \mathbb{N} \tag{07.53}$$

where $\mathrm{R}_n^\circ(\theta_\bullet, \mathfrak{s}_\bullet, \phi)$ *is the oracle rate in a* diSM *§07|01.06 (see Corollary §07|01.59).*

§07|02.58 **Proof** of Corollary §07|02.57. Given in the lecture. □

§07|02.59 **Corollary** (dieMM with noisy operator §07|02.09 continued). *Consider independent noisy versions $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ and $\widehat{\mathfrak{s}}_{\bullet} = \mathfrak{s}_{\bullet} + k^{-1/2}\dot{\eta}_{\bullet}$ defined on $(\mathcal{Z}^{n+k}, \mathscr{Z}^{\otimes(n+k)}, \mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}} = \mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}} \otimes \mathbb{P}^{\otimes k}_{\mathfrak{s}})$ as in Model §07|02.09, where $\psi_{\bullet}, \varphi_{\bullet} \in \mathcal{M}(\mathscr{X} \otimes \mathscr{J})$ satisfies* **(dieMM1)–(dieMM3)** *(Model §07|01.08) and* **(dieMMnO1)–(dieMMnO2)** *(Model §07|02.09) with $\mathbb{v}_{\theta|\mathfrak{s}|\psi}, \mathbb{v}_{\mathfrak{s}|\varphi} \in \mathbb{R}_{\geqslant 1}$, respectively, $\mathfrak{s}_{\bullet} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J}) \cap \mathbb{L}_{\infty}(\nu)$, $\theta_{\bullet} \in \mathbb{J}$ and hence $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}}) \subseteq \ell_2$. Given Assumption §07|02.48 the (infeasible) tOPE $\widehat{\theta}^{m_n^{\circ}}_{\bullet} = \widehat{\mathfrak{s}}^{(k)|\dagger}_{\bullet}\widehat{g}^{m_n^{\circ}}_{\bullet} \in \mathrm{dom}(\phi\nu)$ with oracle dimension m_n° as in (07.29) satisfies*

$$\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m_n^{\circ}}_{\bullet} - \theta_{\bullet})|^2) \leqslant 4\mathbb{v}_{\theta|\mathfrak{s}|\psi}\mathbb{v}_{\mathfrak{s}|\varphi}\,\mathrm{R}^{\circ}_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi_{\bullet}) + 24\mathbb{v}^2_{\mathfrak{s}|\varphi}\|(1 \vee k\mathfrak{s}^2_{\bullet})^{-1/2}\theta_{\bullet}\|^2_{\ell_1(|\phi|)} \quad \forall n, k \in \mathbb{N} \quad (07.54)$$

where $\mathrm{R}^{\circ}_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi_{\bullet})$ is the oracle rate in a dieMM §07|01.08 (see Corollary §07|01.61).

§07|02.60 **Proof** of Corollary §07|02.59. Given in the lecture. □

§07|02.61 **Illustration**. We illustrate the last results considering usual behaviour for $\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi_{\bullet} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$. We distinguish again the two cases **(p)** and **(np)** in Illustration §07|01.63, where in case **(p)** the term $\mathrm{R}^{\circ}_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi_{\bullet})$ is parametric, that is, $n\mathrm{R}^{\circ}_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi_{\bullet}) = \mathrm{O}(1)$, in case **(np)** it is nonparametric, i.e. $\lim_{n\to\infty} n\mathrm{R}^{\circ}_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi_{\bullet}) = \infty$. In case **(np)** we consider again the three specifications **(o-m)**, **(o-s)** and **(s-m)** introduced in Illustration §07|01.63 where also in Table 03 [§07] the order of the oracle dimension m_n° and the oracle rate $\mathrm{R}^{\circ}_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi_{\bullet})$ as $n \to \infty$ are given. The next table depict the oracle rate $\mathrm{R}^{\circ}_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi_{\bullet})$ and the rate of the additional term $\|(1 \vee k\mathfrak{s}^2_{\bullet})^{-1/2}\theta_{\bullet}\|^2_{\ell_1(|\phi|)}$ as $n, k \to \infty$:

Table 07 [§07]

Order of $\mathrm{R}^{\circ}_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi_{\bullet})$ and $\|(1 \vee k\mathfrak{s}^2_{\bullet})^{-1/2}\theta_{\bullet}\|^2_{\ell_1(|\phi|)}$ as $n, k \to \infty$

	$(j \in \mathcal{J})$ $\phi_j = j^{\mathrm{v}-1/2}$	$(\mathrm{a} \in \mathbb{R}_{>0})$ θ_j	$(\mathrm{t} \in \mathbb{R}_{>0})$ $\mathfrak{s}_j$	$\mathrm{R}^{\circ}_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi_{\bullet})$	$\theta_j\phi_j$	$\mathfrak{s}^{\dagger}_j\theta_j\phi_j$	$\|(1 \vee k\mathfrak{s}^2_{\bullet})^{-1/2}\theta_{\bullet}\|^2_{\ell_1(\lvert\phi\rvert)}$
(o-m)	$\mathrm{v} \in (-\mathrm{t}, \mathrm{a})$	$j^{-\mathrm{a}-1/2}$	$j^{-2\mathrm{t}}$	$n^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{a}+\mathrm{t}}}$	$j^{\mathrm{v}-\mathrm{a}-1}$	$j^{\mathrm{t}+\mathrm{v}-\mathrm{a}-1}$	
	$\mathrm{a} - \mathrm{v} < \mathrm{t}$						$k^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
	$\mathrm{a} - \mathrm{v} = \mathrm{t}$						$k^{-1}(\log k)^2$
	$\mathrm{a} - \mathrm{v} > \mathrm{t}$						k^{-1}
(o-s)	$\mathrm{a} - \mathrm{v} \in \mathbb{R}_{>0}$	$j^{-\mathrm{a}-1/2}$	$e^{-j^{2\mathrm{t}}}$	$(\log n)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$	$j^{\mathrm{v}-\mathrm{a}-1}$	$j^{\mathrm{v}-\mathrm{a}-1}e^{j^{2\mathrm{t}}}$	$(\log k)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v} + \mathrm{t} \in \mathbb{R}_{>0}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$\frac{(\log n)^{\frac{(\mathrm{t}+\mathrm{v})}{\mathrm{a}}}}{n}$	$j^{\mathrm{v}-1/2}e^{-j^{2\mathrm{a}}}$	$j^{\mathrm{t}+\mathrm{v}-1/2}e^{-j^{2\mathrm{a}}}$	k^{-1}

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v} < -\mathrm{t}$ the oracle rate $\mathrm{R}^{\circ}_n(\theta_{\bullet}, \mathfrak{s}_{\bullet}, \phi_{\bullet})$ is parametric. □

§07|02|03|02 Maximal local ϕ-risk

§07|02.62 **Assumption**. Consider weights $\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet} \in \mathcal{M}_{>0,\nu}(\mathscr{J})$ and $\phi \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$ (i.e. $\nu(\mathcal{N}_{\mathfrak{a}}) = \nu(\mathcal{N}_{\mathfrak{t}}) = 0 = \nu(\mathcal{N}_{\phi})$), such that $\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet} \in \mathbb{L}_{\infty}(\nu)$, $\mathfrak{a}_{\bullet} \in \mathbb{L}_2(\phi^2\nu)$, and $\mathfrak{t}^{\dagger}_{\bullet}\mathbb{1}^m_{\bullet}, \mathbb{1}^m_{\bullet} \in \mathbb{L}_2(\phi^2\nu)$ for all $m \in \mathbb{N}$. □

§07|02.63 **Reminder**. Under Assumption §07|02.62 we have $\mathbb{J}^{\mathfrak{a}} = \mathbb{L}^{\mathfrak{a}}_2(\nu) = \mathrm{dom}(\mathrm{M}_{\mathfrak{a}^{\dagger}}) = \mathbb{J}\mathfrak{a}_{\bullet} \subseteq \mathbb{J}$ and the three measures ν, $\mathfrak{a}^{2|\dagger}_{\bullet}\nu$ and $|\phi|\nu$ dominate mutually each other, i.e. they share the same null sets (see Property §04|01.02). We consider $\mathbb{J}^{\mathfrak{a}}$ endowed with $\|\cdot\|_{\mathfrak{a}^{\dagger}} = \|\mathrm{M}_{\mathfrak{a}^{\dagger}}\cdot\|_{\mathbb{J}}$ and given a constant $\mathrm{r} \in \mathbb{R}_{>0}$ the ellipsoid $\mathbb{J}^{\mathfrak{a},\mathrm{r}} := \{h_{\bullet} \in \mathbb{J}^{\mathfrak{a}} : \|h_{\bullet}\|_{\mathfrak{a}^{\dagger}} \leqslant \mathrm{r}\} \subseteq \mathbb{J}^{\mathfrak{a}}$. Since $\mathfrak{a}_{\bullet} \in \mathbb{L}_2(\phi^2\nu)$, and hence $\|\mathfrak{a}_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\phi} = \|(\mathfrak{a}\phi)_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\mathbb{J}} \in \mathbb{R}_{\geqslant 0}$ for each $m \in \mathbb{N}$ ($\|\mathfrak{a}_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\phi} = \mathrm{o}(1)$ as $m \to \infty$ by dominated convergence) we have $\mathbb{J}^{\mathfrak{a}} \subseteq \mathrm{dom}(\phi\nu)$ (Property §04|02.23), and $|\phi\nu(\theta_{\bullet}\mathbb{1}^{m|\perp}_{\bullet})| \leqslant \mathrm{r}\,\|\mathfrak{a}_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|_{\phi}$ for all $\theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ (Lemma §04|02.25). Let in addition $\mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ satisfy a link condition as in Definition §04|03.05

with weights $\mathfrak{t}_\bullet \in \mathcal{M}_{>0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$, and radius $\mathrm{d} \in \mathbb{R}_{>0}$. We set $(\mathfrak{t}^\dagger\phi)_\bullet := (\mathfrak{t}_j^\dagger\phi_j)_{j\in\mathcal{J}} = \mathfrak{t}_\bullet^\dagger\phi_\bullet \in \mathscr{J}$. Obviously, for $m \in \mathbb{N}$ the condition $\mathfrak{t}_\bullet^\dagger \mathbb{1}_\bullet^m \in \mathbb{L}_2(\phi^2\nu)$ due to Assumption §07|02.62 implies $\mathfrak{s}_\bullet^\dagger \mathbb{1}_\bullet^m \in \mathbb{L}_2(\phi^2\nu)$ too. Consequently, if Assumption §07|02.62, $\theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ and $\mathrm{M}_\mathfrak{s} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ are satisfied, then Assumption §07|02.48 is also fulfilled. Keep in mind if Assumptions §07|01.48 and §07|01.64 are satisfied, then for $n, m \in \mathbb{N}$ setting

$$\mathrm{R}_n^m(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) := \|\mathfrak{a}_\bullet \mathbb{1}_\bullet^{m|\perp}\|_\phi^2 + n^{-1}\|\mathfrak{t}_\bullet^\dagger \mathbb{1}_\bullet^m\|_\phi^2, \quad m_n^\star := \arg\min\left\{\mathrm{R}_n^m(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) : m \in \mathbb{N}\right\}$$
$$\text{and} \quad \mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) := \mathrm{R}_n^{m_n^\star}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) = \min\left\{\mathrm{R}_n^m(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) : m \in \mathbb{N}\right\} \quad (07.55)$$

for all $\theta_\bullet = \mathfrak{s}_\bullet^\dagger g_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$, *known* $\mathrm{M}_\mathfrak{s} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ and $n \in \mathbb{N}$ the OPE $\widehat{\theta}_\bullet^{m_n^\star} := \mathfrak{s}_\bullet^\dagger \widehat{g}_\bullet^{m_n^\star}$ fulfills

$$\mathbb{E}_{\theta|\mathfrak{s}}^n(|\phi\nu(\widehat{\theta}_\bullet^{m_n^\star} - \theta_\bullet)|^2) \leqslant (\mathrm{d}^2\|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})} \vee \mathrm{r}^2)\, \mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet).$$

due to Corollary §07|01.69. We shall emphasise that Assumptions §07|02.45 and §07|02.62 contains Assumptions §07|01.46 and §07|01.64, respectively. □

§07|02.64 **Proposition** Upper bound. *Let Assumptions §07|02.45 and §07|02.62 be satisfied. If* $\mathrm{M}_\mathfrak{s} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ *with* $\mathrm{d} \in \mathbb{R}_{>0}$, *and* $\theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ *with* $\mathrm{r} \in \mathbb{R}_{>0}$, *then for all* $n, k, m \in \mathbb{N}$ *the thresholded OPE* $\widehat{\theta}_\bullet^m = \widehat{\mathfrak{s}}_\bullet^{(k)|\dagger}\widehat{g}_\bullet^m \in \mathrm{dom}(\phi\nu)$ $\mathbb{P}_{\theta|\mathfrak{s}}^{n\otimes k}$-a.s. *fulfills*

$$\mathbb{E}_{\theta|\mathfrak{s}}^{n\otimes k}(|\phi\nu(\widehat{\theta}_\bullet^m - \theta_\bullet)|^2) \leqslant (3\mathrm{r}^2 \vee 4(\|\Gamma_{\theta|\mathfrak{s}}\|_{\mathbb{L}(\mathbb{J})} \vee 1)\|\mathrm{v}_\bullet^\mathfrak{s} \vee \mathbb{1}_\bullet\|_{\mathbb{L}_\infty(\nu)}\mathrm{d}^2)\mathrm{R}_n^m(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$$
$$+ 6(\|\mathrm{v}_\bullet^{\mathfrak{s}|(2)}\|_{\mathbb{L}_\infty(\nu)} + 3\|\mathrm{v}_\bullet^\mathfrak{s} \vee \mathbb{1}_\bullet\|_{\mathbb{L}_\infty(\nu)})\mathrm{d}^2\mathrm{r}^2\|(1 \vee k\mathfrak{t}_\bullet^2)^{-1/2}\mathfrak{a}_\bullet\mathbb{1}_\bullet^m\|_\phi^2 \quad (07.56)$$
$$\leqslant (3\mathrm{r}^2 \vee 4\mathrm{K}_{\theta|\mathfrak{s}}^2\mathrm{K}_\mathfrak{s}^2\mathrm{d}^2)\mathrm{R}_n^m(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) + 24\mathrm{K}_\mathfrak{s}^4\mathrm{d}^2\mathrm{r}^2\|(1 \vee k\mathfrak{t}_\bullet^2)^{-1/2}\mathfrak{a}_\bullet\mathbb{1}_\bullet^m\|_\phi^2. \quad (07.57)$$

§07|02.65 **Proof** of Proposition §07|02.64. Given in the lecture. □

§07|02.66 **Remark**. Selecting $m_n^\star := \arg\min\left\{\mathrm{R}_n^m(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) : m \in \mathbb{N}\right\}$ and $\mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) = \mathrm{R}_n^{m_n^\star}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$ as in (07.55) (Reminder §07|02.63) we obtain

$$\mathbb{E}_{\theta|\mathfrak{s}}^{n\otimes k}(|\phi\nu(\widehat{\theta}_\bullet^{m_n^\star} - \theta_\bullet)|^2) \leqslant (3\mathrm{r}^2 \vee 4\mathrm{K}_{\theta|\mathfrak{s}}^2\mathrm{K}_\mathfrak{s}^2\mathrm{d}^2)\, \mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) + 24\mathrm{K}_\mathfrak{s}^4\mathrm{d}^2\mathrm{r}^2\|(1 \vee k\mathfrak{t}_\bullet^2)^{-1/2}\mathfrak{a}_\bullet\|_\phi^2 \quad (07.58)$$

where $\mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$ is the rate (Corollary §07|01.69) if $\mathrm{M}_\mathfrak{s} \in \mathbb{L}^{\mathbb{M}}(\mathbb{J})$ is *known* in advance. Furthermore, for each $m \in \mathbb{N}$ we have

$$\|(1 \vee k\mathfrak{t}_\bullet^2)^{-1/2}\mathfrak{a}_\bullet\mathbb{1}_\bullet^m\|_\phi^2 \leqslant \|(1 \vee k\mathfrak{t}_\bullet^2)^{-1/2}\mathfrak{a}_\bullet\|_\phi^2 = \|(1 \vee k\mathfrak{t}_\bullet^2)^{-1/2}\mathfrak{a}_\bullet\mathbb{1}_\bullet^m\|_\phi^2 + \|(1 \vee k\mathfrak{t}_\bullet^2)^{-1/2}\mathfrak{a}_\bullet\mathbb{1}_\bullet^{m|\perp}\|_\phi^2$$
$$\leqslant \|(1 \vee k\mathfrak{t}_\bullet^2)^{-1/2}\mathfrak{a}_\bullet\mathbb{1}_\bullet^m\|_\phi^2 + \|\mathfrak{a}_\bullet\mathbb{1}_\bullet^{m|\perp}\|_\phi^2. \quad (07.59)$$

Consequently, under the assumptions of Proposition §07|02.64 from (??) (Proof §07|02.65) follows immediately

$$\mathbb{E}_{\theta|\mathfrak{s}}^{n\otimes k}(|\phi\nu(\widehat{\theta}_\bullet^m - \theta_\bullet)|^2) \leqslant (3\mathrm{r}^2 \vee 4\mathrm{K}_{\theta|\mathfrak{s}}^2\mathrm{K}_\mathfrak{s}^2\mathrm{d}^2)\mathrm{R}_n^m(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) + 24\mathrm{K}_\mathfrak{s}^4\mathrm{d}^2\mathrm{r}^2\|(1 \vee k\mathfrak{t}_\bullet^2)^{-1/2}\mathfrak{a}_\bullet\mathbb{1}_\bullet^m\|_\phi^2$$
$$\leqslant (3\mathrm{r}^2 \vee 4\mathrm{K}_{\theta|\mathfrak{s}}^2\mathrm{K}_\mathfrak{s}^2\mathrm{d}^2)\mathrm{R}_n^m(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) + 24\mathrm{K}_\mathfrak{s}^4\mathrm{d}^2\mathrm{r}^2\|(1 \vee k\mathfrak{t}_\bullet^2)^{-1/2}\mathfrak{a}_\bullet\|_\phi^2$$
$$\leqslant (3\mathrm{r}^2 \vee 4\mathrm{K}_{\theta|\mathfrak{s}}^2\mathrm{K}_\mathfrak{s}^2\mathrm{d}^2 + 24\mathrm{K}_\mathfrak{s}^4\mathrm{d}^2\mathrm{r}^2)\mathrm{R}_n^m(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) + 24\mathrm{K}_\mathfrak{s}^4\mathrm{d}^2\mathrm{r}^2\|(1 \vee k\mathfrak{t}_\bullet^2)^{-1/2}\mathfrak{a}_\bullet\mathbb{1}_\bullet^m\|_\phi^2$$

We shall emphasise, that the upper bound (07.58) consists (up to the constants) of the sum of the two terms $\mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$ and $\|(1 \vee k\mathfrak{t}_\bullet^2)^{-1/2}\mathfrak{a}_\bullet\|_\phi^2$ depending each on one of the sample sizes n and k only. If in addition there exists $\mathrm{v} \in \mathbb{R}_{>0}$ satisfying $\mathrm{v} \geqslant (\mathrm{K}_{\theta|\mathfrak{s}} \vee \mathrm{K}_\mathfrak{s})$ for all $\theta_\bullet \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}$ and

$\mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathrm{t,d}}$ then for all $n, k \in \mathbb{N}$ the maximal local ϕ-risk of the thresholded OPE $\widehat{\theta}^{m_n^\star}$ with optimally choosen dimension $m_n^\star$ is bounded by

$$\sup\left\{\mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}}(|\phi\nu(\widehat{\theta}^{m_n^\star} - \theta)|^2) : \theta \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathrm{t,d}}\right\} \leqslant (3\mathrm{r}^2 \vee 4\mathrm{K}^2_{\theta|\mathfrak{s}}\mathrm{K}^2_{\mathfrak{s}}\mathrm{d}^2 + 24\mathrm{K}^4_{\mathfrak{s}}\mathrm{d}^2\mathrm{r}^2)\,\mathrm{R}^\star_n(\mathfrak{a},\mathfrak{t},\phi) \vee \|(1 \vee k\mathfrak{t}^2)^{-1/2}\mathfrak{a}\|^2_\phi.$$

Arguing similarly as in Remark §07|01.21 we note that $\mathrm{R}^\star_n(\mathfrak{a},\mathfrak{t},\phi) = \mathrm{o}(1)$ as $n \to \infty$ since $\mathfrak{t}^\dagger\mathbb{1}^m \in \mathbb{L}_2(\phi^2\nu)$ for all $m \in \mathbb{N}$ and $\|\mathfrak{a}\mathbb{1}^{m|\perp}\|^2_\phi = \mathrm{o}(1)$ as $m \to \infty$ by Assumption §07|02.62. Moreover, we have $\|(1 \vee k\mathfrak{t}^2)^{-1/2}\mathfrak{a}\|^2_\phi = \mathrm{o}(1)$ as $k \to \infty$ by dominated convergence. Note that the dimension $m_n^\star := m_n^\star(\mathfrak{a},\mathfrak{t},\phi)$ does depend neither on the unknown parameter of interest θ nor on the unknown operator $\mathrm{M}_{\mathfrak{s}}$ but on the classes $\mathbb{J}^{\mathfrak{a},\mathrm{r}}$ and $\mathbb{M}_{\mathrm{t,d}}$ only, and thus also the statistic $\widehat{\theta}^{m_n^\star}$. In other words, if the regularity of θ and $\mathrm{M}_{\mathfrak{s}}$ is known in advance, then the OPE $\widehat{\theta}^{m_n^\star}$ is a feasible estimator. □

§07|02.67 **Corollary** (GdiSM with noisy operator §07|02.05 continued). *Consider* independent *noisy versions* $(\widehat{g},\widehat{\mathfrak{s}}) = (g + n^{-1/2}\dot{\mathrm{B}}, \mathfrak{s} + k^{-1/2}\dot{\mathrm{W}}) \sim \mathrm{N}^{n\otimes k}_{\theta|\mathfrak{s}} = \mathrm{N}^n_{\theta|\mathfrak{s}} \otimes \mathrm{N}^k_{\mathfrak{s}}$ *as in* Model §07|02.05*, where* $\dot{\mathrm{B}} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$ *and* $\dot{\mathrm{W}} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$ *are* independent*,* $\mathfrak{s} \in \mathbb{R}^{\mathbb{N}}_{\setminus 0} \cap \ell_\infty$*,* $\theta \in \ell_2$*, and hence* $g = \mathfrak{s}\theta \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}'}) \subseteq \ell_2$*. Under* Assumption §07|02.62 *the thresholded OPE* $\widehat{\theta}^{m_n^\star} = \widehat{\mathfrak{s}}^{(k)|\dagger}\widehat{g}^{m_n^\star} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *with dimension* $m_n^\star$ *as in* (07.55) *satisfies*

$$\sup\left\{\mathrm{N}^{n\otimes k}_{\theta|\mathfrak{s}}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m_n^\star} - \theta)|^2) : \theta \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathrm{t,d}}\right\} \leqslant \mathrm{C}_{\mathrm{r,d}}\,\mathrm{R}^\star_n(\mathfrak{a},\mathfrak{t},\phi) \vee \|(1 \vee k\mathfrak{t}^2)^{-1/2}\mathfrak{a}\|^2_\phi. \quad (07.60)$$

with constant $\mathrm{C}_{\mathrm{r,d}} = 3\mathrm{r}^2 \vee 4\mathrm{d}^2 + 36\mathrm{r}^2\mathrm{d}^2$.

§07|02.68 **Proof** of Corollary §07|02.67. Given in the lecture. □

§07|02.69 **Corollary** (diSM with noisy operator §07|02.07 continued). *Consider* independent *noisy versions* $(\widehat{g},\widehat{\mathfrak{s}}) = (g + n^{-1/2}\dot{\varepsilon}, \mathfrak{s} + k^{-1/2}\dot{\eta}) \sim \mathrm{P}^{n\otimes k}_{\theta|\mathfrak{s}|\sigma|\xi} = \mathrm{P}^n_{\theta|\mathfrak{s}|\sigma} \otimes \mathrm{P}^k_{\mathfrak{s}|\xi}$ *as in* Model §07|02.07*, where* $\dot{\varepsilon}$ *and* $\dot{\eta}$ *satisfy* **(iSM1)** *and* **(diSMnO1)** *with* $\mathrm{K}_\sigma := \|\sigma\|_{\ell_\infty} \vee 1 \in \mathbb{R}_{\geqslant 1}$ *and* $\mathrm{K}_\xi := \|\xi\|_{\ell_\infty} \vee 1 \in \mathbb{R}_{\geqslant 1}$*, respectively,* $\mathfrak{s} \in \mathbb{R}^{\mathbb{N}}_{\setminus 0} \cap \ell_\infty$ *and* $\theta \in \ell_2$*, and hence* $g = \mathfrak{s}\theta \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}'}) \subseteq \ell_2$*. Under* Assumption §07|02.62 *the thresholded OPE* $\widehat{\theta}^{m_n^\star} = \widehat{\mathfrak{s}}^{(k)|\dagger}\widehat{g}^{m_n^\star} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *with dimension* $m_n^\star$ *as in* (07.55) *satisfies*

$$\sup\left\{\mathrm{P}^{n\otimes k}_{\theta|\mathfrak{s}|\sigma|\xi}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m_n^\star} - \theta)|^2) : \theta \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathrm{t,d}}\right\} \leqslant \mathrm{C}_{\mathrm{r,d},\sigma,\xi}\,\mathrm{R}^\star_n(\mathfrak{a},\mathfrak{t},\phi) \vee \|(1 \vee k\mathfrak{t}^2)^{-1/2}\mathfrak{a}\|^2_\phi \quad (07.61)$$

with constant $\mathrm{C}_{\mathrm{r,d},\sigma,\xi} = 3\mathrm{r}^2 \vee 4\mathrm{K}^2_\xi\mathrm{K}^2_\sigma\mathrm{d}^2 + 24\mathrm{K}^4_\xi\mathrm{r}^2\mathrm{d}^2$.

§07|02.70 **Proof** of Corollary §07|02.69. Given in the lecture. □

§07|02.71 **Corollary** (dieMM with noisy operator §07|02.09 continued). *Consider* independent *noisy versions* $\widehat{g} = g + n^{-1/2}\dot{\varepsilon}$ *and* $\widehat{\mathfrak{s}} = \mathfrak{s} + k^{-1/2}\dot{\eta}$ *defined on* $(\mathcal{Z}^{n+k}, \mathscr{Z}^{\otimes(n+k)}, \mathbb{P}^{n\otimes k}_{\theta|\mathfrak{s}} = \mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}} \otimes \mathbb{P}^{\otimes k}_{\mathfrak{s}})$ *as in* Model §07|02.09*, where* $\psi, \varphi \in \mathcal{M}(\mathscr{X} \otimes \mathscr{J})$ *satisfies* **(dieMM1)–(dieMM3)** *and* **(dieMMnO1)–(dieMMnO2)** *with* $\mathrm{v}_{\theta|\mathfrak{s}|\psi}, \mathrm{v}_{\mathfrak{s}|\varphi} \in \mathbb{R}_{\geqslant 1}$*, respectively,* $\mathfrak{s} \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$*,* $\theta \in \mathbb{J}$ *and hence* $g = \mathfrak{s}\theta \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}'}) \subseteq \mathbb{J}$*. Under* Assumption §07|02.62 *the thresholded OPE* $\widehat{\theta}^{m_n^\star} = \widehat{\mathfrak{s}}^{(k)|\dagger}\widehat{g}^{m_n^\star} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *with dimension* $m_n^\star$ *as in* (07.55) *satisfies*

$$\sup\left\{\mathrm{P}^{n\otimes k}_{\theta|\mathfrak{s}|\sigma|\xi}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m_n^\star} - \theta)|^2) : \theta \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathrm{t,d}}\right\} \leqslant \mathrm{C}_{\mathrm{r,d}}\,\mathrm{R}^\star_n(\mathfrak{a},\mathfrak{t},\phi) \vee \|(1 \vee k\mathfrak{t}^2)^{-1/2}\mathfrak{a}\|^2_\phi \quad (07.62)$$

with constant $\mathrm{C}_{\mathrm{r,d}} = 3\mathrm{r}^2 \vee 4\mathrm{d}^2 \sup\left\{\mathrm{v}_{\theta|\mathfrak{s}|\psi}\mathrm{v}_{\mathfrak{s}|\varphi} : \theta \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathrm{t,d}}\right\} + 24\mathrm{r}^2\mathrm{d}^2 \sup\left\{\mathrm{v}^2_{\mathfrak{s}|\varphi} : \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathrm{t,d}}\right\}$.

§07|02.72 **Proof** of Corollary §07|02.71. Given in the lecture. □

§07|02.73 **Illustration**. We illustrate the last results considering usual behaviour for $\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet \in \mathcal{M}_{\neq 0,\nu}(\mathscr{J})$. We distinguish again the two cases **(p)** and **(np)** in Illustration §07|02.73 where in case **(p)** the term $\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$ is parametric, that is, $n\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) = \mathrm{O}(1)$, in case **(np)** it is nonparametric, i.e. $\lim_{n\to\infty} n\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) = \infty$. In case **(np)** we consider again the three specifications **(o-m)**, **(o-s)** and **(s-m)** introduced in Illustration §07|01.78 where also in Table 04 [§07] the order of the dimension $m^\star_n$ and the rate $\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$ as $n \to \infty$ are given. The next table depict the rate $\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$ and the additional term $\|(1 \vee k\mathfrak{t}^2_\bullet)^{-1/2}\mathfrak{a}_\bullet\|^2_{\phi_\bullet}$ as $n, k \to \infty$ only:

Table 08 [§07]

Order of $\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$ and $\|(1 \vee k\mathfrak{t}^2_\bullet)^{-1}\mathfrak{a}_\bullet\|^2_{\phi_\bullet}$ as $n, k \to \infty$

	$(j \in \mathscr{J})$ $\phi^2_j = j^{2\mathrm{v}-1}$	$(\mathrm{a} \in \mathbb{R}_{>0})$ $\mathfrak{a}^2_j$	$(\mathrm{t} \in \mathbb{R}_{>0})$ $\mathfrak{t}^2_j$	$\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$	$(\mathfrak{a}\phi)^2_j$	$\mathfrak{t}^{-2}_j(\mathfrak{a}\phi)^2_j$	$\|(1 \vee k\mathfrak{t}^2_\bullet)^{-1/2}\mathfrak{a}_\bullet\|^2_{\phi_\bullet}$
(o-m)	$\mathrm{v} \in (-\mathrm{t}, \mathrm{a})$	$j^{-2\mathrm{a}}$	$j^{-2\mathrm{t}}$	$n^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{a}+\mathrm{t}}}$	$j^{-2(\mathrm{a}-\mathrm{v})-1}$	$j^{2(\mathrm{t}+\mathrm{v}-\mathrm{a})-1}$	
	$\mathrm{a}-\mathrm{v} \leqslant \mathrm{t}$						$k^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
	$\mathrm{a}-\mathrm{v} \geqslant \mathrm{t}$						k^{-1}
(o-s)	$\mathrm{a}-\mathrm{v} \in \mathbb{R}_{>0}$	$j^{-2\mathrm{a}}$	$e^{-j^{2\mathrm{t}}}$	$(\log n)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$	$j^{2(\mathrm{v}-\mathrm{a})-1}$	$j^{2(\mathrm{v}-\mathrm{a})-1}e^{j^{2\mathrm{t}}}$	$(\log k)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v}+\mathrm{t} \in \mathbb{R}_{>0}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$n^{-1}(\log n)^{\frac{\mathrm{t}+\mathrm{v}}{\mathrm{a}}}$	$j^{2\mathrm{v}-1}e^{-j^{2\mathrm{a}}}$	$j^{2(\mathrm{t}+\mathrm{v})-1}e^{-j^{2\mathrm{a}}}$	k^{-1}

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v} < -\mathrm{t}$ the rate $\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$ is parametric. □

§08 (Generalised) Galerkin estimator

§08|00.01 **Notation** (Reminder). Consider $\mathbb{J} = \ell_2 := \mathbb{L}_2(\nu_{\mathbb{N}}) = \mathbb{L}_2(\mathbb{N}, 2^{\mathbb{N}}, \nu_{\mathbb{N}})$ with counting measure $\nu_{\mathbb{N}} := \sum_{j\in\mathbb{N}} \delta_{\{j\}}$, surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \ell_2)$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \ell_2)$. For each $\mathrm{T} \in \mathbb{L}(\mathbb{H}, \mathbb{G})$ and $\mathrm{T}_{\bullet|\bullet} := \mathrm{V}\mathrm{T}\mathrm{U}^\star \in \mathbb{L}(\ell_2) = \mathbb{L}_\bullet(\ell_2)$ (compare Notation §01|04.03) we identify the kernel (infinite dimensional matrix) $\mathrm{T}_{\bullet|\bullet} = (\mathrm{T}_{j,j_\circ})_{j,j_\circ\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}^2}$ and the map from ℓ_2 into itself given by

$$a_\bullet \mapsto \mathrm{T}_{\bullet|\bullet}a_\bullet := \Big(\sum_{j_\circ\in\mathbb{N}} \mathrm{T}_{j|j_\circ}a_{j_\circ} = \langle \mathrm{T}_{j|\bullet}, a_\bullet\rangle_{\ell_2} = \nu_{\mathbb{N}}(\mathrm{T}_{j,\bullet}a_\bullet)\Big)_{j\in\mathbb{N}}$$

(compare Notation §01|05.01). Moreover, we denote by $\mathbb{L}^{\gtrless}_\bullet(\ell_2)$ the subset of all strictly positive definite operator in $\mathbb{L}_\bullet(\ell_2)$. For each $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\gtrless}_\bullet(\ell_2)$ we denote its Moore-Penrose inverse by $\mathrm{T}^\dagger_{\bullet|\bullet} : \ell_2 \supseteq \mathrm{dom}(\mathrm{T}^\dagger_{\bullet|\bullet}) \to \ell_2$ (see Definition §03|00.08). We denote by $\mathbb{L}^{\mathrm{R}}_\bullet(\ell_2)$ the subset of all *injective* $\mathrm{A}_{\bullet|\bullet} \in \mathbb{L}_\bullet(\ell_2)$ such that $[\mathrm{A}_{\bullet|\bullet}]_{\underline{m}} \in \mathbb{R}^{(m,m)}$ is *regular* for all $m \in \mathbb{N}$. For each $m \in \mathbb{N}$ and $\mathrm{A}_{\bullet|\bullet} \in \mathbb{L}^{\mathrm{R}}_\bullet(\ell_2)$, the inverse $[\mathrm{A}_{\bullet|\bullet}]^{-1}_{\underline{m}} \in \mathbb{R}^{(m,m)}$ of $[\mathrm{A}_{\bullet|\bullet}]_{\underline{m}} \in \mathbb{R}^{(m,m)}$ exists. Note that $\mathbb{L}^{\gtrless}_\bullet(\ell_2) \subseteq \mathbb{L}^{\mathrm{R}}_\bullet(\ell_2)$ (Lemma §05|01.22). □

§08|00.02 **Assumption**. For $\mathbb{J} = \ell_2$, surjective partial isometries $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \ell_2)$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \ell_2)$ fixed and presumed to be *known* in advance, the operator $\mathrm{T} \in \mathbb{L}(\mathbb{H}, \mathbb{G})$ satisfies either $\mathrm{T}_{\bullet|\bullet} = \mathrm{V}\mathrm{T}\mathrm{U}^\star \in \mathbb{L}^{\gtrless}_\bullet(\ell_2) \subseteq \mathbb{L}_\bullet(\ell_2) = \mathbb{L}(\ell_2)$ or more generally $\mathrm{T}_{\bullet|\bullet} = \mathrm{V}\mathrm{T}\mathrm{U}^\star \in \mathbb{L}^{\mathrm{R}}_\bullet(\ell_2) \subseteq \mathbb{L}_\bullet(\ell_2) = \mathbb{L}(\ell_2)$. Let $g_\bullet \in \mathrm{dom}(\mathrm{T}^\dagger_{\bullet|\bullet}) = \mathrm{ran}(\mathrm{T}_{\bullet|\bullet})$, and hence $\theta_\bullet = \mathrm{T}^\dagger_{\bullet|\bullet}g_\bullet = \mathrm{T}^{-1}_{\bullet|\bullet}g_\bullet \in \ell_2$. □

§08|00.03 **Reminder**. Under Assumption §08|00.02 we consider $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\gtrless}_\bullet(\ell_2)$ or more generally $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\mathrm{R}}_\bullet(\ell_2)$ and $g_\bullet \in \mathrm{dom}(\mathrm{T}^\dagger_{\bullet|\bullet}) = \mathrm{ran}(\mathrm{T}_{\bullet|\bullet})$, and hence $\theta_\bullet = \mathrm{T}^\dagger_{\bullet|\bullet}g_\bullet = \mathrm{T}^{-1}_{\bullet|\bullet}g_\bullet \in \ell_2$. For each $m \in \mathbb{N}$ and $\mathrm{A}_{\bullet|\bullet} \in \mathbb{L}_\bullet(\ell_2)$ we write $\mathrm{A}^m_{\bullet|\bullet} := \mathrm{M}_{\mathbb{1}^m}\mathrm{A}_{\bullet|\bullet}\mathrm{M}_{\mathbb{1}^m} \in \mathbb{L}_\bullet(\ell_2)$, which restricted to an operator from $\mathbb{R}^m$ ($\mathrm{ran}(\mathrm{M}_{\mathbb{1}^m}) = \ell_2\mathbb{1}^m$) to itself, can be represented by the matrix $[\mathrm{A}_{\bullet|\bullet}]_{\underline{m}} \in \mathbb{R}^{(m,m)}$ (see Notation §05|00.02). If $[\mathrm{A}_{\bullet|\bullet}]^\dagger_{\underline{m}} \in \mathbb{R}^{(m,m)}$

denotes the Moore-Penrose inverse of $[A_{\bullet|\bullet}]_{\underline{m}}$ (as linear map from $\mathbb{R}^m$ into itself), then the Moore-Penrose inverse $A_{\bullet|\bullet}^{m|\dagger} = (A_{\bullet|\bullet}^m)^\dagger \in \mathbb{L}_\bullet(\ell_2)$ of $A_{\bullet|\bullet}^m$ (see Definition §03|00.08), restricted to an operator from $\mathbb{R}^m$ to itself can be represented by the matrix $[A_{\bullet|\bullet}]_{\underline{m}}^\dagger$. In particular, if $A_{\bullet|\bullet} \in \mathbb{L}_\bullet^{\gtrless}(\ell_2)$, i.e. $[A_{\bullet|\bullet}]_{\underline{m}}$ is *regular* (invertible), and hence $[A_{\bullet|\bullet}]_{\underline{m}}^\dagger = [A_{\bullet|\bullet}]_{\underline{m}}^{-1}$, then we have $A_{\bullet|\bullet}^m A_{\bullet|\bullet}^{m|\dagger} = M_{\mathbb{1}^m} = A_{\bullet|\bullet}^{m|\dagger} A_{\bullet|\bullet}^m$. For $T_{\bullet|\bullet} \in \mathbb{L}_\bullet^{\gtrless}(\ell_2)$ and $m \in \mathbb{N}$ we call any element $\theta_\bullet^m \in {}_{\ell_2}\mathbb{1}_\bullet^m$ i.e. $0_\bullet = \theta_\bullet^m(\mathbb{1}_\bullet - \mathbb{1}_\bullet^m) = \theta_\bullet^m \mathbb{1}_\bullet^{m|\perp}$, satisfying

$$\langle \theta_\bullet^m, T_{\bullet|\bullet}\theta_\bullet^m\rangle_{\ell_2} - 2\langle \theta_\bullet^m, g_\bullet\rangle_{\ell_2} \leqslant \langle a_\bullet, T_{\bullet|\bullet} a_\bullet\rangle_{\ell_2} - 2\langle a_\bullet, g_\bullet\rangle_{\ell_2} \quad \text{for all } a_\bullet \in {}_{\ell_2}\mathbb{1}_\bullet^m$$

a *Galerkin solution* in ${}_{\ell_2}\mathbb{1}_\bullet^m$. Since $T_{\bullet|\bullet} \in \mathbb{L}_\bullet^{\gtrless}(\ell_2)$ the Galerkin solution is *uniquely* determined by $[\theta_\bullet^m]_{\underline{m}} = [T_{\bullet|\bullet}]_{\underline{m}}^{-1}[g_\bullet]_{\underline{m}}$, and hence $\theta_\bullet^m = T_{\bullet|\bullet}^{m|\dagger} g_\bullet$ (Lemma §05|01.03). More generally, under Assumption §08|00.02 with $T_{\bullet|\bullet} \in \mathbb{L}_\bullet(\ell_2)$ we call the unique solution $\theta_\bullet^m = T_{\bullet|\bullet}^{m|\dagger} g_\bullet$ of $[T_{\bullet|\bullet}]_{\underline{m}}[\theta_\bullet^m]_{\underline{m}} = [g_\bullet]_{\underline{m}}$ *generalised Galerkin solution* (Definition §05|02.01). Keep in mind that $\mathbb{L}_\bullet^{\gtrless}(\ell_2) \subseteq \mathbb{L}_\bullet^{\mathrm{R}}(\ell_2)$ (Lemma §05|01.22). □

§08|01 Non-diagonal statistical inverse problem

§08|01.01 **Assumption**. Consider a stochastic process $\dot{\varepsilon}_\bullet = (\dot{\varepsilon}_j)_{j\in\mathbb{N}}$ on a probability space $(\Omega, \mathscr{A}, \mathbb{P})$ satisfying Assumption §01|01.04 (i.e. $\dot{\varepsilon}_\bullet \in \mathcal{M}(\mathscr{A} \otimes 2^{\mathbb{N}})$) with *mean zero* (i.e. $\mathbb{P}(\dot{\varepsilon}_\bullet) = (\mathbb{P}(\dot{\varepsilon}_j))_{j\in\mathbb{N}} = 0_\bullet$), a sample size $n \in \mathbb{N}$ and let Assumption §08|00.02 be satisfied where $T_{\bullet|\bullet} \in \mathbb{L}_\bullet^{\gtrless}(\ell_2)$ or $T_{\bullet|\bullet} \in \mathbb{L}_\bullet^{\mathrm{R}}(\ell_2)$ is *known* in advance. For $\theta_\bullet \in \ell_2$ the observable noisy image with mean $g_\bullet = T_{\bullet|\bullet}\theta_\bullet \in \ell_2$ takes the form $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$. We denote by $\mathbb{P}_{\theta|T}^n$ the distribution of $\widehat{g}_\bullet$. In addition

(nSIP) $\dot{\varepsilon}_\bullet$ admits a covariance operator, say $\Gamma_{\theta|T} \in \mathbb{L}_\bullet^{\gtrless}(\ell_2)$ with $\|\Gamma_{\theta|T}\|_{\mathbb{L}(\ell_2)} \leqslant \mathbb{v}_{\theta|T} \in \mathbb{R}_{\geqslant 1}$. □

§08|01.02 **Definition**. Under Assumption §08|01.01 for $\theta_\bullet \in \ell_2$ and $T_{\bullet|\bullet} \in \mathbb{L}_\bullet^{\mathrm{R}}(\ell_2)$ consider a noisy version $\widehat{g}_\bullet \sim \mathbb{P}_{\theta|T}^n$ of $g_\bullet = T_{\bullet|\bullet}\theta_\bullet \in \mathrm{dom}(T_{\bullet|\bullet}^\dagger)$. For each $m \in \mathbb{N}$ we call $\widehat{\theta}_\bullet^m = T_{\bullet|\bullet}^{m|\dagger}\widehat{g}_\bullet = T_{\bullet|\bullet}^{m|\dagger}\widehat{g}_\bullet\mathbb{1}_\bullet^m = T_{\bullet|\bullet}^{m|\dagger}\widehat{g}_\bullet^m$ (generalised) *Galerkin estimator (GE)* of $\theta_\bullet = T_{\bullet|\bullet}^\dagger g_\bullet = T_{\bullet|\bullet}^{-1} g_\bullet \in \ell_2$. □

§08|01.03 **Comment**. The (generalised) Galerkin solution $\theta_\bullet^m = T_{\bullet|\bullet}^{m|\dagger} g_\bullet \in {}_{\ell_2}\mathbb{1}_\bullet^m$ does generally not correspond to the orthogonal projection $\mathbb{1}_\bullet^{m|\perp}\theta_\bullet = (\mathbb{1}_\bullet - \mathbb{1}_\bullet^m)\theta_\bullet$. Moreover, the approximation error $\sup\left\{\|\theta_\bullet^j - \theta_\bullet\|_{\ell_2} : j \in \mathbb{N}_{\geqslant m}\right\}$ does generally not converge to zero as $m \to \infty$ (compare Remark §05|01.05). Here and subsequently, we will restrict ourselves to classes of solutions and operators which ensure the convergence. Obviously, this is a minimal regularity condition for us if we aim to estimate the Galerkin solution.

§08|01|01 Examples

§08|01.04 **GniSM (§01|05.08 continued)**. Let Assumption §08|00.02 be satisfied where $T_{\bullet|\bullet} \in \mathbb{L}_\bullet^{\gtrless}(\ell_2)$ or $T_{\bullet|\bullet} \in \mathbb{L}_\bullet^{\mathrm{R}}(\ell_2)$ is *known* in advance. We illustrate the (generalised) GE in a Gaussian non-diagonal inverse sequence model (GniSM) as in §01|05.08. Here the observable stochastic process $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{B}_\bullet \sim N_{\theta|T}^n$ is a noisy version of $g_\bullet = T_{\bullet|\bullet}\theta_\bullet \in \mathrm{dom}(T_{\bullet|\bullet}^\dagger) \subseteq \ell_2$ with $\theta_\bullet \in \Theta \subseteq \ell_2$ and $\dot{B}_\bullet \sim N_{(0,1)}^{\otimes\mathbb{N}}$. Consequently, $\widehat{g}_\bullet$ admits a $N_{\theta|T}^n$-distribution belonging to the family $N_{\Theta\times\{T_{\bullet|\bullet}\}}^n := (N_{\theta|T}^n)_{\theta_\bullet\in\Theta}$. Summarising the observations satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}^{\otimes\mathbb{N}}, N_{\Theta\times\{T_{\bullet|\bullet}\}}^n)$ where $\Theta \subseteq \ell_2$. □

§08|01.05 **Reminder** (GniSM §08|01.04 continued). Due to Property §07|01.04 the error process $\dot{B}_\bullet \sim N_{(0,1)}^{\otimes\mathbb{N}}$ admits as covariance operator $\Gamma_{\theta|\mathfrak{s}} = \mathrm{id}_{\ell_2} \in \mathbb{L}^{\gtrless}(\ell_2)$ and hence Assumption §08|01.01 is satisfied. □

§08|01.06 **niSM (§01|05.07 continued)**. Let Assumption §08|00.02 be satisfied where $T_{\bullet|\bullet} \in \mathbb{L}_\bullet^{\gtrless}(\ell_2)$ or $T_{\bullet|\bullet} \in \mathbb{L}_\bullet^{\mathrm{R}}(\ell_2)$ is *known* in advanced. We illustrate the (generalised) GE in a Non-diagonal inverse sequence model (niSM) as in §01|05.07. Here the observable stochastic process $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ is a noisy version of $g_\bullet = T_{\bullet|\bullet}\theta_\bullet \in \mathrm{dom}(T_{\bullet|\bullet}^\dagger) \subseteq \ell_2$ with $\theta_\bullet \in \Theta \subseteq \ell_2$ and $\dot{\varepsilon}_\bullet \sim \otimes_{j\in\mathbb{N}}\mathbb{P}^{\dot{\varepsilon}_j}$ satisfying **(iSM1)** and **(iSM2)** introduced in diSM §07|01.06. Under **(iSM1)** $\widehat{g}_\bullet$ admits a $\mathrm{P}^n_{\theta|T|\sigma}$-distribution belonging to the family $\mathrm{P}^n_{\Theta\times\{T_{\bullet|\bullet}\}\times\Sigma} := (\mathrm{P}^n_{\theta|T|\sigma})_{\theta\in\Theta,\sigma\in\Sigma}$. Summarising the observations satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}^{\otimes\mathbb{N}}, \mathrm{P}^n_{\Theta\times\{T_{\bullet|\bullet}\}\times\Sigma})$ where $\Theta \subseteq \ell_2$ and $\Sigma \subseteq \mathbb{R}^{\mathbb{N}}_{>0} \cap \ell_\infty$. □

§08|01.07 **Reminder** (niSM §08|01.06 continued). Due to Property §07|01.07 (i) under **(iSM1)** the process $\dot{\varepsilon}_\bullet \sim \otimes_{j\in\mathbb{N}}\mathrm{P}_{(0,\sigma_j^2)}$ admits as covariance operator $\Gamma_{\theta|\mathfrak{s}} = M_{\sigma^2} \in \mathbb{L}^{\mathbb{M}}(\ell_2)\cap\mathbb{L}^{\geqslant}(\ell_2)$ and hence Assumption §08|01.01 is satisfied. □

§08|01.08 **nieMM (§01|05.05 continued).** Let Assumption §08|00.02 be satisfied where $T_{\bullet|\bullet} \in \mathbb{L}^{\geqslant}_{\bullet}(\ell_2)$ or $T_{\bullet|\bullet} \in \mathbb{L}^{\mathbb{R}}_{\bullet}(\ell_2)$ is *known* in advanced. We illustrate the (generalised) GE in a Non-diagonal inverse empirical mean model (nieMM) as in §01|05.05. Here the observable stochastic process $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ is a noisy version of $g_\bullet = T_{\bullet|\bullet}\theta_\bullet \in \mathrm{dom}(T_{\bullet|\bullet}^\dagger) \subseteq \ell_2$ with $\theta_\bullet = g_\bullet \in \Theta \subseteq \ell_2$, and error process $\dot{\varepsilon}_\bullet = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_\bullet) - \mathbb{P}_{\theta|T}(\psi_\bullet)) \in \mathcal{M}(\mathscr{Z}^{\otimes n} \otimes 2^{\mathbb{N}})$ satisfying Assumption §01|01.04. More precisely, on a measurable space $(\mathcal{Z}, \mathscr{Z})$ for $T_{\bullet|\bullet} \in \mathbb{L}_{\bullet}(\ell_2)$ and for each $\theta_\bullet \in \Theta \subseteq \ell_2$ there is a probability measure $\mathbb{P}_{\theta|T} \in \mathscr{W}(\mathscr{Z})$. Consider a stochastic process $\psi_\bullet = (\psi_j)_{j\in\mathbb{N}} \in \mathcal{M}(\mathscr{Z}\otimes 2^{\mathbb{N}})$ which similar to **(dieMM1)**–**(dieMM4)** introduced in dieMM §07|01.08 in addition for $T_{\bullet|\bullet} \in \mathbb{L}^{\mathbb{R}}_{\bullet}(\ell_2)$ (or $T_{\bullet|\bullet} \in \mathbb{L}^{\geqslant}_{\bullet}(\ell_2)$) and for each $\theta_\bullet \in \Theta \subseteq \ell_2$ satisfies

(nieMM1) $\psi_j \in \mathcal{L}_1(\mathbb{P}_{\theta|T}) := \mathcal{L}_1(\mathcal{Z}, \mathscr{Z}, \mathbb{P}_{\theta|T})$ for all $j \in \mathbb{N}$ and $\mathbb{P}_{\theta|T}(\psi_\bullet) = T_{\bullet|\bullet}\theta_\bullet = g_\bullet$,

(dieMM2) for each $m \in \mathbb{N}$ we have $(\psi_\bullet - \mathbb{P}_{\theta|\mathfrak{s}}(\psi_\bullet))\mathbb{1}^m_\bullet \in \ell_\infty$ $\mathbb{P}_{\theta|T}$-a.s. due to **(nieMM1)**,

(nieMM2) there is $\mathrm{v}_{\theta|T|\psi} \in \mathbb{R}_{\geqslant 1}$ such that $\|\mathbb{P}_{\theta|T}(\psi_\bullet^2)\|_{\ell_\infty} \leqslant \mathrm{v}_{\theta|T|\psi}$ and

$$\mathbb{P}_{\theta|T}\big(|\nu_{\mathbb{N}}(h_\bullet\psi_\bullet)|^2\big) \leqslant \mathrm{v}_{\theta|T|\psi}\|h_\bullet\|^2_{\ell_2}, \quad \forall h_\bullet \in \ell_2,$$

(nieMM3) $\mathrm{v}_\bullet^{\theta|T} := \mathbb{P}_{\theta|T}(\psi_\bullet^2) - |\mathbb{P}_{\theta|T}(\psi_\bullet)|^2 \in \mathbb{R}^{\mathbb{N}}_{>0} \cap \ell_\infty$, $\|(\mathrm{v}_\bullet^{\theta|T})^{-1}\|_{\ell_\infty} \leqslant \mathrm{v}_{\theta|T|\psi}$ and

$$\mathbb{P}_{\theta|T}\big(|\nu_{\mathbb{N}}(h_\bullet\psi_\bullet)|^2\big) \geqslant \mathbb{P}_{\theta|T}\big(|\nu_{\mathbb{N}}(h_\bullet\psi_\bullet)|^2\big) - \big|\mathbb{P}_{\theta|T}\big(\nu_{\mathbb{N}}(h_\bullet\psi_\bullet)\big)\big|^2 \geqslant \mathrm{v}^{-1}_{\theta|T|\psi}\|h_\bullet\|^2_{\ell_2}, \quad \forall h_\bullet \in \ell_2.$$

We consider a statistical product experiment $(\mathcal{Z}^n, \mathscr{Z}^{\otimes n}, \mathbb{P}^{\otimes n}_{\Theta\times\{T_{\bullet|\bullet}\}} = (\mathbb{P}^{\otimes n}_{\theta|T})_{\theta\in\Theta})$ as in an Empirical mean function §01|01.10 where $\Theta \subseteq \ell_2$. □

§08|01.09 **Reminder** (nieMM §08|01.08 continued). Due to Property §07|01.09 (i) under **(nieMM1)** and **(nieMM2)** the error process $\dot{\varepsilon}_\bullet = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_\bullet) - \mathbb{P}_{\theta|T}(\psi_\bullet)) \in \mathcal{M}(\mathscr{Z}^{\otimes n} \otimes 2^{\mathbb{N}})$ admits a covariance operator $\Gamma_{\theta|T} \in \mathbb{L}^{\geqslant}(\ell_2)$ and hence Assumption §08|01.01 is satisfied. □

§08|01|02 Global and maximal global $\mathfrak{v}$-risk

We measure first the accuracy of the (generalised) GE $\widehat{\theta}^m_\bullet := T^{m|\dagger}_{\bullet|\bullet}\widehat{g}_\bullet$ of the (generalised) Galerkin solution $\theta^m_\bullet = T^{m|\dagger}_{\bullet|\bullet}g_\bullet \in {}_{\ell_2}\mathbb{1}^m_\bullet$ with $g_\bullet = T_{\bullet|\bullet}\theta_\bullet \in \mathrm{dom}(T^\dagger_{\bullet|\bullet})$ by the mean of its global $\mathfrak{v}$-error introduced in §05|01|01 and §05|02|01, i.e. its $\mathfrak{v}$-risk.

§08|01.10 **Reminder**. If $\mathfrak{v}_\bullet \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ then we have $\mathfrak{v}_\bullet^2\mathbb{1}^m_\bullet \in \ell_\infty$ and ${}_{\ell_2}\mathbb{1}^m_\bullet \subseteq \ell_2(\mathfrak{v}_\bullet^2)$. Consequently, for each $\theta_\bullet \in \ell_2(\mathfrak{v}_\bullet^2)$ the (generalised) Galerkin solution $\theta^m_\bullet = T^{m|\dagger}_{\bullet|\bullet}g_\bullet \in {}_{\ell_2}\mathbb{1}^m_\bullet$ satisfies $\theta^m_\bullet \in \ell_2(\mathfrak{v}_\bullet^2)$ too. If in addition $C_T := \sup\big\{\|M_{\mathfrak{v}}T^{m|\dagger}_{\bullet|\bullet}T_{\bullet|\bullet}M_{\mathbb{1}^{m|\perp}}\|_{\mathbb{L}(\ell_2)} : m \in \mathbb{N}\big\} \in \mathbb{R}_{\geqslant 0}$ then $\|\theta^m_\bullet - \theta_\bullet\|_{\mathfrak{v}} \leqslant (1 + C_T)\|\mathbb{1}^{m|\perp}_\bullet\theta_\bullet\|_{\ell_2}$ which implies $\sup\big\{\|\theta^j_\bullet - \theta_\bullet\|_{\mathfrak{v}} : j \in \mathbb{N}_{\geqslant m}\big\} = \mathrm{o}(1)$ as $m \to \infty$ (Property §05|01.24 and Property §05|02.08). □

§08|01.11 **Comment**. Under Assumption §08|01.01 since $\theta^m_\bullet, T^{m|\dagger}_{\bullet|\bullet}\mathbb{1}^m_\bullet \in {}_{\ell_2}\mathbb{1}^m_\bullet$ for each $m \in \mathbb{N}$ we have $T^{m|\dagger}_{\bullet|\bullet}\dot{\varepsilon}_\bullet \in {}_{\ell_2}\mathbb{1}^m_\bullet$ $\mathbb{P}^n_{\theta|T}$-a.s.. Indeed, $\dot{\varepsilon}_\bullet \sim \mathrm{P}_{(0,\Gamma_{\theta|T})}$ with $\Gamma_{\theta|T} \in \mathbb{L}^{\geqslant}(\ell_2)$ by Assumption §08|01.01 **(nSIP)** implies

$\mathbb{E}^n_{\theta|T}(\dot{\varepsilon}^2_{\bullet}) \in \ell_\infty$, hence $\dot{\varepsilon}_{\bullet}\mathbb{1}^m_{\bullet} \in \ell_\infty$ $\mathbb{P}^n_{\theta|T}$-a.s. and $\|T^{m|\dagger}_{\bullet|\bullet}\dot{\varepsilon}_{\bullet}\|_{\ell_2} \leqslant \|T^{m|\dagger}_{\bullet|\bullet}\mathbb{1}^m_{\bullet}\|_{\ell_2}\|\dot{\varepsilon}_{\bullet}\mathbb{1}^m_{\bullet}\|_{\ell_\infty} \in \mathbb{R}_{\geqslant 0}$ $\mathbb{P}^n_{\theta|T}$-a.s.. Given $\mathfrak{v}_{\bullet} \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ from ${}_{\ell_2}\mathbb{1}^m_{\bullet} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$ (Reminder §08|01.10) it follows

$$\widehat{\theta}^m_{\bullet} = T^{m|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} = n^{-1/2}T^{m|\dagger}_{\bullet|\bullet}\dot{\varepsilon}_{\bullet} + \theta^m_{\bullet} \in {}_{\ell_2}\mathbb{1}^m_{\bullet} \subseteq \ell_2(\mathfrak{v}^2_{\bullet}) \quad \mathbb{P}^n_{\theta|T}\text{-a.s..} \qquad \square$$

§08|01|02|01 Global $\mathfrak{v}$-risk

§08|01.12 **Assumption**. Let $\mathfrak{v}_{\bullet} \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ and $\theta_{\bullet} \in \ell_2(\mathfrak{v}^2_{\bullet})$ be satisfied. □

§08|01.13 **Definition**. Under Assumptions §08|01.01 and §08|01.12 the *global $\mathfrak{v}$-risk* of a (generalised) GE $\widehat{\theta}^m_{\bullet} = T^{m|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in {}_{\ell_2}\mathbb{1}^m_{\bullet} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$ $\mathbb{P}^n_{\theta|T}$-a.s. satisfies

$$\mathbb{E}^n_{\theta|T}(\|\widehat{\theta}^m_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) = \mathbb{E}^n_{\theta|T}\|T^{m|\dagger}_{\bullet|\bullet}(\widehat{g}_{\bullet} - g_{\bullet})\|^2_{\mathfrak{v}} + \|\theta^m_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}} \qquad (08.01)$$

with *variance* $\mathbb{E}^n_{\theta|T}(\|T^{m|\dagger}_{\bullet|\bullet}(\widehat{g}_{\bullet} - g_{\bullet})\|^2_{\mathfrak{v}}) = n^{-1}\mathbb{E}^n_{\theta|T}(\|T^{m|\dagger}_{\bullet|\bullet}\dot{\varepsilon}_{\bullet}\|^2_{\mathfrak{v}})$ and *bias* $\|\theta^m_{\bullet} - \theta_{\bullet}\|_{\mathfrak{v}}$. □

§08|01.14 **Notation** (Reminder). Let $A \in \mathbb{L}(\ell_2)$ be a *Hilbert-Schmidt operator*, $A \in \mathbb{HS}(\ell_2)$ for short, where $\|A\|^2_{HS} := \operatorname{tr}(A^\star A) = \operatorname{tr}(AA^\star) \in \mathbb{R}_{\geqslant 0}$. If $\Gamma \in \mathbb{L}(\ell_2)$ then $\operatorname{tr}(A^\star\Gamma A) \leqslant \|\Gamma\|_{\mathbb{L}(\ell_2)}\operatorname{tr}(A^\star A) = \|\Gamma\|_{\mathbb{L}(\ell_2)}\|A\|^2_{HS}$. For arbitrary $A \in \mathbb{L}(\ell_2)$ we have $M_{\mathfrak{v}}A^m = M^m_{\mathfrak{v}}A^m \in \mathbb{HS}(\ell_2)$. □

§08|01.15 **Property**. *Under Assumptions §08|01.01 and §08|01.12 we have*

$$\mathbb{E}^n_{\theta|T}(\|T^{m|\dagger}_{\bullet|\bullet}\dot{\varepsilon}_{\bullet}\|^2_{\mathfrak{v}}) = \operatorname{tr}(M_{\mathfrak{v}}T^{m|\dagger}_{\bullet|\bullet}\Gamma_{\theta|T}(T^{m|\dagger}_{\bullet|\bullet})^\star M_{\mathfrak{v}}) = \operatorname{tr}([M_{\mathfrak{v}}]_{\underline{m}}[T_{\bullet|\bullet}]^{-1}_{\underline{m}}[\Gamma_{\theta|T}]_{\underline{m}}([T_{\bullet|\bullet}]^{-1}_{\underline{m}})^\star[M_{\mathfrak{v}}]_{\underline{m}}) \qquad (08.02)$$

and consequently $\mathbb{E}^n_{\theta|T}(\|T^{m|\dagger}_{\bullet|\bullet}(\widehat{g}_{\bullet} - g_{\bullet})\|^2_{\mathfrak{v}}) \leqslant n^{-1}\|\Gamma_{\theta|T}\|_{\mathbb{L}(\ell_2)}\|M_{\mathfrak{v}}T^{m|\dagger}_{\bullet|\bullet}\|^2_{HS} \in \mathbb{R}_{\geqslant 0}$. □

§08|01.16 **Proposition** (Upper bound). *Under Assumptions §08|01.01 and §08|01.12 for all $n, m \in \mathbb{N}$ with (generalised) Galerkin solution $\theta^m_{\bullet} = T^{m|\dagger}_{\bullet|\bullet}g_{\bullet} \in {}_{\ell_2}\mathbb{1}^m_{\bullet}$ setting*

$$R^m_n(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) := \|\theta^m_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}} + n^{-1}\|M_{\mathfrak{v}}T^{m|\dagger}_{\bullet|\bullet}\|^2_{HS}, \quad m^\circ_n := \arg\min\left\{R^m_n(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) : m \in \mathbb{N}\right\}$$
$$\text{and} \quad R^\circ_n(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) := R^{m^\circ_n}_n(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) = \min\left\{R^m_n(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) : m \in \mathbb{N}\right\} \qquad (08.03)$$

we have $\mathbb{E}^n_{\theta|T}(\|\widehat{\theta}^{m^\circ_n}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) \leqslant (1 \vee \|\Gamma_{\theta|T}\|_{\mathbb{L}(\ell_2)})R^\circ_n(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet})$ *for all $n \in \mathbb{N}$.*

§08|01.17 **Proof** of Proposition §08|01.16. Given in the lecture. □

§08|01.18 **Comment**. Let $A \in \mathbb{HS}(\ell_2)$ and $\Gamma \in \mathbb{L}^{\geqslant}(\ell_2)$ be invertible with inverse $\Gamma^{-1} \in \mathbb{L}^{\geqslant}(\ell_2)$. If we set $\mathbb{v} := \max(\|\Gamma\|_{\mathbb{L}(\ell_2)}, \|\Gamma^{-1}\|_{\mathbb{L}(\ell_2)}) \in \mathbb{R}_{>0}$, then we have $\mathbb{v}^{-1}\|A\|^2_{HS} \leqslant \operatorname{tr}(A\Gamma A^\star) \leqslant \mathbb{v}\|A\|^2_{HS}$ by using Notation §08|01.14. □

§08|01.19 **Oracle inequality**. *Under Assumptions §08|01.01 and §08|01.12 if in addition*

$$1 \leqslant \max(\|\Gamma_{\theta|T}\|_{\mathbb{L}(\ell_2)}, \|\Gamma^{-1}_{\theta|T}\|_{\mathbb{L}(\ell_2)}) \leqslant \mathbb{v}_{\theta|T} \in \mathbb{R}_{\geqslant 1}$$

is satisfied, then (08.03) *(and Comment §08|01.18) implies*

$$\mathbb{v}^{-1}_{\theta|T}R^m_n(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) \leqslant \mathbb{E}^n_{\theta|T}\|\widehat{\theta}^m_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}} = n^{-1}\operatorname{tr}(M_{\mathfrak{v}}T^{m|\dagger}_{\bullet|\bullet}\Gamma_{\theta|T}(T^{m|\dagger}_{\bullet|\bullet})^{-1}M_{\mathfrak{v}}) + \|\theta^m_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}$$
$$\leqslant \mathbb{v}_{\theta|T}R^m_n(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) \quad \forall n, m \in \mathbb{N}.$$

As a consequence we immediately obtain the following oracle inequality

$$\mathbb{v}^{-1}_{\theta|T}R^\circ_n(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) \leqslant \inf_{m \in \mathbb{N}}\mathbb{E}^n_{\theta|T}(\|\widehat{\theta}^m_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) \leqslant \mathbb{E}^n_{\theta|T}(\|\widehat{\theta}^{m^\circ_n}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}})$$
$$\leqslant \mathbb{v}_{\theta|T}R^\circ_n(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) \leqslant \mathbb{v}^2_{\theta|T}\inf_{m \in \mathbb{N}}\mathbb{E}^n_{\theta|T}(\|\widehat{\theta}^m_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) \quad \forall n \in \mathbb{N}, \qquad (08.04)$$

and, hence $\mathrm{R}^{\circ}_{n}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet})$, m°_{n} and the statistic $\widehat{\theta}^{m^{\circ}_{n}}_{\bullet}$, respectively, is an oracle bound, an oracle dimension and oracle optimal (up to the constant $\mathrm{v}^{2}_{\theta|\mathrm{T}}$). □

§08|01.20 **Remark**. Arguing similarly as in Remark §07|01.21 we note that $\|\mathrm{M}_{\mathfrak{v}}\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\|_{\mathrm{HS}} \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$ and hence $\mathrm{R}^{\circ}_{n}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) = \mathrm{o}(1)$ as $n \to \infty$, whenever $\|\theta^{m}_{\bullet} - \theta_{\bullet}\|_{\mathfrak{v}} = \mathrm{o}(1)$ as $m \to \infty$ (see Reminder §08|01.10). Note that the oracle dimension $m^{\circ}_{n} := m^{\circ}_{n}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet})$ as defined in Proposition §08|01.16 depends on the unknown parameter of interest $\theta_{\bullet}$, and thus also the oracle optimal statistic $\widehat{\theta}^{m^{\circ}_{n}}_{\bullet}$. In other words $\widehat{\theta}^{m^{\circ}_{n}}_{\bullet}$ is not a feasible estimator. □

§08|01.21 **Corollary** (GniSM §08|01.04 continued). *Consider $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}^{n}_{\theta|\mathrm{T}}$ as in Model §08|01.04, where $\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$, $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\geqslant}_{\bullet}(\ell_2)$ or $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\mathbb{R}}_{\bullet}(\ell_2)$, $\theta_{\bullet} \in \ell_2$, and hence $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) \subseteq \ell_2$. Given Assumption §08|01.12 the (infeasible, generalised) GE $\widehat{\theta}^{m^{\circ}_{n}}_{\bullet} = \mathrm{T}^{m^{\circ}_{n}|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in \ell_2\mathbb{1}^{m^{\circ}_{n}}_{\bullet} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$ with oracle dimension m°_{n} as in (08.03) satisfies*

$$\mathrm{N}^{n}_{\theta|\mathrm{T}}\big(\|\widehat{\theta}^{m^{\circ}_{n}}_{\bullet} - \theta_{\bullet}\|^{2}_{\mathfrak{v}}\big) = \mathrm{R}^{\circ}_{n}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) = \inf_{m\in\mathbb{N}} \mathrm{N}^{n}_{\theta|\mathrm{T}}\big(\|\widehat{\theta}^{m}_{\bullet} - \theta_{\bullet}\|^{2}_{\mathfrak{v}}\big) \quad \forall n \in \mathbb{N},$$

and hence it is oracle optimal (with constant 1).

§08|01.22 **Proof** of Corollary §08|01.21. Given in the lecture. □

§08|01.23 **Corollary** (niSM §08|01.06 continued). *Consider $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} \sim \mathrm{P}^{n}_{\theta|\mathrm{T}|\sigma}$ as in Model §08|01.06, where $\dot{\varepsilon}_{\bullet} \sim \otimes_{j\in\mathbb{N}}\mathrm{P}_{(0,\sigma^2_j)}$ satisfies* **(iSM1)** *and* **(iSM2)** *with $\max(\|\sigma^{-2}_{\bullet}\|_{\ell_\infty}, \|\sigma^{2}_{\bullet}\|_{\ell_\infty}) =: \mathrm{v}_{\sigma} \in \mathbb{R}_{\geqslant 1}$, $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\geqslant}_{\bullet}(\ell_2)$ or $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\mathbb{R}}_{\bullet}(\ell_2)$, $\theta_{\bullet} \in \ell_2$, and hence $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) \subseteq \ell_2$. Given Assumption §08|01.12 the (infeasible, generalised) GE $\widehat{\theta}^{m^{\circ}_{n}}_{\bullet} = \mathrm{T}^{m^{\circ}_{n}|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in \ell_2\mathbb{1}^{m^{\circ}_{n}}_{\bullet} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$ with oracle dimension m°_{n} as in (08.03) satisfies*

$$\mathrm{P}^{n}_{\theta|\mathrm{T}|\sigma}\big(\|\widehat{\theta}^{m^{\circ}_{n}}_{\bullet} - \theta_{\bullet}\|^{2}_{\mathfrak{v}}\big) \leqslant \mathrm{v}_{\sigma}\mathrm{R}^{\circ}_{n}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) \leqslant \mathrm{v}^{2}_{\sigma} \inf_{m\in\mathbb{N}} \mathrm{P}^{n}_{\theta|\mathrm{T}|\sigma}\big(\|\widehat{\theta}^{m}_{\bullet} - \theta_{\bullet}\|^{2}_{\mathfrak{v}}\big) \quad \forall n \in \mathbb{N},$$

and hence it is oracle optimal (with constant v_{σ}).

§08|01.24 **Proof** of Corollary §08|01.23. Given in the lecture. □

§08|01.25 **Corollary** (nieMM §08|01.08 continued). *Let $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ be defined on $(\mathcal{Z}^{n}, \mathscr{Z}^{\otimes n}, \mathbb{P}^{\otimes n}_{\theta|\mathrm{T}})$ as in Model §08|01.08, where $\psi_{\bullet} \in \mathcal{M}(\mathscr{Z} \otimes 2^{\mathbb{N}})$ satisfies* **(nieMM1)–(nieMM3)** *for some $\mathrm{v}_{\theta|\mathrm{T}|\psi} \in \mathbb{R}_{\geqslant 1}$, $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\geqslant}_{\bullet}(\ell_2)$ or $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\mathbb{R}}_{\bullet}(\ell_2)$, $\theta_{\bullet} \in \ell_2$, and hence $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) \subseteq \ell_2$. Given Assumption §08|01.12 the (infeasible, generalised) GE $\widehat{\theta}^{m^{\circ}_{n}}_{\bullet} = \mathrm{T}^{m^{\circ}_{n}|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in \ell_2\mathbb{1}^{m^{\circ}_{n}}_{\bullet} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$ $\mathbb{P}^{\otimes n}_{\theta|\mathrm{T}}$-a.s. with oracle dimension m°_{n} as in (08.03) satisfies*

$$\mathbb{P}^{n}_{\theta|\mathrm{T}}\big(\|\widehat{\theta}^{m^{\circ}_{n}}_{\bullet} - \theta_{\bullet}\|^{2}_{\mathfrak{v}}\big) \leqslant \mathrm{v}_{\theta|\mathrm{T}|\psi}\mathrm{R}^{\circ}_{n}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) \leqslant \mathrm{v}^{2}_{\theta|\mathrm{T}|\psi} \inf_{m\in\mathbb{N}} \mathbb{P}^{n}_{\theta|\mathrm{T}}\big(\|\widehat{\theta}^{m}_{\bullet} - \theta_{\bullet}\|^{2}_{\mathfrak{v}}\big) \quad \forall n \in \mathbb{N},$$

and hence it is oracle optimal (with constant $\mathrm{v}_{\theta|\mathrm{T}|\psi}$).

§08|01.26 **Proof** of Corollary §08|01.25. Given in the lecture. □

§08|01.27 **Illustration**. We distinguish the following two cases

(p) $\sup\big\{\|\mathrm{M}_{\mathfrak{v}}\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\|^{2}_{\mathrm{HS}} : m \in \mathbb{N}\big\} \in \mathbb{R}_{\geqslant 0}$ or $\sup\big\{\|\theta^{m}_{\bullet} - \theta_{\bullet}\|^{2}_{\mathfrak{v}} : m \in \mathbb{N}_{\geqslant K}\big\} = 0$ for some $K \in \mathbb{N}$,

(np) $\sup\big\{\|\mathrm{M}_{\mathfrak{v}}\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\|^{2}_{\mathrm{HS}} : m \in \mathbb{N}\big\} = \infty$ and $\sup\big\{\|\theta^{m}_{\bullet} - \theta_{\bullet}\|^{2}_{\mathfrak{v}} : m \in \mathbb{N}_{\geqslant K}\big\} \in \mathbb{R}_{>0}$ for all $K \in \mathbb{N}$.

Note that $\theta_{\bullet}\mathbb{1}^{K|\perp}_{\bullet} = 0_{\bullet}$ implies the case **(p)**. Interestingly, in case **(p)** the oracle bound is parametric, that is, $n\mathrm{R}^{\circ}_{n}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) = \mathrm{O}(1)$, in case **(np)** the oracle bound is nonparametric, i.e. $\lim_{n\to\infty} n\mathrm{R}^{\circ}_{n}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) = \infty$. In case **(np)** we consider similar to **(o-m)**, **(o-s)** and **(s-m)** in Illustration §07|01.28 the following specifications:

Table 01 [§08]

Order of the oracle rate $R_n^\circ(\theta_\bullet, T_{\bullet|\bullet}, \mathfrak{v}_\bullet)$ as $n \to \infty$

	$(m \in \mathbb{N})$ $(\mathfrak{v}_m = m^{\mathrm{v}})$	(squared bias) $\|\theta_\bullet^m - \theta_\bullet\|_{\mathfrak{v}}^2$ $(\mathrm{a} \in \mathbb{R}_{>0})$	(variance) $\|M_{\mathfrak{v}} T_{\bullet|\bullet}^{m\|\dagger}\|_{HS}^2$ $(\mathrm{t} \in \mathbb{R}_{>0})$	m_n°	$R_n^\circ(\theta_\bullet, T_{\bullet|\bullet}, \mathfrak{v}_\bullet)$
(o-m)	$\mathrm{v} \in (-1/2 - \mathrm{t}, \mathrm{a})$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2(\mathrm{t}+\mathrm{v})+1}$	$n^{\frac{1}{2\mathrm{a}+2\mathrm{t}+1}}$	$n^{-\frac{2(\mathrm{a}-\mathrm{v})}{2\mathrm{a}+2\mathrm{t}+1}}$
	$\mathrm{v} + \mathrm{t} = -1/2$	$m^{-2\mathrm{a}-2\mathrm{t}-1}$	$\log m$	$\left(\frac{n}{\log n}\right)^{\frac{1}{2\mathrm{a}+2\mathrm{t}+1}}$	$\frac{\log n}{n}$
(o-s)	$\mathrm{a} - \mathrm{v} \in \mathbb{R}_{>0}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{(1-2(\mathrm{t}-\mathrm{v}))_+} e^{m^{2\mathrm{t}}}$	$(\log n)^{\frac{1}{2\mathrm{t}}}$	$(\log n)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v} + \mathrm{t} + 1/2 \in \mathbb{R}_{>0}$	$m^{(1-2(\mathrm{a}-\mathrm{v}))_+} e^{-m^{2\mathrm{a}}}$	$m^{2(\mathrm{t}+\mathrm{v})+1}$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{(\log n)^{\frac{2\mathrm{t}+2\mathrm{v}+1}{2\mathrm{a}}}}{n}$
	$\mathrm{v} + \mathrm{t} = -1/2$	$e^{-m^{2\mathrm{a}}}$	$\log m$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{\log \log n}{n}$

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v} + \mathrm{t} < -1/2$ the oracle rate $R_n^\circ(\theta_\bullet, T_{\bullet|\bullet}, \mathfrak{v}_\bullet)$ is parametric. □

§08|01|02|02 Maximal global $\mathfrak{v}$-risk

§08|01.28 **Notation** (Reminder). For sequences $a_\bullet, b_\bullet \in (\mathbb{K})^{\mathbb{N}}$ taking its values in $\mathbb{K} \in \{\mathbb{R}, \mathbb{R}_{>0}, \mathbb{Z}, \dots\}$ we write $a_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\nearrow}$ and $b_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\searrow}$ if $a_\bullet$ and $b_\bullet$, respectively, is monotonically *non-decreasing* and *non-increasing*. If in addition $a_n \to \infty$ and $b_n \to 0$ as $n \to \infty$, then we write $a_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\uparrow\infty}$ and $b_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\downarrow 0}$ for short. For $\mathrm{w}_\bullet \in \ell_\infty$ we set $\mathrm{w}_{(0)} := \|\mathrm{w}_\bullet\|_{\ell_\infty}$ and $\mathrm{w}_{(\bullet)} = (\mathrm{w}_{(j)} := \|\mathrm{w}_\bullet \mathbb{1}_\bullet^{j|\perp}\|_{\ell_\infty})_{j\in\mathbb{N}}$, where by construction $\mathrm{w}_{(\bullet)} \in (\mathbb{R}_{\geqslant 0})^{\mathbb{N}}_{\searrow}$. □

§08|01.29 **Assumption**. Consider weights $\mathfrak{t}_\bullet, \mathfrak{a}_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ and $\mathfrak{v}_\bullet \in \mathbb{R}^{\mathbb{N}}_{>0}$ such that $(\mathfrak{a}\mathfrak{v})_\bullet := \mathfrak{a}_\bullet \mathfrak{v}_\bullet \in \ell_\infty$, $(\mathfrak{a}\mathfrak{v})_{(\bullet)} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$, and $(\mathfrak{t}/\mathfrak{v})_\bullet = \mathfrak{t}_\bullet \mathfrak{v}_\bullet^{-1} \in \ell_\infty$ are satisfied. In addition there exists $C_{(\mathfrak{t}/\mathfrak{v})} \in (0, 1]$ such that for all $m \in \mathbb{N}$

$$(\mathfrak{t}/\mathfrak{v})^2_{(m-1)} \geqslant \min\left\{(\mathfrak{t}/\mathfrak{v})^2_j : j \in [\![m]\!]\right\} \geqslant C_{(\mathfrak{t}/\mathfrak{v})} (\mathfrak{t}/\mathfrak{v})^2_{(m)} \tag{08.05}$$

or in equal $C_{(\mathfrak{t}/\mathfrak{v})} \|(\mathfrak{t}/\mathfrak{v})_\bullet^{-2} \mathbb{1}_\bullet^m\|_{\ell_\infty} \leqslant (\mathfrak{t}/\mathfrak{v})^{-2}_{(m)}$. □

§08|01.30 **Reminder**. Under Assumption §08|01.29 we have $\ell_2^{\mathfrak{a}} = \mathrm{dom}(M_{\mathfrak{a}^{-1}}) = \ell_2 \mathfrak{a}_\bullet \subseteq \ell_2$ and the three measures $\nu_{\mathbb{N}}$, $\mathfrak{a}_\bullet^{-2}\nu_{\mathbb{N}}$ and $\mathfrak{v}_\bullet^2 \nu_{\mathbb{N}}$ dominate mutually each other, i.e. they share the same null sets (see Property §04|01.02). We consider $\ell_2^{\mathfrak{a}}$ endowed with $\|\cdot\|_{\mathfrak{a}^{-1}} = \|M_{\mathfrak{a}^{-1}} \cdot\|_{\ell_2}$ and given a constant $\mathrm{r} \in \mathbb{R}_{>0}$ the ellipsoid $\ell_2^{\mathfrak{a},\mathrm{r}} := \{a_\bullet \in \ell_2^{\mathfrak{a}} : \|a_\bullet\|_{\mathfrak{a}^{-1}} \leqslant \mathrm{r}\} \subseteq \ell_2^{\mathfrak{a}}$. Since $(\mathfrak{a}\mathfrak{v})_\bullet \in \ell_\infty$, and hence $(\mathfrak{a}\mathfrak{v})_{(m)} := \|(\mathfrak{a}\mathfrak{v})_\bullet \mathbb{1}_\bullet^{m|\perp}\|_{\ell_\infty} \in \mathbb{R}_{\geqslant 0}$ for each $m \in \mathbb{N}$ we have $\ell_2^{\mathfrak{a}} \subseteq \ell_2(\mathfrak{v}_\bullet^2)$ (Property §04|02.11). Consequently, if Assumption §08|01.29 and $\theta_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}$ are satisfied, then Assumption §08|01.12 is also fulfilled. Since $\mathfrak{v}_\bullet, \mathfrak{t}_\bullet \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ under Assumption §08|01.29, we have $\|\mathfrak{t}_\bullet^{-1}\mathbb{1}_\bullet^m\|_{\mathfrak{v}} = \|(\mathfrak{v}/\mathfrak{t})_\bullet \mathbb{1}_\bullet^m\|_{\ell_2} \in \mathbb{R}_{\geqslant 0}$ for each $m \in \mathbb{N}$. Under the Assumptions §08|00.02 and §08|01.29 considering the generalised link condition $T_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}$ with band $\mathrm{D} \in [1, \infty)$ and $\mathrm{d} \in [1, \mathrm{D}]$ as in Definition §05|02.05 we have $\sup_{m\in\mathbb{N}}\{\|[M_{\mathfrak{t}}]_{\underline{m}} [T_{\bullet|\bullet}]_{\underline{m}}^{-1}\|_{\mathrm{spec}}\} \leqslant \mathrm{D}$, and hence

$$\begin{aligned}\|M_{\mathfrak{v}} T_{\bullet|\bullet}^{m|\dagger}\|_{HS}^2 &= \mathrm{tr}(M_{\mathfrak{v}}^m T_{\bullet|\bullet}^{m|\dagger} (T_{\bullet|\bullet}^{m|\dagger})^\star M_{\mathfrak{v}}^m) = \mathrm{tr}([M_{\mathfrak{v}}^m]_{\underline{m}} [T_{\bullet|\bullet}]_{\underline{m}}^{-1} ([T_{\bullet|\bullet}]_{\underline{m}}^{-1})^\star [M_{\mathfrak{v}}^m]_{\underline{m}}) \\ &= \mathrm{tr}([M_{(\mathfrak{v}/\mathfrak{t})}^m]_{\underline{m}} [M_{\mathfrak{t}}]_{\underline{m}} [T_{\bullet|\bullet}]_{\underline{m}}^{-1} ([T_{\bullet|\bullet}]_{\underline{m}}^{-1})^\star [M_{\mathfrak{t}}]_{\underline{m}} [M_{(\mathfrak{v}/\mathfrak{t})}^m]_{\underline{m}}) \leqslant \|[M_{\mathfrak{t}}]_{\underline{m}} [T_{\bullet|\bullet}]_{\underline{m}}^{-1}\|_{\mathrm{spec}}^2 \,\mathrm{tr}([M_{(\mathfrak{v}/\mathfrak{t})}^m]_{\underline{m}}^2) \\ &\leqslant \mathrm{D}^2 \|\mathfrak{t}_\bullet^{-1} \mathbb{1}_\bullet^m\|_{\mathfrak{v}}^2\end{aligned} \tag{08.06}$$

using $\mathrm{tr}([M_{(\mathfrak{v}/\mathfrak{t})}^m]_{\underline{m}}^2) = \|(\mathfrak{v}/\mathfrak{t})_\bullet \mathbb{1}_\bullet^m\|_{\ell_2}^2 = \|\mathfrak{t}_\bullet^{-1}\mathbb{1}_\bullet^m\|_{\mathfrak{v}}^2$. Moreover, for each $m \in \mathbb{N}$ the generalised

Galerkin solution $\theta_{\bullet}^{m} := \mathrm{T}_{\bullet|\bullet}^{m|\dagger} g_{\bullet} \in {}_{\ell_2}\mathbb{1}_{\bullet}^{m}$ of $\theta_{\bullet} = \mathrm{T}_{\bullet|\bullet}^{\dagger} g_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}$ satisfies (Lemma §05|02.09)

$$\|\theta_{\bullet} - \theta_{\bullet}^{m}\|_{\mathfrak{v}}^2 \leqslant (\mathrm{D}^2\mathrm{d}^2\mathrm{C}_{(\mathfrak{t}/\mathfrak{v})}^{-2} + 1)(\mathfrak{a}\mathfrak{v})_{(m)}^2 \mathrm{r}^2.$$

Note that under Assumptions §08|00.02 and §08|01.29 the link condition $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d}}^{\gtrless}$ with band $\mathrm{d} \in \mathbb{R}_{\geqslant 1}$ as in Definition §05|01.08 implies $\sup_{m\in\mathbb{N}}\{\|[\mathrm{M}_{\mathfrak{t}}]_{\underline{m}}[\mathrm{T}_{\bullet|\bullet}]_{\underline{m}}^{-1}\|_{\mathrm{spec}}\} \leqslant 3\mathrm{d}^2$ (Lemma §05|01.22), and hence for each $m \in \mathbb{N}$ we have (08.06) with $\mathrm{D} = 3\mathrm{d}^2$ and the Galerkin solution $\theta_{\bullet}^{m} := \mathrm{T}_{\bullet|\bullet}^{m|\dagger} g_{\bullet} \in {}_{\ell_2}\mathbb{1}_{\bullet}^{m}$ of $\theta_{\bullet} = \mathrm{T}_{\bullet|\bullet}^{\dagger} g_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}$ satisfies $\|\theta_{\bullet} - \theta_{\bullet}^{m}\|_{\mathfrak{v}}^2 \leqslant (9\mathrm{d}^6\mathrm{C}_{(\mathfrak{t}/\mathfrak{v})}^{-2} + 1)(\mathfrak{a}\mathfrak{v})_{(m)}^2 \mathrm{r}^2$ (Lemma §05|01.28). □

§08|01.31 **Proposition.** *Under Assumptions §08|01.01 and §08|01.29 setting for* $n, m \in \mathbb{N}$

$$\mathrm{R}_n^m(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) := [(\mathfrak{a}\mathfrak{v})_{(m)}^2 \vee n^{-1}\|\mathfrak{t}_{\bullet}^{-1}\mathbb{1}_{\bullet}^{m}\|_{\mathfrak{v}}^2], \quad m_n^{\star} := \arg\min\left\{\mathrm{R}_n^m(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) : m \in \mathbb{N}\right\}$$
$$\text{and} \quad \mathrm{R}_n^{\star}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) := \mathrm{R}_n^{m_n^{\star}}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) = \min\left\{\mathrm{R}_n^m(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) : m \in \mathbb{N}\right\} \quad (08.07)$$

and $\|\Gamma_{\theta|\mathrm{T}}\|_{\mathbb{L}(\ell_2)} =: \mathrm{v}_{\theta|\mathrm{T}} \in \mathbb{R}_{\geqslant 0}$*, for* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}$ *and for all* $\theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}$*, hence* $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{T}_{\bullet|\bullet}^{\dagger}) \subseteq \ell_2$*, we have*

$$\mathbb{E}_{\theta|\mathrm{T}}^{n}(\|\widehat{\theta}_{\bullet}^{m_n^{\star}} - \theta_{\bullet}\|_{\mathfrak{v}}^2) \leqslant (\mathrm{D}^2\mathrm{v}_{\theta|\mathrm{T}} + 2\mathrm{C}_{(\mathfrak{t}/\mathfrak{v})}^{-2}\mathrm{D}^2\mathrm{d}^2\mathrm{r}^2)\,\mathrm{R}_n^{\star}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) \quad \forall n \in \mathbb{N}$$

(or for $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d}}^{\gtrless}$ *with* $\mathrm{D} = 3\mathrm{d}^2$*).*

§08|01.32 **Proof** of Proposition §08|01.31. Given in the lecture. □

§08|01.33 **Remark.** Under the assumptions of Proposition §08|01.31 if there exists in addition $\mathrm{v} \in \mathbb{R}_{>0}$ satisfying $\|\Gamma_{\theta|\mathrm{T}}\|_{\mathbb{L}(\ell_2)} \leqslant \mathrm{v}$ for all $\theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}$ and $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}$ (or $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d}}^{\gtrless}$), then we have

$$\sup\left\{\mathbb{E}_{\theta|\mathrm{T}}^{n}(\|\widehat{\theta}_{\bullet}^{m_n^{\star}} - \theta_{\bullet}\|_{\mathfrak{v}}^2) : \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}\right\} \leqslant (\mathrm{D}^2\mathrm{v} + 2\mathrm{C}_{(\mathfrak{t}/\mathfrak{v})}^{-2}\mathrm{D}^2\mathrm{d}^2\mathrm{r}^2)\,\mathrm{R}_n^{\star}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) \quad \forall n \in \mathbb{N}.$$

Arguing similarly as in Remark §07|01.21 we note that $\mathrm{R}_n^{\star}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) = \mathrm{o}(1)$ as $n \to \infty$ since $\|\mathfrak{t}_{\bullet}^{-1}\mathbb{1}_{\bullet}^{m}\|_{\mathfrak{v}} \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$, and $(\mathfrak{a}\mathfrak{v})_{(m)} = \mathrm{o}(1)$ as $m \to \infty$ by Assumption §08|01.29. The latter is satisfied, for example, if $(\mathfrak{a}\mathfrak{v})_{\bullet} \in \ell_2$ (in equal $\mathfrak{a}_{\bullet} \in \ell_2(\mathfrak{v}_{\bullet}^2)$). Note that the dimension $m_n^{\star} := m_n^{\star}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet})$ as defined in (08.07) does not depend on the unknown parameter of interest $\theta_{\bullet}$ but on the classes $\ell_2^{\mathfrak{a},\mathrm{r}}$ and $\mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}$ only, and thus also the statistic $\widehat{\theta}_{\bullet}^{m_n^{\star}}$. In other words, if the regularity of $\theta_{\bullet}$ and $\mathrm{T}_{\bullet|\bullet}$ is known in advance, then the (generalised) GE $\widehat{\theta}_{\bullet}^{m_n^{\star}}$ is a feasible estimator. □

§08|01.34 **Corollary** (GniSM §08|01.04 continued). *Consider* $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}_{\theta|\mathrm{T}}^{n}$ *as in Model §08|01.04, where* $\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}$*,* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}_{\bullet}^{\gtrless}(\ell_2)$ *or* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}_{\bullet}^{\mathrm{R}}(\ell_2)$*,* $\theta_{\bullet} \in \ell_2$ *and hence* $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{T}_{\bullet|\bullet}^{\dagger}) \subseteq \ell_2$*. Under Assumption §08|01.29 the (generalised) GE* $\widehat{\theta}_{\bullet}^{m_n^{\star}} = \mathrm{T}_{\bullet|\bullet}^{m_n^{\star}|\dagger}\widehat{g}_{\bullet} \in {}_{\ell_2}\mathbb{1}_{\bullet}^{m_n^{\star}} \subseteq \ell_2(\mathfrak{v}_{\bullet}^2)$ *with dimension* $m_n^{\star}$ *as in (08.07) satisfies*

$$\sup\left\{\mathrm{N}_{\theta|\mathrm{T}}^{n}(\|\widehat{\theta}_{\bullet}^{m_n^{\star}} - \theta_{\bullet}\|_{\mathfrak{v}}^2) : \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}\right\} \leqslant \mathrm{C}\,\mathrm{R}_n^{\star}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) \quad \forall n \in \mathbb{N} \quad (08.08)$$

with constant $\mathrm{C} = \mathrm{D}^2 + 2\mathrm{C}_{(\mathfrak{t}/\mathfrak{v})}^{-2}\mathrm{D}^2\mathrm{d}^2\mathrm{r}^2$ *(for* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d}}^{\gtrless}$ *with* $\mathrm{D} = 3\mathrm{d}^2$*).*

§08|01.35 **Proof** of Corollary §08|01.34. Given in the lecture. □

§08|01.36 **Corollary** (niSM §08|01.06 continued). *Consider* $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} \sim \mathrm{P}_{\theta|\mathrm{T}|\sigma}^{n}$ *as in Model §08|01.06, where* $\dot{\varepsilon}_{\bullet} \sim \otimes_{j\in\mathbb{N}}\mathrm{P}_{(0,\sigma_j^2)}$ *satisfies* **(iSM1)** *with* $\|\sigma_{\bullet}^2\|_{\ell_\infty} =: \mathrm{v}_{\sigma} \in \mathbb{R}_{>0}$*,* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}_{\bullet}^{\gtrless}(\ell_2)$ *or* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}_{\bullet}^{\mathrm{R}}(\ell_2)$*,* $\theta_{\bullet} \in \ell_2$ *and hence* $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{T}_{\bullet|\bullet}^{\dagger}) \subseteq \ell_2$*. Under Assumption §08|01.29 the (generalised) GE* $\widehat{\theta}_{\bullet}^{m_n^{\star}} = \mathrm{T}_{\bullet|\bullet}^{m_n^{\star}|\dagger}\widehat{g}_{\bullet} \in {}_{\ell_2}\mathbb{1}_{\bullet}^{m_n^{\star}} \subseteq \ell_2(\mathfrak{v}_{\bullet}^2)$ *with dimension* $m_n^{\star}$ *as in (08.07) satisfies*

$$\sup\left\{\mathrm{P}_{\theta|\mathrm{T}|\sigma}^{n}(\|\widehat{\theta}_{\bullet}^{m_n^{\circ}} - \theta_{\bullet}\|_{\mathfrak{v}}^2) : \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}\right\} \leqslant \mathrm{C}\,\mathrm{R}_n^{\star}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) \quad \forall n \in \mathbb{N}$$

with constant $\mathrm{C} = \mathrm{D}^2\mathrm{v}_{\sigma} + 2\mathrm{C}_{(\mathfrak{t}/\mathfrak{v})}^{-2}\mathrm{D}^2\mathrm{d}^2\mathrm{r}^2$ *(for* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d}}^{\gtrless}$ *with* $\mathrm{D} = 3\mathrm{d}^2$*).*

§08|01.37 **Proof** of Corollary §08|01.36. Given in the lecture. □

§08|01.38 **Corollary** (nieMM §08|01.08 continued). *Let $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ be defined on $(\mathcal{Z}^n, \mathscr{Z}^{\otimes n}, \mathbb{P}^{\otimes n}_{\theta|\mathrm{T}})$ as in Model §08|01.08, where $\psi_{\bullet} \in \mathcal{M}(\mathscr{X} \otimes 2^{\mathbb{N}})$ satisfies* **(nieMM1)** *and* **(nieMM2)** *for some $\mathrm{v}_{\theta|\mathrm{T}|\psi} \in \mathbb{R}_{\geqslant 1}$, $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\geqslant}_{\bullet}(\ell_2)$ or $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\mathbb{R}}_{\bullet}(\ell_2)$, $\theta_{\bullet} \in \ell_2$ and hence $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) \subseteq \ell_2$. Under Assumption §08|01.29 the (generalised) GE $\widehat{\theta}^{m^\star_n}_{\bullet} = \mathrm{T}^{m^\star_n|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in {}_{\ell_2}\mathbb{1}^{m^\star_n}_{\bullet} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$ with dimension $m^\star_n$ as in (08.07) satisfies*

$$\sup\left\{\mathbb{P}^{\otimes n}_{\theta|\mathfrak{s}}\left(\|\widehat{\theta}^{m^\circ_n}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}\right): \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}\right\} \leqslant \mathrm{C}_{\mathfrak{a},\mathrm{r},\mathfrak{t},\mathrm{d},\mathrm{D}}\, \mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) \quad \forall n \in \mathbb{N}$$

with constant $\mathrm{C}_{\mathfrak{a},\mathrm{r},\mathfrak{t},\mathrm{d},\mathrm{D}} = \mathrm{D}^2 \sup\left\{\mathrm{v}_{\theta|\mathrm{T}|\psi}: \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}\right\} + 2\mathrm{C}^{-2}_{(\mathfrak{t}/\mathfrak{v})}\mathrm{D}^2\mathrm{d}^2\mathrm{r}^2$ (for $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}^{\gtrless}_{\mathfrak{t},\mathrm{d}}$ with $\mathrm{D} = 3\mathrm{d}^2$).

§08|01.39 **Proof** of Corollary §08|01.38. Given in the lecture. □

§08|01.40 **Illustration**. We distinguish the following two cases **(p)** $(\mathfrak{v}/\mathfrak{t})_{\bullet} \in \ell_2$, and **(np)** $(\mathfrak{v}/\mathfrak{t})_{\bullet} \notin \ell_2$. Interestingly, in case **(p)** the bound in Proposition §08|01.31 is parametric, that is, $n\mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) = \mathrm{O}(1)$, in case **(np)** the bound is nonparametric, i.e. $\lim_{n\to\infty} n\mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) = \infty$. In case **(np)** we consider similar to **(o-m)**, **(o-s)** and **(s-m)** in Illustration §07|01.44 the following three specifications:

Table 02 [§08]

Order of the rate $\mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet})$ as $n \to \infty$

	$(j \in \mathbb{N})$ $\mathfrak{v}^2_j = j^{2v}$	$(\mathrm{a} \in \mathbb{R}_{>0})$ $\mathfrak{a}^2_j$	$(\mathrm{t} \in \mathbb{R}_{>0})$ $\mathfrak{t}^2_j$	(squared bias) $(\mathfrak{a}\mathfrak{v})^2_{(m)}$	(variance) $\|\mathfrak{t}^{-1}_{\bullet}\mathbb{1}^m_{\bullet}\|^2_{\mathfrak{v}}$	$m^\star_n$	$\mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet})$
(o-m)	$\mathrm{v} \in (-1/2 - \mathrm{t}, \mathrm{a})$	$j^{-2\mathrm{a}}$	$j^{-2\mathrm{t}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2\mathrm{v}+2\mathrm{t}+1}$	$n^{\frac{1}{2\mathrm{a}+2\mathrm{t}+1}}$	$n^{-\frac{2(\mathrm{a}-\mathrm{v})}{2\mathrm{a}+2\mathrm{t}+1}}$
	$\mathrm{v} + \mathrm{t} = -1/2$	$j^{-2\mathrm{a}}$	$j^{-2\mathrm{t}}$	$m^{-2\mathrm{a}-2\mathrm{t}-1}$	$\log m$	$\left(\frac{n}{\log n}\right)^{\frac{1}{2\mathrm{a}+2\mathrm{t}+1}}$	$\frac{\log n}{n}$
(o-s)	$\mathrm{a} - \mathrm{v} \in \mathbb{R}_{>0}$	$j^{-2\mathrm{a}}$	$e^{-j^{2\mathrm{t}}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{(1-2(\mathrm{t}-\mathrm{v}))_+}e^{m^{2\mathrm{t}}}$	$(\log n)^{\frac{1}{2\mathrm{t}}}$	$(\log n)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v} + \mathrm{t} + 1/2 \in \mathbb{R}_{>0}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$m^{2\mathrm{v}}e^{-m^{2\mathrm{a}}}$	$m^{2\mathrm{v}+2\mathrm{t}+1}$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{(\log n)^{\frac{2\mathrm{t}+2\mathrm{v}+1}{2\mathrm{a}}}}{n}$
	$\mathrm{v} + \mathrm{t} = -1/2$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$m^{2\mathrm{v}}e^{-m^{2\mathrm{a}}}$	$\log m$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{\log\log n}{n}$

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v} + \mathrm{t} < -1/2$ the rate $\mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet})$ is parametric. □

§08|01|03 Local and maximal local ϕ-risk

Secondly, we measure the accuracy of the (generalised) GE $\widehat{\theta}^m_{\bullet} := \mathrm{T}^{m|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet}$ of the (generalised) Galerkin solution $\theta^m_{\bullet} = \mathrm{T}^{m|\dagger}_{\bullet|\bullet}g_{\bullet} \in {}_{\ell_2}\mathbb{1}^m_{\bullet}$ with $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet})$ by the mean of its local ϕ-error introduced in §05|01|02 and §05|02|02, i.e. its ϕ-risk.

§08|01.41 **Reminder**. If $\phi_{\bullet} \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ then we have $\phi^2_{\bullet}\mathbb{1}^m_{\bullet} \in \ell_2$ and ${}_{\ell_2}\mathbb{1}^m_{\bullet} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$. Consequently, for each $\theta_{\bullet} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ the (generalised) Galerkin solution $\theta^m_{\bullet} = \mathrm{T}^{m|\dagger}_{\bullet|\bullet}g_{\bullet} \in {}_{\ell_2}\mathbb{1}^m_{\bullet}$ satisfies $\theta^m_{\bullet} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ too. If in addition $\mathrm{C}_{\mathrm{T}} := \sup\left\{\|\mathrm{M}_{\mathbb{1}^{m|\perp}_{\bullet}}\mathrm{T}^\star_{\bullet|\bullet}(\mathrm{T}^{m|\dagger}_{\bullet|\bullet})^\star\phi_{\bullet}\|_{\ell_2}: m \in \mathbb{N}\right\} \in \mathbb{R}_{\geqslant 0}$ then $|\phi\nu_{\mathbb{N}}(\theta^m_{\bullet} - \theta_{\bullet})| \leqslant (1 + \mathrm{C}_{\mathrm{T}})\|\mathbb{1}^{m|\perp}_{\bullet}\theta_{\bullet}\|_{\ell_2}$ which implies $\sup\left\{|\phi\nu_{\mathbb{N}}(\theta^j_{\bullet} - \theta_{\bullet})|: j \in \mathbb{N}_{\geqslant m}\right\} = \mathrm{o}(1)$ as $m \to \infty$ (Property §05|01.31 and Property §05|02.12). □

§08|01.42 **Comment**. Under Assumption §08|01.01 since $\theta^m_{\bullet}, \mathrm{T}^{m|\dagger}_{\bullet|\bullet}\mathbb{1}^m_{\bullet} \in {}_{\ell_2}\mathbb{1}^m_{\bullet}$ for each $m \in \mathbb{N}$ we have $\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\dot{\varepsilon}_{\bullet} \in {}_{\ell_2}\mathbb{1}^m_{\bullet}$ $\mathbb{P}^n_{\theta|\mathrm{T}}$-a.s.. Indeed, $\dot{\varepsilon}_{\bullet} \sim \mathrm{P}_{(0_{\bullet},\Gamma_{\theta|\mathrm{T}})}$ with $\Gamma_{\theta|\mathrm{T}} \in \mathbb{L}^{\geqslant}(\ell_2)$ by Assumption §08|01.01 **(nSIP)** implies $\mathbb{P}^n_{\theta|\mathrm{T}}(\dot{\varepsilon}^2_{\bullet}) \in \ell_\infty$, hence $\dot{\varepsilon}_{\bullet}\mathbb{1}^m_{\bullet} \in \ell_\infty$ $\mathbb{P}^n_{g}$-a.s. and $\|\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\dot{\varepsilon}_{\bullet}\|_{\ell_2} \leqslant \|\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\mathbb{1}^m_{\bullet}\|_{\ell_2}\|\dot{\varepsilon}_{\bullet}\mathbb{1}^m_{\bullet}\|_{\ell_\infty} \in \mathbb{R}_{\geqslant 0}$ $\mathbb{P}^n_{\theta|\mathrm{T}}$-a.s..

Given $\phi_\bullet \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ from $\ell_2 \mathbb{1}^m_\bullet \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ (Reminder §08|01.41) it follows

$$\widehat{\theta}^m_\bullet = \mathrm{T}^{m|\dagger}_{\bullet|\bullet} \widehat{g}_\bullet = n^{-1/2}\mathrm{T}^{m|\dagger}_{\bullet|\bullet} \dot{\varepsilon}_\bullet + \theta^m_\bullet \in \ell_2 \mathbb{1}^m_\bullet \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}}) \quad \mathbb{P}^n_{\theta|\mathrm{T}}\text{-a.s..}$$ □

§08|01|03|01 Local ϕ-risk

§08|01.43 **Assumption**. Let $\phi_\bullet \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ and $\theta_\bullet \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ be satisfied. □

§08|01.44 **Definition**. Under Assumptions §08|01.01 and §08|01.43 the *local ϕ-risk* of a (generalised) GE $\widehat{\theta}^m_\bullet = \mathrm{T}^{m|\dagger}_{\bullet|\bullet}\widehat{g}_\bullet \in \ell_2\mathbb{1}^m_\bullet \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ $\mathbb{P}^n_{\theta|\mathrm{T}}$-a.s. satisfies

$$\mathbb{E}^n_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^m_\bullet - \theta_\bullet)|^2) = \mathbb{E}^n_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\mathrm{T}^{m|\dagger}_{\bullet|\bullet}(\widehat{g}_\bullet - g_\bullet))|^2) + |\phi\nu_{\mathbb{N}}(\theta^m_\bullet - \theta_\bullet)|^2 \tag{08.09}$$

with *variance* $\mathbb{E}^n_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\mathrm{T}^{m|\dagger}_{\bullet|\bullet}(\widehat{g}_\bullet - g_\bullet))|^2) = n^{-1}\mathbb{E}^n_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\dot{\varepsilon}_\bullet)|^2)$ and *bias* $|\phi\nu_{\mathbb{N}}(\theta^m_\bullet - \theta_\bullet)|$. □

§08|01.45 **Property**. *Under Assumptions §08|01.01 and §08|01.43 we have*

$$\mathbb{E}^n_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\dot{\varepsilon}_\bullet)|^2) = \mathbb{E}^n_{\theta|\mathrm{T}}(|\langle \phi\mathbb{1}^m_\bullet, \mathrm{T}^{m|\dagger}_{\bullet|\bullet}\dot{\varepsilon}_\bullet\rangle_{\ell_2}|^2) = \mathbb{E}^n_{\theta|\mathrm{T}}(|\langle (\mathrm{T}^{m|\dagger}_{\bullet|\bullet})^\star \phi^m_\bullet, \dot{\varepsilon}_\bullet\rangle_{\ell_2}|^2)$$
$$= \langle \Gamma_{\theta|\mathrm{T}}(\mathrm{T}^{m|\dagger}_{\bullet|\bullet})^\star\phi^m_\bullet, (\mathrm{T}^{m|\dagger}_{\bullet|\bullet})^\star\phi^m_\bullet\rangle_{\ell_2} = \|(\mathrm{T}^{m|\dagger}_{\bullet|\bullet})^\star\phi^m_\bullet\|^2_{\Gamma_{\theta|\mathrm{T}}} \tag{08.10}$$

and consequently $\mathbb{E}^n_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\mathrm{T}^{m|\dagger}_{\bullet|\bullet}(\widehat{g}_\bullet - g_\bullet))|^2) \leqslant n^{-1}\|\Gamma_{\theta|\mathrm{T}}\|_{\mathbb{L}(\ell_2)}\|(\mathrm{T}^{m|\dagger}_{\bullet|\bullet})^\star\phi^m_\bullet\|^2_{\ell_2} \in \mathbb{R}_{\geqslant 0}$. □

§08|01.46 **Proposition** (Upper bound). *Under Assumptions §08|01.01 and §08|01.43 for all $n, m \in \mathbb{N}$ with (generalised) Galerkin solution $\theta^m_\bullet = \mathrm{T}^{m|\dagger}_{\bullet|\bullet} g_\bullet \in \ell_2\mathbb{1}^m_\bullet$ setting*

$$\mathrm{R}^m_n(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet) := |\phi\nu_{\mathbb{N}}(\theta^m_\bullet - \theta_\bullet)|^2 + n^{-1}\|(\mathrm{T}^{m|\dagger}_{\bullet|\bullet})^\star\phi^m_\bullet\|^2_{\ell_2}, \quad m^\circ_n := \arg\min\left\{\mathrm{R}^m_n(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet) : m \in \mathbb{N}\right\}$$
$$\textit{and} \quad \mathrm{R}^\circ_n(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet) := \mathrm{R}^{m^\circ_n}_n(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet) = \min\left\{\mathrm{R}^m_n(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet) : m \in \mathbb{N}\right\} \tag{08.11}$$

we have $\mathbb{E}^n_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m^\circ_n}_\bullet - \theta_\bullet)|^2) \leqslant (1 \vee \|\Gamma_{\theta|\mathrm{T}}\|_{\mathbb{L}(\ell_2)})\mathrm{R}^\circ_n(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet)$ *for all $n \in \mathbb{N}$.*

§08|01.47 **Proof** of Proposition §08|01.46. Given in the lecture. □

§08|01.48 **Reminder**. If $\Gamma_{\theta|\mathrm{T}} \in \mathbb{L}^{\geqslant}(\ell_2)$ is invertible with inverse $\Gamma^{-1}_{\theta|\mathrm{T}} \in \mathbb{L}(\ell_2)$, i.e. $\Gamma_{\theta|\mathrm{T}}\Gamma^{-1}_{\theta|\mathrm{T}} = \mathrm{id}_{\ell_2} = \Gamma^{-1}_{\theta|\mathrm{T}}\Gamma_{\theta|\mathrm{T}}$, then we write shortly $1 \leqslant \max(\|\Gamma_{\theta|\mathrm{T}}\|_{\mathbb{L}(\ell_2)}, \|\Gamma^{-1}_{\theta|\mathrm{T}}\|_{\mathbb{L}(\ell_2)}) \leqslant \mathrm{v}_{\theta|\mathrm{T}} \in \mathbb{R}_{\geqslant 1}$. In this situation for all $h_\bullet \in \ell_2$ we have $\mathrm{v}^{-1}_{\theta|\mathrm{T}}\|h_\bullet\|^2_{\ell_2} \leqslant \|h_\bullet\|^2_{\Gamma_{\theta|\mathrm{T}}} = \langle \Gamma_{\theta|\mathrm{T}}h_\bullet, h_\bullet\rangle_{\mathbb{J}} \leqslant \mathrm{v}_{\theta|\mathrm{T}}\|h_\bullet\|^2_{\ell_2}$. □

§08|01.49 **Oracle inequality**. *Under Assumptions §08|01.01 and §08|01.43 if in addition*

$$1 \leqslant \max(\|\Gamma_{\theta|\mathrm{T}}\|_{\mathbb{L}(\ell_2)}, \|\Gamma^{-1}_{\theta|\mathrm{T}}\|_{\mathbb{L}(\ell_2)}) \leqslant \mathrm{v}_{\theta|\mathrm{T}} \in \mathbb{R}_{\geqslant 1}$$

is satisfied then (08.11) *(and Reminder §08|01.48) implies*

$$\mathrm{v}^{-1}_{\theta|\mathrm{T}}\mathrm{R}^m_n(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet) \leqslant \mathbb{E}^n_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^m_\bullet - \theta_\bullet)|^2) = n^{-1}\|(\mathrm{T}^{m|\dagger}_{\bullet|\bullet})^\star\phi^m_\bullet\|^2_{\Gamma_{\theta|\mathrm{T}}} + |\phi\nu_{\mathbb{N}}(\theta^m_\bullet - \theta_\bullet)|^2$$
$$\leqslant \mathrm{v}_{\theta|\mathrm{T}}\mathrm{R}^m_n(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet) \quad \forall n, m \in \mathbb{N}.$$

As a consequence we immediately obtain the following oracle inequality

$$\mathrm{v}^{-1}_{\theta|\mathrm{T}}\mathrm{R}^\circ_n(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet) \leqslant \inf_{m\in\mathbb{N}} \mathbb{E}^n_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^m_\bullet - \theta_\bullet)|^2) \leqslant \mathbb{E}^n_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m^\circ_n}_\bullet - \theta_\bullet)|^2)$$
$$\leqslant \mathrm{v}_{\theta|\mathrm{T}}\mathrm{R}^\circ_n(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet) \leqslant \mathrm{v}^2_{\theta|\mathrm{T}} \inf_{m\in\mathbb{N}} \mathbb{E}^n_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^m_\bullet - \theta_\bullet)|^2) \quad \forall n \in \mathbb{N}, \tag{08.12}$$

and hence $\mathrm{R}^\circ_n(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet)$, m°_n and the statistic $\widehat{\theta}^{m^\circ_n}_\bullet$, respectively, is an oracle bound, an oracle dimension and oracle optimal (up to the constant $\mathrm{v}^2_{\theta|\mathrm{T}}$). □

§08|01.50 **Remark**. Arguing similarly as in Remark §07|01.21 we note that $\|(T_{\bullet|\bullet}^{m|\dagger})^\star \phi^m\|_{\ell_2}^2 \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$ and hence $R_n^\circ(\theta, T_{\bullet|\bullet}, \phi) = o(1)$ as $n \to \infty$, whenever $|\phi\nu_{\mathbb{N}}(\theta^m - \theta)|^2 = o(1)$ as $m \to \infty$ (see Reminder §08|01.41). Note that the oracle dimension $m_n^\circ := m_n^\circ(\theta, T_{\bullet|\bullet}, \phi)$ as defined in (08.11) depends on the unknown parameter of interest θ, and thus also the oracle optimal statistic $\widehat{\theta}^{m_n^\circ}$. In other words $\widehat{\theta}^{m_n^\circ}$ is not a feasible estimator. □

§08|01.51 **Corollary** (GniSM §08|01.04 continued). *Consider $\widehat{g} = g + n^{-1/2}\dot{B} \sim N_{\theta|T}^n$ as in Model §08|01.04, where $\dot{B} \sim N_{(0,1)}^{\otimes\mathbb{N}}$, $T_{\bullet|\bullet} \in \mathbb{L}^{\geqslant}(\ell_2)$ or $T_{\bullet|\bullet} \in \mathbb{L}^{R}(\ell_2)$, $\theta \in \ell_2$, and hence $g = T_{\bullet|\bullet}\theta \in \mathrm{dom}(T_{\bullet|\bullet}^\dagger) \subseteq \ell_2$. Given Assumption §08|01.43 the (infeasible, generalised) GE $\widehat{\theta}^{m_n^\circ} = T_{\bullet|\bullet}^{m_n^\circ|\dagger}\widehat{g} \in \ell_2 \mathbb{1}^{m_n^\circ} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ with oracle dimension m_n° as in (08.11) satisfies*

$$N_{\theta|T}^n\big(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m_n^\circ} - \theta)|^2\big) = R_n^\circ(\theta, T_{\bullet|\bullet}, \phi) = \inf_{m \in \mathbb{N}} N_{\theta|T}^n\big(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^m - \theta)|^2\big),$$

and hence it is oracle optimal (with constant 1).

§08|01.52 **Proof** of Corollary §08|01.51. Given in the lecture. □

§08|01.53 **Corollary** (niSM §08|01.06 continued). *Consider $\widehat{g} = g + n^{-1/2}\dot{\varepsilon} \sim P_{\theta|T|\sigma}^n$ as in Model §08|01.06, where $\dot{\varepsilon} \sim \otimes_{j\in\mathbb{N}} P_{(0,\sigma^2)}$ satisfies (iSM1) and (iSM2) with $\max(\|\sigma^{-2}\|_{\ell_\infty}, \|\sigma^2\|_{\ell_\infty}) =: v_\sigma \in \mathbb{R}_{\geqslant 1}$, $T_{\bullet|\bullet} \in \mathbb{L}^{\geqslant}(\ell_2)$ or $T_{\bullet|\bullet} \in \mathbb{L}^{R}(\ell_2)$, $\theta \in \ell_2$, and hence $g = T_{\bullet|\bullet}\theta \in \mathrm{dom}(T_{\bullet|\bullet}^\dagger) \subseteq \ell_2$. Given Assumption §08|01.43 the (infeasible, generalised) GE $\widehat{\theta}^{m_n^\circ} = T_{\bullet|\bullet}^{m_n^\circ|\dagger}\widehat{g} \in \ell_2 \mathbb{1}^{m_n^\circ} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ with oracle dimension m_n° as in (08.11) satisfies*

$$P_{\theta|T|\sigma}^n\big(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m_n^\circ} - \theta)|^2\big) \leqslant v_\sigma R_n^\circ(\theta, T, \phi) \leqslant v_\sigma^2 \inf_{m \in \mathbb{N}} P_{\theta|T|\sigma}^n\big(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^m - \theta)|^2\big),$$

and hence it is oracle optimal (with constant v_σ).

§08|01.54 **Proof** of Corollary §08|01.53. Given in the lecture. □

§08|01.55 **Corollary** (nieMM §08|01.08 continued). *Let $\widehat{g} = g + n^{-1/2}\dot{\varepsilon}$ be defined on $(\mathcal{Z}^n, \mathscr{Z}^{\otimes n}, P_{\theta|T}^{\otimes n})$ as in Model §08|01.08, where $\psi \in \mathcal{M}(\mathscr{Z} \otimes 2^{\mathbb{N}})$ satisfies (nieMM1)–(nieMM3) for some $v_{\theta|T|\psi} \in \mathbb{R}_{\geqslant 1}$, $T_{\bullet|\bullet} \in \mathbb{L}^{\geqslant}(\ell_2)$ or $T_{\bullet|\bullet} \in \mathbb{L}^{R}(\ell_2)$, $\theta \in \ell_2$, and hence $g = T_{\bullet|\bullet}\theta \in \mathrm{dom}(T_{\bullet|\bullet}^\dagger) \subseteq \ell_2$. Given Assumption §08|01.43 the (infeasible, generalised) GE $\widehat{\theta}^{m_n^\circ} = T_{\bullet|\bullet}^{m_n^\circ|\dagger}\widehat{g} \in \ell_2 \mathbb{1}^{m_n^\circ} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ with oracle dimension m_n° as in (08.11) satisfies*

$$\mathbb{P}_{\theta|T}^n\big(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m_n^\circ} - \theta)|^2\big) \leqslant v_{\theta|T|\psi} R_n^\circ(\theta, T_{\bullet|\bullet}, \mathfrak{v}) \leqslant v_{\theta|T|\psi}^2 \inf_{m \in \mathbb{N}} \mathbb{P}_{\theta|T}^n\big(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^m - \theta)|^2\big),$$

and hence it is oracle optimal (with constant $v_{\theta|T|\psi}$).

§08|01.56 **Proof** of Corollary §08|01.55. Given in the lecture. □

§08|01.57 **Illustration**. We distinguish the following two cases

(p) $\sup\big\{\|(T_{\bullet|\bullet}^{m|\dagger})^\star\phi^m\|_{\ell_2}^2 : m \in \mathbb{N}\big\} \in \mathbb{R}_{\geqslant 0}$ or $\sup\big\{|\phi\nu(\widehat{\theta}^m - \theta)|^2 : m \in \mathbb{N}_{\geqslant K}\big\} = 0$ for $K \in \mathbb{N}$,

(np) $\sup\big\{\|(T_{\bullet|\bullet}^{m|\dagger})^\star\phi^m\|_{\ell_2}^2 : m \in \mathbb{N}\big\} = \infty$ and $\sup\big\{|\phi\nu(\widehat{\theta}^m - \theta)|^2 : m \in \mathbb{N}_{\geqslant K}\big\} \in \mathbb{R}_{>0}$ for all $K \in \mathbb{N}$.

Note that $\theta\mathbb{1}^{K|\perp} = 0$ implies the case **(p)**. Interestingly, in case **(p)** the oracle bound is parametric, that is, $nR_n^\circ(\theta, T_{\bullet|\bullet}, \phi) = O(1)$, in case **(np)** the oracle bound is nonparametric, i.e. $\lim_{n\to\infty} nR_n^\circ(\theta, T_{\bullet|\bullet}, \phi) = \infty$. In case **(np)** we consider similar to **(o-m)**, **(o-s)** and **(s-m)** in Illustration §07|01.63 the following specifications:

Table 03 [§08]

Order of the oracle rate $\mathrm{R}^\circ_n(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet)$ as $n \to \infty$

	$(m \in \mathbb{N})$ $(\phi_{m} = m^{\mathrm{v}-1/2})$	(squarred bias) $\lvert\phi\nu_{\mathbb{N}}(\widehat{\theta}^m_\bullet - \theta_\bullet)\rvert^2$ $(\mathrm{a} \in \mathbb{R}_{>0})$	(variance) $\lVert(\mathrm{T}^{m\mid\dagger}_{\bullet|\bullet})^\star\phi^m_\bullet\rVert^2_{\ell_2}$ $(\mathrm{t} \in \mathbb{R}_{>0})$	m°_n	$\mathrm{R}^\circ_n(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet)$
(o-m)	$\mathrm{v} \in (-\mathrm{t}, \mathrm{a})$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2\mathrm{t}+2\mathrm{v}}$	$n^{\frac{1}{2\mathrm{a}+2\mathrm{t}}}$	$n^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{a}+\mathrm{t}}}$
	$\mathrm{v} = -\mathrm{t}$	$m^{-2(\mathrm{a}+\mathrm{t})}$	$\log m$	$\left(\frac{n}{\log n}\right)^{\frac{1}{2(\mathrm{a}+\mathrm{t})}}$	$\frac{\log n}{n}$
(o-s)	$\mathrm{a} - \mathrm{v} \in \mathbb{R}_{>0}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2(\mathrm{v}-\mathrm{t})_+}e^{m^{2\mathrm{t}}}$	$(\log n)^{\frac{1}{2\mathrm{t}}}$	$(\log n)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v} + \mathrm{t} \in \mathbb{R}_{>0}$	$m^{(1-4\mathrm{a}+2\mathrm{v})_+}e^{-m^{2\mathrm{a}}}$	$m^{2\mathrm{t}+\mathrm{v}}$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{(\log n)^{\frac{\mathrm{t}+\mathrm{v}}{\mathrm{a}}}}{n}$
	$\mathrm{v} = -\mathrm{t}$	$m^{(1-4\mathrm{a}-2\mathrm{t})_+}e^{-m^{2\mathrm{a}}}$	$\log m$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{\log\log n}{n}$

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v} < -\mathrm{t}$ the oracle rate $\mathrm{R}^\circ_n(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet)$ is parametric. □

§08|01|03|02 Maximal local ϕ-risk

§08|01.58 **Assumption**. Consider weights $\mathfrak{t}_\bullet, \mathfrak{a}_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ and $\phi_\bullet \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ such that $(\mathfrak{a}\phi)_\bullet := \mathfrak{a}_\bullet\phi_\bullet \in \ell_2$ and $(\mathfrak{a}\mathfrak{t})_\bullet := \mathfrak{a}_\bullet\mathfrak{t}_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$. □

§08|01.59 **Comment**. Assuming $\mathfrak{t}_\bullet, \mathfrak{a}_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ and hence $(\mathfrak{a}\mathfrak{t})^2_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ is rather weak. If in addition $\liminf_{j\to\infty} (\mathfrak{a}\mathfrak{t})^2_j \geqslant c \in \mathbb{R}_{>0}$ is satisfied, and hence $(\mathfrak{a}\mathfrak{t})^2_\bullet, \mathfrak{a}^2_\bullet, \mathfrak{t}^2_\bullet \notin (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$, then $\mathfrak{a}^2_\bullet \notin (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$ and the assumption $(\mathfrak{a}\phi)_\bullet \in \ell_2$ implies $\phi_\bullet \in \ell_2$, which together with $\mathfrak{t}^2_\bullet \notin (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$ implies $(\phi/\mathfrak{t})_\bullet \in \ell_2$, and thus the rate $\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$ is parametric (Illustration §08|01.72). Since we are interested in the case of a non-parametric rate, the additional assumption $(\mathfrak{a}\mathfrak{t})^2_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$ imposes a rather weak condition satisfied also in Illustration §08|01.72. □

§08|01.60 **Reminder**. Under Assumption §08|01.58 we have $\ell^{\mathfrak{a}}_2 = \mathrm{dom}(\mathrm{M}_{\mathfrak{a}^{-1}_\bullet}) = \ell_2\mathfrak{a}_\bullet \subseteq \ell_2$ and the three measures $\nu_{\mathbb{N}}$, $\mathfrak{a}^{-2}_\bullet\nu_{\mathbb{N}}$ and $\lvert\phi_\bullet\rvert\nu_{\mathbb{N}}$ dominate mutually each other, i.e. they share the same null sets (see Property §04|01.02). We consider $\ell^{\mathfrak{a}}_2$ endowed with $\lVert\cdot\rVert_{\mathfrak{a}^{-1}} = \lVert\mathrm{M}_{\mathfrak{a}^{-1}_\bullet}\cdot\rVert_{\ell_2}$ and given a constant $\mathrm{r} \in \mathbb{R}_{>0}$ the ellipsoid $\ell^{\mathfrak{a},\mathrm{r}}_2 := \{a_\bullet \in \ell^{\mathfrak{a}_\bullet}_2 : \lVert a_\bullet\rVert_{\mathfrak{a}^{-1}} \leqslant \mathrm{r}\} \subseteq \ell^{\mathfrak{a}}_2$. Since $(\mathfrak{a}\phi)_\bullet \in \ell_2$ we have $\ell^{\mathfrak{a}}_2 \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ (Property §04|02.23). Consequently, if Assumption §08|01.58 and $\theta_\bullet \in \ell^{\mathfrak{a},\mathrm{r}}_2$ are satisfied, then Assumption §08|01.43 is also fulfilled. Moreover, from $(\mathfrak{a}\phi)_\bullet \in \ell_2$ follows $\lVert\mathfrak{a}_\bullet\mathbb{1}^{m\mid\perp}_\bullet\rVert_\phi = \lVert(\mathfrak{a}\phi)_\bullet\mathbb{1}^{m\mid\perp}_\bullet\rVert_{\ell_2} = \mathrm{o}(1)$ as $m \to \infty$. For $s \in [0,1]$ from $(\mathfrak{a}\mathfrak{t}^s)_\bullet = \mathfrak{a}_\bullet\mathfrak{t}^s_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ follows $(\mathfrak{a}\mathfrak{t}^s)_{(\bullet)} = \big((\mathfrak{a}\mathfrak{t}^s)_{(m)} := (\mathfrak{a}\mathfrak{t}^s)_{m+1} = \lVert(\mathfrak{a}\mathfrak{t}^s)_\bullet\mathbb{1}^{m\mid\perp}_\bullet\rVert_{\ell_\infty}\big)_{m\in\mathbb{N}} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$. Since $\phi_\bullet, \mathfrak{t}_\bullet \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ under Assumption §08|01.58, we have $\ell_2\mathbb{1}^m_\bullet \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ and $\lVert\mathfrak{t}^{-1}_\bullet\mathbb{1}^m_\bullet\rVert_\phi = \lVert(\phi/\mathfrak{t})_\bullet\mathbb{1}^m_\bullet\rVert_{\ell_2} \in \mathbb{R}_{>0}$ for each $m \in \mathbb{N}$. Under the Assumptions §08|00.02 and §08|01.58 considering the generalised link condition $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}$ with band $\mathrm{D} \in \mathbb{R}_{\geqslant 1}$ and $\mathrm{d} \in [1, \mathrm{D}]$ as in Definition §05|02.05 we have $\sup_{m\in\mathbb{N}}\{\lVert([\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}})^\star[\mathrm{M}_{\mathfrak{t}}]_{\underline{m}}\rVert_{\mathrm{spec}}\} \leqslant \mathrm{D}$, and hence

$$\begin{aligned}\lVert(\mathrm{T}^{m\mid\dagger}_{\bullet|\bullet})^\star\phi^m_\bullet\rVert_{\ell_2} &= \lVert([\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}})^\star[\phi_\bullet]_{\underline{m}}\rVert = \lVert([\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}})^\star[\mathrm{M}_{\mathfrak{t}}]_{\underline{m}}[\mathrm{M}_{\mathfrak{t}}]^{-1}_{\underline{m}}[\phi_\bullet]_{\underline{m}}\rVert \\ &\leqslant \lVert([\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}})^\star[\mathrm{M}_{\mathfrak{t}}]_{\underline{m}}\rVert_{\mathrm{spec}}\lVert[\mathrm{M}_{\mathfrak{t}^{-1}}]_{\underline{m}}[\phi_\bullet]_{\underline{m}}\rVert \leqslant \mathrm{D}\lVert\mathfrak{t}^{-1}_\bullet\mathbb{1}^m_\bullet\rVert_\phi \end{aligned} \quad (08.13)$$

using $\lVert[\mathrm{M}_{\mathfrak{t}^{-1}}]_{\underline{m}}[\phi_\bullet]_{\underline{m}}\rVert = \lVert\mathfrak{t}^{-1}_\bullet\mathbb{1}^m_\bullet\rVert_\phi$. Moreover, for each $m \in \mathbb{N}$ the generalised Galerkin solution $\theta^m_\bullet := \mathrm{T}^{m\mid\dagger}_{\bullet|\bullet}g_\bullet \in \ell_2\mathbb{1}^m_\bullet$ of $\theta_\bullet = \mathrm{T}^\dagger_{\bullet|\bullet}g_\bullet \in \ell^{\mathfrak{a},\mathrm{r}}_2$ satisfies (Lemma §05|02.14)

$$\lvert\phi\nu_{\mathbb{N}}(\theta^m_\bullet - \theta_\bullet)\rvert^2 \leqslant \mathrm{Dd}(\mathrm{Dd}+1)\mathrm{r}^2\big(\lVert\mathfrak{a}_\bullet\mathbb{1}^{m\mid\perp}_\bullet\rVert^2_\phi + (\mathfrak{a}\mathfrak{t})^2_{(m)}\lVert\mathfrak{t}^{-1}_\bullet\mathbb{1}^m_\bullet\rVert^2_\phi\big). \quad (08.14)$$

Under Assumptions §08|00.02 and §08|01.58 the link condition $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}}$ with band $\mathrm{d} \in \mathbb{R}_{\geqslant 1}$ as in Definition §05|01.08 implies $\sup_{m\in\mathbb{N}}\{\|([\mathrm{T}_{\bullet|\bullet}]_{\underline{m}}^{-1})^{\star}[\mathrm{M}_{\mathfrak{t}}]_{\underline{m}}\|_{\mathrm{spec}}\} \leqslant 3\mathrm{d}^2$ (Lemma §05|01.22), and hence for each $m \in \mathbb{N}$ we have (08.13) with $\mathrm{D} = 3\mathrm{d}^2$ and the Galerkin solution $\theta^{m}_{\bullet} := \mathrm{T}^{m|\dagger}_{\bullet|\bullet} g_{\bullet} \in {}_{\ell_2}\mathbb{1}^{m}_{\bullet}$ of $\theta_{\bullet} = \mathrm{T}^{\dagger}_{\bullet|\bullet} g_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}$ satisfies (08.14) with $\mathrm{D} = 3\mathrm{d}^2$ (Lemma §05|01.34). □

§08|01.61 **Lemma.** *Under Assumption §08|01.58 setting for $n, m \in \mathbb{N}$*

$$\mathrm{R}^{m}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet}) := \|\mathfrak{a}_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|^2_{\phi_{\bullet}} + n^{-1}\|\mathfrak{t}^{-1}_{\bullet}\mathbb{1}^{m}_{\bullet}\|^2_{\phi_{\bullet}}, \quad m^{\star}_{n} := \arg\min\left\{\mathrm{R}^{m}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet}) : m \in \mathbb{N}\right\}$$

$$\text{and} \quad \mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet}) := \mathrm{R}^{m^{\star}_{n}}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet}) = \min\left\{\mathrm{R}^{m}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet}) : m \in \mathbb{N}\right\} \quad (08.15)$$

we have $(\mathfrak{a}\mathfrak{t})^2_{m^{\star}_{n}} > n^{-1} \geqslant (\mathfrak{a}\mathfrak{t})^2_{m^{\star}_{n}+1} = (\mathfrak{a}\mathfrak{t})^2_{(m^{\star}_{n})}$ for all $n \in \mathbb{N}_{>(\mathfrak{a}\mathfrak{t})_2^{-2}}$, i.e. $(\mathfrak{a}\mathfrak{t})^2_2 > n^{-1}$ is satisfied.

§08|01.62 **Proof** of Lemma §08|01.61. Given in the lecture. □

§08|01.63 **Proposition** (Upper bound). *Under Assumptions §08|01.01 and §08|01.58 setting $m^{\star}_{n}$ and $\mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet})$ for $n \in \mathbb{N}$ as in (08.15) and $\|\Gamma_{\theta|\mathrm{T}}\|_{\mathbb{L}(\ell_2)} =: \mathrm{v}_{\theta|\mathrm{T}} \in \mathbb{R}_{\geqslant 0}$, for $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d,D}}$ and for all $\theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}$, hence $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) \subseteq \ell_2$, we have*

$$\mathbb{P}^{n}_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m^{\star}_{n}}_{\bullet} - \theta_{\bullet})|^2) \leqslant \mathrm{D}^2(\mathrm{v}_{\theta|\mathrm{T}} + 2\mathrm{d}^2\mathrm{r}^2)\,\mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet}) \quad \forall n \in \mathbb{N}_{>(\mathfrak{a}\mathfrak{t})_2^{-2}}$$

(or for $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}}$ with $\mathrm{D} = 3\mathrm{d}^2$).

§08|01.64 **Proof** of Proposition §08|01.63. Given in the lecture. □

§08|01.65 **Remark.** Under the assumptions of Proposition §08|01.63 if there exists in addition $\mathrm{v} \in \mathbb{R}_{>0}$ satisfying $\|\Gamma_{\theta|\mathrm{T}}\|_{\mathbb{L}(\ell_2)} \leqslant \mathrm{v}$ for all $\theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}$ and $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d,D}}$ (or $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}}$), then we have

$$\sup\left\{\mathbb{P}^{n}_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m^{\star}_{n}}_{\bullet} - \theta_{\bullet})|^2) : \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d,D}}\right\} \leqslant \mathrm{D}^2(\mathrm{v} + 2\mathrm{d}^2\mathrm{r}^2)\,\mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet}) \quad \forall n \in \mathbb{N}_{>(\mathfrak{a}\mathfrak{t})_2^{-2}}$$

Arguing similarly as in Remark §07|01.56 we note that $\|\mathfrak{t}^{-1}_{\bullet}\mathbb{1}^{m}_{\bullet}\|_{\phi_{\bullet}} \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$ and $(\|\phi_{\bullet}\mathbb{1}^{m|\perp}_{\bullet}\|^2_{\mathfrak{a}_{\bullet}} = \mathrm{o}(1)$ as $m \to \infty$ (since $(\mathfrak{a}\phi)_{\bullet} \in \ell_2$), and hence $\mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet}) = \mathrm{o}(1)$ as $n \to \infty$. Note that the dimension $m^{\star}_{n} := m^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet})$ as defined in (08.15) does not depend on the unknown parameter of interest $\theta_{\bullet}$ but on the classes $\ell_2^{\mathfrak{a},\mathrm{r}}$ and $\mathbb{T}_{\mathrm{t,d,D}}$ only, and thus also the statistic $\widehat{\theta}^{m^{\star}_{n}}_{\bullet}$. In other words, if the regularity of $\theta_{\bullet}$ and $\mathrm{T}_{\bullet|\bullet}$ is known in advance, then the (generalised) GE $\widehat{\theta}^{m^{\star}_{n}}_{\bullet}$ is a feasible estimator. □

§08|01.66 **Corollary** (GniSM §08|01.04 continued). *Consider $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}^{n}_{\theta|\mathrm{T}}$ as in Model §08|01.04, where $\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$, $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\gtrless}_{\bullet}(\ell_2)$ or $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\mathrm{R}}_{\bullet}(\ell_2)$, $\theta_{\bullet} \in \ell_2$ and hence $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) \subseteq \ell_2$. Under Assumption §08|01.58 the (generalised) GE $\widehat{\theta}^{m^{\star}_{n}}_{\bullet} = \mathrm{T}^{m^{\star}_{n}|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in {}_{\ell_2}\mathbb{1}^{m^{\star}_{n}}_{\bullet} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ with dimension $m^{\star}_{n}$ as in (08.15) satisfies*

$$\sup\left\{\mathrm{N}^{n}_{\theta|\mathrm{T}}\left(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m^{\star}_{n}}_{\bullet} - \theta_{\bullet})|^2\right) : \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d,D}}\right\} \leqslant \mathrm{C}_{\mathrm{r,d,D}}\,\mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet}) \quad \forall n \in \mathbb{N}_{>(\mathfrak{a}\mathfrak{t})_2^{-2}} \quad (08.16)$$

with constant $\mathrm{C}_{\mathrm{r,d,D}} = \mathrm{D}^2(1 + 2\mathrm{d}^2\mathrm{r}^2)$ (for $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}}$ with $\mathrm{D} = 3\mathrm{d}^2$).

§08|01.67 **Proof** of Corollary §08|01.66. Given in the lecture. □

§08|01.68 **Corollary** (niSM §08|01.06 continued). *Consider $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet} \sim \mathrm{P}^{n}_{\theta|\mathrm{T}|\sigma}$ as in Model §08|01.06, where $\dot{\varepsilon}_{\bullet} \sim \otimes_{j\in\mathbb{N}}\mathrm{P}_{(0,\sigma_j^2)}$ satisfies **(iSM1)** with $\|\sigma^2_{\bullet}\|_{\ell_\infty} =: \mathrm{v}_{\sigma} \in \mathbb{R}_{>0}$, $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\gtrless}_{\bullet}(\ell_2)$ or $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\mathrm{R}}_{\bullet}(\ell_2)$, $\theta_{\bullet} \in \ell_2$ and hence $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) \subseteq \ell_2$. Under Assumption §08|01.58 the (generalised) GE $\widehat{\theta}^{m^{\star}_{n}}_{\bullet} = \mathrm{T}^{m^{\star}_{n}|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in {}_{\ell_2}\mathbb{1}^{m^{\star}_{n}}_{\bullet} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ with dimension $m^{\star}_{n}$ as in (08.15) satisfies*

$$\sup\left\{\mathrm{P}^{n}_{\theta|\mathrm{T}|\sigma}\left(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m^{\star}_{n}}_{\bullet} - \theta_{\bullet})|^2\right) : \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d,D}}\right\} \leqslant \mathrm{C}_{\mathrm{r,d,D},\sigma}\,\mathrm{R}^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet}) \quad \forall n \in \mathbb{N}_{>(\mathfrak{a}\mathfrak{t})_2^{-2}}$$

with constant $\mathrm{C}_{\mathrm{r,d,D},\sigma} = \mathrm{D}^2(\|\sigma^2_{\bullet}\|_{\ell_\infty} + 2\mathrm{d}^2\mathrm{r}^2)$ (for $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}}$ with $\mathrm{D} = 3\mathrm{d}^2$).

§08|01.69 **Proof** of Corollary §08|01.68. Given in the lecture. □

§08|01.70 **Corollary** (nieMM §08|01.08 continued). *Let* $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}$ *be defined on* $(\mathcal{Z}^n, \mathscr{Z}^{\otimes n}, \mathbb{P}^{\otimes n}_{\theta|\mathrm{T}})$ *as in Model §08|01.08, where* $\psi_{\bullet} \in \mathcal{M}(\mathscr{Z} \otimes 2^{\mathbb{N}})$ *satisfies* **(nieMM1)** *and* **(nieMM2)** *for some* $\mathrm{v}_{\theta|\mathrm{T}|\psi} \in \mathbb{R}_{\geqslant 1}$, $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\geqslant}_{\bullet}(\ell_2)$ *or* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\mathbb{R}}_{\bullet}(\ell_2)$, $\theta_{\bullet} \in \ell_2$ *and hence* $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) \subseteq \ell_2$. *Under Assumption §08|01.58 the (generalised) GE* $\widehat{\theta}^{m_n^\star}_{\bullet} = \mathrm{T}^{m_n^\star|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in \ell_2\mathbb{1}^{m_n^\star}_{\bullet} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *with dimension* $m_n^\star$ *as in* (08.15) *satisfies*

$$\sup\left\{\mathbb{P}^{\otimes n}_{\theta|\mathrm{T}}\left(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m_n^\circ}_{\bullet} - \theta_{\bullet})|^2\right) : \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}\right\} \leqslant \mathrm{C}_{\mathrm{r},\mathfrak{a}_\bullet,\mathrm{d},\mathrm{D},\mathfrak{t}_\bullet}\,\mathrm{R}^\star_n(\mathfrak{a}_\bullet,\mathfrak{t}_\bullet,\phi_\bullet) \quad \forall n \in \mathbb{N}_{>(\mathfrak{a}\mathfrak{t})_1^{-2}}$$

with constant $\mathrm{C}_{\mathrm{r},\mathfrak{a}_\bullet,\mathrm{d},\mathrm{D},\mathfrak{t}_\bullet} = \mathrm{D}^2(\sup\left\{\mathrm{v}_{\theta|\mathrm{T}|\psi} : \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}\right\} + 2\mathrm{d}^2\mathrm{r}^2)$ *(for* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}^{\geqslant}_{\mathfrak{t},\mathrm{d}}$ *with* $\mathrm{D} = 3\mathrm{d}^2$*).*

§08|01.71 **Proof** of Corollary §08|01.70. Given in the lecture. □

§08|01.72 **Illustration**. We distinguish the following two cases **(p)** $(\phi/\mathfrak{t})_{\bullet} \in \ell_2$, and **(np)** $(\phi/\mathfrak{t})_{\bullet} \notin \ell_2$. Interestingly, in case **(p)** the bound in Proposition §08|01.63 is parametric, that is, $n\mathrm{R}^\star_n(\mathfrak{a}_\bullet,\mathfrak{t}_\bullet,\phi_\bullet) = \mathrm{O}(1)$, in case **(p)** the bound is nonparametric, i.e. $\lim_{n\to\infty} n\mathrm{R}^\star_n(\mathfrak{a}_\bullet,\mathfrak{t}_\bullet,\phi_\bullet) = \infty$. In case **(p)** we consider similar to **(o-m)**, **(o-s)** and **(s-m)** in Illustration §07|01.78 the following specifications:

Table 04 [§08]

Order of the rate $\mathrm{R}^\star_n(\mathfrak{a}_\bullet,\mathfrak{t}_\bullet,\phi_\bullet)$ as $n \to \infty$

	$(j \in \mathbb{N})$ $\phi_j^2 = j^{2\mathrm{v}-1}$	$(\mathrm{a} \in \mathbb{R}_{>0})$ $\mathfrak{a}_j^2$	$(\mathrm{t} \in \mathbb{R}_{>0})$ $\mathfrak{t}_j^2$	(squarred bias) $\|\mathfrak{a}_\bullet\mathbb{1}^{m\|\perp}_\bullet\|^2_{\phi}$	(variance) $\|\mathfrak{t}^{-1}_\bullet\mathbb{1}^m_\bullet\|^2_{\phi}$	$m_n^\star$	$\mathrm{R}^\star_n(\mathfrak{a}_\bullet,\mathfrak{t}_\bullet,\phi_\bullet)$
(o-m)	$\mathrm{v} \in (-\mathrm{t},\mathrm{a})$	$j^{-2\mathrm{a}}$	$j^{-2\mathrm{t}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2\mathrm{v}+2\mathrm{t}}$	$n^{\frac{1}{2\mathrm{a}+2\mathrm{t}}}$	$n^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{a}+\mathrm{t}}}$
	$\mathrm{v} = -\mathrm{t}$	$j^{-2\mathrm{a}}$	$j^{-2\mathrm{t}}$	$m^{-2(\mathrm{a}+\mathrm{t})}$	$\log m$	$\left(\frac{n}{\log n}\right)^{\frac{1}{2(\mathrm{a}+\mathrm{t})}}$	$\frac{\log n}{n}$
(o-s)	$\mathrm{a}-\mathrm{v} \in \mathbb{R}_{>0}$	$j^{-2\mathrm{a}}$	$e^{-j^{2\mathrm{t}}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2(\mathrm{v}-\mathrm{t})_+}e^{m^{2\mathrm{t}}}$	$(\log n)^{\frac{1}{2\mathrm{t}}}$	$(\log n)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v}+\mathrm{t} \in \mathbb{R}_{>0}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$e^{-m^{2\mathrm{a}}}$	$m^{2\mathrm{v}+2\mathrm{t}}$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{(\log n)^{\frac{\mathrm{t}+\mathrm{v}}{\mathrm{a}}}}{n}$
	$\mathrm{v} = -\mathrm{t}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$e^{-m^{2\mathrm{a}}}$	$\log m$	$(\log n)^{\frac{1}{2\mathrm{a}}}$	$\frac{\log\log n}{n}$

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v} < -\mathrm{t}$ the rate $\mathrm{R}^\star_n(\mathfrak{a}_\bullet,\mathfrak{t}_\bullet,\phi_\bullet)$ is parametric. □

§08|02 Non-diagonal statistical inverse problem with noisy operator

§08|02.01 **Notation** Reminder. For $\mathrm{A}_{\bullet|\bullet} = (\mathrm{A}_{j|j_o})_{j,j_o\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}^2}$ we denote by $\mathrm{A}^m_{\bullet|\bullet} := \mathrm{M}_{\mathbb{1}^m}\mathrm{A}_{\bullet|\bullet}\mathrm{M}_{\mathbb{1}^m} \in \mathbb{L}(\ell_2)$ with

$$a_{\bullet} \mapsto \mathrm{A}^m_{\bullet|\bullet}a_{\bullet} = (\mathbb{1}^m_j \sum_{j_o\in[\![m]\!]} \mathrm{A}_{j,j_o}a_{j_o} = \mathbb{1}^m_j\langle \mathrm{A}_{j|\bullet}\mathbb{1}^m_\bullet, a_\bullet\rangle_{\ell_2} = \mathbb{1}^m_j\nu_{\mathbb{N}}(\mathrm{A}_{j,\bullet}a_\bullet\mathbb{1}^m_\bullet))_{j\in\mathbb{N}}$$

the operator which restricted to a linear map from $\mathbb{R}^m$ $(\mathrm{ran}(\mathrm{M}_{\mathbb{1}^m}) = \ell_2\mathbb{1}^m_\bullet)$ into itself is represented by the sub-matrix $[\mathrm{A}_{\bullet|\bullet}]_{\underline{m}} := (\mathrm{A}_{j|j_o})_{j,j_o\in[\![m]\!]} \in \mathbb{R}^{(m,m)}$ (compare Notation §05|00.02). Moreover, $\|\cdot\|$ and $\|A\|_{\mathrm{spec}} := \sup\{\|Ax\| : \|x\| \leqslant 1\}$ denotes, respectively, the Euclidean norm of a vector and the spectral norm of a matrix A. Clearly, we have $\|\mathrm{A}^m_{\bullet|\bullet}\|_{\mathbb{L}(\ell_2)} = \|\mathrm{M}_{\mathbb{1}^m}\mathrm{A}_{\bullet|\bullet}\mathrm{M}_{\mathbb{1}^m}\|_{\mathbb{L}(\ell_2)} = \|[\mathrm{A}_{\bullet|\bullet}]_{\underline{m}}\|_{\mathrm{spec}}$. Furthermore, $\mathrm{A}^m_{\bullet|\bullet} \in \mathbb{L}(\ell_2)$ is a Hilbert-Schmidt operator (§08|01.14), i.e. $\mathrm{A}^m_{\bullet|\bullet} \in \mathbb{HS}(\ell_2)$, and $\mathrm{M}_{\mathrm{w}_\bullet}\mathrm{A}^m_{\bullet|\bullet} = \mathrm{M}^m_{\mathrm{w}}\mathrm{A}^m \in \mathbb{HS}(\ell_2)$ for arbitrary $\mathrm{w}_\bullet \in \mathbb{R}^{\mathbb{N}}$. Moreover, introduce $\ell_p(\mathbb{N}^2) := \mathbb{L}_p(\mathbb{N}^2, 2^{\mathbb{N}^2}, \nu_{\mathbb{N}^2})$ for $p \in \overline{\mathbb{R}}_{\geqslant 1}$. □

§08|02.02 **Assumption**. Consider a stochastic process $\dot{\varepsilon}_\bullet = (\dot{\varepsilon}_j)_{j\in\mathbb{N}}$ satisfying Assumption §01|01.04 with *mean zero* and a sample size $n \in \mathbb{N}$, and in addition a stochastic process $\dot{\boldsymbol{\eta}}_{\bullet|\bullet} = (\dot{\boldsymbol{\eta}}_{j|j_o})_{j,j_o\in\mathbb{N}}$ satisfying Assumption §02|01.02, with *mean zero* and a sample size $k \in \mathbb{N}$. Let Assumption §08|00.02

be satisfied where $\mathrm{T}_{\cdot|\cdot} \in \mathbb{L}^{\geqslant}(\ell_2)$ or $\mathrm{T}_{\cdot|\cdot} \in \mathbb{L}^{\mathbb{R}}(\ell_2)$ is *not known* anymore. For $\theta_\cdot \in \ell_2$ the observable noisy image with mean $g_\cdot = \mathrm{T}_{\cdot|\cdot}\theta_\cdot \in \ell_2$ and the observable noisy non-diagonal operator with mean kernel $\mathrm{T}_{\cdot|\cdot} \in \mathbb{R}^{\mathbb{N}^2}$ takes the form $\widehat{g}_\cdot = g_\cdot + n^{-1/2}\dot{\varepsilon}_\cdot$ and $\widehat{\mathrm{T}}_{\cdot|\cdot} = \mathrm{T}_{\cdot|\cdot} + k^{-1/2}\dot{\eta}_{\cdot|\cdot}$, respectively. We denote by $\mathbb{P}^{n,k}_{\theta|\mathrm{T}}$ the joint distribution of $(\widehat{g}_\cdot, \widehat{\mathrm{T}}_{\cdot|\cdot})$. Denoting by $\mathbb{P}^{n}_{\theta|\mathrm{T}}$ and $\mathbb{P}^{k}_{\mathrm{T}}$ the marginal distribution of $\widehat{g}_\cdot$ and $\widehat{\mathrm{T}}_{\cdot|\cdot}$, respectively, if $\dot{\varepsilon}_\cdot$ and $\dot{\eta}_{\cdot|\cdot}$ are *independent* then we write $\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}} = \mathbb{P}^{n}_{\theta|\mathrm{T}} \otimes \mathbb{P}^{k}_{\mathrm{T}}$ for the joint product distribution of $(\widehat{g}_\cdot, \widehat{\mathrm{T}}_{\cdot|\cdot})$. In addition $\dot{\varepsilon}_\cdot$ satisfies **(nSIP)** (Assumption §08|01.01) with $\mathrm{v}_{\theta|\mathrm{T}} \in \mathbb{R}_{\geqslant 1}$ and $\dot{\eta}_{\cdot|\cdot}$ fulfils

(nSIPnO1) there is $\mathrm{v}_{\mathrm{T}} \in \mathbb{R}_{\geqslant 1}$ such that $\dot{\eta}_{\cdot|\cdot}$ for all $m \in \mathbb{N}$ and $a_\cdot, b_\cdot \in \ell_2$ satisfies

$$\mathbb{P}^{k}_{\mathrm{T}}\left(|\langle b_\cdot, \dot{\eta}^{m}_{\cdot|\cdot} a_\cdot\rangle_{\ell_2}|^2\right) \leqslant \mathrm{v}_{\mathrm{T}}\|a_\cdot\|^2_{\ell_2}\|b_\cdot\|^2_{\ell_2}$$

implying $\mathbb{P}^{k}_{\mathrm{T}}(\dot{\eta}^2_{j|j_\circ}) =: \mathrm{v}^{\mathrm{T}}_{j|j_\circ} \leqslant \mathrm{v}_{\mathrm{T}}$ for all $j, j_\circ \in \mathbb{N}$, and hence $1 \vee \|\mathrm{v}^{\mathrm{T}}_{\cdot|\cdot}\|_{\ell_\infty(\mathbb{N}^2)} \leqslant \mathrm{v}_{\mathrm{T}}$;

(nSIPnO2) there is $l \in \mathbb{N}$ and $\mathrm{K}^2_{\mathrm{T}} \in \mathbb{R}_{\geqslant \mathrm{v}_{\mathrm{T}}}$ such that $\dot{\eta}_{\cdot|\cdot}$ satisfies

$$\mathrm{v}^{\mathrm{T}|(l)}_{\cdot|\cdot} := \mathbb{P}^{k}_{\mathrm{T}}(\dot{\eta}^{2l}_{\cdot|\cdot}) := (\mathrm{v}^{\mathrm{T}|(l)}_{j|j_\circ} := \mathbb{P}^{k}_{\mathrm{T}}(\dot{\eta}^{2l}_{j|j_\circ}))_{j,j_\circ\in\mathbb{N}} \in \ell_\infty(\mathbb{N}^2),$$

and $1 \vee \|\mathrm{v}^{\mathrm{T}|(l)}_{\cdot|\cdot}\|_{\ell_\infty(\mathbb{N}^2)} \leqslant \mathrm{K}^{2l}_{\mathrm{T}}$ where $1 \vee \|\mathrm{v}^{\mathrm{T}}_{\cdot|\cdot}\|_{\ell_\infty(\mathbb{N}^2)} \leqslant \mathrm{v}_{\mathrm{T}} \leqslant \mathrm{K}^2_{\mathrm{T}}$. □

§08|02.03 **Lemma**. *Let Assumption §08|02.02* **(nSIPnO1)** *and* **(nSIPnO2)** *be satisfied, and let* $m \in \mathbb{N}$.

(i) *Under* **(nSIPnO1)** *for any* $\mathrm{A}_{\cdot|\cdot} \in \mathbb{HS}(\ell_2)$ *and* $a_\cdot \in \ell_2$ *we have*

$$\mathbb{P}^{k}_{\mathrm{T}}\left(\|\mathrm{A}_{\cdot|\cdot}\dot{\eta}^{m}_{\cdot|\cdot}a_\cdot\|^2_{\ell_2}\right) \leqslant \mathrm{v}_{\mathrm{T}}\|\mathrm{A}_{\cdot|\cdot}\|^2_{\mathrm{HS}}\|a_\cdot\|^2_{\ell_2}$$

and in particular, $m^{-1}\mathbb{P}^{k}_{\mathrm{T}}\left(\|\dot{\eta}^{m}_{\cdot|\cdot}a_\cdot\|^2_{\ell_2}\right) \leqslant \mathrm{v}_{\mathrm{T}}\|a_\cdot\|^2_{\ell_2}$ *by using* $\|\mathrm{id}^{m}_{\cdot|\cdot}\|^2_{\mathrm{HS}} = m$.

(ii) *Under* **(nSIPnO2)** *for all* $x \in \mathbb{R}_{>0}$ *we have*

$$\mathbb{P}^{k}_{\mathrm{T}}\left(\|\dot{\eta}^{m}_{\cdot|\cdot}\|_{\mathbb{L}(\ell_2)} \geqslant x\right) \leqslant m^{2l}x^{-2l}\mathrm{K}^{2l}_{\mathrm{T}}$$

and $\mathbb{P}^{k}_{\mathrm{T}}\left(\|\dot{\eta}^{m}_{\cdot|\cdot}\|^2_{\mathbb{L}(\ell_2)}\mathbb{1}_{\{\|\dot{\eta}^{m}_{\cdot|\cdot}\|_{\mathbb{L}(\ell_2)} \geqslant x\}}\right) \leqslant m^{2l}x^{-2(l-1)}\mathrm{K}^{2l}_{\mathrm{T}}$.

§08|02.04 **Proof** of Lemma §08|02.03. Given in the lecture. □

§08|02.05 **Notation**. For each $m \in \mathbb{N}$ and $\mathrm{T}_{\cdot|\cdot} \in \mathbb{L}^{\mathbb{R}}(\ell_2)$ we introduce below an observable event $\Omega_{m,k\wedge n}$ and its complement $\Omega^c_{m,k\wedge n}$ such that on the event $\Omega_{m,k\wedge n}$ the random matrix $[\widehat{\mathrm{T}}_{\cdot|\cdot}]_{\underline{m}} \in \mathbb{R}^{(m,m)}$ is regular and hence its inverse $[\widehat{\mathrm{T}}_{\cdot|\cdot}]^{-1}_{\underline{m}} \in \mathbb{R}^{(m,m)}$ always exists. We denote by $\mathbb{1}_{\Omega_{m,k\wedge n}}$ the observable elementary random variable which takes the value 1 on the event $\Omega_{m,k\wedge n}$ and zero otherwise. We denote by $\widehat{\mathrm{T}}^{m}_{\cdot|\cdot} := \mathrm{M}_{\mathbb{1}^m}\widehat{\mathrm{T}}_{\cdot|\cdot}\mathrm{M}_{\mathbb{1}^m} \in \mathbb{L}(\ell_2)$ the random operator which restricted to a linear map from $\mathbb{R}^m$ into itself can be represented by the random matrix $[\widehat{\mathrm{T}}_{\cdot|\cdot}]_{\underline{m}}$. Note that its Moore-Penrose inverse $\widehat{\mathrm{T}}^{m|\dagger}_{\cdot|\cdot} \in \mathbb{L}(\ell_2)$ restricted to a linear map from $\mathbb{R}^m$ into itself can be represented by the Moore-Penrose inverse matrix $[\widehat{\mathrm{T}}_{\cdot|\cdot}]^{\dagger}_{\underline{m}}$ of $[\widehat{\mathrm{T}}_{\cdot|\cdot}]_{\underline{m}}$ (see Definition §03|00.08). On the event $\Omega_{m,k\wedge n}$ we have $\widehat{\mathrm{T}}^{m}_{\cdot|\cdot}\widehat{\mathrm{T}}^{m|\dagger}_{\cdot|\cdot} = \widehat{\mathrm{T}}^{m|\dagger}_{\cdot|\cdot}\widehat{\mathrm{T}}^{m}_{\cdot|\cdot} = \mathrm{M}_{\mathbb{1}^m}$. Let $0_{\cdot|\cdot} \in \mathbb{L}(\ell_2)$ be the zero operator mapping ℓ_2 to $\{0_\cdot\}$. The random operator, which equals $\mathrm{A}_{\cdot|\cdot} \in \mathbb{L}(\ell_2)$ on $\Omega_{m,k\wedge n}$ and $0_{\cdot|\cdot}$ on $\Omega^c_{m,k\wedge n}$, is denoted by $\mathrm{A}^{(k\wedge n)}_{\cdot|\cdot} := \mathrm{A}_{\cdot|\cdot}\mathbb{1}_{\Omega_{m,k\wedge n}} \in \mathbb{L}(\ell_2)$. Let $\widehat{\mathrm{T}}^{m|(k\wedge n)}_{\cdot|\cdot} := \widehat{\mathrm{T}}^{m}_{\cdot|\cdot}\mathbb{1}_{\Omega_{m,k\wedge n}} \in \mathbb{L}(\ell_2)$ and denote its Moore-Penrose inverse by $\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\cdot|\cdot} \in \mathbb{L}(\ell_2)$ where trivially $\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\cdot|\cdot} = \widehat{\mathrm{T}}^{m|\dagger}_{\cdot|\cdot}\mathbb{1}_{\Omega_{m,k\wedge n}}$. We eventually use the elementary identity $\widehat{\mathrm{T}}^{m|\dagger}_{\cdot|\cdot}\widehat{\mathrm{T}}^{m}_{\cdot|\cdot}\mathbb{1}_{\Omega_{m,k\wedge n}} = \widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\cdot|\cdot}\widehat{\mathrm{T}}^{m|(k\wedge n)}_{\cdot|\cdot} = \widehat{\mathrm{T}}^{m|(k\wedge n)}_{\cdot|\cdot}\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\cdot|\cdot} = \mathrm{M}_{\mathbb{1}^m}\mathbb{1}_{\Omega_{m,k\wedge n}}$. □

§08|02.06 **Definition**. Under Assumption §08|02.02 let $(\widehat{g}_\cdot, \widehat{\mathrm{T}}_{\cdot|\cdot}) \sim \mathbb{P}^{n,k}_{\theta|\mathrm{T}}$ be noisy versions of $g_\cdot \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\cdot|\cdot})$ and $\mathrm{T}_{\cdot|\cdot} \in \mathbb{L}^{\mathbb{R}}(\ell_2)$ (or $\mathrm{T}_{\cdot|\cdot} \in \mathbb{L}^{\geqslant}(\ell_2)$). For each $m \in \mathbb{N}$ we call $\widehat{\theta}^{m}_\cdot = \widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\cdot|\cdot}\widehat{g}_\cdot = \widehat{\mathrm{T}}^{m|\dagger}_{\cdot|\cdot}\mathbb{1}_{\Omega_{m,k\wedge n}}\widehat{g}_\cdot\mathbb{1}^m_\cdot = \widehat{\mathrm{T}}^{m|\dagger}_{\cdot|\cdot}\mathbb{1}_{\Omega_{m,k\wedge n}}\widehat{g}^m_\cdot$ (generalised) thresholded *Galerkin estimator (tGE)* of $\theta_\cdot = \mathrm{T}^{\dagger}_{\cdot|\cdot}g_\cdot \in \ell_2$. □

§08|02.07 **Remark.** Under Assumption §08|02.02 we have $\dot{\varepsilon}_{\centerdot}\mathbb{1}^m_{\centerdot} \in \ell_\infty$ $\mathbb{P}^n_{\theta|\mathrm{T}}$-a.s. and $\widehat{\mathrm{T}}^m_{\centerdot|\centerdot} \in \mathbb{L}(\ell_2)$ with $\mathrm{ran}(\widehat{\mathrm{T}}^m_{\centerdot|\centerdot}) \subseteq \ell_2\mathbb{1}^m_{\centerdot}$ $\mathbb{P}^k_{\mathrm{T}}$-a.s. for each $m \in \mathbb{N}$. Consequently, $\mathrm{ran}(\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\centerdot|\centerdot}) \subseteq \ell_2\mathbb{1}^m_{\centerdot}$ $\mathbb{P}^k_{\mathrm{T}}$-a.s., and $\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\centerdot|\centerdot}\dot{\varepsilon}_{\centerdot} \in \ell_2\mathbb{1}^m_{\centerdot}$ $\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}$-a.s., and hence

$$\widehat{\theta}^m_{\centerdot} = \widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\centerdot|\centerdot}\widehat{g}_{\centerdot} = n^{-1/2}\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\centerdot|\centerdot}\dot{\varepsilon}_{\centerdot} + \widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\centerdot|\centerdot} g_{\centerdot} \in \ell_2\mathbb{1}^m_{\centerdot} \quad \mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}\text{-a.s.}.$$

Let us recall that the (generalised) Galerkin solution $\theta^m_{\centerdot} \in \ell_2\mathbb{1}^m_{\centerdot}$ does generally not correspond to the orthogonal projection $\mathbb{1}^{m|\perp}_{\centerdot}\theta_{\centerdot} = (\mathbb{1}_{\centerdot} - \mathbb{1}^m_{\centerdot})\theta_{\centerdot}$. Moreover, the approximation error $\sup\{\|\theta^m_{\centerdot} - \theta_{\centerdot}\|_{\ell_2} : m \geqslant n\}$ does generally not converge to zero as $n \to \infty$ (compare Remark §05|01.05). Here and subsequently, we will restrict ourselves to classes of solutions and operators which ensure the convergence. Obviously, this is a minimal regularity condition for us if we aim to estimate the Galerkin solution. □

§08|02|01 Examples

§08|02.08 **GniSM with noisy operator** (§02|02.06 continued). Let Assumption §08|00.02 be satisfied where $\mathrm{T}_{\centerdot|\centerdot} \in \mathbb{T} \subseteq \mathbb{L}^{\geqslant}_{\centerdot}(\ell_2)$ or $\mathrm{T}_{\centerdot|\centerdot} \in \mathbb{T} \subseteq \mathbb{L}^{\mathrm{R}}_{\centerdot}(\ell_2)$ is *not known* anymore. We illustrate the (generalised) tGE in a Gaussian non-diagonal inverse sequence model (GniSM) with noisy operator as in §02|02.06. Here the observable stochastic processes $\widehat{\mathrm{T}}_{\centerdot|\centerdot} = \mathrm{T}_{\centerdot|\centerdot} + k^{-1/2}\dot{\mathrm{W}}_{\centerdot|\centerdot} \sim \mathrm{N}^k_{\mathrm{T}}$ and $\widehat{g}_{\centerdot} = g_{\centerdot} + n^{-1/2}\dot{\mathrm{B}}_{\centerdot} \sim \mathrm{N}^n_{\theta|\mathrm{T}}$ are noisy version of $\mathrm{T}_{\centerdot|\centerdot} \in \mathbb{T}$ and $g_{\centerdot} = \mathrm{T}_{\centerdot|\centerdot}\theta_{\centerdot} \in \mathrm{dom}(\mathrm{T}^\dagger_{\centerdot|\centerdot}) \subseteq \ell_2$ with $\theta_{\centerdot} \in \Theta \subseteq \ell_2$, respectively, where $\dot{\mathrm{W}}_{\centerdot|\centerdot} := (\dot{\mathrm{W}}_{j|j_\circ})_{j,j_\circ\in\mathbb{N}} \sim \mathrm{N}^{\otimes\mathbb{N}^2}_{(0,1)}$ and $\dot{\mathrm{B}}_{\centerdot} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$ are *independent*. Consequently, $(\widehat{g}_{\centerdot}, \widehat{\mathrm{T}}_{\centerdot|\centerdot})$ admits a joint $\mathrm{N}^{n\otimes k}_{\theta|\mathrm{T}} = \mathrm{N}^n_{\theta|\mathrm{T}} \otimes \mathrm{N}^k_{\mathrm{T}}$-distribution belonging to the family $\mathrm{N}^{n\otimes k}_{\Theta\times\mathbb{T}} := (\mathrm{N}^n_{\theta|\mathrm{T}} \otimes \mathrm{N}^k_{\mathrm{T}})_{\theta_{\centerdot}\in\Theta,\mathrm{T}_{\centerdot|\centerdot}\in\mathbb{T}}$. Summarising the observations satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}^3}, \mathscr{B}^{\otimes\mathbb{N}^3}, \mathrm{N}^{n\otimes k}_{\Theta\times\mathbb{T}})$ where $\Theta \subseteq \ell_2$ and $\mathbb{T} \subseteq \mathbb{L}^{\geqslant}_{\centerdot}(\ell_2)$ or $\mathbb{T} \subseteq \mathbb{L}^{\mathrm{R}}_{\centerdot}(\ell_2)$. □

§08|02.09 **Property** (GniSM with noisy operator §08|02.08 continued). *Let* $\dot{\mathrm{W}}_{\centerdot|\centerdot} \sim \mathrm{N}^{\otimes\mathbb{N}^2}_{(0,1)}$ *and* $\dot{\mathrm{B}}_{\centerdot} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$ *be* independent *as in Model §08|02.08. Then Assumption §08|02.02 is satisfied:*

(i) *Due to Property §07|01.04* $\dot{\mathrm{B}}_{\centerdot}$ *admits* $\mathrm{id}_{\ell_2} \in \mathbb{L}^{\geqslant}(\ell_2)$ *as covariance operator with* $\|\mathrm{id}_{\ell_2}\|_{\mathbb{L}(\ell_2)} = 1$, *i.e.* (nSIP) *is fulfilled with* $\mathrm{v}_{\theta|\mathrm{T}} = 1$. *For all* $h_{\centerdot} \in \ell_2$ *we have* $\|h_{\centerdot}\|^2_{\ell_2} = \|h_{\centerdot}\|^2_{\mathrm{id}_{\ell_2}} = \langle \mathrm{id}_{\ell_2}h_{\centerdot}, h_{\centerdot}\rangle_{\ell_2}$.

(ii) *For all* $m \in \mathbb{N}$ *and* $a_{\centerdot}, b_{\centerdot} \in \ell_2$ *we have* $\dot{\mathrm{W}}^m_{\centerdot|\centerdot}a_{\centerdot} \sim \|a^m_{\centerdot}\|_{\ell_2}\dot{\mathrm{B}}^m_{\centerdot}$, $\langle b_{\centerdot}, \dot{\mathrm{B}}^m_{\centerdot}\rangle_{\ell_2} \sim \mathrm{N}_{(0,\|b^m_{\centerdot}\|^2_{\ell_2})}$, *and hence*

$$\mathrm{N}^{\otimes\mathbb{N}^2}_{(0,1)}\big(|\langle b_{\centerdot}, \dot{\mathrm{W}}^m_{\centerdot|\centerdot}a_{\centerdot}\rangle_{\ell_2}|^2\big) = \|a^m_{\centerdot}\|^2_{\ell_2}\|b^m_{\centerdot}\|^2_{\ell_2} \leqslant \|a_{\centerdot}\|^2_{\ell_2}\|b_{\centerdot}\|^2_{\ell_2},$$

i.e. (nSIPnO1) *is satisfied with* $\mathrm{v}_{\mathrm{T}} = 1$.

(iii) *For any* $l \in \mathbb{N}$ *setting* $\mathrm{K}^{2l}_{2l} := \prod_{j\in[\![l]\!]}(2j-1) =: (2l-1)!!$ *we have* $\mathrm{K}^2_{2l} \geqslant 1$ *and* $1 \vee \|\mathrm{v}^{\mathrm{T}|(l)}_{\centerdot|\centerdot}\|_{\ell_\infty(\mathbb{N}^2)} = \mathrm{K}^{2l}_{2l}$, *i.e.* (nSIPnO2) *is satisfied with* $\mathrm{K}_{\mathrm{T}} = \mathrm{K}_{2l}$. □

§08|02.10 **niSM with noisy operator** (§02|02.05 continued). Let Assumption §08|00.02 be satisfied where $\mathrm{T}_{\centerdot|\centerdot} \in \mathbb{T} \subseteq \mathbb{L}^{\geqslant}_{\centerdot}(\ell_2)$ or $\mathrm{T}_{\centerdot|\centerdot} \in \mathbb{T} \subseteq \mathbb{L}^{\mathrm{R}}_{\centerdot}(\ell_2)$ is *not known* anymore. We illustrate the (generalised) GE in a Non-diagonal inverse sequence model (niSM) with noisy operator as in §02|02.05. Here the observable stochastic process $\widehat{\mathrm{T}}_{\centerdot|\centerdot} = \mathrm{T}_{\centerdot|\centerdot} + k^{-1/2}\dot{\eta}_{\centerdot|\centerdot} \sim \mathbb{P}^k_{\mathrm{T}}$ and $\widehat{g}_{\centerdot} = g_{\centerdot} + n^{-1/2}\dot{\varepsilon}_{\centerdot} \sim \mathbb{P}^n_{\theta|\mathrm{T}}$ is a noisy version of $\mathrm{T}_{\centerdot|\centerdot} \in \mathbb{T}$ and $g_{\centerdot} = \mathrm{T}_{\centerdot|\centerdot}\theta_{\centerdot} \in \mathrm{dom}(\mathrm{T}^\dagger_{\centerdot|\centerdot}) \subseteq \ell_2$ with $\theta_{\centerdot} \in \Theta \subseteq \ell_2$, respectively, where $\dot{\varepsilon}_{\centerdot} \sim \otimes_{j\in\mathbb{N}}\mathbb{P}^{\dot{\varepsilon}_j}$ and $\dot{\eta}_{\centerdot|\centerdot} \sim \otimes_{j,j_\circ\in\mathbb{N}}\mathbb{P}^{\dot{\eta}_{j|j_\circ}}$ are *independent*. In addition, let $\dot{\varepsilon}_{\centerdot}$ satisfy (iSM1) of Model §07|01.06 for $\sigma_{\centerdot} \in \Sigma \subseteq \mathbb{R}^{\mathbb{N}}_{\geqslant 0} \cap \ell_\infty$ and let $\dot{\eta}_{\centerdot|\centerdot}$ fulfill

(niSMnO1) for $\xi_{\centerdot|\centerdot} \in \Xi \subseteq (\mathbb{R}_{>0})^{\mathbb{N}^2} \cap \ell_\infty(\mathbb{N}^2)$ we have $\dot{\eta}_{j|j_\circ} \sim \mathrm{P}_{(0,\xi^2_{j|j_\circ})} \in \mathscr{W}_2(\mathscr{B})$, for all $j, j_\circ \in \mathbb{N}$,

(niSMnO2) for $l \in \mathbb{N}$ and $\xi^{(2l)}_{\centerdot|\centerdot} \in \Xi^{2l} \subseteq (\mathbb{R}_{>0})^{\mathbb{N}^2} \cap \ell_\infty(\mathbb{N}^2)$ we have $\xi^{(2l)}_{\centerdot|\centerdot} := (\xi^{(2l)}_{j|j_\circ} := \mathrm{P}_{(0,\xi^2_{j|j_\circ})}(\dot{\eta}^{2l}_{j|j_\circ}))_{j,j_\circ\in\mathbb{N}}$.

Under **(iSM1)** $\widehat{g}_{\bullet}$ admits a $\mathrm{P}^n_{\theta|\mathrm{T}|\sigma}$-distribution belonging to the family $\mathrm{P}^n_{\Theta\times\mathbb{T}\times\Sigma} := (\mathrm{P}^n_{\theta|\mathrm{T}|\sigma})_{\theta_{\bullet}\in\Theta,\mathrm{T}_{\bullet|\bullet}\in\mathbb{T},\sigma_{\bullet}\in\Sigma}$ and under **(niSMnO1)**, **(niSMnO2)** $\widehat{\mathrm{T}}_{\bullet|\bullet}$ admits a $\mathrm{P}^k_{\mathrm{T}|\xi|\xi^{(2l)}}$-distribution belonging to the family $\mathrm{P}^k_{\mathbb{T}\times\Xi\times\Xi^{2l}} := (\mathrm{P}^k_{\mathrm{T}|\xi|\xi^{(2l)}})_{\mathrm{T}_{\bullet|\bullet}\in\mathbb{T},\xi_{\bullet|\bullet}\in\Xi,\xi^{(2l)}_{\bullet|\bullet}\in\Xi^{2l}}$. Summarising $(\widehat{g}_{\bullet},\widehat{\mathrm{T}}_{\bullet|\bullet})$ admits a joint $\mathrm{P}^{n\otimes k}_{\theta|\mathrm{T}|\sigma|\xi|\xi^{(2l)}} = \mathrm{P}^n_{\theta|\mathrm{T}|\sigma}\otimes\mathrm{P}^k_{\mathrm{T}|\xi|\xi^{(2l)}}$ distribution belonging to the family $\mathrm{P}^{n\otimes k}_{\Theta\times\mathbb{T}\times\Sigma\times\Xi\times\Xi^{2l}} := (\mathrm{P}^n_{\theta|\mathrm{T}|\sigma}\otimes\mathrm{P}^k_{\mathrm{T}|\xi|\xi^{(2l)}})_{\theta_{\bullet}\in\Theta,\mathrm{T}_{\bullet|\bullet}\in\mathbb{T},\sigma_{\bullet}\in\Sigma,\xi_{\bullet|\bullet}\in\Xi,\xi^{(2l)}_{\bullet|\bullet}\in\Xi^{2l}}$ and the observations satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}^3},\mathscr{B}^{\otimes\mathbb{N}^3},\mathrm{P}^{n\otimes k}_{\Theta\times\mathbb{T}\times\Sigma\times\Xi\times\Xi^{(2l)}})$ where $\Theta\subseteq\ell_2$ and $\mathbb{T}\subseteq\mathbb{L}^{\gtrless}_{\bullet}(\ell_2)$ or $\mathbb{T}\subseteq\mathbb{L}^{\mathrm{R}}_{\bullet}(\ell_2)$. □

§08|02.11 **Lemma** (niSM with noisy operator §08|01.06 continued). *Consider error processes $\dot{\eta}_{\bullet|\bullet}$ and $\dot{\varepsilon}_{\bullet}$ as in Model §08|02.10 satisfying* **(iSM1)**, **(niSMnO1)** *and* **(niSMnO2)**. *Then Assumption §08|02.02 is satisfied:*

(i) *Due to Property §07|01.07 (i) under* **(iSM1)**, *$\dot{\varepsilon}_{\bullet}$ admits $\Gamma_{\theta|\mathfrak{s}} = \mathrm{M}_{\sigma^2}\in\mathbb{L}^{\mathrm{M}}(\ell_2)\cap\mathbb{L}^{\mathrm{p}}(\ell_2)$ as covariance operator with $\|\mathrm{M}_{\sigma^2}\|_{\mathbb{L}(\ell_2)} = \|\sigma^2_{\bullet}\|_{\ell_\infty}\leqslant\|\sigma^2_{\bullet}\|_{\ell_\infty}\vee 1 =: \mathrm{v}_\sigma\in\mathbb{R}_{\geqslant 1}$, i.e.* **(nSIP)** *is fulfilled with $\mathrm{v}_{\theta|\mathrm{T}} = \mathrm{v}_\sigma$. For all $h_{\bullet}\in\ell_2$ we have $\|h_{\bullet}\|^2_{\mathrm{M}_{\sigma^2}} = \langle\mathrm{M}_{\sigma^2}h_{\bullet},h_{\bullet}\rangle_{\ell_2}\leqslant\mathrm{v}_\sigma\|h_{\bullet}\|^2_{\ell_2}$.*

(ii) *Under* **(niSMnO1)** *for all $m\in\mathbb{N}$ and $a_{\bullet},b_{\bullet}\in\ell_2$ with $1\vee\|\xi^2_{\bullet|\bullet}\|_{\ell_\infty(\mathbb{N}^2)} =: \mathrm{v}_\xi\in\mathbb{R}_{\geqslant 1}$ we have*

$$\mathrm{P}^k_{\mathrm{T}|\xi|\xi^{(2l)}}\big(|\langle b_{\bullet},\dot{\eta}^{m}_{\bullet|\bullet}a_{\bullet}\rangle_{\ell_2}|^2\big)\leqslant\mathrm{v}_\xi\|a_{\bullet}\|^2_{\ell_2}\|b_{\bullet}\|^2_{\ell_2}$$

i.e. **(nSIPnO1)** *is satisfied with $\mathrm{v}_{\mathrm{T}} = \mathrm{v}_\xi$.*

(iii) *Under* **(niSMnO2)** *setting $1\vee\|\xi^{(2l)}_{\bullet|\bullet}\|_{\ell_\infty(\mathbb{N}^2)} =: \mathrm{K}^{2l}_{\xi^{(2l)}}\in\mathbb{R}_{\geqslant 1}$ we have $\mathrm{K}^2_{\xi^{(2l)}}\geqslant\mathrm{v}_\xi$ and $1\vee\|\mathrm{v}^{\mathrm{T}|(l)}_{\bullet|\bullet}\|_{\ell_\infty(\mathbb{N}^2)}\leqslant\mathrm{K}^{2l}_{\xi^{(2l)}}$, i.e.* **(nSIPnO2)** *is satisfied with $\mathrm{K}_{\mathrm{T}} = \mathrm{K}_{\xi^{(2l)}}$.*

§08|02.12 **Proof** of Lemma §08|02.11. Given in the lecture. □

§08|02.13 **nieMM with noisy operator** (**§02|02.04 continued**). Let Assumption §08|00.02 be satisfied where $\mathrm{T}_{\bullet|\bullet}\in\mathbb{T}\subseteq\mathbb{L}^{\gtrless}_{\bullet}(\ell_2)$ or $\mathrm{T}_{\bullet|\bullet}\in\mathbb{T}\subseteq\mathbb{L}^{\mathrm{R}}_{\bullet}(\ell_2)$ is *not known* anymore. We illustrate the (generalised) tGE in a Non-diagonal inverse empirical mean model (nieMM) with noisy operator as in §02|02.04. Here the observable stochastic processes $\widehat{\mathrm{T}}_{\bullet|\bullet} = \mathrm{T}_{\bullet|\bullet}+k^{-1/2}\dot{\eta}_{\bullet|\bullet}\sim\mathrm{P}^k_{\mathrm{T}}$ and $\widehat{g}_{\bullet} = g_{\bullet}+n^{-1/2}\dot{\varepsilon}_{\bullet}$ are noisy version of $\mathrm{T}_{\bullet|\bullet}\in\mathbb{T}$ and $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet}\in\mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet})\subseteq\ell_2$ with $\theta_{\bullet}\in\Theta\subseteq\ell_2$, and *independent* error processes $\dot{\varepsilon}_{\bullet} = n^{1/2}(\widehat{\mathbb{P}}_n(\psi_{\bullet})-\mathbb{P}_{\theta|\mathrm{T}}(\psi_{\bullet}))\in\mathcal{M}(\mathscr{Z}^{\otimes n}\otimes 2^{\mathbb{N}})$ and $\dot{\eta}_{\bullet|\bullet} = k^{1/2}(\widehat{\mathbb{P}}_k(\varphi_{\bullet|\bullet})-\mathbb{P}_{\mathrm{T}}(\varphi_{\bullet|\bullet}))\in\mathcal{M}(\mathscr{Z}^{\otimes k}\otimes 2^{\mathbb{N}^2})$ satisfying Assumption §01|01.04 and Assumption §02|01.02. More precisely, on a measurable space $(\mathcal{Z},\mathscr{Z})$ for each $\theta_{\bullet}\in\Theta\subseteq\ell_2$ and $\mathrm{T}_{\bullet|\bullet}\in\mathbb{T}$ there are probability measures $\mathbb{P}_{\theta|\mathrm{T}},\mathbb{P}_{\mathrm{T}}\in\mathscr{W}(\mathscr{Z})$. Similar to Model §02|02.04 consider stochastic processes $\psi_{\bullet}\in\mathcal{M}(\mathscr{Z}\otimes 2^{\mathbb{N}})$ and $\varphi_{\bullet|\bullet}\in\mathcal{M}(\mathscr{Z}\otimes 2^{\mathbb{N}^2})$. In addition for all $\theta_{\bullet}\in\Theta$ and $\mathrm{T}_{\bullet|\bullet}\in\mathbb{T}$ the process $\psi_{\bullet}\in\mathcal{M}(\mathscr{Z}\otimes 2^{\mathbb{N}})$ satisfies **(nieMM1)** and **(nieMM2)** of Model §08|01.08 for $\mathrm{v}_{\theta|\mathrm{T}|\psi}\in\mathbb{R}_{\geqslant 1}$ and the process $\varphi_{\bullet|\bullet}\in\mathcal{M}(\mathscr{Z}\otimes 2^{\mathbb{N}^2})$ fulfils

(nieMMnO1) $\varphi_{j|j_\circ}\in\mathcal{L}_1(\mathbb{P}_{\mathrm{T}}) := \mathcal{L}_1(\mathcal{Z},\mathscr{Z},\mathbb{P}_{\mathrm{T}})$ for all $j,j_\circ\in\mathbb{N}$ and $\mathbb{P}_{\mathrm{T}}(\varphi_{\bullet|\bullet}) = \mathrm{T}_{\bullet|\bullet}$,

(nieMMnO2) there is $\mathrm{v}_{\mathrm{T}|\varphi}\in\mathbb{R}_{\geqslant 1}$ such that $\varphi_{\bullet|\bullet}$ for all $m\in\mathbb{N}$ and $a_{\bullet},b_{\bullet}\in\ell_2$ satisfies

$$\mathbb{P}_{\mathrm{T}}\big(|\langle b_{\bullet},\varphi^{m}_{\bullet|\bullet}a_{\bullet}\rangle_{\ell_2}|^2\big)\leqslant\mathrm{v}_{\mathrm{T}|\psi}\|a_{\bullet}\|^2_{\ell_2}\|b_{\bullet}\|^2_{\ell_2}.$$

(nieMMnO3) there is $l\in\mathbb{N}$ and $\mathrm{K}^2_{\mathrm{T}|\varphi}\in\mathbb{R}_{\geqslant\mathrm{v}_{\mathrm{T}|\varphi}}$ such that $\varphi_{\bullet|\bullet}$ satisfies

$$\mathbb{P}_{\mathrm{T}}(\varphi^{2l}_{\bullet|\bullet}) := (\mathbb{P}_{\mathrm{T}}(\varphi^{2l}_{j|j_\circ}))_{j,j_\circ\in\mathbb{N}}\in\ell_\infty(\mathbb{N}^2),$$

and $1\vee\|\mathbb{P}_{\mathrm{T}}(|\varphi_{\bullet|\bullet}-\mathbb{P}_{\mathrm{T}}\varphi_{\bullet|\bullet}|^{2l})\|_{\ell_\infty(\mathbb{N}^2)}\leqslant\mathrm{K}^{2l}_{\mathrm{T}|\varphi}$.

We consider a statistical product experiment $(\mathcal{Z}^{n+k},\mathscr{Z}^{\otimes(n+k)},\mathbb{P}^{n\otimes k}_{\Theta\times\mathbb{T}} = (\mathbb{P}^{\otimes n}_{\theta|\mathrm{T}}\otimes\mathbb{P}^{\otimes k}_{\mathrm{T}})_{\theta_{\bullet}\in\Theta,\mathrm{T}_{\bullet|\bullet}\in\mathbb{T}})$ as in an Empirical mean function §01|01.10 where $\Theta\subseteq\ell_2$ and $\mathbb{T}\subseteq\mathbb{L}^{\gtrless}_{\bullet}(\ell_2)$ or $\mathbb{T}\subseteq\mathbb{L}^{\mathrm{R}}_{\bullet}(\ell_2)$. □

§08|02.14 **Lemma** (nieMM with noisy operator §08|01.08 continued). *Consider error processes* $\dot{\eta}_{\cdot|\cdot}$ *and* $\dot{\varepsilon}_{\cdot}$ *as in Model §08|02.13 where* $\psi_{\cdot} \in \mathcal{M}(\mathscr{X} \otimes 2^{\mathbb{N}})$ *satisfies* (nieMM1) *and* (nieMM2) *and* $\varphi_{\cdot|\cdot} \in \mathcal{M}(\mathscr{X} \otimes 2^{\mathbb{N}^2})$ *fulfils* (nieMMnO1)-(nieMMnO3). *Then Assumption §08|02.02 is satisfied:*

(i) *Due to Property §07|01.09 (i) under* (nieMM1) *and* (nieMM2) $\dot{\varepsilon}_{\cdot}$ *admits a covariance operator* $\Gamma_{\theta|\mathrm{T}} \in \mathbb{L}^{\geqslant}(\ell_2)$ *satisfying* $\|\Gamma_{\theta|\mathrm{T}}\|_{\mathbb{L}(\ell_2)} \leqslant \mathrm{v}_{\theta|\mathrm{T}|\psi}$. *i.e.* (nSIP) *is fulfilled with* $\mathrm{v}_{\theta|\mathrm{T}} = \mathrm{v}_{\theta|\mathrm{T}|\psi}$. *For all* $h_{\cdot} \in \ell_2$ *we have* $\|h_{\cdot}\|^2_{\Gamma_{\theta|\mathrm{T}}} = \langle \Gamma_{\theta|\mathrm{T}} h_{\cdot}, h_{\cdot}\rangle_{\ell_2} \leqslant \mathrm{v}_{\theta|\mathrm{T}|\psi}\|h_{\cdot}\|^2_{\ell_2}$.

(ii) *Under* (nieMMnO1) *and* (nieMMnO2) *for all* $m \in \mathbb{N}$ *and* $a_{\cdot}, b_{\cdot} \in \ell_2$ *we have*

$$\mathbb{P}^{\otimes k}_{\mathrm{T}}\big(|\langle b_{\cdot}, \dot{\eta}^m_{\cdot|\cdot} a_{\cdot}\rangle_{\ell_2}|^2\big) \leqslant \mathrm{v}_{\mathrm{T}|\varphi}\|a_{\cdot}\|^2_{\ell_2}\|b_{\cdot}\|^2_{\ell_2}$$

i.e. (nSIPnO1) *is satisfied with* $\mathrm{v}_{\mathrm{T}} = \mathrm{v}_{\mathrm{T}|\varphi}$.

(iii) *Under* (nieMMnO1) *and* (nieMMnO3) *there exists a constant* $\mathrm{C}_{2l} \in \mathbb{R}_{\geqslant 1}$ *depending on* $l \in \mathbb{N}$ *only such that we have* $1 \vee \|\mathrm{v}^{\mathrm{T}|(l)}_{\cdot|\cdot}\|_{\ell_\infty(\mathbb{N}^2)} \leqslant \mathrm{C}_{2l}\mathrm{K}^{2l}_{\mathrm{T}|\varphi}$, *i.e.* (nSIPnO2) *is satisfied with* $\mathrm{K}_{\mathrm{T}} = \mathrm{C}^{1/2l}_{2l}\mathrm{K}_{\mathrm{T}|\varphi} \in \mathbb{R}_{\geqslant 1}$.

§08|02.15 **Proof** of Lemma §08|02.14. Given in the lecture. □

§08|02|02 Global and maximal global $\mathfrak{v}$-risk

We measure first the accuracy of the thresholded (generalised) GE $\widehat{\theta}^m_{\cdot} := \widehat{\mathrm{T}}^{m|(n\wedge k)|\dagger}_{\cdot|\cdot}\widehat{g}_{\cdot}$ of the (generalised) Galerkin solution $\theta^m_{\cdot} = \mathrm{T}^{m|\dagger}_{\cdot|\cdot} g_{\cdot} \in \ell_2 \mathbb{1}^m_{\cdot}$ with $g_{\cdot} = \mathrm{T}_{\cdot|\cdot}\theta_{\cdot} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\cdot|\cdot})$ and $\mathrm{T}_{\cdot|\cdot} \in \mathbb{L}^{\mathrm{R}}(\ell_2)$ by the mean of its global $\mathfrak{v}$-error introduced in §05|01|01 and §05|02|01, i.e. its $\mathfrak{v}$-risk.

§08|02.16 **Reminder**. If $\mathfrak{v}_{\cdot} \in (\mathbb{R}_{\setminus 0})^{\mathbb{N}}$ then we have $\mathfrak{v}^2_{\cdot}\mathbb{1}^m_{\cdot} \in \ell_\infty$ and $\ell_2\mathbb{1}^m_{\cdot} \subseteq \ell_2(\mathfrak{v}^2_{\cdot})$. Consequently, for each $\theta_{\cdot} \in \ell_2(\mathfrak{v}^2_{\cdot})$ the (generalised) Galerkin solution $\theta^m_{\cdot} = \mathrm{T}^{m|\dagger}_{\cdot|\cdot} g_{\cdot} \in \ell_2\mathbb{1}^m_{\cdot}$ satisfies $\theta^m_{\cdot} \in \ell_2(\mathfrak{v}^2_{\cdot})$ too. If in addition $\mathrm{C}_{\mathrm{T}} := \sup\big\{\|\mathrm{M}_{\mathfrak{v}}\mathrm{T}^{m|\dagger}_{\cdot|\cdot}\mathrm{T}_{\cdot|\cdot}\mathrm{M}_{\mathbb{1}^{m|\perp}_{\cdot}}\|_{\mathbb{L}(\ell_2)} : m \in \mathbb{N}\big\} \in \mathbb{R}_{\geqslant 0}$ then $\|\theta^m_{\cdot} - \theta_{\cdot}\|_{\mathfrak{v}} \leqslant (1 + \mathrm{C}_{\mathrm{T}})\|\mathbb{1}^{m|\perp}_{\cdot}\theta_{\cdot}\|_{\ell_2}$ which implies $\sup\big\{\|\theta^j_{\cdot} - \theta_{\cdot}\|_{\mathfrak{v}} : j \in \mathbb{N}_{\geqslant m}\big\} = \mathrm{o}(1)$ as $m \to \infty$ (Property §05|01.24 and Property §05|02.08). □

§08|02.17 **Comment**. Under Assumption §08|02.02 we have $\dot{\varepsilon}_{\cdot}\mathbb{1}^m_{\cdot} \in \ell_\infty$ $\mathbb{P}^n_{\theta|\mathrm{T}}$-a.s. and $\widehat{\mathrm{T}}^m_{\cdot|\cdot} \in \mathbb{L}(\ell_2)$ with $\mathrm{ran}(\widehat{\mathrm{T}}^m_{\cdot|\cdot}) \subseteq \ell_2\mathbb{1}^m_{\cdot}$ $\mathbb{P}^k_{\mathrm{T}}$-a.s. for each $m \in \mathbb{N}$. Consequently, $\mathrm{ran}(\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\cdot|\cdot}) \subseteq \ell_2\mathbb{1}^m_{\cdot}$ $\mathbb{P}^k_{\mathrm{T}}$-a.s., and $\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\cdot|\cdot}\dot{\varepsilon}_{\cdot} \in \ell_2\mathbb{1}^m_{\cdot}$ $\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}$-a.s., and hence

$$\widehat{\theta}^m_{\cdot} = \widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\cdot|\cdot}\widehat{g}_{\cdot} = n^{-1/2}\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\cdot|\cdot}\dot{\varepsilon}_{\cdot} + \widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\cdot|\cdot} g_{\cdot} \in \ell_2\mathbb{1}^m_{\cdot} \subseteq \ell_2(\mathfrak{v}^2_{\cdot}) \quad \mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}\text{-a.s.}.$$ □

§08|02|02|01 Global $\mathfrak{v}$-risk

§08|02.18 **Assumption**. Let $\mathfrak{v}_{\cdot} \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ and $\theta_{\cdot} \in \ell_2(\mathfrak{v}^2_{\cdot})$ be satisfied. □

§08|02.19 **Definition**. Under Assumptions §08|02.02 and §08|02.18 the *global* $\mathfrak{v}$*-risk* of a (generalised) tGE $\widehat{\theta}^m_{\cdot} = \widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\cdot|\cdot}\widehat{g}_{\cdot} \in \ell_2\mathbb{1}^m_{\cdot} \subseteq \ell_2(\mathfrak{v}^2_{\cdot})$ $\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}$-a.s. satisfies

$$\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}\big(\|\widehat{\theta}^m_{\cdot} - \theta_{\cdot}\|^2_{\mathfrak{v}}\big) = \mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}\big(\|\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\cdot|\cdot}(\widehat{g}_{\cdot} - g_{\cdot})\|^2_{\mathfrak{v}}\big) + \mathbb{P}^k_{\mathrm{T}}\big(\|\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\cdot|\cdot} g_{\cdot} - \theta_{\cdot}\|^2_{\mathfrak{v}}\big) \qquad (08.17)$$

with $\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}\big(\|\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\cdot|\cdot}(\widehat{g}_{\cdot} - g_{\cdot})\|^2_{\mathfrak{v}}\big) = n^{-1}\mathbb{P}^k_{\mathrm{T}}\Big(\mathrm{tr}\big(\mathrm{M}_{\mathfrak{v}}\widehat{\mathrm{T}}^{m|\dagger}_{\cdot|\cdot}\Gamma_{\theta|\mathrm{T}}(\widehat{\mathrm{T}}^{m|\dagger}_{\cdot|\cdot})^{\star}\mathrm{M}_{\mathfrak{v}}\big)\mathbb{1}_{\Omega_{m,k\wedge n}}\Big)$ (see Property §08|01.15). □

§08|02.20 **Property**. *Under Assumption §08|02.02 we have*

$$\begin{aligned}\mathbb{P}^k_{\mathrm{T}}\big(\|\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\cdot|\cdot} g_{\cdot} - \theta_{\cdot}\|^2_{\mathfrak{v}}\big) &= \mathbb{P}^k_{\mathrm{T}}\big(\|\widehat{\mathrm{T}}^{m|\dagger}_{\cdot|\cdot}(\mathrm{T}^m_{\cdot|\cdot} - \widehat{\mathrm{T}}^m_{\cdot|\cdot})\theta^m_{\cdot} + (\theta^m_{\cdot} - \theta_{\cdot})\|^2_{\mathfrak{v}}\mathbb{1}_{\Omega_{m,k\wedge n}}\big) + \|\theta_{\cdot}\|^2_{\mathfrak{v}}\mathbb{P}^k_{\mathrm{T}}(\Omega^c_{m,k\wedge n})\\ &\leqslant 2\mathbb{P}^k_{\mathrm{T}}\big(\|\widehat{\mathrm{T}}^{m|\dagger}_{\cdot|\cdot}(\mathrm{T}^m_{\cdot|\cdot} - \widehat{\mathrm{T}}^m_{\cdot|\cdot})\theta^m_{\cdot}\|^2_{\mathfrak{v}}\mathbb{1}_{\Omega_{m,k\wedge n}}\big) + 2\|\theta^m_{\cdot} - \theta_{\cdot}\|^2_{\mathfrak{v}} + \|\theta_{\cdot}\|^2_{\mathfrak{v}}\mathbb{P}^k_{\mathrm{T}}(\Omega^c_{m,k\wedge n})\end{aligned}$$

(since $\widehat{\mathrm{T}}^{m|\dagger}_{\bullet|\bullet}\widehat{\mathrm{T}}^{m}_{\bullet|\bullet}\mathbb{1}_{\Omega_{m,k\wedge n}} = \widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{\mathrm{T}}^{m|(k\wedge n)}_{\bullet|\bullet} = \mathrm{M}_{\mathbb{1}^m_\bullet}\mathbb{1}_{\Omega_{m,k\wedge n}}$*).* □

§08|02.21 **Notation** (Reminder). Let $\mathrm{A} \in \mathbb{L}(\ell_2)$ be a *Hilbert-Schmidt operator*, $\mathrm{A} \in \mathbb{HS}(\ell_2)$ for short, where $\|\mathrm{A}\|^2_{\mathrm{HS}} := \mathrm{tr}(\mathrm{A}^\star\mathrm{A}) = \mathrm{tr}(\mathrm{A}\mathrm{A}^\star) \in \mathbb{R}_{\geqslant 0}$. If $\Gamma \in \mathbb{L}(\ell_2)$ then $\mathrm{tr}(\mathrm{A}^\star\Gamma\mathrm{A}) \leqslant \|\Gamma\|_{\mathbb{L}(\ell_2)}\mathrm{tr}(\mathrm{A}^\star\mathrm{A}) = \|\Gamma\|_{\mathbb{L}(\ell_2)}\|\mathrm{A}\|^2_{\mathrm{HS}}$. For arbitrary $\mathrm{A} \in \mathbb{L}(\ell_2)$ we have $\mathrm{M}_{\mathfrak{v}}\mathrm{A}^m = \mathrm{M}^m_{\mathfrak{v}}\mathrm{A}^m \in \mathbb{HS}(\ell_2)$. □

§08|02.22 **Notation**. For each $m \in \mathbb{N}$ and $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\mathbb{R}}(\ell_2)$ we consider the observable event and its complement

$$\Omega_{m,k\wedge n} := \{\|[\mathrm{M}_{\mathfrak{v}}]_{\underline{m}}[\widehat{\mathrm{T}}_{\bullet|\bullet}]^{-1}_{\underline{m}}\|^2_{\mathrm{HS}} \leqslant k\wedge n\} \quad\text{and}\quad \Omega^c_{m,k\wedge n} := \{\|[\mathrm{M}_{\mathfrak{v}}]_{\underline{m}}[\widehat{\mathrm{T}}_{\bullet|\bullet}]^{-1}_{\underline{m}}\|^2_{\mathrm{HS}} > k\wedge n\}. \quad (08.18)$$

On the event $\Omega_{m,k\wedge n}$ the random matrix $[\widehat{\mathrm{T}}_{\bullet|\bullet}]_{\underline{m}} \in \mathbb{R}^{(m,m)}$ is regular with inverse $[\widehat{\mathrm{T}}_{\bullet|\bullet}]^{-1}_{\underline{m}} \in \mathbb{R}^{(m,m)}$. Moreover, setting $\mathrm{A}^m_{\bullet|\bullet} := \dot{\boldsymbol{\eta}}^m_{\bullet|\bullet}\mathrm{T}^{m|\dagger}_{\bullet|\bullet}$ we introduce an unobserved event and its complement

$$\mho_{m,k} := \{4m\|\mathrm{A}^m_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)} \leqslant k\} \quad\text{and}\quad \mho^c_{m,k} := \{4m\|\mathrm{A}^m_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)} > k\}. \quad (08.19)$$

Note that $\mathbb{1}_{\mho_{m,k}} = \mathbb{1}_{\{4m\|\mathrm{A}^m_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)} \leqslant k\}}$ denotes an unobserved elementary random variable. □

§08|02.23 **Lemma**. *Under Assumptions §08|02.02 and §08|02.18 for all $m,k,n \in \mathbb{N}$ we have*

(i) *if* $4\|\mathrm{M}_{\mathfrak{v}}\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\|^2_{\mathrm{HS}} \leqslant k\wedge n$ *then* $\mho_{m,k} \subseteq \Omega_{m,k\wedge n}$,

(ii) $\mathbb{P}^k_{\mathrm{T}}\big(\mathrm{tr}(\mathrm{M}_{\mathfrak{v}}\widehat{\mathrm{T}}^{m|\dagger}_{\bullet|\bullet}\Gamma_{\theta|\mathrm{T}}(\widehat{\mathrm{T}}^{m|\dagger}_{\bullet|\bullet})^\star\mathrm{M}_{\mathfrak{v}})\mathbb{1}_{\Omega_{m,k\wedge n}}\big) \leqslant \mathbb{v}_{\theta|\mathrm{T}}(4\|\mathrm{M}_{\mathfrak{v}}\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\|^2_{\mathrm{HS}} + (k\wedge n)\mathbb{P}^k_{\mathrm{T}}(\mho^c_{m,k}))$*, and*

(iii) $\mathbb{P}^k_{\mathrm{T}}(\|\widehat{\mathrm{T}}^{m|\dagger}_{\bullet|\bullet}(\mathrm{T}^m_{\bullet|\bullet} - \widehat{\mathrm{T}}^m_{\bullet|\bullet})\theta^m_\bullet\|^2_{\mathfrak{v}}\mathbb{1}_{\Omega_{m,k\wedge n}}) \leqslant 4k^{-1}\mathbb{v}_{\mathrm{T}}\|\mathrm{M}_{\mathfrak{v}}\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\|^2_{\mathrm{HS}}\|\theta^m_\bullet\|^2_{\ell_2} + \mathbb{P}^k_{\mathrm{T}}(\|\dot{\boldsymbol{\eta}}^m_{\bullet|\bullet}\theta^m_\bullet\|^2_{\ell_2}\mathbb{1}_{\mho^c_{m,k}})$.

with $\Omega_{m,k\wedge n}$ and $\mho_{m,k}$ as in (08.18) and (08.19), respectively.

§08|02.24 **Proof** of Lemma §08|02.23. Given in the lecture. □

§08|02.25 **Proposition** (Upper bound). *Under Assumptions §08|02.02 and §08|02.18 for all $n,m \in \mathbb{N}$ with (generalised) Galerkin solution $\theta^m_\bullet = \mathrm{T}^{m|\dagger}_{\bullet|\bullet}g_\bullet \in {}_{\ell_2}\mathbb{1}^m_\bullet$ setting similar to (08.03)*

$$\mathrm{R}^m_{n\wedge k}(\theta_\bullet,\mathrm{T}_{\bullet|\bullet},\mathfrak{v}_\bullet) := \|\theta^m_\bullet - \theta_\bullet\|^2_{\mathfrak{v}} + (n\wedge k)^{-1}\|\mathrm{M}_{\mathfrak{v}}\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\|^2_{\mathrm{HS}},$$

$$m^\circ_{n\wedge k} := \arg\min\big\{\mathrm{R}^m_{n\wedge k}(\theta_\bullet,\mathrm{T}_{\bullet|\bullet},\mathfrak{v}_\bullet) : m\in\mathbb{N}\big\} \quad \textit{and}$$

$$\mathrm{R}^\circ_{n\wedge k}(\theta_\bullet,\mathrm{T}_{\bullet|\bullet},\mathfrak{v}_\bullet) := \mathrm{R}^{m^\circ_{n\wedge k}}_{n\wedge k}(\theta_\bullet,\mathrm{T}_{\bullet|\bullet},\mathfrak{v}_\bullet) = \min\big\{\mathrm{R}^m_{n\wedge k}(\theta_\bullet,\mathrm{T}_{\bullet|\bullet},\mathfrak{v}_\bullet) : m\in\mathbb{N}\big\} \quad (08.20)$$

for all $m \in \mathbb{N}$ the (generalised) tGE $\widehat{\theta}^m_\bullet = \widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{g}_\bullet \in {}_{\ell_2}\mathbb{1}^m_\bullet$ $\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}$-a.s. satisfies

$$\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}(\|\widehat{\theta}^m_\bullet - \theta_\bullet\|^2_{\mathfrak{v}}) \leqslant (4\mathbb{v}_{\theta|\mathrm{T}} + 8\mathbb{v}_{\mathrm{T}}\|\theta^m_\bullet\|^2_{\ell_2})\,\mathrm{R}^m_{n\wedge k}(\theta_\bullet,\mathrm{T}_{\bullet|\bullet},\mathfrak{v}_\bullet)$$
$$+ \mathbb{v}_{\theta|\mathrm{T}}\mathbb{P}^k_{\mathrm{T}}(\mho^c_{m,k}) + 2\mathbb{P}^k_{\mathrm{T}}(\|\dot{\boldsymbol{\eta}}^m_{\bullet|\bullet}\theta^m_\bullet\|^2_{\ell_2}\mathbb{1}_{\mho^c_{m,k}}) + \|\theta_\bullet\|^2_{\mathfrak{v}}\mathbb{P}^k_{\mathrm{T}}(\Omega^c_{m,k\wedge n}) \quad (08.21)$$

with $\Omega_{m,k\wedge n}$ and $\mho_{m,k}$ as in (08.18) and (08.19), respectively.

§08|02.26 **Proof** of Proposition §08|02.25. Given in the lecture. □

§08|02.27 **Corollary**. *Under the assumptions of Proposition §08|02.25 the (infeasible, generalised) tGE $\widehat{\theta}^{m^\circ_{n\wedge k}}_\bullet = \widehat{\mathrm{T}}^{m^\circ_{n\wedge k}|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{g}_\bullet \in {}_{\ell_2}\mathbb{1}^{m^\circ_{n\wedge k}}_\bullet \subseteq \ell_2(\mathfrak{v}^2_\bullet)$ $\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}$-a.s. with $\Omega_{m^\circ_{n\wedge k},k\wedge n}$ as in (08.18) and (infeasible) dimension $m^\circ_{n\wedge k}$ as in (08.20) for each $k,n \in \mathbb{N}$ with $\mathrm{R}^\circ_{n\wedge k}(\theta_\bullet,\mathrm{T}_{\bullet|\bullet},\mathfrak{v}_\bullet) \leqslant 1/4$ satisfies*

$$\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}(\|\widehat{\theta}^{m^\circ_{n\wedge k}}_\bullet - \theta_\bullet\|^2_{\mathfrak{v}}) \leqslant (4\mathbb{v}_{\theta|\mathrm{T}} + 8\mathbb{v}_{\mathrm{T}}\|\theta^{m^\circ_{n\wedge k}}_\bullet\|^2_{\ell_2})\,\mathrm{R}^\circ_{n\wedge k}(\theta_\bullet,\mathrm{T}_{\bullet|\bullet},\mathfrak{v}_\bullet)$$
$$+ 2^{2l}\mathrm{K}^{2l}_{\mathrm{T}}\big\{\big(\mathbb{v}_{\theta|\mathrm{T}} + \|\theta_\bullet\|^2_{\mathfrak{v}}\big)\,k^{-1}m^\circ_{n\wedge k}\|\mathrm{T}^{m^\circ_{n\wedge k}|\dagger}_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)} + \|\theta^{m^\circ_{n\wedge k}}_\bullet\|^2_{\mathbb{L}(\ell_2)}/4\big\}$$
$$\times\,(m^\circ_{n\wedge k})^2\big(k^{-1}(m^\circ_{n\wedge k})^3\|\mathrm{T}^{m^\circ_{n\wedge k}|\dagger}_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)}\big)^{l-1} \quad (08.22)$$

and if in addition

$$(m^{\circ}_{n\wedge k})^2 (k^{-1}(m^{\circ}_{n\wedge k})^3 \|T_{\bullet|\bullet}^{m^{\circ}_{n\wedge k}|\dagger}\|^2_{\mathbb{L}(\ell_2)})^{l-1} \leqslant R^{\circ}_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) \leqslant 1/4 \quad (08.23)$$

then we have

$$\begin{aligned}\mathbb{E}^{n\otimes k}_{\theta|T}(\|\widehat{\theta}^{m^{\circ}_{n\wedge k}}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) &\leqslant R^{\circ}_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) \\ &\quad \times \{(4 + 2^{2l}K^{2l}_{T}(m^{\circ}_{n\wedge k})^{-2})\mathrm{v}_{\theta|T} + (8\mathrm{v}_{T} + 2^{2l-2}K^{2l}_{T})\|\theta^{m^{\circ}_{n\wedge k}}_{\bullet}\|^2_{\ell_2} + 2^{2l}K^{2l}_{T}(m^{\circ}_{n\wedge k})^{-2}\|\theta_{\bullet}\|^2_{\mathfrak{v}}\} \\ &\leqslant 2^{2l+2}K^{2l}_{T}\left(\mathrm{v}_{\theta|T} + \|\theta^{m^{\circ}_{n\wedge k}}_{\bullet}\|^2_{\ell_2} + \|\theta_{\bullet}\|^2_{\mathfrak{v}}\right) R^{\circ}_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) \quad (08.24)\end{aligned}$$

§08|02.28 **Proof** of Corollary §08|02.27. Given in the lecture. □

§08|02.29 **Remark**. Consider $m^{\circ}_{n\wedge k} = \arg\min\{R^{m}_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) : m \in \mathbb{N}\}$ and $R^{\circ}_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) = R^{m^{\circ}_{n\wedge k}}_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet})$ as in (08.20). Arguing similarly as in Remark §07|01.21 we note that $\|M_{\mathfrak{v}} T^{m|\dagger}_{\bullet|\bullet}\|_{HS} \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$ and hence $R^{\circ}_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) = o(1)$ as $n \wedge k \to \infty$, whenever $\|\theta^{m}_{\bullet} - \theta_{\bullet}\|_{\mathfrak{v}} = o(1)$ as $m \to \infty$ (c.f. Remark §05|01.05). In this situation if $\sup\{\|\theta^{m}_{\bullet}\|^2_{\ell_2} : m \in \mathbb{N}\} \leqslant K^2_{\theta|T} \in \mathbb{R}_{\geqslant 0}$ then from (08.24) in Corollary §08|02.27 follows

$$\mathbb{E}^{n\otimes k}_{\theta|T}(\|\widehat{\theta}^{m^{\circ}_{n\wedge k}}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) \leqslant 2^{2l+2}K^{2l}_{T}\left(\mathrm{v}_{\theta|T} + K^2_{\theta|T} + \|\theta_{\bullet}\|^2_{\mathfrak{v}}\right) R^{\circ}_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet}).$$

However, the dimension $m^{\circ}_{n\wedge k} = m^{\circ}_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet})$ as defined in (08.03) depends on the unknown parameter of interest $\theta_{\bullet}$ and the nuissance parameter $T_{\bullet|\bullet}$, and thus also the statistic $\widehat{\theta}^{m^{\circ}_{n\wedge k}}_{\bullet}$. In other words $\widehat{\theta}^{m^{\circ}_{n\wedge k}}_{\bullet}$ is not a feasible estimator. □

§08|02.30 **Corollary** (GniSM with noisy operator §08|02.08 continued). *Consider independent noisy versions* $(\widehat{g}_{\bullet}, \widehat{T}_{\bullet|\bullet}) = (g_{\bullet} + n^{-1/2}\dot{B}_{\bullet}, T_{\bullet|\bullet} + k^{-1/2}\dot{W}_{\bullet|\bullet}) \sim N^{n\otimes k}_{\theta|T} = N^{n}_{\theta|T} \otimes N^{k}_{T}$ *as in Model §08|02.08, where* $\dot{B}_{\bullet} \sim N^{\otimes\mathbb{N}}_{(0,1)}$ *and* $\dot{W}_{\bullet|\bullet} \sim N^{\otimes\mathbb{N}^2}_{(0,1)}$ *are independent,* $T_{\bullet|\bullet} \in \mathbb{T}$ *and* $\theta_{\bullet} \in \ell_2$*, and hence* $g_{\bullet} = T_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(T^{\dagger}_{\bullet|\bullet}) \subseteq \ell_2$*. Given Assumption §08|02.18 for each* $k, n \in \mathbb{N}$ *fulfilling* (08.23) *the (infeasible, generalised) tGE* $\widehat{\theta}^{m^{\circ}_{n\wedge k}}_{\bullet} = \widehat{T}^{m^{\circ}_{n\wedge k}|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in {}_{\ell_2}\mathbb{1}^{m^{\circ}_{n\wedge k}} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$ *satisfies*

$$N^{n\otimes k}_{\theta|T}(\|\widehat{\theta}^{m^{\circ}_{n\wedge k}}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) \leqslant 2^{2l+2}((2l-1)!!)\left(1 + \|\theta^{m^{\circ}_{n\wedge k}}_{\bullet}\|^2_{\ell_2} + \|\theta_{\bullet}\|^2_{\mathfrak{v}}\right) R^{\circ}_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet})$$

where $R^{\circ}_{n}(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet})$ *is the oracle rate in a* GniSM *§08|01.04 (see Corollary §08|01.21).*

§08|02.31 **Proof** of Corollary §08|02.30. Given in the lecture. □

§08|02.32 **Corollary** (niSM with noisy operator §08|02.10 continued). *Consider independent noisy versions* $(\widehat{g}_{\bullet}, \widehat{\mathfrak{s}}_{\bullet}) = (g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}, T_{\bullet|\bullet} + k^{-1/2}\dot{\eta}_{\bullet|\bullet}) \sim P^{n\otimes k}_{\theta|T|\sigma|\xi|\xi^{(2l)}}$ *as in Model §08|02.10, where* $\dot{\varepsilon}_{\bullet}$ *and* $\dot{\eta}_{\bullet|\bullet}$ *satisfies* **(iSM1)** *with* $\mathrm{v}_{\sigma} = \|\sigma^2_{\bullet}\|_{\ell_\infty} \vee 1$ *and* **(niSMnO1)–(niSMnO2)** *with* $K^{2l}_{\xi^{(2l)}} := 1 \vee \|\xi^{(2l)}_{\bullet|\bullet}\|_{\ell_\infty(\mathbb{N}^2)}$*, respectively,* $T_{\bullet|\bullet} \in \mathbb{T}$ *and* $\theta_{\bullet} \in \ell_2$*, and hence* $g_{\bullet} = T_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(T^{\dagger}_{\bullet|\bullet}) \subseteq \ell_2$*. Given Assumption §08|02.18 for each* $k, n \in \mathbb{N}$ *fulfilling* (08.23) *the (infeasible, generalised) tGE* $\widehat{\theta}^{m^{\circ}_{n\wedge k}}_{\bullet} = \widehat{T}^{m^{\circ}_{n\wedge k}|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in {}_{\ell_2}\mathbb{1}^{m^{\circ}_{n\wedge k}} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$ *satisfies*

$$P^{n\otimes k}_{\theta|T|\sigma|\xi|\xi^{(2l)}}(\|\widehat{\theta}^{m^{\circ}_{n\wedge k}}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) \leqslant 2^{2l+2}K^{2l}_{\xi^{(2l)}}\left(\mathrm{v}_{\sigma} + \|\theta^{m^{\circ}_{n\wedge k}}_{\bullet}\|^2_{\ell_2} + \|\theta_{\bullet}\|^2_{\mathfrak{v}}\right) R^{\circ}_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet})$$

where $R^{\circ}_{n}(\theta_{\bullet}, T_{\bullet|\bullet}, \mathfrak{v}_{\bullet})$ *is the oracle rate in a* niSM *§08|01.06 (see Corollary §08|01.23).*

§08|02.33 **Proof** of Corollary §08|02.32. Given in the lecture. □

§08|02.34 **Corollary** (nieMM with noisy operator §08|02.13 continued). *Consider independent noisy versions* $(\widehat{g}_{\bullet}, \widehat{\mathrm{T}}_{\bullet|\bullet}) = (g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}, \mathrm{T}_{\bullet|\bullet} + k^{-1/2}\dot{\eta}_{\bullet|\bullet})$ *defined on* $(\mathcal{Z}^{n+k}, \mathscr{Z}^{\otimes(n+k)}, \mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}})$ *as in Model §08|02.13, where* $\psi_{\bullet} \in \mathcal{M}(\mathscr{X} \otimes 2^{\mathbb{N}})$ *and* $\varphi_{\bullet|\bullet} \in \mathcal{M}(\mathscr{X} \otimes 2^{\mathbb{N}^2})$ *satisfies* **(nieMM1)-(nieMM2)** *for* $\mathrm{v}_{\theta|\mathrm{T}|\psi} \in \mathbb{R}_{\geqslant 1}$ *and* **(nieMMnO1)-(nieMMnO3)** *for* $\mathrm{K}_{\mathrm{T}|\varphi} \in \mathbb{R}_{\geqslant 1}$, *respectively,* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}$ *and* $\theta_{\bullet} \in \ell_2$, *and hence* $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) \subseteq \ell_2$. *Given Assumption §08|02.18 for each* $k, n \in \mathbb{N}$ *fulfilling* (08.23) *the (infeasible, generalised) tGE* $\widehat{\theta}^{m^{\circ}_{n\wedge k}}_{\bullet} = \widehat{\mathrm{T}}^{m^{\circ}_{n\wedge k}|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in \ell_2 \mathbb{1}^{m^{\circ}_{n\wedge k}}_{\bullet} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$ *satisfies*

$$\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}(\|\widehat{\theta}^{m^{\circ}_{n\wedge k}}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) \leqslant \mathrm{C}_{2l}\mathrm{K}^{2l}_{\mathrm{T}|\varphi}\big(\mathrm{v}_{\theta|\mathrm{T}|\psi} + \|\theta^{m^{\circ}_{n\wedge k}}_{\bullet}\|^2_{\ell_2} + \|\theta_{\bullet}\|^2_{\mathfrak{v}}\big)\, \mathrm{R}^{\circ}_{n\wedge k}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet})$$

where $\mathrm{C}_{2l} \in \mathbb{R}_{\geqslant 1}$ *is a constant depending on* $l \in \mathbb{N}$ *only and* $\mathrm{R}^{\circ}_{n}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet})$ *is the oracle rate in a* nieMM *§08|01.08 (see Corollary §08|01.25).*

§08|02.35 **Proof** of Corollary §08|02.34. Given in the lecture. □

§08|02.36 **Illustration**. We distinguish as in Illustration §08|01.27 the two cases **(p)** and **(np)**, where $\theta_{\bullet}\mathbb{1}^{K|\perp}_{\bullet} = 0_{\bullet}$ implies the case **(p)**. In case **(p)** the oracle bound is parametric, that is, $n\mathrm{R}^{\circ}_{n}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) = \mathrm{O}(1)$, in case **(np)** the oracle bound is nonparametric, i.e. $\lim_{n\to\infty} n\mathrm{R}^{\circ}_{n}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) = \infty$. In case **(np)** consider similar to **(o-m)**, **(o-s)** and **(s-m)** in Illustration §08|01.27 the following specifications:

Table 05 [§08]

Order of the rate $\mathrm{R}^{\circ}_{n\wedge k}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet})$ as $n \wedge k \to \infty$

	$(m \in \mathbb{N})$ $(\mathfrak{v}_m = m^{\mathrm{v}})$	(squared bias) $\|\theta^{m}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}$ $(\mathrm{a} \in \mathbb{R}_{>0})$	(variance) $\|\mathrm{M}_{\mathfrak{v}}\mathrm{T}^{m\|\dagger}_{\bullet\|\bullet}\|^2_{\mathrm{HS}}$ $(\mathrm{t} \in \mathbb{R}_{>0})$	$m^{\circ}_{n\wedge k}$	$\mathrm{R}^{\circ}_{n\wedge k}(\theta_{\bullet}, \mathrm{T}_{\bullet\|\bullet}, \mathfrak{v}_{\bullet})$
(o-m)	$\mathrm{v} \in (-1/2 - \mathrm{t}, \mathrm{a})$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2(\mathrm{t}+\mathrm{v})+1}$	$(n \wedge k)^{\frac{1}{2\mathrm{a}+2\mathrm{t}+1}}$	$(n \wedge k)^{-\frac{2(\mathrm{a}-\mathrm{v})}{2\mathrm{a}+2\mathrm{t}+1}}$
	$\mathrm{v} + \mathrm{t} = -1/2$	$m^{-2\mathrm{a}-2\mathrm{t}-1}$	$\log m$	$\big(\frac{n\wedge k}{\log n\wedge k}\big)^{\frac{1}{2\mathrm{a}+2\mathrm{t}+1}}$	$\frac{\log n\wedge k}{n\wedge k}$
(o-s)	$\mathrm{a} - \mathrm{v} \in \mathbb{R}_{>0}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{(1-2(\mathrm{t}-\mathrm{v}))_+}e^{m^{2\mathrm{t}}}$	$(\log n \wedge k)^{\frac{1}{2\mathrm{t}}}$	$(\log n \wedge k)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v} + \mathrm{t} + 1/2 \in \mathbb{R}_{>0}$	$m^{(1-2(\mathrm{a}-\mathrm{v}))_+}e^{-m^{2\mathrm{a}}}$	$m^{2(\mathrm{t}+\mathrm{v})+1}$	$(\log n \wedge k)^{\frac{1}{2\mathrm{a}}}$	$\frac{(\log n\wedge k)^{\frac{2\mathrm{t}+2\mathrm{v}+1}{2\mathrm{a}}}}{n\wedge k}$
	$\mathrm{v} + \mathrm{t} = -1/2$	$e^{-m^{2\mathrm{a}}}$	$\log m$	$(\log n \wedge k)^{\frac{1}{2\mathrm{a}}}$	$\frac{\log\log n\wedge k}{n\wedge k}$

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v} + \mathrm{t} < -1/2$ the rate $\mathrm{R}^{\circ}_{n\wedge k}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet})$ is parametric. The tGE attains the rate $\mathrm{R}^{\circ}_{n\wedge k} := \mathrm{R}^{\circ}_{n\wedge k}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet})$ due to Corollary §08|02.27 under the additional condition

$$\big(k^{-1}(m^{\circ}_{n\wedge k})^3\|\mathrm{T}^{m^{\circ}_{n\wedge k}|\dagger}_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)}\big)^{l-1} \leqslant (m^{\circ}_{n\wedge k})^{-2}\mathrm{R}^{\circ}_{n\wedge k}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet}). \tag{08.25}$$

Since $(m^{\circ}_{n\wedge k})^{-2}\mathrm{R}^{\circ}_{n\wedge k}(\theta_{\bullet}, \mathrm{T}_{\bullet|\bullet}, \mathfrak{v}_{\bullet}) = \mathrm{o}(1)$ also $k^{-1}(m^{\circ}_{n\wedge k})^3\|\mathrm{T}^{m^{\circ}_{n\wedge k}|\dagger}_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)} = \mathrm{o}(1)$ is necessary as $n \wedge k \to \infty$. The next table depicts the order of both terms in case **(o-m)**, **(o-s)** and **(s-m)**.

Table 06 [§08]

Order as $n \wedge k \to \infty$

	(o-m) $\mathrm{v} \in (-1/2 - \mathrm{t}, \mathrm{a})$	**(o-s)** $\mathrm{a} - \mathrm{v} \in \mathbb{R}_{>0}$	**(s-m)** $\mathrm{v} + \mathrm{t} + 1/2 \in \mathbb{R}_{>0}$
$(m^{\circ}_{n\wedge k})^{-2}\mathrm{R}^{\circ}_{n\wedge k}(\theta_{\bullet}, \mathrm{T}_{\bullet\|\bullet}, \mathfrak{v}_{\bullet})$	$(n \wedge k)^{-\frac{2(\mathrm{a}-\mathrm{v})+2}{2\mathrm{a}+2\mathrm{t}+1}}$	$(\log n \wedge k)^{-\frac{2\mathrm{a}-2\mathrm{v}+2}{2\mathrm{t}}}$	$\frac{(\log n\wedge k)^{\frac{2\mathrm{t}+2\mathrm{v}-1}{2\mathrm{a}}}}{n\wedge k}$
$(n \wedge k)^{-1}(m^{\circ}_{n\wedge k})^3\|\mathrm{T}^{m^{\circ}_{n\wedge k}\|\dagger}_{\bullet\|\bullet}\|^2_{\mathbb{L}(\ell_2)}$	$(n \wedge k)^{-\frac{2\mathrm{a}-2}{2\mathrm{a}+2\mathrm{t}+1}}$	$(n \wedge k)^{-c}$	$\frac{(\log n\wedge k)^{\frac{2\mathrm{t}+3}{2\mathrm{a}}}}{n\wedge k}$

In case **(o-s)** a value $l \geqslant 2$ and **(s-m)** a value $l \geqslant 3$ is sufficient to ensure (08.25) as $n \wedge k \to \infty$. In case **(o-m)** assuming $\mathrm{a} > 1$ we have $k^{-1}(m^{\circ}_{n\wedge k})^3\|\mathrm{T}^{m^{\circ}_{n\wedge k}|\dagger}_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)} = \mathrm{o}(1)$ as $n \wedge k \to \infty$. In this situation we have (08.25) if $2(\mathrm{a} - 1)(l - 1) > 2(\mathrm{a} - \mathrm{v}) + 2$ or in equal $l > (2\mathrm{a} - \mathrm{v})/(\mathrm{a} - 1)$. □

§08|02|02|02 Maximal global $\mathfrak{v}$-risk

§08|02.37 **Notation** (Reminder). For sequences $a_\bullet, b_\bullet \in (\mathbb{K})^{\mathbb{N}}$ taking its values in $\mathbb{K} \in \{\mathbb{R}, \mathbb{R}_{>0}, \mathbb{Z}, \dots\}$ we write $a_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\nearrow}$ and $b_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\searrow}$ if $a_\bullet$ and $b_\bullet$, respectively, is monotonically *non-decreasing* and *non-increasing*. If in addition $a_n \to \infty$ and $b_n \to 0$ as $n \to \infty$, then we write $a_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\uparrow\infty}$ and $b_\bullet \in (\mathbb{K})^{\mathbb{N}}_{\downarrow 0}$ for short. For $\mathrm{w}_\bullet \in \ell_\infty$ we set $\mathrm{w}_{(0)} := \|\mathrm{w}_\bullet\|_{\ell_\infty}$ and $\mathrm{w}_{(\bullet)} = (\mathrm{w}_{(j)} := \|\mathrm{w}_\bullet \mathbb{1}^{j|\perp}_\bullet\|_{\ell_\infty})_{j\in\mathbb{N}}$, where by construction $\mathrm{w}_{(\bullet)} \in (\mathbb{R}_{\geqslant 0})^{\mathbb{N}}_{\searrow}$. □

§08|02.38 **Assumption**. Consider weights $\mathfrak{t}_\bullet, \mathfrak{a}_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ and $\mathfrak{v}_\bullet \in \mathbb{R}^{\mathbb{N}}_{>0}$ such that $(\mathfrak{a}\mathfrak{v})_\bullet := \mathfrak{a}_\bullet\mathfrak{v}_\bullet \in \ell_\infty$, $(\mathfrak{a}\mathfrak{v})_{(\bullet)} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$, and $(\mathfrak{t}/\mathfrak{v})_\bullet = \mathfrak{t}_\bullet\mathfrak{v}_\bullet^{-1} \in \ell_\infty$ are satisfied. In addition there exists $\mathrm{C}_{(\mathfrak{t}/\mathfrak{v})} \in (0,1]$ such that for all $m \in \mathbb{N}$

$$(\mathfrak{t}/\mathfrak{v})^2_{(m-1)} \geqslant \min\left\{(\mathfrak{t}/\mathfrak{v})^2_j \colon j \in [\![m]\!]\right\} \geqslant \mathrm{C}_{(\mathfrak{t}/\mathfrak{v})}(\mathfrak{t}/\mathfrak{v})^2_{(m)} \tag{08.26}$$

or in equal $\mathrm{C}_{(\mathfrak{t}/\mathfrak{v})}\|(\mathfrak{t}/\mathfrak{v})^{-2}_\bullet\mathbb{1}^m_\bullet\|_{\ell_\infty} \leqslant (\mathfrak{t}/\mathfrak{v})^{-2}_{(m)}$. □

§08|02.39 **Reminder**. Under Assumption §08|02.38 we have $\ell_2^{\mathfrak{a}} = \mathrm{dom}(\mathrm{M}_{\mathfrak{a}^{-1}_\bullet}) = \ell_2\mathfrak{a}_\bullet \subseteq \ell_2$ and the three measures $\nu_{\mathbb{N}}$, $\mathfrak{a}^{-2}_\bullet\nu_{\mathbb{N}}$ and $\mathfrak{v}^2_\bullet\nu_{\mathbb{N}}$ dominate mutually each other, i.e. they share the same null sets (see Property §04|01.02). We consider $\ell_2^{\mathfrak{a}}$ endowed with $\|\cdot\|_{\mathfrak{a}^{-1}} = \|\mathrm{M}_{\mathfrak{a}^{-1}_\bullet}\cdot\|_{\ell_2}$ and given a constant $\mathrm{r} \in \mathbb{R}_{>0}$ the ellipsoid $\ell_2^{\mathfrak{a},\mathrm{r}} := \{a_\bullet \in \ell_2^{\mathfrak{a}} : \|a_\bullet\|_{\mathfrak{a}^{-1}} \leqslant \mathrm{r}\} \subseteq \ell_2^{\mathfrak{a}}$. Since $(\mathfrak{a}\mathfrak{v})_\bullet \in \ell_\infty$, and hence $(\mathfrak{a}\mathfrak{v})_{(m)} := \|(\mathfrak{a}\mathfrak{v})_\bullet\mathbb{1}^{m|\perp}_\bullet\|_{\ell_\infty} \in \mathbb{R}_{\geqslant 0}$ for each $m \in \mathbb{N}$ we have $\ell_2^{\mathfrak{a}} \subseteq \ell_2(\mathfrak{v}^2_\bullet)$ (Property §04|02.11). Consequently, if Assumption §08|02.38 and $\theta_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}$ are satisfied, then Assumption §08|02.18 is also fulfilled. Since $\mathfrak{v}_\bullet, \mathfrak{t}_\bullet \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ under Assumption §08|02.38, we have $\|\mathfrak{t}^{-1}_\bullet\mathbb{1}^m_\bullet\|_{\mathfrak{v}} = \|(\mathfrak{v}/\mathfrak{t})_\bullet\mathbb{1}^m_\bullet\|_{\ell_2} \in \mathbb{R}_{\geqslant 0}$ for each $m \in \mathbb{N}$. Under the Assumptions §08|00.02 and §08|02.38 considering the generalised link condition $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d,D}}$ with band $\mathrm{D} \in [1,\infty)$ and $\mathrm{d} \in [1,\mathrm{D}]$ as in Definition §05|02.05 we have $\sup_{m\in\mathbb{N}}\{\|[\mathrm{M}_{\mathfrak{t}}]_{\underline{m}}[\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}}\|_{\mathrm{spec}}\} \leqslant \mathrm{D}$, and hence $\|\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\|_{\mathbb{L}(\ell_2)} \leqslant \mathrm{D}\mathfrak{t}^{-1}_m$ and $\|\mathrm{M}_{\mathfrak{v}}\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\|_{\mathrm{HS}} \leqslant \mathrm{D}\|\mathfrak{t}^{-1}_\bullet\mathbb{1}^m_\bullet\|_{\mathfrak{v}}$ shown in (08.06) using $\mathrm{tr}([\mathrm{M}^m_{(\mathfrak{v}/\mathfrak{t})}]^2_{\underline{m}}) = \|(\mathfrak{v}/\mathfrak{t})_\bullet\mathbb{1}^m_\bullet\|^2_{\ell_2} = \|\mathfrak{t}^{-1}_\bullet\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}}$. Moreover, for each $m \in \mathbb{N}$ the generalised Galerkin solution $\theta^m_\bullet := \mathrm{T}^{m|\dagger}_{\bullet|\bullet}g_\bullet \in \ell_2\mathbb{1}^m_\bullet$ of $\theta_\bullet = \mathrm{T}^{\dagger}_{\bullet|\bullet}g_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}$ satisfies (Lemma §05|02.09)

$$\|\theta^m_\bullet\|_{\ell_2} \leqslant \mathfrak{a}_1\|\theta^m_\bullet\|_{\mathfrak{a}^{-1}} \leqslant \mathfrak{a}_1\mathrm{D}\mathrm{d}\mathrm{r} \quad \text{and} \quad \|\theta_\bullet - \theta^m_\bullet\|^2_{\mathfrak{v}} \leqslant (\mathrm{D}^2\mathrm{d}^2\mathrm{C}^{-2}_{(\mathfrak{t}/\mathfrak{v})} + 1)(\mathfrak{a}\mathfrak{v})^2_{(m)}\mathrm{r}^2.$$

Note that under Assumptions §08|00.02 and §08|02.38 the link condition $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}^{\gtrsim}_{\mathrm{t,d}}$ with band $\mathrm{d} \in \mathbb{R}_{\geqslant 1}$ as in Definition §05|01.08 implies $\sup_{m\in\mathbb{N}}\{\|[\mathrm{M}_{\mathfrak{t}}]_{\underline{m}}[\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}}\|_{\mathrm{spec}}\} \leqslant 3\mathrm{d}^2$ (Lemma §05|01.22), and hence for each $m \in \mathbb{N}$ we have $\|\mathrm{M}_{\mathfrak{v}}\mathrm{T}^{m|\dagger}_{\bullet|\bullet}\|^2_{\mathrm{HS}} \leqslant 9\mathrm{d}^4\|\mathfrak{t}^{-1}_\bullet\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}}$ and the Galerkin solution $\theta^m_\bullet := \mathrm{T}^{m|\dagger}_{\bullet|\bullet}g_\bullet \in \ell_2\mathbb{1}^m_\bullet$ of $\theta_\bullet = \mathrm{T}^{\dagger}_{\bullet|\bullet}g_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}$ satisfies $\|\theta^m_\bullet\|_{\ell_2} \leqslant 3\mathfrak{a}_1\mathrm{d}^3\mathrm{r}$ and $\|\theta_\bullet - \theta^m_\bullet\|^2_{\mathfrak{v}} \leqslant (9\mathrm{d}^6\mathrm{C}^{-2}_{(\mathfrak{t}/\mathfrak{v})} + 1)(\mathfrak{a}\mathfrak{v})^2_{(m)}\mathrm{r}^2$ (Lemma §05|01.28). □

§08|02.40 **Corollary**. *Under Assumptions §08|02.02 and §08|02.38 let* $\theta_\bullet := \mathrm{T}^{\dagger}_{\bullet|\bullet}g_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}$ *and* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d,D}}$ *(or* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}^{\gtrsim}_{\mathrm{t,d}}$ *with* $\mathrm{D} = 3\mathrm{d}^2$*), for all* $m, k, n \in \mathbb{N}$ *we have*

(i) *if* $4\mathrm{D}^2\|\mathfrak{t}^{-1}_\bullet\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}} \leqslant k \wedge n$ *then* $\mho_{m,k} \subseteq \Omega_{m,k\wedge n}$,

(ii) $\mathbb{P}^k_{\mathrm{T}}\left(\mathrm{tr}(\mathrm{M}_{\mathfrak{v}}\widehat{\mathrm{T}}^{m|\dagger}_{\bullet|\bullet}\Gamma_{\theta|\mathrm{T}}(\widehat{\mathrm{T}}^{m|\dagger}_{\bullet|\bullet})^\star\mathrm{M}_{\mathfrak{v}})\mathbb{1}_{\Omega_{m,k\wedge n}}\right) \leqslant \mathrm{v}_{\theta|\mathrm{T}}(4\mathrm{D}^2\|\mathfrak{t}^{-1}_\bullet\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}} + (k\wedge n)\mathbb{P}^k_{\mathrm{T}}(\mho^c_{m,k}))$, *and*

(iii) $\mathbb{P}^k_{\mathrm{T}}(\|\widehat{\mathrm{T}}^{m|\dagger}_{\bullet|\bullet}(\mathrm{T}^m_{\bullet|\bullet} - \widehat{\mathrm{T}}^m_{\bullet|\bullet})\theta^m_\bullet\|^2_{\mathfrak{v}}\mathbb{1}_{\Omega_{m,k\wedge n}}) \leqslant 4k^{-1}\mathrm{v}_{\mathrm{T}}\mathrm{D}^2\|\mathfrak{t}^{-1}_\bullet\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}}\|\theta^m_\bullet\|^2_{\ell_2} + \mathbb{P}^k_{\mathrm{T}}(\|\dot{\eta}^m_{\bullet|\bullet}\theta^m_\bullet\|^2_{\ell_2}\mathbb{1}_{\mho^c_{m,k}})$.

with $\Omega_{m,k\wedge n}$ *and* $\mho_{m,k}$ *as in* (08.18) *and* (08.19)*, respectively.*

§08|02.41 **Proof** of Corollary §08|02.40. Given in the lecture. □

§08|02.42 **Proposition** (Upper bound). *Under Assumptions §08|02.02 and §08|02.38 let* $\theta_\bullet := \mathrm{T}^{\dagger}_{\bullet|\bullet}g_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}$ *and*

$\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d,D}}$ *(or* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}^{\gtrless}_{\mathrm{t,d}}$ *with* $\mathrm{D} = 3\mathrm{d}^2$*) for all* $n, m \in \mathbb{N}$ *setting similar to* (08.07)

$$\mathrm{R}^m_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) := [(\mathfrak{a}\mathfrak{v})^2_{(m)} \vee (n\wedge k)^{-1}\|\mathfrak{t}_\bullet^{-1}\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}_\bullet}],$$
$$m^\star_{n\wedge k} := \arg\min\{\mathrm{R}^m_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) : m \in \mathbb{N}\} \quad \textit{and}$$
$$\mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) := \mathrm{R}^{m^\star_{n\wedge k}}_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) = \min\{\mathrm{R}^m_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) : m \in \mathbb{N}\} \tag{08.27}$$

for all $m \in \mathbb{N}$ *the (generalised) tGE* $\widehat{\theta}^m_\bullet = \widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{g}_\bullet \in \ell_2\mathbb{1}^m_\bullet$ $\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}$-a.s. *satisfies*

$$\mathbb{E}^{n\otimes k}_{\theta|\mathrm{T}}(\|\widehat{\theta}^m_\bullet - \theta_\bullet\|^2_{\mathfrak{v}_\bullet}) \leqslant 2\mathrm{D}^2(\mathrm{C}^{-2}_{(\mathfrak{t}/\mathfrak{v})}\mathrm{d}^2\mathrm{r}^2 + 2\mathrm{v}_{\theta|\mathrm{T}} + 4\mathrm{v}_{\mathrm{T}}\mathfrak{a}_1^2\mathrm{D}^2\mathrm{d}^2\mathrm{r}^2)\,\mathrm{R}^m_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet)$$
$$+ \mathrm{v}_{\theta|\mathrm{T}}\mathbb{P}^k_{\mathrm{T}}(\mho^c_{m,k}) + 2\mathbb{E}^k_{\mathrm{T}}(\|\dot{\eta}^m_{\bullet|\bullet}\theta^m_\bullet\|^2_{\ell_2}\mathbb{1}_{\mho^c_{m,k}}) + \|\theta_\bullet\|^2_{\mathfrak{v}_\bullet}\mathbb{P}^k_{\mathrm{T}}(\Omega^c_{m,k\wedge n}) \tag{08.28}$$

with $\Omega_{m,k\wedge n}$ *and* $\mho_{m,k}$ *as in* (08.18) *and* (08.19)*, respectively.*

§08|02.43 **Proof** of Proposition §08|02.42. Given in the lecture. □

§08|02.44 **Corollary**. *Under the assumptions of Proposition §08|02.42 for* $k, n \in \mathbb{N}$ *with* $\mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) \leqslant 1/(4\mathrm{D}^2)$ *the (generalised) tGE* $\widehat{\theta}^{m^\star_{n\wedge k}}_\bullet = \widehat{\mathrm{T}}^{m^\star_{n\wedge k}|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{g}_\bullet \in \ell_2\mathbb{1}^{m^\star_{n\wedge k}}_\bullet \subseteq \ell_2(\mathfrak{v}^2_\bullet)$ $\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}$-a.s. *with* $\Omega_{m^\star_{n\wedge k},k\wedge n}$ *as in* (08.18) *and dimension* $m^\star_{n\wedge k}$ *as in* (08.27) *satisfies*

$$\mathbb{E}^{n\otimes k}_{\theta|\mathrm{T}}(\|\widehat{\theta}^{m^\star_{n\wedge k}}_\bullet - \theta_\bullet\|^2_{\mathfrak{v}_\bullet}) \leqslant 2\mathrm{D}^2(\mathrm{C}^{-2}_{(\mathfrak{t}/\mathfrak{v})}\mathrm{d}^2\mathrm{r}^2 + 2\mathrm{v}_{\theta|\mathrm{T}} + 4\mathrm{v}_{\mathrm{T}}\mathfrak{a}_1^2\mathrm{D}^2\mathrm{d}^2\mathrm{r}^2)\,\mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet)$$
$$+ 2^{2l}\mathrm{K}^{2l}_{\mathrm{T}}\mathrm{D}^{2(l-1)}\Big\{\mathrm{D}^2\big(\mathrm{v}_{\theta|\mathrm{T}} + (\mathfrak{a}\mathfrak{v})^2_{(0)}\mathrm{r}^2\big)\,k^{-1}m^\star_{n\wedge k}\mathfrak{t}^{-2}_{m^\star_{n\wedge k}} + \mathfrak{a}_1^2\mathrm{D}^2\mathrm{d}^2\mathrm{r}^2/4\Big\}$$
$$\times (m^\star_{n\wedge k})^2(k^{-1}(m^\star_{n\wedge k})^3\mathfrak{t}^{-2}_{m^\star_{n\wedge k}})^{l-1}. \tag{08.29}$$

and if in addition

$$(m^\star_{n\wedge k})^2(k^{-1}(m^\star_{n\wedge k})^3\mathfrak{t}^{-2}_{m^\star_{n\wedge k}})^{l-1} \leqslant \mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) \leqslant 1/(4\mathrm{D}^2) \tag{08.30}$$

then we have

$$\mathbb{E}^{n\otimes k}_{\theta|\mathrm{T}}(\|\widehat{\theta}^{m^\circ_{n\wedge k}}_\bullet - \theta_\bullet\|^2_{\mathfrak{v}_\bullet}) \leqslant 2^{2l+2}\mathrm{K}^{2l}_{\mathrm{T}}\mathrm{D}^{2l}(\mathrm{v}_{\theta|\mathrm{T}} + (\mathfrak{a}\mathfrak{v})^2_{(0)}\mathrm{r}^2 + (\mathrm{C}^{-2}_{(\mathfrak{t}/\mathfrak{v})} + \mathfrak{a}_1^2)\mathrm{d}^2\mathrm{r}^2)\,\mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet). \tag{08.31}$$

§08|02.45 **Proof** of Corollary §08|02.44. Given in the lecture. □

§08|02.46 **Remark**. Arguing similarly as in Remark §07|01.21 we note that $\|\mathfrak{t}_\bullet^{-1}\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}_\bullet} \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$ and hence $\mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) = \mathrm{o}(1)$ as $n \wedge k \to \infty$, whenever $(\mathfrak{a}\mathfrak{v})_{(m)} = \mathrm{o}(1)$ as $m \to \infty$, i.e. $(\mathfrak{a}\mathfrak{v})_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$. If there is in addition $\mathrm{C} \in \mathbb{R}_{\geqslant 1}$ such that $\mathrm{K}^{2l}_{\mathrm{T}} \leqslant \mathrm{C}$ and $\mathrm{v}_{\theta|\mathrm{T}} \leqslant \mathrm{C}$ for all $\theta_\bullet := \mathrm{T}^\dagger_{\bullet|\bullet}g_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}$ and $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d,D}}$ then from the bound (08.31) Corollary §08|02.44 follows immediately

$$\sup\{\mathbb{E}^{n\otimes k}_{\theta|\mathrm{T}}(\|\widehat{\theta}^{m^\star_{n\wedge k}}_\bullet - \theta_\bullet\|^2_{\mathfrak{v}_\bullet})\colon \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t,d,D}}, \theta_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}\} \leqslant \mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet)$$
$$\times\, 2^{2l+2}\mathrm{C}\mathrm{D}^{2l}(\mathrm{C} + (\mathfrak{a}\mathfrak{v})^2_{(0)}\mathrm{r}^2 + (\mathrm{C}^{-2}_{(\mathfrak{t}/\mathfrak{v})} + \mathfrak{a}_1^2)\mathrm{d}^2\mathrm{r}^2).$$

Note that the dimension $m^\star_{n\wedge k} := m^\star_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet)$ does not depend on the unknown parameter of interest $\theta_\bullet$ but on the classes $\ell_2^{\mathfrak{a},\mathrm{r}}$ and $\mathbb{T}_{\mathrm{t,d,D}}$ only, and thus also the statistic $\widehat{\theta}^{m^\star_n}_\bullet$. In other words, if the regularity of $\theta_\bullet$ and $\mathrm{T}_{\bullet|\bullet}$ is known in advance, then the thresholded GE $\widehat{\theta}^{m^\star_n}_\bullet$ is a feasible estimator. □

§08|02.47 **Corollary** (GniSM with noisy operator §08|02.08 continued). *Consider* independent *noisy versions* $(\widehat{g}_\bullet, \widehat{\mathrm{T}}_{\bullet|\bullet}) = (g_\bullet + n^{-1/2}\dot{\mathrm{B}}_\bullet, \mathrm{T}_{\bullet|\bullet} + k^{-1/2}\dot{\mathrm{W}}_{\bullet|\bullet}) \sim \mathrm{N}^{n\otimes k}_{\theta|\mathrm{T}} = \mathrm{N}^n_{\theta|\mathrm{T}} \otimes \mathrm{N}^k_{\mathrm{T}}$ *as in Model §08|02.08, where* $\dot{\mathrm{B}}_\bullet \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$ *and* $\dot{\mathrm{W}}_{\bullet|\bullet} \sim \mathrm{N}^{\otimes\mathbb{N}^2}_{(0,1)}$ *are* independent, $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}$ *and* $\theta_\bullet \in \ell_2$*, and hence* $g_\bullet = \mathrm{T}_{\bullet|\bullet}\theta_\bullet \in$

$\mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) \subseteq \ell_2$. *Given* Assumption §08|02.38 *for each* $k, n \in \mathbb{N}$ *fulfilling* (08.30) *the (generalised) tGE* $\widehat{\theta}^{m^\star_{n\wedge k}}_{\bullet} = \widehat{\mathrm{T}}^{m^\star_{n\wedge k}|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in \ell_2\mathbb{1}^{m^\star_{n\wedge k}}_{\bullet} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$ *satisfies*

$$\sup\left\{\mathrm{N}^{n\otimes k}_{\theta|\mathrm{T}}(\|\widehat{\theta}^{m^\star_{n\wedge k}}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) \colon \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}, \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}\right\} \leqslant \mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) \\ \times 2^{2l+2}((2l-1)!!)\mathrm{D}^{2l}(1 + (\mathfrak{a}\mathfrak{v})^2_{(0)}\mathrm{r}^2 + (\mathrm{C}^{-2}_{(\mathfrak{t}/\mathfrak{v})} + \mathfrak{a}^2_1)\mathrm{d}^2\mathrm{r}^2)$$

where $\mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet})$ *is the rate in a* GniSM *§08|01.04 (see* Corollary *§08|01.34).*

§08|02.48 **Proof** of Corollary §08|02.47. Given in the lecture. □

§08|02.49 **Corollary** (niSM with noisy operator §08|02.10 continued). *Consider* independent *noisy versions* $(\widehat{g}_{\bullet}, \widehat{\mathfrak{s}}_{\bullet}) = (g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}, \mathrm{T}_{\bullet|\bullet} + k^{-1/2}\dot{\eta}_{\bullet|\bullet}) \sim \mathrm{P}^{n\otimes k}_{\theta|\mathrm{T}|\sigma|\xi|\xi^{(2l)}}$ *as in* Model *§08|02.10, where* $\dot{\varepsilon}_{\bullet}$ *and* $\dot{\eta}_{\bullet|\bullet}$ *satisfies* **(iSM1)** *with* $\mathbb{v}_{\sigma} = \|\sigma^2_{\bullet}\|_{\ell_\infty} \vee 1$ *and* **(niSMnO1)–(niSMnO2)** *with* $\mathrm{K}^{2l}_{\xi^{(2l)}} := 1 \vee \|\xi^{(2l)}_{\bullet|\bullet}\|_{\ell_\infty(\mathbb{N}^2)}$*, respectively,* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}$ *and* $\theta_{\bullet} \in \ell_2$*, and hence* $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) \subseteq \ell_2$*. Given* Assumption *§08|02.38 for each* $k, n \in \mathbb{N}$ *fulfilling* (08.30) *the (generalised) tGE* $\widehat{\theta}^{m^\star_{n\wedge k}}_{\bullet} = \widehat{\mathrm{T}}^{m^\star_{n\wedge k}|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in \ell_2\mathbb{1}^{m^\star_{n\wedge k}}_{\bullet} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$ *satisfies*

$$\sup\left\{\mathrm{P}^{n\otimes k}_{\theta|\mathrm{T}|\sigma|\xi|\xi^{(2l)}}(\|\widehat{\theta}^{m^\star_{n\wedge k}}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) \colon \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}, \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}\right\} \leqslant \mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) \\ \times 2^{2l+2}\mathrm{K}^{2l}_{\xi^{(2l)}}\mathrm{D}^{2l}(\mathbb{v}_{\sigma} + (\mathfrak{a}\mathfrak{v})^2_{(0)}\mathrm{r}^2 + (\mathrm{C}^{-2}_{(\mathfrak{t}/\mathfrak{v})} + \mathfrak{a}^2_1)\mathrm{d}^2\mathrm{r}^2)$$

where $\mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet})$ *is the rate in a* niSM *§08|01.06 (see* Corollary *§08|01.36).*

§08|02.50 **Proof** of Corollary §08|02.49. Given in the lecture. □

§08|02.51 **Corollary** (nieMM with noisy operator §08|02.13 continued). *Consider* independent *noisy versions* $(\widehat{g}_{\bullet}, \widehat{\mathrm{T}}_{\bullet|\bullet}) = (g_{\bullet} + n^{-1/2}\dot{\varepsilon}_{\bullet}, \mathrm{T}_{\bullet|\bullet} + k^{-1/2}\dot{\eta}_{\bullet|\bullet})$ *defined on* $(\mathcal{Z}^{n+k}, \mathscr{Z}^{\otimes(n+k)}, \mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}})$ *as in* Model *§08|02.13, where* $\psi_{\bullet} \in \mathcal{M}(\mathscr{Z} \otimes 2^{\mathbb{N}})$ *and* $\varphi_{\bullet|\bullet} \in \mathcal{M}(\mathscr{Z} \otimes 2^{\mathbb{N}^2})$ *satisfies* **(nieMM1)-(nieMM2)** *for* $\mathbb{v}_{\theta|\mathrm{T}|\psi} \in \mathbb{R}_{\geqslant 1}$ *and* **(nieMMnO1)-(nieMMnO3)** *for* $\mathrm{K}_{\mathrm{T}|\varphi} \in \mathbb{R}_{\geqslant 1}$*, respectively,* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}$ *and* $\theta_{\bullet} \in \ell_2$*, and hence* $g_{\bullet} = \mathrm{T}_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{T}^{\dagger}_{\bullet|\bullet}) \subseteq \ell_2$*. Given* Assumption *§08|02.38 for each* $k, n \in \mathbb{N}$ *fulfilling* (08.30) *the (generalised) tGE* $\widehat{\theta}^{m^\circ_{n\wedge k}}_{\bullet} = \widehat{\mathrm{T}}^{m^\circ_{n\wedge k}|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in \ell_2\mathbb{1}^{m^\circ_{n\wedge k}}_{\bullet} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$ *satisfies*

$$\sup\left\{\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}(\|\widehat{\theta}^{m^\star_{n\wedge k}}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}) \colon \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}, \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}\right\} \leqslant \mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) \\ \times \mathrm{C}_{2l}\sup\left\{\mathrm{K}^{2l}_{\mathrm{T}|\varphi} \colon \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}\right\}\mathrm{D}^{2l}(\sup\left\{\mathbb{v}_{\theta|\mathrm{T}|\psi} \colon \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}, \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}\right\} + (\mathfrak{a}\mathfrak{v})^2_{(0)}\mathrm{r}^2 + (\mathrm{C}^{-2}_{(\mathfrak{t}/\mathfrak{v})} + \mathfrak{a}^2_1)\mathrm{d}^2\mathrm{r}^2)$$

where $\mathrm{C}_{2l} \in \mathbb{R}_{\geqslant 1}$ *is a constant depending on* $l \in \mathbb{N}$ *only and* $\mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet})$ *is the rate in a* nieMM *§08|01.08 (see* Corollary *§08|01.38).*

§08|02.52 **Proof** of Corollary §08|02.51. Given in the lecture. □

§08|02.53 **Illustration**. We distinguish again the two cases **(p)** and **(np)** given in Illustration §08|01.40 where in case **(p)** the bound in Corollary §08|02.44 is parametric, that is, $(n \wedge k)\mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) = \mathrm{O}(1)$, in case **(np)** the bound is nonparametric, i.e. $\lim_{n\to\infty}(n \wedge k)\mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) = \infty$. In case **(np)** we consider similar to **(o-m)**, **(o-s)** and **(s-m)** in Illustration §07|01.44 the following three specifications:

Table 07 [§08]

Order of the rate $R^{\star}_{n\wedge k}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet})$ as $n \wedge k \to \infty$

	$(j \in \mathbb{N})$ $\mathfrak{v}_j^2 = j^{2v}$	$(\mathrm{a} \in \mathbb{R}_{>0})$ $\mathfrak{a}_j^2$	$(\mathrm{t} \in \mathbb{R}_{>0})$ $\mathfrak{t}_j^2$	(squared bias) $(\mathfrak{a}\mathfrak{v})^2_{(m)}$	(variance) $\|\mathfrak{t}_{\bullet}^{-1}\mathbb{1}_{\bullet}^{m}\|^2_{\mathfrak{v}}$	$m^{\star}_{n\wedge k}$	$R^{\star}_{n\wedge k}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet})$
(o-m)	$\mathrm{v} \in (-1/2 - \mathrm{t}, \mathrm{a})$	$j^{-2\mathrm{a}}$	$j^{-2\mathrm{t}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2\mathrm{v}+2\mathrm{t}+1}$	$(n \wedge k)^{\frac{1}{2\mathrm{a}+2\mathrm{t}+1}}$	$(n \wedge k)^{-\frac{2(\mathrm{a}-\mathrm{v})}{2\mathrm{a}+2\mathrm{t}+1}}$
	$\mathrm{v} + \mathrm{t} = -1/2$	$j^{-2\mathrm{a}}$	$j^{-2\mathrm{t}}$	$m^{-2\mathrm{a}-2\mathrm{t}-1}$	$\log m$	$\left(\frac{n\wedge k}{\log n\wedge k}\right)^{\frac{1}{2\mathrm{a}+2\mathrm{t}+1}}$	$\frac{\log n\wedge k}{n\wedge k}$
(o-s)	$\mathrm{a} - \mathrm{v} \in \mathbb{R}_{>0}$	$j^{-2\mathrm{a}}$	$e^{-j^{2\mathrm{t}}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{(1-2(\mathrm{t}-\mathrm{v}))_+}e^{m^{2\mathrm{t}}}$	$(\log n \wedge k)^{\frac{1}{2\mathrm{t}}}$	$(\log n \wedge k)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v} + \mathrm{t} + 1/2 \in \mathbb{R}_{>0}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$m^{2\mathrm{v}}e^{-m^{2\mathrm{a}}}$	$m^{2\mathrm{v}+2\mathrm{t}+1}$	$(\log n \wedge k)^{\frac{1}{2\mathrm{a}}}$	$\frac{(\log n\wedge k)^{\frac{2\mathrm{t}+2\mathrm{v}+1}{2\mathrm{a}}}}{n\wedge k}$
	$\mathrm{v} + \mathrm{t} = -1/2$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$m^{2\mathrm{v}}e^{-m^{2\mathrm{a}}}$	$\log m$	$(\log n \wedge k)^{\frac{1}{2\mathrm{a}}}$	$\frac{\log\log n\wedge k}{n\wedge k}$

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v} + \mathrm{t} < -1/2$ the rate $R^{\star}_{n}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet})$ is parametric. The tGE attains the rate $R^{\star}_{n\wedge k} := R^{\star}_{n\wedge k}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet})$ due to Corollary §08|02.44 under the additional condition

$$(k^{-1}(m^{\star}_{n\wedge k})^3\mathfrak{t}^{-2}_{m^{\star}_{n\wedge k}})^{l-1} \leqslant (m^{\star}_{n\wedge k})^{-2}R^{\star}_{n\wedge k}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}). \tag{08.32}$$

Since $(m^{\star}_{n\wedge k})^{-2}R^{\circ}_{n\wedge k}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet}) = o(1)$ also $k^{-1}(m^{\star}_{n\wedge k})^3\mathfrak{t}^{-2}_{m^{\star}_{n\wedge k}} = o(1)$ is necessary as $n \wedge k \to \infty$. The next table depicts the order of both terms in case **(o-m)**, **(o-s)** and **(s-m)**.

Table 08 [§08]

Order as $n \wedge k \to \infty$

	(o-m) $\mathrm{v} \in (-1/2 - \mathrm{t}, \mathrm{a})$	**(o-s)** $\mathrm{a} - \mathrm{v} \in \mathbb{R}_{>0}$	**(s-m)** $\mathrm{v} + \mathrm{t} + 1/2 \in \mathbb{R}_{>0}$
$(m^{\circ}_{n\wedge k})^{-2}R^{\star}_{n\wedge k}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \mathfrak{v}_{\bullet})$	$(n \wedge k)^{-\frac{2(\mathrm{a}-\mathrm{v})+2}{2\mathrm{a}+2\mathrm{t}+1}}$	$(\log n \wedge k)^{-\frac{2\mathrm{a}-2\mathrm{v}+2}{2\mathrm{t}}}$	$\frac{(\log n\wedge k)^{\frac{2\mathrm{t}+2\mathrm{v}-1}{2\mathrm{a}}}}{n\wedge k}$
$(n \wedge k)^{-1}(m^{\star}_{n\wedge k})^3\mathfrak{t}^{-2}_{m^{\star}_{n\wedge k}}$	$(n \wedge k)^{-\frac{2\mathrm{a}-2}{2\mathrm{a}+2\mathrm{t}+1}}$	$(n \wedge k)^{-c}$	$\frac{(\log n\wedge k)^{\frac{2\mathrm{t}+3}{2\mathrm{a}}}}{n\wedge k}$

In case **(o-s)** a value $l \geqslant 2$ and **(s-m)** a value $l \geqslant 3$ is sufficient to ensure (08.32) as $n \wedge k \to \infty$. In case **(o-m)** assuming $\mathrm{a} > 1$ we have $k^{-1}(m^{\circ}_{n\wedge k})^3\mathfrak{t}^{-2}_{m^{\circ}_{n\wedge k}} = o(1)$ as $n \wedge k \to \infty$. In this situation we have (08.32) if $2(\mathrm{a} - 1)(l - 1) > 2(\mathrm{a} - \mathrm{v}) + 2$ or in equal $l > (2\mathrm{a} - \mathrm{v})/(\mathrm{a} - 1)$. □

§08|02|03 Local and maximal local ϕ-risk

We measure the accuracy of the (generalised) tGE $\widehat{\theta}^m_{\bullet} := \widehat{T}^{m|(n\wedge k)|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet}$ of the (generalised) Galerkin solution $\theta^m_{\bullet} = T^{m|\dagger}_{\bullet|\bullet}g_{\bullet} \in \ell_2\mathbb{1}^m_{\bullet}$ with $g_{\bullet} = T_{\bullet|\bullet}\theta_{\bullet} \in \mathrm{dom}(T^{\dagger}_{\bullet|\bullet})$ by the mean of its local ϕ-error introduced in §05|01|02 and §05|02|02, i.e. its ϕ-risk.

§08|02.54 **Reminder**. If $\phi_{\bullet} \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ then we have $\phi^2_{\bullet}\mathbb{1}^m_{\bullet} \in \ell_2$ and $\ell_2\mathbb{1}^m_{\bullet} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$. Consequently, for each $\theta_{\bullet} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ the (generalised) Galerkin solution $\theta^m_{\bullet} = T^{m|\dagger}_{\bullet|\bullet}g_{\bullet} \in \ell_2\mathbb{1}^m_{\bullet}$ satisfies $\theta^m_{\bullet} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ too. If in addition $C_{T} := \sup\left\{\|M_{\mathbb{1}^{m|\perp}_{\bullet}}T^{\star}_{\bullet|\bullet}(T^{m|\dagger}_{\bullet|\bullet})^{\star}\phi_{\bullet}\|_{\ell_2} : m \in \mathbb{N}\right\} \in \mathbb{R}_{\geqslant 0}$ then $|\phi\nu_{\mathbb{N}}(\theta^m_{\bullet} - \theta_{\bullet})| \leqslant (1 + C_{T})\|\mathbb{1}^{m|\perp}_{\bullet}\theta_{\bullet}\|_{\ell_2}$ which implies $\sup\left\{|\phi\nu_{\mathbb{N}}(\theta^j_{\bullet} - \theta_{\bullet})| : j \in \mathbb{N}_{\geqslant m}\right\} = o(1)$ as $m \to \infty$ (Property §05|01.31 and Property §05|02.12). □

§08|02.55 **Comment**. Under Assumption §08|02.02 we have $\dot{\varepsilon}_{\bullet}\mathbb{1}^m_{\bullet} \in \ell_\infty$ $\mathbb{P}^n_{\theta|T}$-a.s. and $\widehat{T}^m_{\bullet|\bullet} \in \mathbb{L}(\ell_2)$ with $\mathrm{ran}(\widehat{T}^m_{\bullet|\bullet}) \subseteq \ell_2\mathbb{1}^m_{\bullet}$ $\mathbb{P}^k_{T}$-a.s. for each $m \in \mathbb{N}$. Consequently, $\mathrm{ran}(\widehat{T}^{m|(k\wedge n)|\dagger}_{\bullet|\bullet}) \subseteq \ell_2\mathbb{1}^m_{\bullet}$ $\mathbb{P}^k_{T}$-a.s., and $\widehat{T}^{m|(k\wedge n)|\dagger}_{\bullet|\bullet}\dot{\varepsilon}_{\bullet} \in \ell_2\mathbb{1}^m_{\bullet}$ $\mathbb{P}^{n\otimes k}_{\theta|T}$-a.s., and hence

$$\widehat{\theta}^m_{\bullet} = \widehat{T}^{m|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} = n^{-1/2}\widehat{T}^{m|(k\wedge n)|\dagger}_{\bullet|\bullet}\dot{\varepsilon}_{\bullet} + \widehat{T}^{m|(k\wedge n)|\dagger}_{\bullet|\bullet}g_{\bullet} \in \ell_2\mathbb{1}^m_{\bullet} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}}) \quad \mathbb{P}^{n\otimes k}_{\theta|T}\text{-a.s..}$$ □

§08|02|03|01 Local ϕ-risk

§08|02.56 **Assumption**. Let $\phi \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ and $\theta \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ be satisfied. □

§08|02.57 **Definition**. Under Assumptions §08|02.02 and §08|02.56 the *local ϕ-risk* of a (generalised) tGE $\widehat{\theta}^m = \widehat{\mathrm{T}}^{m|(n\wedge k)|\dagger}_{\bullet|\bullet}\widehat{g} \in \ell_2\mathbb{1}^m_\bullet \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ $\mathbb{P}^n_{\theta|\mathrm{T}}$-a.s. satisfies

$$\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^m-\theta)|^2) = \mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\bullet|\bullet}(\widehat{g}-g))|^2)+\mathbb{P}^k_{\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\bullet|\bullet}g-\theta)|^2) \quad (08.33)$$

with $\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\bullet|\bullet}(\widehat{g}-g))|^2) = n^{-1}\mathbb{P}^k_{\mathrm{T}}\|(\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\bullet|\bullet})^\star\phi^m_\bullet\|^2_{\Gamma_{\theta|\mathrm{T}}}$ (see Property §08|01.45). □

§08|02.58 **Property**. *Under Assumption §08|02.02 we have*

$$\begin{aligned}\mathbb{P}^k_{\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\bullet|\bullet}g-\theta)|^2) &= \mathbb{P}^k_{\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\mathrm{T}}^{m|\dagger}_{\bullet|\bullet}(\mathrm{T}^m_{\bullet|\bullet}-\widehat{\mathrm{T}}^m_{\bullet|\bullet})\theta^m+(\theta^m_\bullet-\theta))|^2\mathbb{1}_{\Omega_{m,k\wedge n}})+|\phi\nu_{\mathbb{N}}(\theta)|^2\mathbb{P}^k_{\mathrm{T}}(\Omega^c_{m,k\wedge n})\\ &\leqslant 2\mathbb{P}^k_{\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\mathrm{T}}^{m|\dagger}_{\bullet|\bullet}(\mathrm{T}^m_{\bullet|\bullet}-\widehat{\mathrm{T}}^m_{\bullet|\bullet})\theta^m)|^2\mathbb{1}_{\Omega_{m,k\wedge n}})+2|\phi\nu_{\mathbb{N}}(\theta^m_\bullet-\theta)|^2+|\phi\nu_{\mathbb{N}}(\theta)|^2\mathbb{P}^k_{\mathrm{T}}(\Omega^c_{m,k\wedge n})\end{aligned}$$

(since $\widehat{\mathrm{T}}^{m|\dagger}_{\bullet|\bullet}\widehat{\mathrm{T}}^m_{\bullet|\bullet}\mathbb{1}_{\Omega_{m,k\wedge n}} = \widehat{\mathrm{T}}^{m|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{\mathrm{T}}^{m|(k\wedge n)}_{\bullet|\bullet} = \mathrm{M}_{\mathbb{1}^m_\bullet}\mathbb{1}_{\Omega_{m,k\wedge n}}$). □

§08|02.59 **Notation** (Reminder). Let $\mathrm{A} \in \mathbb{L}(\ell_2)$ be a *Hilbert-Schmidt operator*, $\mathrm{A} \in \mathbb{HS}(\ell_2)$ for short, where $\|\mathrm{A}\|^2_{\mathrm{HS}} := \mathrm{tr}(\mathrm{A}^\star\mathrm{A}) = \mathrm{tr}(\mathrm{A}\mathrm{A}^\star) \in \mathbb{R}_{\geqslant 0}$. If $\Gamma \in \mathbb{L}(\ell_2)$ then $\mathrm{tr}(\mathrm{A}^\star\Gamma\mathrm{A}) \leqslant \|\Gamma\|_{\mathbb{L}(\ell_2)}\mathrm{tr}(\mathrm{A}^\star\mathrm{A}) = \|\Gamma\|_{\mathbb{L}(\ell_2)}\|\mathrm{A}\|^2_{\mathrm{HS}}$. For arbitrary $\mathrm{A} \in \mathbb{L}(\ell_2)$ we have $\mathrm{M}_{\mathfrak{v}}\mathrm{A}^m = \mathrm{M}^m_{\mathfrak{v}}\mathrm{A}^m \in \mathbb{HS}(\ell_2)$. □

§08|02.60 **Notation**. For each $m \in \mathbb{N}$ and $\mathrm{T}_{\bullet|\bullet} \in \mathbb{L}^{\mathbb{R}}(\ell_2)$ we consider the observable event and its complement

$$\Omega_{m,k\wedge n} := \{\|([\widehat{\mathrm{T}}_{\bullet|\bullet}]^{-1}_{\underline{m}})^\star[\phi]_{\underline{m}}\|^2 \leqslant k\wedge n\} \quad \text{and} \quad \Omega^c_{m,k\wedge n} := \{\|([\widehat{\mathrm{T}}_{\bullet|\bullet}]^{-1}_{\underline{m}})^\star[\phi]_{\underline{m}}\|^2 > k\wedge n\}. \quad (08.34)$$

On the event $\Omega_{m,k\wedge n}$ the random matrix $[\widehat{\mathrm{T}}_{\bullet|\bullet}]_{\underline{m}} \in \mathbb{R}^{(m,m)}$ is regular with inverse $[\widehat{\mathrm{T}}_{\bullet|\bullet}]^{-1}_{\underline{m}} \in \mathbb{R}^{(m,m)}$. Moreover, setting $\mathrm{A}^m_{\bullet|\bullet} := \dot{\boldsymbol{\eta}}^m_{\bullet|\bullet}\mathrm{T}^{m|\dagger}_{\bullet|\bullet}$ we introduce an unobserved event and its complement

$$\mho_{m,k} := \{4m\|\mathrm{A}^m_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)} \leqslant k\} \quad \text{and} \quad \mho^c_{m,k} := \{4m\|\mathrm{A}^m_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)} > k\}. \quad (08.35)$$

Note that $\mathbb{1}_{\mho_{m,k}} = \mathbb{1}_{\{4m\|\mathrm{A}^m_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)}\leqslant k\}}$ denotes an unobserved elementary random variable. □

§08|02.61 **Lemma**. *Under Assumptions §08|02.02 and §08|02.56 for all $m,k,n \in \mathbb{N}$ we have*

(i) *if $4\|(\mathrm{T}^{m|\dagger}_{\bullet|\bullet})^\star\phi^m_\bullet\|^2_{\ell_2} \leqslant k\wedge n$ then $\mho_{m,k} \subseteq \Omega_{m,k\wedge n}$,*

(ii) $\mathbb{P}^k_{\mathrm{T}}(\|(\widehat{\mathrm{T}}^{m|\dagger}_{\bullet|\bullet})^\star\phi^m_\bullet\|^2_{\Gamma_{\theta|\mathrm{T}}}\mathbb{1}_{\Omega_{m,k\wedge n}}) \leqslant \mathrm{v}_{\theta|\mathrm{T}}(4\|(\mathrm{T}^{m|\dagger}_{\bullet|\bullet})^\star\phi^m_\bullet\|^2_{\ell_2}+(k\wedge n)\mathbb{P}^k_{\mathrm{T}}(\mho^c_{m,k}))$, *and*

(iii) $\mathbb{P}^k_{\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\mathrm{T}}^{m|\dagger}_{\bullet|\bullet}(\mathrm{T}^m_{\bullet|\bullet}-\widehat{\mathrm{T}}^m_{\bullet|\bullet})\theta^m_\bullet)|^2\mathbb{1}_{\Omega_{m,k\wedge n}}) \leqslant 4k^{-1}\mathrm{v}_{\mathrm{T}}\|(\mathrm{T}^{m|\dagger}_{\bullet|\bullet})^\star\phi^m_\bullet\|^2_{\ell_2}\|\theta^m_\bullet\|^2_{\ell_2}+\mathbb{P}^k_{\mathrm{T}}(\|\dot{\boldsymbol{\eta}}^m_{\bullet|\bullet}\theta^m_\bullet\|^2_{\ell_2}\mathbb{1}_{\mho^c_{m,k}})$.

with $\Omega_{m,k\wedge n}$ and $\mho_{m,k}$ as in (08.34) and (08.35), respectively.

§08|02.62 **Proof** of Lemma §08|02.61. Given in the lecture. □

§08|02.63 **Proposition** (Upper bound). *Under Assumptions §08|02.02 and §08|02.56 for all $n,m \in \mathbb{N}$ with (generalised) Galerkin solution $\theta^m_\bullet = \mathrm{T}^{m|\dagger}_{\bullet|\bullet}g \in \ell_2\mathbb{1}^m_\bullet \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ setting similar to (08.11)*

$$\begin{aligned}\mathrm{R}^m_{n\wedge k}(\theta,\mathrm{T}_{\bullet|\bullet},\phi) &:= |\phi\nu_{\mathbb{N}}(\theta^m_\bullet-\theta)|^2+(n\wedge k)^{-1}\|(\mathrm{T}^{m|\dagger}_{\bullet|\bullet})^\star\phi^m_\bullet\|^2_{\ell_2},\\ m^\circ_{n\wedge k} &:= \arg\min\left\{\mathrm{R}^m_{n\wedge k}(\theta,\mathrm{T}_{\bullet|\bullet},\phi) : m \in \mathbb{N}\right\} \quad \textit{and}\\ \mathrm{R}^\circ_{n\wedge k}(\theta,\mathrm{T}_{\bullet|\bullet},\phi) &:= \mathrm{R}^{m^\circ_{n\wedge k}}_{n\wedge k}(\theta,\mathrm{T}_{\bullet|\bullet},\phi) = \min\left\{\mathrm{R}^m_{n\wedge k}(\theta,\mathrm{T}_{\bullet|\bullet},\phi) : m \in \mathbb{N}\right\} \quad (08.36)\end{aligned}$$

for all $m \in \mathbb{N}$ *the (generalised) tGE* $\widehat{\theta}^m_{\bullet} = \widehat{T}^{m|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in \ell_2\mathbb{1}^m_{\bullet} \subseteq \operatorname{dom}(\phi\nu_{\mathbb{N}})$ $\mathbb{P}^{n\otimes k}_{\theta|T}$-a.s. *satisfies*

$$\mathbb{P}^{n\otimes k}_{\theta|T}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^m_{\bullet} - \theta_{\bullet})|^2) \leqslant (4\mathbb{v}_{\theta|T} + 8\mathbb{v}_T\|\theta^m_{\bullet}\|^2_{\ell_2})\,\mathrm{R}^m_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \phi) + \mathbb{v}_{\theta|T}\mathbb{P}^k_T(\mho^c_{m,k}) + 2\mathbb{P}^k_T(\|\dot{\eta}^m_{\bullet|\bullet}\theta^m_{\bullet}\|^2_{\ell_2}\mathbb{1}_{\mho^c_{m,k}}) + |\phi\nu_{\mathbb{N}}(\theta_{\bullet})|^2\mathbb{P}^k_T(\Omega^c_{m,k\wedge n}) \quad (08.37)$$

with $\Omega_{m,k\wedge n}$ *and* $\mho_{m,k}$ *as in* (08.34) *and* (08.35)*, respectively.*

§08|02.64 **Proof** of Proposition §08|02.63. Given in the lecture. □

§08|02.65 **Corollary**. *Under the assumptions of Proposition §08|02.63 the (infeasible, generalised) tGE* $\widehat{\theta}^{m^\circ_{n\wedge k}}_{\bullet} = \widehat{T}^{m^\circ_{n\wedge k}|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in \ell_2\mathbb{1}^{m^\circ_{n\wedge k}}_{\bullet} \subseteq \operatorname{dom}(\phi\nu_{\mathbb{N}})$ $\mathbb{P}^{n\otimes k}_{\theta|T}$-a.s. *with* $\Omega_{m^\circ_{n\wedge k},k\wedge n}$ *as in* (08.34) *and (infeasible) dimension* $m^\circ_{n\wedge k}$ *as in* (08.36) *for each* $k, n \in \mathbb{N}$ *with* $\mathrm{R}^\circ_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \phi) \leqslant 1/4$ *satisfies*

$$\mathbb{P}^{n\otimes k}_{\theta|T}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m^\circ_{n\wedge k}}_{\bullet} - \theta_{\bullet})|^2) \leqslant (4\mathbb{v}_{\theta|T} + 8\mathbb{v}_T\|\theta^{m^\circ_{n\wedge k}}_{\bullet}\|^2_{\ell_2})\,\mathrm{R}^\circ_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \phi) + 2^{2l}\mathrm{K}^{2l}_T\Big\{\big(\mathbb{v}_{\theta|T} + |\phi\nu_{\mathbb{N}}(\theta_{\bullet})|^2\big)\,k^{-1}m^\circ_{n\wedge k}\|T^{m^\circ_{n\wedge k}|\dagger}_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)} + \|\theta^{m^\circ_{n\wedge k}}_{\bullet}\|^2_{\mathbb{L}(\ell_2)}/4\Big\} \times (m^\circ_{n\wedge k})^2\big(k^{-1}(m^\circ_{n\wedge k})^3\|T^{m^\circ_{n\wedge k}|\dagger}_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)}\big)^{l-1} \quad (08.38)$$

and if in addition

$$(m^\circ_{n\wedge k})^2\big(k^{-1}(m^\circ_{n\wedge k})^3\|T^{m^\circ_{n\wedge k}|\dagger}_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)}\big)^{l-1} \leqslant \mathrm{R}^\circ_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \phi) \leqslant 1/4 \quad (08.39)$$

then we have

$$\mathbb{P}^{n\otimes k}_{\theta|T}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m^\circ_{n\wedge k}}_{\bullet} - \theta_{\bullet})|^2) \leqslant \mathrm{R}^\circ_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \phi) \times \Big\{(4 + 2^{2l}\mathrm{K}^{2l}_T(m^\circ_{n\wedge k})^{-2})\mathbb{v}_{\theta|T} + (8\mathbb{v}_T + 2^{2l-2}\mathrm{K}^{2l}_T)\|\theta^{m^\circ_{n\wedge k}}_{\bullet}\|^2_{\ell_2} + 2^{2l}\mathrm{K}^{2l}_T(m^\circ_{n\wedge k})^{-2}|\phi\nu_{\mathbb{N}}(\theta_{\bullet})|^2\Big\} \leqslant 2^{2l+2}\mathrm{K}^{2l}_T\big(\mathbb{v}_{\theta|T} + \|\theta^{m^\circ_{n\wedge k}}_{\bullet}\|^2_{\ell_2} + |\phi\nu_{\mathbb{N}}(\theta_{\bullet})|^2\big)\,\mathrm{R}^\circ_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \phi) \quad (08.40)$$

§08|02.66 **Proof** of Corollary §08|02.65. Given in the lecture. □

§08|02.67 **Remark**. Consider $m^\circ_{n\wedge k} = \arg\min\big\{\mathrm{R}^m_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \phi) : m \in \mathbb{N}\big\}$ and $\mathrm{R}^\circ_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \phi) = \mathrm{R}^{m^\circ_{n\wedge k}}_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \phi)$ as in (08.36). Arguing similarly as in Remark §07|01.21 we note that $\|(T^{m|\dagger}_{\bullet|\bullet})^\star\phi^m_{\bullet}\|^2_{\ell_2} \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$ and hence $\mathrm{R}^\circ_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \phi) = \mathrm{o}(1)$ as $n \wedge k \to \infty$, whenever $|\phi\nu_{\mathbb{N}}(\theta^m_{\bullet} - \theta_{\bullet})| = \mathrm{o}(1)$ as $m \to \infty$ (c.f. Remark §05|01.05). In this situation if $\sup\{\|\theta^m_{\bullet}\|^2_{\ell_2} : m \in \mathbb{N}\} \leqslant \mathrm{K}^2_{\theta|T} \in \mathbb{R}_{\geqslant 0}$ then from (08.40) in Corollary §08|02.65 follows

$$\mathbb{P}^{n\otimes k}_{\theta|T}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m^\circ_{n\wedge k}}_{\bullet} - \theta_{\bullet})|^2) \leqslant 2^{2l+2}\mathrm{K}^{2l}_T\big(\mathbb{v}_{\theta|T} + \mathrm{K}^2_{\theta|T} + |\phi\nu_{\mathbb{N}}(\theta_{\bullet})|^2\big)\,\mathrm{R}^\circ_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \phi).$$

However, the dimension $m^\circ_{n\wedge k} = m^\circ_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \phi)$ as defined in (08.11) depends on the unknown parameter of interest $\theta_{\bullet}$ and the nuissance parameter $T_{\bullet|\bullet}$, and thus also the statistic $\widehat{\theta}^{m^\circ_{n\wedge k}}_{\bullet}$. In other words $\widehat{\theta}^{m^\circ_{n\wedge k}}_{\bullet}$ is not a feasible estimator. □

§08|02.68 **Corollary** (GniSM with noisy operator §08|02.08 continued). *Consider* independent *noisy versions* $(\widehat{g}_{\bullet}, \widehat{T}_{\bullet|\bullet}) = (g_{\bullet} + n^{-1/2}\dot{B}_{\bullet}, T_{\bullet|\bullet} + k^{-1/2}\dot{W}_{\bullet|\bullet}) \sim \mathrm{N}^{n\otimes k}_{\theta|T} = \mathrm{N}^n_{\theta|T}\otimes\mathrm{N}^k_T$ *as in Model §08|02.08, where* $\dot{B}_{\bullet} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$ *and* $\dot{W}_{\bullet|\bullet} \sim \mathrm{N}^{\otimes\mathbb{N}^2}_{(0,1)}$ *are* independent, $T_{\bullet|\bullet} \in \mathbb{T}$ *and* $\theta_{\bullet} \in \ell_2$, *and hence* $g_{\bullet} = T_{\bullet|\bullet}\theta_{\bullet} \in \operatorname{dom}(T^\dagger_{\bullet|\bullet}) \subseteq \ell_2$. *Given Assumption §08|02.18 for each* $k, n \in \mathbb{N}$ *fulfilling* (08.23) *the (infeasible, generalised) tGE* $\widehat{\theta}^{m^\circ_{n\wedge k}}_{\bullet} = \widehat{T}^{m^\circ_{n\wedge k}|(k\wedge n)|\dagger}_{\bullet|\bullet}\widehat{g}_{\bullet} \in \ell_2\mathbb{1}^{m^\circ_{n\wedge k}}_{\bullet} \subseteq \operatorname{dom}(\phi\nu_{\mathbb{N}})$ *satisfies*

$$\mathrm{N}^{n\otimes k}_{\theta|T}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m^\circ_{n\wedge k}}_{\bullet} - \theta_{\bullet})|^2) \leqslant 2^{2l+2}((2l-1)!!)\big(1 + \|\theta^{m^\circ_{n\wedge k}}_{\bullet}\|^2_{\ell_2} + |\phi\nu_{\mathbb{N}}(\theta_{\bullet})|^2\big)\,\mathrm{R}^\circ_{n\wedge k}(\theta_{\bullet}, T_{\bullet|\bullet}, \phi)$$

where $\mathrm{R}^\circ_n(\theta_{\bullet}, T_{\bullet|\bullet}, \phi)$ *is the oracle rate in a* GniSM *§08|01.04 (see Corollary §08|01.51).*

§08|02.69 **Proof** of Corollary §08|02.68. Given in the lecture. □

§08|02.70 **Corollary** (niSM with noisy operator §08|02.10 continued). *Consider independent noisy versions* $(\widehat{g}, \widehat{\mathfrak{s}}) = (g + n^{-1/2}\dot{\varepsilon}, \mathrm{T} + k^{-1/2}\dot{\eta})\sim \mathrm{P}^{n\otimes k}_{\theta|\mathrm{T}|\sigma|\xi|\xi^{(2l)}}$ *as in Model §08|02.10, where* $\dot{\varepsilon}$ *and* $\dot{\eta}$ *satisfies* **(iSM1)** *with* $\mathrm{v}_\sigma = \|\sigma^2\|_{\ell_\infty} \vee 1$ *and* **(niSMnO1)**–**(niSMnO2)** *with* $\mathrm{K}^{2l}_{\xi^{(2l)}} := 1 \vee \|\xi^{(2l)}\|_{\ell_\infty(\mathbb{N}^2)}$, *respectively,* $\mathrm{T} \in \mathbb{T}$ *and* $\theta \in \ell_2$, *and hence* $g = \mathrm{T}\theta \in \mathrm{dom}(\mathrm{T}^\dagger) \subseteq \ell_2$. *Given Assumption §08|02.56 for each* $k, n \in \mathbb{N}$ *fulfilling* (08.39) *the (infeasible, generalised) tGE* $\widehat{\theta}^{m^\circ_{n\wedge k}} = \widehat{\mathrm{T}}^{m^\circ_{n\wedge k}|(k\wedge n)|\dagger}\widehat{g} \in \ell_2 \mathbb{1}^{m^\circ_{n\wedge k}} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *satisfies*

$$\mathrm{P}^{n\otimes k}_{\theta|\mathrm{T}|\sigma|\xi|\xi^{(2l)}}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m^\circ_{n\wedge k}} - \theta)|^2) \leqslant 2^{2l+2}\mathrm{K}^{2l}_{\xi^{(2l)}}\left(\mathrm{v}_\sigma + \|\theta^{m^\circ_{n\wedge k}}\|^2_{\ell_2} + |\phi\nu_{\mathbb{N}}(\theta)|^2\right)\mathrm{R}^\circ_{n\wedge k}(\theta, \mathrm{T}, \phi)$$

where $\mathrm{R}^\circ_n(\theta, \mathrm{T}, \phi)$ *is the oracle rate in a* niSM *§08|01.06 (see Corollary §08|01.53).*

§08|02.71 **Proof** of Corollary §08|02.70. Given in the lecture. □

§08|02.72 **Corollary** (nieMM with noisy operator §08|02.13 continued). *Consider independent noisy versions* $(\widehat{g}, \widehat{\mathrm{T}}) = (g + n^{-1/2}\dot{\varepsilon}, \mathrm{T} + k^{-1/2}\dot{\eta})$ *defined on* $(\mathcal{Z}^{n+k}, \mathscr{Z}^{\otimes(n+k)}, \mathrm{P}^{n\otimes k}_{\theta|\mathrm{T}})$ *as in Model §08|02.13, where* $\psi \in \mathcal{M}(\mathscr{Z}\otimes 2^{\mathbb{N}})$ *and* $\varphi \in \mathcal{M}(\mathscr{Z}\otimes 2^{\mathbb{N}^2})$ *satisfies* **(nieMM1)-(nieMM2)** *for* $\mathrm{v}_{\theta|\mathrm{T}|\psi} \in \mathbb{R}_{\geqslant 1}$ *and* **(nieMMnO1)-(nieMMnO3)** *for* $\mathrm{K}_{\mathrm{T}|\varphi} \in \mathbb{R}_{\geqslant 1}$, *respectively,* $\mathrm{T} \in \mathbb{T}$ *and* $\theta \in \ell_2$, *and hence* $g = \mathrm{T}\theta \in \mathrm{dom}(\mathrm{T}^\dagger) \subseteq \ell_2$. *Given Assumption §08|02.56 for each* $k, n \in \mathbb{N}$ *fulfilling* (08.39) *the (infeasible, generalised) tGE* $\widehat{\theta}^{m^\circ_{n\wedge k}} = \widehat{\mathrm{T}}^{m^\circ_{n\wedge k}|(k\wedge n)|\dagger}\widehat{g} \in \ell_2\mathbb{1}^{m^\circ_{n\wedge k}} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *satisfies*

$$\mathbb{P}^{n\otimes k}_{\theta|\mathrm{T}}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m^\circ_{n\wedge k}} - \theta)|^2) \leqslant \mathrm{C}_{2l}\mathrm{K}^{2l}_{\mathrm{T}|\varphi}\left(\mathrm{v}_{\theta|\mathrm{T}|\psi} + \|\theta^{m^\circ_{n\wedge k}}\|^2_{\ell_2} + |\phi\nu_{\mathbb{N}}(\theta)|^2\right)\mathrm{R}^\circ_{n\wedge k}(\theta, \mathrm{T}, \phi)$$

where $\mathrm{C}_{2l} \in \mathbb{R}_{\geqslant 1}$ *is a constant depending on* $l \in \mathbb{N}$ *only and* $\mathrm{R}^\circ_n(\theta, \mathrm{T}, \phi)$ *is the oracle rate in a* nieMM *§08|01.08 (see Corollary §08|01.55).*

§08|02.73 **Proof** of Corollary §08|02.72. Given in the lecture. □

§08|02.74 **Illustration**. We distinguish as in Illustration §08|01.57 the two cases **(p)** and **(np)**, where **(p)** is implied by $\theta\mathbb{1}^{K|\perp} = 0$. In case **(p)** the oracle bound is parametric, that is, $n\mathrm{R}^\circ_n(\theta, \mathrm{T}, \phi) = \mathrm{O}(1)$, in case **(np)** the oracle bound is nonparametric, i.e. $\lim_{n\to\infty} n\mathrm{R}^\circ_n(\theta, \mathrm{T}, \phi) = \infty$. In case **(np)** consider similar to **(o-m)**, **(o-s)** and **(s-m)** in Illustration §08|01.57 the following specifications:

Table 09 [§08]

Order of the rate $\mathrm{R}^\circ_{n\wedge k}(\theta, \mathrm{T}, \phi)$ as $n\wedge k \to \infty$

	$(m\in\mathbb{N})$ $(\phi_m = m^{\mathrm{v}-1/2})$	(squarred bias) $\lvert\phi\nu_{\mathbb{N}}(\widehat{\theta}^m - \theta)\rvert^2$ $(\mathrm{a}\in\mathbb{R}_{>0})$	(variance) $\lVert(\mathrm{T}^{m\lvert\dagger})^\star\phi^m\rVert^2_{\ell_2}$ $(\mathrm{t}\in\mathbb{R}_{>0})$	$m^\circ_{n\wedge k}$	$\mathrm{R}^\circ_{n\wedge k}(\theta, \mathrm{T}, \phi)$
(o-m)	$\mathrm{v}\in(-\mathrm{t},\mathrm{a})$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2\mathrm{t}+2\mathrm{v}}$	$(n\wedge k)^{\frac{1}{2\mathrm{a}+2\mathrm{t}}}$	$(n\wedge k)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{a}+\mathrm{t}}}$
	$\mathrm{v}=-\mathrm{t}$	$m^{-2(\mathrm{a}+\mathrm{t})}$	$\log m$	$\left(\frac{n\wedge k}{\log n\wedge k}\right)^{\frac{1}{2(\mathrm{a}+\mathrm{t})}}$	$\frac{\log n\wedge k}{n\wedge k}$
(o-s)	$\mathrm{a}-\mathrm{v}\in\mathbb{R}_{>0}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2(\mathrm{v}-\mathrm{t})_+}e^{m^{2\mathrm{t}}}$	$(\log n\wedge k)^{\frac{1}{2\mathrm{t}}}$	$(\log n\wedge k)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v}+\mathrm{t}\in\mathbb{R}_{>0}$	$m^{(1-4\mathrm{a}+2\mathrm{v})_+}e^{-m^{2\mathrm{a}}}$	$m^{2\mathrm{t}+\mathrm{v}}$	$(\log n\wedge k)^{\frac{1}{2\mathrm{a}}}$	$\frac{(\log n\wedge k)^{\frac{\mathrm{t}+\mathrm{v}}{\mathrm{a}}}}{n\wedge k}$
	$\mathrm{v}=-\mathrm{t}$	$m^{(1-4\mathrm{a}-2\mathrm{t})_+}e^{-m^{2\mathrm{a}}}$	$\log m$	$(\log n\wedge k)^{\frac{1}{2\mathrm{a}}}$	$\frac{\log\log n\wedge k}{n\wedge k}$

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v} < -\mathrm{t}$ the rate $\mathrm{R}^\circ_{n\wedge k}(\theta, \mathrm{T}, \phi)$ is parametric. The tGE attains the rate $\mathrm{R}^\circ_{n\wedge k} := \mathrm{R}^\circ_{n\wedge k}(\theta, \mathrm{T}, \phi)$ due to Corollary §08|02.65 under the additional condition

$$\left(k^{-1}(m^\circ_{n\wedge k})^3\|\mathrm{T}^{m^\circ_{n\wedge k}|\dagger}\|^2_{\mathbb{L}(\ell_2)}\right)^{l-1} \leqslant (m^\circ_{n\wedge k})^{-2}\mathrm{R}^\circ_{n\wedge k}(\theta, \mathrm{T}, \phi). \qquad (08.41)$$

Since $(m^\circ_{n\wedge k})^{-2}\mathrm{R}^\circ_{n\wedge k}(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet) = \mathrm{o}(1)$ also $k^{-1}(m^\circ_{n\wedge k})^3\|\mathrm{T}^{m^\circ_{n\wedge k}|\dagger}_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)} = \mathrm{o}(1)$ is necessary as $n\wedge k \to \infty$. The next table depicts the order of both terms in case **(o-m)**, **(o-s)** and **(s-m)**.

Table 10 [§08]

Order as $n \wedge k \to \infty$

	(o-m) $\mathrm{v} \in (-\mathrm{t}, \mathrm{a})$	**(o-s)** $\mathrm{a} - \mathrm{v} \in \mathbb{R}_{>0}$	**(s-m)** $\mathrm{v} + \mathrm{t} \in \mathbb{R}_{>0}$
$(m^\circ_{n\wedge k})^{-2}\mathrm{R}^\circ_{n\wedge k}(\theta_\bullet, \mathrm{T}_{\bullet|\bullet}, \phi_\bullet)$	$(n\wedge k)^{-\frac{2(\mathrm{a}-\mathrm{v})+2}{2\mathrm{a}+2\mathrm{t}}}$	$(\log n\wedge k)^{-\frac{2\mathrm{a}-2\mathrm{v}+2}{2\mathrm{t}}}$	$\frac{(\log n\wedge k)^{\frac{2\mathrm{t}+2\mathrm{v}-2}{2\mathrm{a}}}}{n\wedge k}$
$(n\wedge k)^{-1}(m^\circ_{n\wedge k})^3\|\mathrm{T}^{m^\circ_{n\wedge k}|\dagger}_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)}$	$(n\wedge k)^{-\frac{2\mathrm{a}-3}{2\mathrm{a}+2\mathrm{t}}}$	$(n\wedge k)^{-c}$	$\frac{(\log n\wedge k)^{\frac{2\mathrm{t}+3}{2\mathrm{a}}}}{n\wedge k}$

In case **(o-s)** a value $l \geqslant 2$ and **(s-m)** a value $l \geqslant 3$ is sufficient to ensure (08.41) as $n\wedge k \to \infty$. In case **(o-m)** assuming $\mathrm{a} > 3/2$ we have $k^{-1}(m^\circ_{n\wedge k})^3\|\mathrm{T}^{m^\circ_{n\wedge k}|\dagger}_{\bullet|\bullet}\|^2_{\mathbb{L}(\ell_2)} = \mathrm{o}(1)$ as $n\wedge k\to\infty$. In this situation we have (08.41) if $(2\mathrm{a}-3)(l-1) > 2(\mathrm{a}-\mathrm{v})+2$ or in equal $l > (4\mathrm{a}-2\mathrm{v}-1)/(2\mathrm{a}-3)$. □

§08|02|03|02 Maximal local ϕ-risk

§08|02.75 **Assumption**. Consider weights $\mathfrak{t}_\bullet, \mathfrak{a}_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ and $\phi_\bullet \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ such that $(\mathfrak{a}\phi)_\bullet := \mathfrak{a}_\bullet\phi_\bullet \in \ell_2$ and $(\mathfrak{at})_\bullet := \mathfrak{a}_\bullet\mathfrak{t}_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$. □

§08|02.76 **Comment**. Assuming $\mathfrak{t}_\bullet, \mathfrak{a}_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ and hence $(\mathfrak{at})^2_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ is rather weak. If in addition $\liminf_{j\to\infty} (\mathfrak{at})^2_j \geqslant c \in \mathbb{R}_{>0}$ is satisfied, and hence $(\mathfrak{at})^2_\bullet, \mathfrak{a}^2_\bullet, \mathfrak{t}^2_\bullet \notin (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$, then $\mathfrak{a}^2_\bullet \notin (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$ and the assumption $(\mathfrak{a}\phi)_\bullet \in \ell_2$ implies $\phi_\bullet \in \ell_2$, which together with $\mathfrak{t}^2_\bullet \notin (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$ implies $(\phi/\mathfrak{t})_\bullet \in \ell_2$, and thus the rate $\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$ is parametric (Illustration §08|01.72). Since we are interested in the case of a non-parametric rate, the additional assumption $(\mathfrak{at})^2_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$ imposes a rather weak condition satisfied also in Illustration §08|01.72. □

§08|02.77 **Reminder**. Under Assumption §08|02.75 we have $\ell_2^{\mathfrak{a}} = \mathrm{dom}(\mathrm{M}_{\mathfrak{a}^{-1}}) = \ell_2\mathfrak{a}_\bullet \subseteq \ell_2$ and the three measures $\nu_{\mathbb{N}}$, $\mathfrak{a}_\bullet^{-2}\nu_{\mathbb{N}}$ and $|\phi_\bullet|\nu_{\mathbb{N}}$ dominate mutually each other, i.e. they share the same null sets (see Property §04|01.02). We consider $\ell_2^{\mathfrak{a}}$ endowed with $\|\cdot\|_{\mathfrak{a}^{-1}} = \|\mathrm{M}_{\mathfrak{a}^{-1}}\cdot\|_{\ell_2}$ and given a constant $\mathrm{r} \in \mathbb{R}_{>0}$ the ellipsoid $\ell_2^{\mathfrak{a},\mathrm{r}} := \{a_\bullet \in \ell_2^{\mathfrak{a}} : \|a_\bullet\|_{\mathfrak{a}^{-1}} \leqslant \mathrm{r}\} \subseteq \ell_2^{\mathfrak{a}}$. Since $(\mathfrak{a}\phi)_\bullet \in \ell_2$ we have $\ell_2^{\mathfrak{a}} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ (Property §04|02.23). Consequently, if Assumption §08|02.75 and $\theta_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}$ are satisfied, then Assumption §08|02.56 is also fulfilled. Moreover, from $(\mathfrak{a}\phi)_\bullet \in \ell_2$ follows $\|\mathfrak{a}_\bullet\mathbb{1}^{m|\perp}_\bullet\|_\phi = \|(\mathfrak{a}\phi)_\bullet\mathbb{1}^{m|\perp}_\bullet\|_{\ell_2} = \mathrm{o}(1)$ as $m\to\infty$. For $s \in [0,1]$ from $(\mathfrak{at}^s)_\bullet = \mathfrak{a}_\bullet\mathfrak{t}^s_\bullet \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ follows $(\mathfrak{at}^s)_{(\bullet)} = ((\mathfrak{at}^s)_{(m)} := (\mathfrak{at}^s)_{m+1} = \|(\mathfrak{at}^s)_\bullet\mathbb{1}^{m|\perp}_\bullet\|_{\ell_\infty})_{m\in\mathbb{N}} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$. Since $\phi_\bullet, \mathfrak{t}_\bullet \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ under Assumption §08|02.75, we have $\ell_2\mathbb{1}^m_\bullet \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ and $\|\mathfrak{t}_\bullet^{-1}\mathbb{1}^m_\bullet\|_\phi = \|(\phi/\mathfrak{t})_\bullet\mathbb{1}^m_\bullet\|_{\ell_2} \in \mathbb{R}_{\geqslant 0}$ for each $m\in\mathbb{N}$. Under the Assumptions §08|00.02 and §08|02.75 considering the generalised link condition $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}$ with band $\mathrm{D} \in \mathbb{R}_{\geqslant 1}$ and $\mathrm{d} \in [1, \mathrm{D}]$ as in Definition §05|02.05 we have $\sup_{m\in\mathbb{N}}\{\|([\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}})^\star[\mathrm{M}_{\mathfrak{t}}]_{\underline{m}}\|_{\mathrm{spec}}\} \leqslant \mathrm{D}$, and hence

$$\|(\mathrm{T}^{m|\dagger}_{\bullet|\bullet})^\star\phi^m_\bullet\|_{\ell_2} = \|([\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}})^\star[\phi]_{\underline{m}}\| = \|([\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}})^\star[\mathrm{M}_{\mathfrak{t}}]_{\underline{m}}[\mathrm{M}_{\mathfrak{t}}]^{-1}_{\underline{m}}[\phi]_{\underline{m}}\| \leqslant \|([\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}})^\star[\mathrm{M}_{\mathfrak{t}}]_{\underline{m}}\|_{\mathrm{spec}}\|[\mathrm{M}_{\mathfrak{t}^{-1}}]_{\underline{m}}[\phi]_{\underline{m}}\| \leqslant \mathrm{D}\|\mathfrak{t}_\bullet^{-1}\mathbb{1}^m_\bullet\|_\phi \quad (08.42)$$

using $\|[\mathrm{M}_{\mathfrak{t}^{-1}}]_{\underline{m}}[\phi]_{\underline{m}}\| = \|\mathfrak{t}_\bullet^{-1}\mathbb{1}^m_\bullet\|_\phi$. Moreover, for each $m\in\mathbb{N}$ the generalised Galerkin solution $\theta^m_\bullet := \mathrm{T}^{m|\dagger}_{\bullet|\bullet}g_\bullet \in \ell_2\mathbb{1}^m_\bullet$ of $\theta_\bullet = \mathrm{T}^\dagger_{\bullet|\bullet}g_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}$ satisfies (Lemma §05|02.14)

$$|\phi\nu_{\mathbb{N}}(\theta^m_\bullet - \theta_\bullet)|^2 \leqslant \mathrm{Dd}(\mathrm{Dd}+1)\mathrm{r}^2\big(\|\mathfrak{a}_\bullet\mathbb{1}^{m|\perp}_\bullet\|^2_\phi + (\mathfrak{at})^2_{(m)}\|\mathfrak{t}_\bullet^{-1}\mathbb{1}^m_\bullet\|^2_\phi\big). \quad (08.43)$$

Under Assumptions §08|00.02 and §08|02.75 the link condition $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}^{\gtrless}_{\mathfrak{t},\mathrm{d}}$ with band $\mathrm{d} \in \mathbb{R}_{\geqslant 1}$ as in Definition §05|01.08 implies $\sup_{m\in\mathbb{N}}\{\|([\mathrm{T}_{\bullet|\bullet}]^{-1}_{\underline{m}})^\star[\mathrm{M}_{\mathfrak{t}}]_{\underline{m}}\|_{\mathrm{spec}}\} \leqslant 3\mathrm{d}^2$ (Lemma §05|01.22), and hence for each $m \in \mathbb{N}$ we have (08.42) with $\mathrm{D} = 3\mathrm{d}^2$ and the Galerkin solution $\theta_{\bullet}^{m} := \mathrm{T}_{\bullet|\bullet}^{m|\dagger} g_{\bullet} \in \ell_2 \mathbb{1}_{\bullet}^{m}$ of $\theta_{\bullet} = \mathrm{T}_{\bullet|\bullet}^{\dagger} g_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}$ satisfies (08.43) with $\mathrm{D} = 3\mathrm{d}^2$ (Lemma §05|01.34). □

§08|02.78 **Corollary**. *Under Assumptions §08|02.02 and §08|02.75 let* $\theta_{\bullet} := \mathrm{T}_{\bullet|\bullet}^{\dagger} g_{\bullet} \in \ell_2^{\mathfrak{a}_{\bullet},\mathrm{r}}$ *and* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}$ *(or* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d}}^{\gtrless}$ *with* $\mathrm{D} = 3\mathrm{d}^2$*), for all* $m, k, n \in \mathbb{N}$ *we have*

(i) *if* $4\mathrm{D}^2 \|\mathfrak{t}_{\bullet}^{-1} \mathbb{1}_{\bullet}^{m}\|_{\phi}^2 \leqslant k \wedge n$ *then* $\mho_{m,k} \subseteq \Omega_{m,k\wedge n}$,

(ii) $\mathbb{P}_{\mathrm{T}}^{k}\big(\|(\widehat{\mathrm{T}}_{\bullet|\bullet}^{m|\dagger})^{\star} \phi_{\bullet}^{m}\|_{\Gamma_{\theta|\mathrm{T}}}^2 \mathbb{1}_{\Omega_{m,k\wedge n}}\big) \leqslant \mathbb{v}_{\theta|\mathrm{T}}(4\mathrm{D}^2 \|\mathfrak{t}_{\bullet}^{-1}\mathbb{1}_{\bullet}^{m}\|_{\phi}^2 + (k\wedge n)\mathbb{P}_{\mathrm{T}}^{k}(\mho_{m,k}^{c}))$, *and*

(iii) $\mathbb{P}_{\mathrm{T}}^{k}(|\phi\nu_{\mathbb{N}}(\widehat{\mathrm{T}}_{\bullet|\bullet}^{m|\dagger}(\mathrm{T}_{\bullet|\bullet}^{m} - \widehat{\mathrm{T}}_{\bullet|\bullet}^{m})\theta_{\bullet}^{m})|^2 \mathbb{1}_{\Omega_{m,k\wedge n}}) \leqslant 4k^{-1}\mathbb{v}_{\mathrm{T}}\mathrm{D}^2 \|\mathfrak{t}_{\bullet}^{-1}\mathbb{1}_{\bullet}^{m}\|_{\phi}^2 \|\theta_{\bullet}^{m}\|_{\ell_2}^2 + \mathbb{P}_{\mathrm{T}}^{k}(\|\dot{\boldsymbol{\eta}}_{\bullet|\bullet}^{m}\theta_{\bullet}^{m}\|_{\ell_2}^2 \mathbb{1}_{\mho_{m,k}^{c}})$.

with $\Omega_{m,k\wedge n}$ *and* $\mho_{m,k}$ *as in* (08.34) *and* (08.35)*, respectively.*

§08|02.79 **Proof** of Corollary §08|02.78. Given in the lecture. □

§08|02.80 **Proposition** (Upper bound). *Under Assumptions §08|02.02 and §08|02.75 let* $\theta_{\bullet} := \mathrm{T}_{\bullet|\bullet}^{\dagger} g_{\bullet} \in \ell_2^{\mathfrak{a}_{\bullet},\mathrm{r}}$ *and* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d},\mathrm{D}}$ *(or* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathfrak{t},\mathrm{d}}^{\gtrless}$ *with* $\mathrm{D} = 3\mathrm{d}^2$*) for all* $n, m \in \mathbb{N}$ *setting similar to* (08.15)

$$\mathrm{R}_{n\wedge k}^{m}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi) := \|\mathfrak{a}_{\bullet}\mathbb{1}_{\bullet}^{m|\perp}\|_{\phi}^2 + (n\wedge k)^{-1}\|\mathfrak{t}_{\bullet}^{-1}\mathbb{1}_{\bullet}^{m}\|_{\phi}^2,$$
$$m_{n\wedge k}^{\star} := \arg\min\left\{\mathrm{R}_{n\wedge k}^{m}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi) : m \in \mathbb{N}\right\} \quad \text{and}$$
$$\mathrm{R}_{n\wedge k}^{\star}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi) := \mathrm{R}_{n\wedge k}^{m_{n\wedge k}^{\star}}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi) = \min\left\{\mathrm{R}_{n\wedge k}^{m}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi) : m \in \mathbb{N}\right\} \quad (08.44)$$

for all $m \in \mathbb{N}$ *the (generalised) tGE* $\widehat{\theta}_{\bullet}^{m} = \widehat{\mathrm{T}}_{\bullet|\bullet}^{m|(k\wedge n)|\dagger}\widehat{g}_{\bullet} \in \ell_2\mathbb{1}_{\bullet}^{m}$ $\mathbb{P}_{\theta|\mathrm{T}}^{n\otimes k}$-a.s. *satisfies*

$$\mathbb{P}_{\theta|\mathrm{T}}^{n\otimes k}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_{\bullet}^{m} - \theta_{\bullet})|^2) \leqslant 4\mathrm{D}^2((1\vee(\mathfrak{a}\mathfrak{t})_{(m)}^2(n\wedge k))\mathrm{d}^2\mathrm{r}^2 + \mathbb{v}_{\theta|\mathrm{T}} + 2\mathbb{v}_{\mathrm{T}}\mathfrak{a}_1^2\mathrm{D}^2\mathrm{d}^2\mathrm{r}^2)\,\mathrm{R}_{n\wedge k}^{m}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi)$$
$$+ \mathbb{v}_{\theta|\mathrm{T}}\mathbb{P}_{\mathrm{T}}^{k}(\mho_{m,k}^{c}) + 2\mathbb{P}_{\mathrm{T}}^{k}(\|\dot{\boldsymbol{\eta}}_{\bullet|\bullet}^{m}\theta_{\bullet}^{m}\|_{\ell_2}^2\mathbb{1}_{\mho_{m,k}^{c}}) + |\phi\nu_{\mathbb{N}}(\theta_{\bullet})|^2\mathbb{P}_{\mathrm{T}}^{k}(\Omega_{m,k\wedge n}^{c}) \quad (08.45)$$

with $\Omega_{m,k\wedge n}$ *and* $\mho_{m,k}$ *as in* (08.34) *and* (08.35)*, respectively.*

§08|02.81 **Proof** of Proposition §08|02.80. Given in the lecture. □

§08|02.82 **Corollary**. *Under the assumptions of Proposition §08|02.80 for* $k, n \in \mathbb{N}$ *with* $\mathrm{R}_{n\wedge k}^{\star}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi) \leqslant 1/(4\mathrm{D}^2)$ *the (generalised) tGE* $\widehat{\theta}_{\bullet}^{m_{n\wedge k}^{\star}} = \widehat{\mathrm{T}}_{\bullet|\bullet}^{m_{n\wedge k}^{\star}|(k\wedge n)|\dagger}\widehat{g}_{\bullet} \in \ell_2\mathbb{1}_{\bullet}^{m_{n\wedge k}^{\star}} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ $\mathbb{P}_{\theta|\mathrm{T}}^{n\otimes k}$-a.s. *with* $\Omega_{m_{n\wedge k}^{\star},k\wedge n}$ *as in* (08.34) *and dimension* $m_{n\wedge k}^{\star}$ *as in* (08.44) *satisfies*

$$\mathbb{P}_{\theta|\mathrm{T}}^{n\otimes k}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_{\bullet}^{m_{n\wedge k}^{\star}} - \theta_{\bullet})|^2) \leqslant 4\mathrm{D}^2(\mathrm{d}^2\mathrm{r}^2 + \mathbb{v}_{\theta|\mathrm{T}} + 2\mathbb{v}_{\mathrm{T}}\mathfrak{a}_1^2\mathrm{D}^2\mathrm{d}^2\mathrm{r}^2)\,\mathrm{R}_{n\wedge k}^{\star}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi)$$
$$+ 2^{2l}\mathrm{K}_{\mathrm{T}}^{2l}\mathrm{D}^{2(l-1)}\left\{\mathrm{D}^2\left(\mathbb{v}_{\theta|\mathrm{T}} + \|\mathfrak{a}_{\bullet}\|_{\phi}^2\mathrm{r}^2\right)k^{-1}m_{n\wedge k}^{\star}\mathfrak{t}_{m_{n\wedge k}^{\star}}^{-2} + \mathfrak{a}_1^2\mathrm{D}^2\mathrm{d}^2\mathrm{r}^2/4\right\}$$
$$\times (m_{n\wedge k}^{\star})^2(k^{-1}(m_{n\wedge k}^{\star})^3\mathfrak{t}_{m_{n\wedge k}^{\star}}^{-2})^{l-1}. \quad (08.46)$$

and if in addition

$$(m_{n\wedge k}^{\star})^2(k^{-1}(m_{n\wedge k}^{\star})^3\mathfrak{t}_{m_{n\wedge k}^{\star}}^{-2})^{l-1} \leqslant \mathrm{R}_{n\wedge k}^{\star}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi) \leqslant 1/(4\mathrm{D}^2) \quad (08.47)$$

then we have

$$\mathbb{P}_{\theta|\mathrm{T}}^{n\otimes k}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_{\bullet}^{m_{n\wedge k}^{\star}} - \theta_{\bullet})|^2) \leqslant 2^{2l+2}\mathrm{K}_{\mathrm{T}}^{2l}\mathrm{D}^{2l}(\mathbb{v}_{\theta|\mathrm{T}} + \|\mathfrak{a}_{\bullet}\|_{\phi}^2\mathrm{r}^2 + (1\vee\mathfrak{a}_1^2)\mathrm{d}^2\mathrm{r}^2)\mathrm{R}_{n\wedge k}^{\star}(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi). \quad (08.48)$$

§08|02.83 **Proof** of Corollary §08|02.82. Given in the lecture. □

§08|02.84 **Remark**. Arguing similarly as in Remark §07|01.56 we note that $\|\mathrm{t}_{\bullet}^{-1}\mathbb{1}_{\bullet}^{m}\|_{\phi} \in \mathbb{R}_{\geqslant 0}$ for all $m \in \mathbb{N}$ and $(\|\phi\mathbb{1}_{\bullet}^{m|\perp}\|_{\mathfrak{a}_\bullet}^2 = \mathrm{o}(1)$ as $m \to \infty$ (since $(\mathfrak{a}\phi)_\bullet \in \ell_2$), and hence $\mathrm{R}_n^{\star}(\mathfrak{a}_\bullet, \mathrm{t}_\bullet, \phi) = \mathrm{o}(1)$ as $n \to \infty$. If there is in addition $\mathrm{C} \in \mathbb{R}_{\geqslant 1}$ such that $\mathrm{K}_{\mathrm{T}}^{2l} \leqslant \mathrm{C}$ and $\mathbb{v}_{\theta|\mathrm{T}} \leqslant \mathrm{C}$ for all $\theta_\bullet := \mathrm{T}_{\bullet|\bullet}^{\dagger} g_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}$ and $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t},\mathrm{d},\mathrm{D}}$ then from the bound (08.48) Corollary §08|02.82 follows immediately

$$\sup\left\{\mathbb{P}_{\theta|\mathrm{T}}^{n\otimes k}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}_\bullet^{m_{n\wedge k}^{\star}} - \theta_\bullet)|^2) \colon \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t},\mathrm{d},\mathrm{D}}, \theta_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}\right\} \leqslant \mathrm{R}_{n\wedge k}^{\star}(\mathfrak{a}_\bullet, \mathrm{t}_\bullet, \phi) \\ \times 2^{2l+2}\mathrm{C}\mathrm{D}^{2l}(\mathrm{C} + \|\mathfrak{a}_\bullet\|_{\phi}^2\mathrm{r}^2 + (1 \vee \mathfrak{a}_1^2)\mathrm{d}^2\mathrm{r}^2).$$

Note that the dimension $m_{n\wedge k}^{\star} := m_{n\wedge k}^{\star}(\mathfrak{a}_\bullet, \mathrm{t}_\bullet, \phi)$ as defined in (08.44) does not depend on the unknown parameter of interest $\theta_\bullet$ but on the classes $\ell_2^{\mathfrak{a},\mathrm{r}}$ and $\mathbb{T}_{\mathrm{t},\mathrm{d},\mathrm{D}}$ only, and thus also the statistic $\widehat{\theta}_\bullet^{m_{n\wedge k}^{\star}}$. In other words, if the regularity of $\theta_\bullet$ and $\mathrm{T}_{\bullet|\bullet}$ is known in advance, then the thresholded GE $\widehat{\theta}_\bullet^{m_{n\wedge k}^{\star}}$ is a feasible estimator. □

§08|02.85 **Corollary** (GniSM with noisy operator §08|02.08 continued). *Consider* independent *noisy versions* $(\widehat{g}_\bullet, \widehat{\mathrm{T}}_{\bullet|\bullet}) = (g_\bullet + n^{-1/2}\dot{\mathrm{B}}_\bullet, \mathrm{T}_{\bullet|\bullet} + k^{-1/2}\dot{\mathrm{W}}_{\bullet|\bullet}) \sim \mathrm{N}_{\theta|\mathrm{T}}^{n\otimes k} = \mathrm{N}_{\theta|\mathrm{T}}^{n} \otimes \mathrm{N}_{\mathrm{T}}^{k}$ *as in Model §08|02.08, where* $\dot{\mathrm{B}}_\bullet \sim \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}$ *and* $\dot{\mathrm{W}}_{\bullet|\bullet} \sim \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}^2}$ *are* independent, $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}$ *and* $\theta_\bullet \in \ell_2$, *and hence* $g_\bullet = \mathrm{T}_{\bullet|\bullet}\theta_\bullet \in \mathrm{dom}(\mathrm{T}_{\bullet|\bullet}^{\dagger}) \subseteq \ell_2$. *Given Assumption §08|02.75 for each* $k, n \in \mathbb{N}$ *fulfilling* (08.47) *the (generalised) tGE* $\widehat{\theta}_\bullet^{m_{n\wedge k}^{\star}} = \widehat{\mathrm{T}}_{\bullet|\bullet}^{m_{n\wedge k}^{\star}|(k\wedge n)|\dagger}\widehat{g}_\bullet \in \ell_2\mathbb{1}_\bullet^{m_{n\wedge k}^{\star}} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *satisfies*

$$\sup\left\{\mathrm{N}_{\theta|\mathrm{T}}^{n\otimes k}(\|\widehat{\theta}_\bullet^{m_{n\wedge k}^{\star}} - \theta_\bullet\|_{\mathfrak{v}}^2) \colon \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t},\mathrm{d},\mathrm{D}}, \theta_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}\right\} \leqslant \mathrm{R}_{n\wedge k}^{\star}(\mathfrak{a}_\bullet, \mathrm{t}_\bullet, \phi) \\ \times 2^{2l+2}((2l-1)!!)\mathrm{D}^{2l}(1 + \|\mathfrak{a}_\bullet\|_{\phi}^2\mathrm{r}^2 + (1 \vee \mathfrak{a}_1^2)\mathrm{d}^2\mathrm{r}^2)$$

where $\mathrm{R}_n^{\star}(\mathfrak{a}_\bullet, \mathrm{t}_\bullet, \phi)$ *is the rate in a* GniSM *§08|01.04 (see Corollary §08|01.66).*

§08|02.86 **Proof** of Corollary §08|02.85. Given in the lecture. □

§08|02.87 **Corollary** (niSM with noisy operator §08|02.10 continued). *Consider* independent *noisy versions* $(\widehat{g}_\bullet, \widehat{\mathfrak{s}}_\bullet) = (g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet, \mathrm{T}_{\bullet|\bullet} + k^{-1/2}\dot{\eta}_{\bullet|\bullet}) \sim \mathrm{P}_{\theta|\mathrm{T}|\sigma|\xi|\xi^{(2l)}}^{n\otimes k}$ *as in Model §08|02.10, where* $\dot{\varepsilon}_\bullet$ *and* $\dot{\eta}_{\bullet|\bullet}$ *satisfies* **(iSM1)** *with* $\mathbb{v}_{\sigma} = \|\sigma_\bullet^2\|_{\ell_\infty} \vee 1$ *and* **(niSMnO1)**–**(niSMnO2)** *with* $\mathrm{K}_{\xi^{(2l)}}^{2l} := 1 \vee \|\xi_{\bullet|\bullet}^{(2l)}\|_{\ell_\infty(\mathbb{N}^2)}$, *respectively,* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}$ *and* $\theta_\bullet \in \ell_2$, *and hence* $g_\bullet = \mathrm{T}_{\bullet|\bullet}\theta_\bullet \in \mathrm{dom}(\mathrm{T}_{\bullet|\bullet}^{\dagger}) \subseteq \ell_2$. *Given Assumption §08|02.75 for each* $k, n \in \mathbb{N}$ *fulfilling* (08.47) *the (generalised) tGE* $\widehat{\theta}_\bullet^{m_{n\wedge k}^{\star}} = \widehat{\mathrm{T}}_{\bullet|\bullet}^{m_{n\wedge k}^{\star}|(k\wedge n)|\dagger}\widehat{g}_\bullet \in \ell_2\mathbb{1}_\bullet^{m_{n\wedge k}^{\star}} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *satisfies*

$$\sup\left\{\mathrm{P}_{\theta|\mathrm{T}|\sigma|\xi|\xi^{(2l)}}^{n\otimes k}(\|\widehat{\theta}_\bullet^{m_{n\wedge k}^{\star}} - \theta_\bullet\|_{\mathfrak{v}}^2) \colon \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t},\mathrm{d},\mathrm{D}}, \theta_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}\right\} \leqslant \mathrm{R}_{n\wedge k}^{\star}(\mathfrak{a}_\bullet, \mathrm{t}_\bullet, \phi) \\ \times 2^{2l+2}\mathrm{K}_{\xi^{(2l)}}^{2l}\mathrm{D}^{2l}(\mathbb{v}_{\sigma} + \|\mathfrak{a}_\bullet\|_{\phi}^2\mathrm{r}^2 + (1 \vee \mathfrak{a}_1^2)\mathrm{d}^2\mathrm{r}^2)$$

where $\mathrm{R}_n^{\star}(\mathfrak{a}_\bullet, \mathrm{t}_\bullet, \phi)$ *is the rate in a* niSM *§08|01.06 (see Corollary §08|01.68).*

§08|02.88 **Proof** of Corollary §08|02.87. Given in the lecture. □

§08|02.89 **Corollary** (nieMM with noisy operator §08|02.13 continued). *Consider* independent *noisy versions* $(\widehat{g}_\bullet, \widehat{\mathrm{T}}_{\bullet|\bullet}) = (g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet, \mathrm{T}_{\bullet|\bullet} + k^{-1/2}\dot{\eta}_{\bullet|\bullet})$ *defined on* $(\mathcal{Z}^{n+k}, \mathscr{Z}^{\otimes(n+k)}, \mathrm{P}_{\theta|\mathrm{T}}^{n\otimes k})$ *as in Model §08|02.13, where* $\psi_\bullet \in \mathcal{M}(\mathscr{Z} \otimes 2^{\mathbb{N}})$ *and* $\varphi_{\bullet|\bullet} \in \mathcal{M}(\mathscr{Z} \otimes 2^{\mathbb{N}^2})$ *satisfies* **(nieMM1)**-**(nieMM2)** *for* $\mathbb{v}_{\theta|\mathrm{T}|\psi} \in \mathbb{R}_{\geqslant 1}$ *and* **(nieMMnO1)**-**(nieMMnO3)** *for* $\mathrm{K}_{\mathrm{T}|\varphi} \in \mathbb{R}_{\geqslant 1}$, *respectively,* $\mathrm{T}_{\bullet|\bullet} \in \mathbb{T}$ *and* $\theta_\bullet \in \ell_2$, *and hence* $g_\bullet = \mathrm{T}_{\bullet|\bullet}\theta_\bullet \in \mathrm{dom}(\mathrm{T}_{\bullet|\bullet}^{\dagger}) \subseteq \ell_2$. *Given Assumption §08|02.75 for each* $k, n \in \mathbb{N}$ *fulfilling* (08.47) *the (generalised) tGE* $\widehat{\theta}_\bullet^{m_{n\wedge k}^{\circ}} = \widehat{\mathrm{T}}_{\bullet|\bullet}^{m_{n\wedge k}^{\circ}|(k\wedge n)|\dagger}\widehat{g}_\bullet \in \ell_2\mathbb{1}_\bullet^{m_{n\wedge k}^{\circ}} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ *satisfies*

$$\sup\left\{\mathbb{P}_{\theta|\mathrm{T}}^{n\otimes k}(\|\widehat{\theta}_\bullet^{m_{n\wedge k}^{\star}} - \theta_\bullet\|_{\mathfrak{v}}^2) \colon \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t},\mathrm{d},\mathrm{D}}, \theta_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}\right\} \leqslant \mathrm{R}_{n\wedge k}^{\star}(\mathfrak{a}_\bullet, \mathrm{t}_\bullet, \phi) \\ \times \mathrm{C}_{2l}\sup\left\{\mathrm{K}_{\mathrm{T}|\varphi}^{2l} \colon \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t},\mathrm{d},\mathrm{D}}\right\}\mathrm{D}^{2l}(\sup\left\{\mathbb{v}_{\theta|\mathrm{T}|\psi} \colon \mathrm{T}_{\bullet|\bullet} \in \mathbb{T}_{\mathrm{t},\mathrm{d},\mathrm{D}}, \theta_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}\right\} + \|\mathfrak{a}_\bullet\|_{\phi}^2\mathrm{r}^2 + (1 \vee \mathfrak{a}_1^2)\mathrm{d}^2\mathrm{r}^2)$$

where $\mathrm{C}_{2l} \in \mathbb{R}_{\geqslant 1}$ *is a constant depending on* $l \in \mathbb{N}$ *only and* $\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$ *is the rate in a* nieMM *§08|01.08 (see Corollary §08|01.70).*

§08|02.90 **Proof** of Corollary §08|02.89. Given in the lecture. □

§08|02.91 **Illustration**. We distinguish as in Illustration §08|01.72 the two cases **(p)** and **(np)**. Interestingly, in case **(p)** the bound is parametric, that is, $n\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) = \mathrm{O}(1)$, in case **(np)** the bound is nonparametric, i.e. $\lim_{n\to\infty} n\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) = \infty$. In case **(np)** consider similar to **(o-m)**, **(o-s)** and **(s-m)** in Illustration §08|01.72 the following specifications:

Table 11 [§08]

Order of the rate $\mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$ as $n \wedge k \to \infty$

	($j \in \mathbb{N}$) $\phi^2_j = j^{2\mathrm{v}-1}$	($\mathrm{a} \in \mathbb{R}_{>0}$) $\mathfrak{a}^2_j$	($\mathrm{t} \in \mathbb{R}_{>0}$) $\mathfrak{t}^2_j$	(squarred bias) $\Vert\mathfrak{a}_\bullet \mathbb{1}^{m\vert\perp}_\bullet\Vert^2_{\phi_\bullet}$	(variance) $\Vert\mathfrak{t}^{-1}_\bullet \mathbb{1}^{m}_\bullet\Vert^2_{\phi_\bullet}$	$m^\star_{n\wedge k}$	$\mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$
(o-m)	$\mathrm{v} \in (-\mathrm{t}, \mathrm{a})$	$j^{-2\mathrm{a}}$	$j^{-2\mathrm{t}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2\mathrm{v}+2\mathrm{t}}$	$(n \wedge k)^{\frac{1}{2\mathrm{a}+2\mathrm{t}}}$	$(n \wedge k)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{a}+\mathrm{t}}}$
	$\mathrm{v} = -\mathrm{t}$	$j^{-2\mathrm{a}}$	$j^{-2\mathrm{t}}$	$m^{-2(\mathrm{a}+\mathrm{t})}$	$\log m$	$\left(\frac{n\wedge k}{\log n\wedge k}\right)^{\frac{1}{2(\mathrm{a}+\mathrm{t})}}$	$\frac{\log n\wedge k}{n\wedge k}$
(o-s)	$\mathrm{a} - \mathrm{v} \in \mathbb{R}_{>0}$	$j^{-2\mathrm{a}}$	$e^{-j^{2\mathrm{t}}}$	$m^{-2(\mathrm{a}-\mathrm{v})}$	$m^{2(\mathrm{v}-\mathrm{t})_+} e^{m^{2\mathrm{t}}}$	$(\log n \wedge k)^{\frac{1}{2\mathrm{t}}}$	$(\log n \wedge k)^{-\frac{\mathrm{a}-\mathrm{v}}{\mathrm{t}}}$
(s-m)	$\mathrm{v} + \mathrm{t} \in \mathbb{R}_{>0}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$e^{-m^{2\mathrm{a}}}$	$m^{2\mathrm{v}+2\mathrm{t}}$	$(\log n \wedge k)^{\frac{1}{2\mathrm{a}}}$	$\frac{(\log n\wedge k)^{\frac{\mathrm{t}+\mathrm{v}}{\mathrm{a}}}}{n\wedge k}$
	$\mathrm{v} = -\mathrm{t}$	$e^{-j^{2\mathrm{a}}}$	$j^{-2\mathrm{t}}$	$e^{-m^{2\mathrm{a}}}$	$\log m$	$(\log n \wedge k)^{\frac{1}{2\mathrm{a}}}$	$\frac{\log\log n\wedge k}{n\wedge k}$

We note that in case **(o-m)** and **(s-m)** for $\mathrm{v} < -\mathrm{t}$ the rate $\mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$ is parametric. The tGE attains the rate $\mathrm{R}^\star_{n\wedge k} := \mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet)$ due to Corollary §08|02.82 under the additional condition

$$(k^{-1}(m^\star_{n\wedge k})^3 \mathfrak{t}^{-2}_{m^\star_{n\wedge k}})^{l-1} \leqslant (m^\star_{n\wedge k})^{-2} \mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet). \tag{08.49}$$

Since $(m^\star_{n\wedge k})^{-2}\mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi_\bullet) = \mathrm{o}(1)$ also $k^{-1}(m^\star_{n\wedge k})^3 \mathfrak{t}^{-2}_{m^\star_{n\wedge k}} = \mathrm{o}(1)$ is necessary as $n \wedge k \to \infty$. The next table depicts the order of both terms in case **(o-m)**, **(o-s)** and **(s-m)**.

Table 12 [§08]

Order as $n \wedge k \to \infty$

	(o-m) $\mathrm{v} \in (-\mathrm{t}, \mathrm{a})$	**(o-s)** $\mathrm{a} - \mathrm{v} \in \mathbb{R}_{>0}$	**(s-m)** $\mathrm{v} + \mathrm{t} \in \mathbb{R}_{>0}$
$(m^\circ_{n\wedge k})^{-2}\mathrm{R}^\star_{n\wedge k}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet)$	$(n \wedge k)^{-\frac{2(\mathrm{a}-\mathrm{v})+2}{2\mathrm{a}+2\mathrm{t}}}$	$(\log n \wedge k)^{-\frac{2\mathrm{a}-2\mathrm{v}+2}{2\mathrm{t}}}$	$\frac{(\log n\wedge k)^{\frac{2\mathrm{t}+2\mathrm{v}-2}{2\mathrm{a}}}}{n\wedge k}$
$(n \wedge k)^{-1}(m^\star_{n\wedge k})^3 \mathfrak{t}^{-2}_{m^\star_{n\wedge k}}$	$(n \wedge k)^{-\frac{2\mathrm{a}-3}{2\mathrm{a}+2\mathrm{t}}}$	$(n \wedge k)^{-c}$	$\frac{(\log n\wedge k)^{\frac{2\mathrm{t}+3}{2\mathrm{a}}}}{n\wedge k}$

In case **(o-s)** a value $l \geqslant 2$ and **(s-m)** a value $l \geqslant 3$ is sufficient to ensure (08.49) as $n \wedge k \to \infty$. In case **(o-m)** assuming $\mathrm{a} > 3/2$ we have $k^{-1}(m^\circ_{n\wedge k})^3 \mathfrak{t}^{-2}_{m^\star_{n\wedge k}} = \mathrm{o}(1)$ as $n \wedge k \to \infty$. In this situation we have (08.49) if $(2\mathrm{a} - 3)(l - 1) > 2(\mathrm{a} - \mathrm{v}) + 2$ or in equal $l > (4\mathrm{a} - 2\mathrm{v} - 1)/(2\mathrm{a} - 3)$. □

§09 Spectral regularisation estimator

§09|00.01 **Notation**. Consider the measure space $(\mathcal{J}, \mathscr{J}, \nu)$ and the Hilbert space $\mathbb{J} = \mathbb{L}_2(\nu)$ as in Notation §01|01.01. We suppose that $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \mathbb{J})$ are surjective partial isometries, hence $\mathrm{V}\mathrm{V}^\star = \mathrm{id}_{\mathbb{J}} = \mathrm{U}\mathrm{U}^\star$. As in Definition §03|00.08 we denote for $\mathrm{A} := \mathrm{V}\mathrm{T}\mathrm{U}^\star \in \mathbb{L}(\mathbb{J})$ its Moore-Penrose inverse by $\mathrm{A}^\dagger : \mathbb{J} \supseteq \mathrm{dom}(\mathrm{A}^\dagger) \to \mathbb{J}$. □

§09|00.02 **Assumption**. For $\mathbb{J} = \mathbb{L}_2(\nu)$ let $\mathrm{U} \in \mathbb{L}(\mathbb{H}, \mathbb{J})$ and $\mathrm{V} \in \mathbb{L}(\mathbb{G}, \mathbb{J})$ be surjective partial isometries fixed and presumed to be *known* in advance, let $\mathrm{T} \in \mathbb{L}(\mathbb{H}, \mathbb{G})$, hence $\mathrm{A} = \mathrm{V}\mathrm{T}\mathrm{U}^\star \in \mathbb{L}(\mathbb{J})$ with Moore-Penrose inverse $\mathrm{A}^\dagger : \mathbb{J} \supseteq \mathrm{dom}(\mathrm{A}^\dagger) \to \mathbb{J}$ and let $g_\bullet \in \mathrm{dom}(\mathrm{A}^\dagger)$, and hence $\theta_\bullet = \mathrm{A}^\dagger g_\bullet \in \mathbb{J}$. □

§09|00.03 **Reminder**. Under Assumption §09|00.02 we consider $\theta_{\circ} \in \mathbb{J}$ and $A \in \mathbb{L}(\mathbb{J})$ and hence $g_{\circ} = A\theta_{\circ} \in \operatorname{ran}(A) \subseteq \operatorname{dom}(A^{\dagger})$. Let $\{r_{\alpha} : \alpha \in (0,1)\}$ be a collection of real-valued Borel-measurable functions defined on $[0, \|T_{\bullet|\bullet}\|^2_{\mathbb{L}(\mathbb{J})}]$ satisfying (see §06|02.01)

(sR1) for all $\alpha \in (0,1)$ there exists $C_{\alpha} \in \mathbb{R}_{\geqslant 0}$ such that $|r_{\alpha}(x)| \leqslant C_{\alpha}$ for all $x \in [0, \|A\|^2_{\mathbb{L}(\mathbb{J})}]$,

(sR2) for all $x \in (0, \|A\|^2_{\mathbb{L}(\mathbb{J})}]$ holds $|1 - x r_{\alpha}(x)| = o(1)$ as $\alpha \to 0$, and

(sR3) there is $K \in \mathbb{R}_{>0}$ such that $|x r_{\alpha}(x)| \leqslant K$ for all $x \in [0, \|A\|^2_{\mathbb{L}(\mathbb{J})}]$ and $\alpha \in (0,1)$,

Then the collection $\{R_{\alpha} := r_{\alpha}(A^{\star}A)A^{\star} \in \mathbb{L}(\mathbb{J}) : \alpha \in (0,1)\}$ of operators is called *spectral regularisation* of $A^{\dagger} : \mathbb{J} \supseteq \operatorname{dom}(A^{\dagger}) \to \mathbb{J}$. □

§09|01 Statistical inverse problem

§09|01.01 **Assumption**. Consider a random function $\widetilde{g}_{\circ} \in \mathcal{M}(\mathscr{A}, \mathscr{B}_{\mathbb{J}})$ on a measurable space $(\Omega, \mathscr{A})$ with values in $\mathbb{J}$ (Definition §01|01.17). Let Assumption §08|00.02 be satisfied where $A \in \mathbb{L}(\mathbb{J})$ is *known* in advance. For $\theta_{\circ} \in \mathbb{J}$, hence image $g_{\circ} = A\theta_{\circ} \in \mathbb{J}$, and probability measure $\mathbb{P}_{\theta|A} \in \mathscr{W}(\mathscr{A})$ on $(\Omega, \mathscr{A})$ the random function $\widetilde{g}_{\circ}$ has a finite second moment (i.e. $\mathbb{E}_{\theta|A}(\|\widetilde{g}_{\circ}\|^2_{\mathbb{J}}) \in \mathbb{R}_{\geqslant 0}$). □

§09|01.02 **Definition**. Under Assumption §09|01.01 for $\theta_{\circ} \in \mathbb{J}$, $A \in \mathbb{L}(\mathbb{J})$, and a continuous spectral regularisation $\{R_{\alpha} := r_{\alpha}(A^{\star}A)A^{\star} \in \mathbb{L}(\mathbb{J}) : \alpha \in (0,1)\}$ of $A^{\dagger}$ as in Definition §06|02.01 we call $\widetilde{\theta}_{\circ}^{\alpha} = R_{\alpha}\widetilde{g}_{\circ}$ *spectral regularisation estimator (sRE)* of $\theta_{\circ}$. □

§09|01.03 **Comment**. Since $g_{\circ} = A\theta_{\circ} \in \operatorname{dom}(A^{\dagger})$ and hence $\theta_{\circ} = A^{\dagger}g_{\circ}$ the spectral regularised approximation $\theta_{\circ}^{\alpha} := R_{\alpha}g_{\circ} = r_{\alpha}(A^{\star}A)A^{\star}g_{\circ} \in \mathbb{J}$ converges to $\theta_{\circ}$ as $\alpha \to 0$, i.e. the approximation error $\|\theta_{\circ}^{\alpha} - \theta_{\circ}\|_{\mathbb{J}}$ converges to zero as $\alpha \to 0$ (compare Proposition §06|02.02). □

§09|01|01 Global risk

§09|01.04 **Lemma** ($\mathbb{J}$-consistency). *Let $\{R_{\alpha} := r_{\alpha}(A^{\star}A)A^{\star} \in \mathbb{L}(\mathbb{J}) : \alpha \in (0,1)\}$ be a continuous spectral regularisation of $A^{\dagger}$ as in Definition §06|02.01. Assume Definition §06|02.01 (sR1) and (sR2), and in addition replace (sR3) by*

(sR3a) *for all $s \in [0,1]$ there exists $K_s \in \mathbb{R}_{>0}$ such that $x^s|r_{\alpha}(x)| \leqslant C_s\alpha^{s-1}$ for all $x \in [0, \|A\|^2_{\mathbb{L}(\ell_2)}]$ and $\alpha \in (0,1)$.*

Under Assumption §09|01.01 a sRE $\widetilde{\theta}_{\circ}^{\alpha} = R_{\alpha}\widetilde{g}_{\circ}$ of $\theta_{\circ} = A^{\dagger}g_{\circ} \in \mathbb{J}$ satisfies for all $\alpha \in (0,1)$

$$\mathbb{E}_{\theta|A}(\|\widetilde{\theta}_{\circ}^{\alpha} - \theta_{\circ}\|^2_{\mathbb{J}}) \leqslant 2K^2_{1/2}\alpha^{-1}\mathbb{E}_{\theta|A}(\|\widetilde{g}_{\circ} - g_{\circ}\|^2_{\mathbb{J}}) + 2\|\theta_{\circ}^{\alpha} - \theta_{\circ}\|^2_{\mathbb{J}} \tag{09.01}$$

If $\widetilde{g}_{\circ}$ is a $\mathbb{J}$-consistent estimator of $g_{\circ}$, that is $\mathbb{E}^n_{\theta|A}(\|\widetilde{g}_{\circ} - g_{\circ}\|^2_{\mathbb{J}}) = o(1)$ as $n \to \infty$, then

$$\mathbb{E}^n_{\theta|A}(\|\widetilde{\theta}_{\circ}^{\alpha_n} - \theta\|^2_{\mathbb{J}}) = o(1) \quad \text{as } n \to \infty$$

for any sequence $(\alpha_n)_{n\in\mathbb{N}}$ such that $\alpha_n = o(1)$ and $\alpha_n^{-1}\mathbb{E}^n_{\theta|A}(\|\widetilde{g}_{\circ} - g_{\circ}\|^2_{\mathbb{J}}) = o(1)$ as $n \to \infty$.

§09|01.05 **Proof** of Lemma §09|01.04. Given in the lecture. □

§09|01.06 **Reminder**. Given $A \in \mathbb{L}(\mathbb{J})$ let $\{R_{\alpha} := r_{\alpha}(A^{\star}A)A^{\star} \in \mathbb{L}(\mathbb{J}) : \alpha \in (0,1)\}$ be a spectral regularisation of $A^{\dagger} : \mathbb{J} \supseteq \operatorname{dom}(A^{\dagger}) \to \mathbb{J}$ as in Definition §06|02.01. Assume Definition §06|02.01 (sR1), and (sR3), and in addition replace (sR2) by

(sR2a) there are $s_{\circ} \in \mathbb{R}_{\geqslant 1}$ $C_s \in \mathbb{R}_{\geqslant 0}$ for all $s \in [0, s_{\circ}]$ such that $x^s|1 - x r_{\alpha}(x)| \leqslant C_s\alpha^s$ for all $x \in [0, \|A\|^2_{\mathbb{L}(\mathbb{J})}]$ and $\alpha \in (0,1)$.

For $\theta_\circ \in \mathbb{J}$, $g_\circ = A\theta_\circ \in \mathrm{dom}(A^\dagger)$, and $\alpha \in (0,1)$ consider $\theta_\circ^\alpha = R_\alpha g_\circ = r_\alpha(A^\star A)A^\star g_\circ \in \mathbb{J}$. If $\theta_\circ = A^\dagger g_\circ \in \mathbb{J}$ fulfills a source condition as in Definition §06|02.05, that is, there are $s \in [0, 2s_\circ]$ and $h_\circ \in \mathbb{J}$ such that $\theta_\circ = (A^\star A)^{s/2}h_\circ$ or in equal $\theta_\circ \in \mathrm{ran}((A^\star A)^{s/2})$, then we have

$$\|\theta_\circ^\alpha - \theta_\circ\|_{\mathbb{J}} \leqslant C_{s/2}\alpha^{s/2}\|h_\circ\|_{\mathbb{J}} \quad \forall \alpha \in (0,1) \tag{09.02}$$

due to Proposition §06|02.06. □

§09|01.07 **Corollary**. *Let the assumptions of Lemma §09|01.04 be satisfied and in addition let (sR2) be replaced by (sR2a). If $\theta_\circ = A^\dagger g_\circ \in \mathbb{J}$ fulfills a source condition as in Definition §06|02.05, that is, there are $s \in [0, 2s_\circ]$ and $h_\circ \in \mathbb{J}$ such that $\theta_\circ = (A^\star A)^{s/2}h_\circ$ then the sRE $\widetilde{\theta}_\circ^{\alpha_n} = R_{\alpha_n}\widetilde{g}_\circ$ of $\theta_\circ = A^\dagger g_\circ \in \mathbb{J}$ with $\alpha_\circ := \big(\mathbb{E}_{\theta|A}(\|\widetilde{g}_\circ - g_\circ\|_{\mathbb{J}}^2)\big)^{1/(1+s)}$ fulfills*

$$\mathbb{E}_{\theta|A}(\|\widetilde{\theta}_\circ^{\alpha_\circ} - \theta_\circ\|_{\mathbb{J}}^2) \leqslant 2(K_{1/2}^2 + C_{s/2}^2\|h_\circ\|_{\mathbb{J}}^2)\big(\mathbb{E}_{\theta|A}(\|\widetilde{g}_\circ - g_\circ\|_{\mathbb{J}}^2)\big)^{s/(1+s)}. \tag{09.03}$$

§09|01.08 **Proof** of Corollary §09|01.07. Given in the lecture. □

§09|01.09 **Reminder**. Given $A \in \mathbb{L}^{\geqslant}(\mathbb{J})$, i.e., A is positive definite, we eventually consider as in Notation §06|02.19 a spectral regularisation $\{R_\alpha := r_\alpha(A) \in \mathbb{L}(\mathbb{J}) \colon \alpha \in (0,1)\}$ of $A^\dagger$ for a given collection $\{r_\alpha \colon \alpha \in (0,1)\}$ of real-valued Borel-measurable functions defined on $[0, \|A\|_{\mathbb{L}(\mathbb{J})}]$ satisfying

(sR1') for all $\alpha \in (0,1)$ there exists $C_\alpha \in \mathbb{R}_{\geqslant 0}$ such that $|r_\alpha(x)| \leqslant C_\alpha$ for all $x \in [0, \|A\|_{\mathbb{L}(\mathbb{J})}]$,

(sR2'a) there are $s_\circ \in [1, \infty)$ and $C_s \in \mathbb{R}_{\geqslant 0}$ for all $s \in [0, s_\circ]$ such that $x^s|1 - xr_\alpha(x)| \leqslant C_s\alpha^s$ for all $x \in [0, \|A\|_{\mathbb{L}(\mathbb{J})}]$ and $\alpha \in (0,1)$,

(sR3') there is $K \in \mathbb{R}_{>0}$ such that $|xr_\alpha(x)| \leqslant K$ for all $x \in [0, \|A\|_{\mathbb{L}(\mathbb{J})}]$ and $\alpha \in (0,1)$.

We consider the spectral regularised approximation $\theta_\circ^\alpha = R_\alpha g_\circ = r_\alpha(A)g_\circ \in \mathbb{J}$ of $\theta_\circ := A^\dagger g_\circ \in \mathbb{J}$ for $g_\circ \in \mathrm{dom}(A^\dagger)$. Under Assumption §09|01.01 we call $\widetilde{\theta}_\circ^\alpha = R_\alpha\widetilde{g}_\circ$ *spectral regularisation estimator (sRE)* of $\theta_\circ$. □

§09|01.10 **Lemma** ($\mathbb{J}$-consistency). *Given $A \in \mathbb{L}^{\geqslant}(\mathbb{J})$ let $\{R_\alpha := r_\alpha(A) \in \mathbb{L}(\mathbb{J}) \colon \alpha \in (0,1)\}$ be a continuous spectral regularisation of $A^\dagger$ as in Notation §06|02.19. Assume Notation §06|02.19 (sR1') and (sR2'a), and in addition replace (sR3') by*

(sR3'a) *for all $s \in [0,1]$ there exists $K_s \in \mathbb{R}_{\geqslant 0}$ such that $x^s|r_\alpha(x)| \leqslant K_s\alpha^{s-1}$ for all $x \in [0, \|A\|_{\mathbb{L}(\mathbb{J})}]$ and $\alpha \in (0,1)$.*

Under Assumption §09|01.01 a sRE $\widetilde{\theta}_\circ^\alpha = R_\alpha\widetilde{g}_\circ$ of $\theta_\circ = A^\dagger g_\circ \in \mathbb{J}$ satisfies for all $\alpha \in (0,1)$

$$\mathbb{E}_{\theta|A}(\|\widetilde{\theta}_\circ^\alpha - \theta_\circ\|_{\mathbb{J}}^2) \leqslant 2K_0^2\alpha^{-2}\mathbb{E}_{\theta|A}(\|\widetilde{g}_\circ - g_\circ\|_{\mathbb{J}}^2) + 2\|\theta_\circ^\alpha - \theta_\circ\|_{\mathbb{J}}^2 \tag{09.04}$$

If $\widetilde{g}_\circ$ is a $\mathbb{J}$-consistent *estimator of $g_\circ$, that is $\mathbb{E}_{\theta|A}^n(\|\widetilde{g}_\circ - g_\circ\|_{\mathbb{J}}^2) = o(1)$ as $n \to \infty$, then*

$$\mathbb{E}_{\theta|A}^n(\|\widetilde{\theta}_\circ^{\alpha_n} - \theta\|_{\mathbb{J}}^2) = o(1) \quad \text{as } n \to \infty$$

for any sequence $(\alpha_n)_{n\in\mathbb{N}}$ such that $\alpha_n = o(1)$ and $\alpha_n^{-2}\mathbb{E}_{\theta|A}^n(\|\widetilde{g}_\circ - g_\circ\|_{\ell_2}^2) = o(1)$ as $n \to \infty$.

§09|01.11 **Proof** of Lemma §09|01.10. Given in the lecture. □

§09|01.12 **Reminder**. Given $A \in \mathbb{L}^{\geqslant}(\mathbb{J})$ let $\{R_\alpha := r_\alpha(A) \in \mathbb{L}(\mathbb{J}) \colon \alpha \in (0,1)\}$ be a spectral regularisation of $A^\dagger : \mathbb{J} \supseteq \mathrm{dom}(A^\dagger) \to \mathbb{J}$ as in Notation §06|02.19. Assume Notation §06|02.19 (sR1'), (sR2'a), and (sR3'). For $\theta_\circ \in \mathbb{J}$, $g_\circ = A\theta_\circ \in \mathrm{dom}(A^\dagger)$, and $\alpha \in (0,1)$ consider $\theta_\circ^\alpha = R_\alpha g_\circ = r_\alpha(A)g_\circ \in \mathbb{J}$. If

$\theta_{\circ} = A^{\dagger} g_{\circ} \in \mathbb{J}$ fulfills a source condition as in Definition §06|02.05, that is, there are $s \in [0, 2s_{\circ}]$ and $h_{\circ} \in \mathbb{J}$ such that $\theta_{\circ} = A^{s} h_{\circ}$ or in equal $\theta_{\circ} \in \operatorname{ran}(A^{s})$, then we have

$$\|\theta_{\circ}^{\alpha} - \theta_{\circ}\|_{\mathbb{J}} \leqslant C_{s} \alpha^{s} \|h_{\circ}\|_{\mathbb{J}} \quad \forall \alpha \in (0, 1) \tag{09.05}$$

due to Proposition §06|02.20. □

§09|01.13 **Corollary**. *Let the assumptions of Lemma §09|01.10 be satisfied. If $\theta_{\circ} = A^{\dagger} g_{\circ} \in \mathbb{J}$ fulfills a source condition as in Definition §06|02.05, that is, there are $s \in [0, 2s_{\circ}]$ and $h_{\circ} \in \mathbb{J}$ such that $\theta_{\circ} = (A)^{s} h_{\circ}$ then the sRE $\widetilde{\theta}_{\circ}^{\alpha_{n}} = R_{\alpha_{n}} \widetilde{g}_{\circ}$ of $\theta_{\circ} = A^{\dagger} g_{\circ} \in \mathbb{J}$ with $\alpha_{\circ} := \big(\mathbb{E}_{\theta|A}(\|\widetilde{g}_{\circ} - g_{\circ}\|_{\mathbb{J}}^{2})\big)^{1/(2+2s)}$ fulfills*

$$\mathbb{E}_{\theta|A}(\|\widetilde{\theta}_{\circ}^{\alpha_{\circ}} - \theta_{\circ}\|_{\mathbb{J}}^{2}) \leqslant 2(K_{0}^{2} + C_{s}^{2}\|h_{\circ}\|_{\mathbb{J}}^{2})\big(\mathbb{E}_{\theta|A}(\|\widetilde{g}_{\circ} - g_{\circ}\|_{\mathbb{J}}^{2})\big)^{s/(1+s)}. \tag{09.06}$$

§09|01.14 **Proof** of Corollary §09|01.13. Given in the lecture. □

§09|01|02 Maximal global $\mathfrak{v}$-risk

§09|01.15 **Assumption**. Consider the separable Hilbert space $\mathbb{J} = \mathbb{L}_{2}(\mathcal{J}, \mathscr{J}, \nu)$ with σ-algebra $\mathscr{J}$ over $\mathcal{J}$ containing all elementary events $\{j\}$, $j \in \mathcal{J}$, and all events $[\![m]\!] := [-m, m] \cap \mathcal{J}$, $m \in \mathbb{N}$, and with σ-finite measure $\nu \in \mathscr{M}_{\sigma}(\mathscr{J})$ such that $\nu([\![m]\!]) \in \mathbb{R}_{\geqslant 0}$, for all $m \in \mathbb{N}$. Let Assumption §09|00.02 be satisfied where $A \in \mathbb{L}(\mathbb{J})$ is *known* in advance. For $\theta_{\circ} \in \mathbb{J}$, and hence image $g_{\circ} = A\theta_{\circ} \in \mathbb{J}$, let $\mathbb{P}_{\theta|A} \in \mathscr{W}(\mathscr{A})$ be a probability measure on $(\Omega, \mathscr{A})$. Consider a stochastic process $\dot{\varepsilon}_{\circ} = (\dot{\varepsilon}_{j})_{j \in \mathcal{J}}$ on $(\Omega, \mathscr{A})$ satisfying Assumption §01|01.04 (i.e. $\dot{\varepsilon}_{\circ} \in \mathcal{M}(\mathscr{A} \otimes \mathscr{J})$) which for each $\theta_{\circ} \in \mathbb{J}$ and $A \in \mathbb{L}(\mathbb{J})$ in addition fulfills

(SIPg1) $\dot{\varepsilon}_{j} \in \mathcal{L}_{1}(\mathbb{P}_{\theta|A}) := \mathcal{L}_{1}(\Omega, \mathscr{A}, \mathbb{P}_{\theta|A})$ for all $j \in \mathcal{J}$ and $\mathbb{P}_{\theta|A}(\dot{\varepsilon}_{\circ}) = (\mathbb{P}_{\theta|A}(\dot{\varepsilon}_{j}))_{j \in \mathcal{J}} = 0_{\circ}$,

(SIPg2) $\mathbb{v}_{\circ}^{\theta|A} := \mathbb{P}_{\theta|A}(\dot{\varepsilon}_{\circ}^{2}) := (\mathbb{v}_{j}^{\theta|A} := \mathbb{P}_{\theta|A}(\dot{\varepsilon}_{j}^{2}))_{j \in \mathcal{J}} \in \mathbb{L}_{\infty}(\nu)$ and

(SIPg3) $\dot{\varepsilon}_{\circ}\mathbb{1}_{\circ}^{m} \in \mathbb{L}_{\infty}(\nu)$ $\mathbb{P}_{\theta|A}$-a.s., for each $m \in \mathbb{N}$.

Given a sample size $n \in \mathbb{N}$ the observable noisy image with mean $g_{\circ} = A\theta_{\circ} \in \mathbb{J}$ takes the form $\widehat{g}_{\circ} = g_{\circ} + n^{-1/2}\dot{\varepsilon}_{\circ}$. We denote by $\mathbb{P}_{\theta|A}^{n}$ the distribution of $\widehat{g}_{\circ}$. □

§09|01.16 **Comment**. Under Assumption §09|01.15 we have $\dot{\varepsilon}_{\circ}\mathbb{1}_{\circ}^{m} \in \mathbb{J}$ $\mathbb{P}_{\theta|A}$-a.s.. Since $g_{\circ} \in \mathbb{J}$, and hence $g_{\circ}^{m} = g_{\circ}\mathbb{1}_{\circ}^{m} \in \mathbb{J}$ (Property §04|03.09), it follows

$$\widehat{g}_{\circ}^{m} = \widehat{g}_{\circ}\mathbb{1}_{\circ}^{m} = n^{-1/2}\dot{\varepsilon}_{\circ}\mathbb{1}_{\circ}^{m} + g_{\circ}^{m} \in \mathbb{J} \quad \mathbb{P}_{\theta|A}^{n}\text{-a.s..} \tag{09.07}$$

If $\mathcal{J} \subseteq \mathbb{Z}$ (at most countable) and $\nu_{\mathcal{J}}$ is the counting measure over the index set $\mathcal{J}$ then Assumption §01|01.04 and (SIPg2) $\mathbb{v}_{\circ}^{\theta|A} = \mathbb{P}_{\theta|A}(\dot{\varepsilon}_{\circ}^{2}) \in \mathbb{L}_{\infty}(\nu_{\mathcal{J}})$ implies the additional assumption (SIPg3) $\dot{\varepsilon}_{\circ}\mathbb{1}_{\circ}^{m} \in \mathbb{L}_{\infty}(\nu_{\mathcal{J}})$ $\mathbb{P}_{\theta|s}^{n}$-a.s.. However, the last implication does generally not hold, if $\mathcal{J} \in \{\mathbb{R}, \mathbb{R}_{\geqslant 0}\}$ for example. □

§09|01.17 **Assumption**. Consider $\mathfrak{v}_{\circ} \in \mathcal{M}_{>0,\nu}(\mathscr{J}) \cap \mathbb{L}_{\infty}(\nu)$, and for $t \in \mathbb{R}_{>0}$, $a \in (0, t]$ set $\mathfrak{t}_{\circ} := \mathfrak{v}_{\circ}^{t}$ and $\mathfrak{a}_{\circ} := \mathfrak{v}_{\circ}^{a}$ where $\mathfrak{t}_{\circ}, \mathfrak{a}_{\circ} \in \mathcal{M}_{>0,\nu}(\mathscr{J}) \cap \mathbb{L}_{\infty}(\nu)$ and hence $\nu(\mathcal{N}_{\mathfrak{v}}) = \nu(\mathcal{N}_{\mathfrak{a}}) = \nu(\mathcal{N}_{\mathfrak{t}}) = 0$. □

§09|01.18 **Reminder**. Under Assumption §09|01.17 we have $\mathbb{J}^{\mathfrak{a}} = \mathbb{L}_{2}^{\mathfrak{a}}(\nu) = \operatorname{dom}(M_{\mathfrak{a}^{-1}}) = \mathbb{J}\mathfrak{a}_{\circ} = \mathbb{L}_{2}(\mathfrak{a}_{\circ}^{-2}\nu)$ and the measures ν, $\mathfrak{v}_{\circ}^{2}\nu$, $\mathfrak{t}_{\circ}^{2}\nu$ and $\mathfrak{a}_{\circ}^{-2}\nu$ dominate mutually each other (see Property §04|01.02). Consequently, $\mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{J} = \mathbb{L}_{2}(\nu)$ and $\mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{L}_{2}(\mathfrak{v}_{\circ}^{2}\nu)$ (Property §04|02.11) since $(\mathfrak{a}\mathfrak{v})_{\circ} = \mathfrak{v}_{\circ}^{1+a} \in \mathbb{L}_{\infty}(\nu)$. We assume in the following that $\theta_{\circ} \in \mathbb{J}$ satisfies an abstract smoothness condition (Definition §04|02.12), i.e., there is $r \in \mathbb{R}_{>0}$ such that $\theta_{\circ} \in \mathbb{J}^{\mathfrak{a},r} = \{h_{\circ} \in \mathbb{J}^{\mathfrak{a}} : \|h_{\circ}\|_{\mathfrak{a}^{-1}} \leqslant r\} \subseteq \mathbb{J}^{\mathfrak{a}} \subseteq \mathbb{J}$. Under Assumption §06|02.11 by Corollary §05|01.14 (see Comment §05|01.16) if $A \in \mathbb{T}_{t,d}$ (or in equal $(A^{\star}A)^{1/2} \in \mathbb{T}_{t,d}^{\gtrless}$) then (i) for any $\theta_{\circ} \in \mathbb{J}^{\mathfrak{a}}$ we have $\theta_{\circ} = (A^{\star}A)^{a/(2t)} h_{\circ}$ with $\|h_{\circ}\|_{\mathbb{J}} \leqslant d^{a/t}\|\theta_{\circ}\|_{\mathfrak{a}^{-1}}$, and conversely

(ii) for any $\theta_\centerdot = (A^\star A)^{a/(2t)} h_\centerdot$ with $h_\centerdot \in \mathbb{L}_2(\nu)$ we obtain $\theta_\centerdot \in \mathbb{J}^a$ with $\|\theta_\centerdot\|_{\mathfrak{a}^{-1}} \leqslant d^{a/t}\|h_\centerdot\|_{\mathbb{J}}$. In particular since $(\mathfrak{ta})_\centerdot = \mathfrak{v}_\centerdot^{t+a} \in \mathcal{M}_{>0,\nu}(\mathscr{J}) \cap \mathbb{L}_\infty(\nu)$ if $A \in \mathbb{T}_{t,d}$ and $\theta_\centerdot \in \mathbb{J}^{a,r}$, then due to Corollary §06|02.13 we have $g_\centerdot = A\theta_\centerdot \in \mathbb{J}^{(\mathfrak{ta}),dr}$. □

§09|01.19 **Notation** (Reminder). For sequences $a_\centerdot, b_\centerdot \in (\mathbb{K})^{\mathbb{N}}$ taking its values in $\mathbb{K} \in \{\mathbb{R}, \mathbb{R}_{>0}, \mathbb{Q}, \mathbb{Z}, \dots\}$ we write $a_\centerdot \in (\mathbb{K})^{\mathbb{N}}_{\nearrow}$ and $b_\centerdot \in (\mathbb{K})^{\mathbb{N}}_{\searrow}$ if $a_\centerdot$ and $b_\centerdot$, respectively, is monotonically *non-decreasing* and *non-increasing*. If in addition $a_n \to \infty$ and $b_n \to 0$ as $n \to \infty$, then we write $a_\centerdot \in (\mathbb{K})^{\mathbb{N}}_{\uparrow\infty}$ and $b_\centerdot \in (\mathbb{K})^{\mathbb{N}}_{\downarrow 0}$ for short. For $w_\centerdot \in \mathbb{L}_\infty(\nu)$ we set $w_{(0)} := \|w_\centerdot\|_{\mathbb{L}_\infty(\nu)}$ and $w_{(\centerdot)} = (w_{(j)} := \|w_\centerdot \mathbb{1}_\centerdot^{j|\perp}\|_{\mathbb{L}_\infty(\nu)})_{j\in\mathbb{N}}$, where by construction $w_{(\centerdot)} \in (\mathbb{R}_{\geqslant 0})^{\mathbb{N}}_{\searrow}$. □

§09|01.20 **Corollary**. *Under Assumptions §09|01.15 and §09|01.17 setting for $n, m \in \mathbb{N}$*

$$R_n^m((\mathfrak{at})_\centerdot) := [(\mathfrak{at})_{(m)} \vee n^{-1}m], \quad m_n^\star := \arg\min\{R_n^m((\mathfrak{at})_\centerdot) : m \in \mathbb{N}\}$$
$$\text{and} \quad R_n^\star((\mathfrak{at})_\centerdot) := R_n^{m_n^\star}((\mathfrak{at})_\centerdot) = \min\{R_n^m((\mathfrak{at})_\centerdot) : m \in \mathbb{N}\} \quad (09.08)$$

and $\|\mathbb{v}_\centerdot^{\theta|A}\|_{\mathbb{L}_\infty(\nu)} \leqslant \mathbb{v}_{\theta|A} \in \mathbb{R}_{\geqslant 0}$, for $A \in \mathbb{T}_{t,d}$ and for all $\theta_\centerdot \in \mathbb{J}^{a,r}$, hence $g_\centerdot = A\theta_\centerdot \in \mathrm{dom}(A^\dagger) \subseteq \mathbb{J}$, the orthogonal projection estimator (OPE) $\widehat{g}_\centerdot^m := \widehat{g}_\centerdot \mathbb{1}_\centerdot^m$ fulfills

$$\mathbb{E}_{\theta|A}^n(\|\widehat{g}_\centerdot^m - g_\centerdot\|_{\mathbb{J}}^2) \leqslant (\mathbb{v}_{\theta|A} + d^2r^2)\, R_n^m((\mathfrak{at})_\centerdot) \quad \forall m, n \in \mathbb{N}$$

and hence $\mathbb{E}_{\theta|A}^n(\|\widehat{g}_\centerdot^{m_n^\star} - g_\centerdot\|_{\mathbb{J}}^2) \leqslant (\mathbb{v}_{\theta|A} + d^2r^2)\, R_n^\star((\mathfrak{at})_\centerdot)$.

§09|01.21 **Proof** of Corollary §09|01.20. Given in the lecture. □

Consider the OPE $\widehat{g}_\centerdot^m := \widehat{g}_\centerdot \mathbb{1}_\centerdot^m$ for the orthogonal projection $g_\centerdot^m = g_\centerdot \mathbb{1}_\centerdot^m \in \mathbb{J}\mathbb{1}_\centerdot^m$ of $g_\centerdot = A\theta_\centerdot \in \mathbb{J}$. Given a continuous spectral regularisation $\{R_\alpha := r_\alpha(A^\star A)A^\star \in \mathbb{L}(\mathbb{J}) : \alpha \in (0,1)\}$ of $A^\dagger$ as in Definition §06|02.01 We measure the accuracy of the sRE $\widehat{\theta}_\centerdot^{\alpha,m} = R_\alpha \widehat{g}_\centerdot^m$ of $\theta_\centerdot = A^\dagger g_\centerdot \in \mathbb{J}$ by the mean of its global $\mathfrak{v}$-error introduced in §04|03|01, i.e. its $\mathfrak{v}$-risk.

§09|01.22 **Proposition**. *Under Assumptions §09|01.15 and §09|01.17 with $\|\mathbb{v}_\centerdot^{\theta|A}\|_{\mathbb{L}_\infty(\nu)} \leqslant \mathbb{v}_{\theta|A} \in \mathbb{R}_{\geqslant 0}$ let $\{R_\alpha := r_\alpha(A^\star A)A^\star \in \mathbb{L}(\mathbb{J}) : \alpha \in (0,1)\}$ be a continuous spectral regularisation of $A^\dagger$ as in Definition §06|02.01 and in addition replace (sR2) and (sR3) by (sR2a) and (sR3a), respectively. Consider for $m \in \mathbb{N}$ and $\alpha \in (0,1)$ the sRE $\widehat{\theta}_\centerdot^{\alpha,m} = R_\alpha \widehat{g}_\centerdot^m$. If $\theta_\centerdot \in \mathbb{J}^{a,r}$ and $A \in \mathbb{T}_{t,d}$ then for all $m \in \mathbb{N}$ and $\alpha \in (0,1)$ we have*

$$\mathbb{E}_{\theta|A}^n(\|\widehat{\theta}_\centerdot^{\alpha,m} - \theta_\centerdot\|_{\mathfrak{v}^q}^2) \leqslant \big[\alpha^{(a+q)/t} \vee \alpha^{(q-t)/t}\, R_n^m(\mathfrak{v}_\centerdot^{a+t})\big]$$
$$\times\, 2d^{2|q|/t}\big\{C_{(q+a)/(2t)}^2\, d^{2(a+|q|)/t}\, r^2 + K_{(q+t)/(2t)}^2\, (\mathbb{v}_{\theta|A} + d^2r^2)\big\}. \quad (09.09)$$

§09|01.23 **Proof** of Proof §09|01.23. Given in the lecture. □

§09|01.24 **Corollary**. *Under the assumptions of Proposition §09|01.22 the SRE $\widehat{\theta}_\centerdot^{\alpha_n^\star,m_n^\star} := R_{\alpha_n^\star}\widehat{g}_\centerdot^{m_n^\star}$ with $m_n^\star$ and $R_n^\star(\mathfrak{v}_\centerdot^{p+a})$ as in (09.08) (Corollary §09|01.20 using $(\mathfrak{at})_\centerdot = \mathfrak{v}_\centerdot^{a+t}$) and $\alpha_n^\star := (R_n^\star(\mathfrak{v}_\centerdot^{a+t}))^{t/(a+t)}$ for all n satisfies*

$$\mathbb{E}_{\theta|A}^n(\|\widehat{\theta}_\centerdot^{\alpha,m} - \theta_\centerdot\|_{\mathfrak{v}^q}^2) \leqslant (R_n^\star(\mathfrak{v}_\centerdot^{a+t}))^{(a+q)/(a+t)}$$
$$\times\, 2d^{2|q|/t}\big\{C_{(q+a)/(2t)}^2\, d^{2(a+|q|)/t}\, r^2 + K_{(q+t)/(2t)}^2\, (\mathbb{v}_{\theta|A} + d^2r^2)\big\}. \quad (09.10)$$

§09|01.25 **Proof** of Corollary §09|01.24. Given in the lecture. □

Chapter 4

Minimax optimal estimation

We present a general approach to derive lower bounds and thus in combination with the upper bounds Chapter 3 establish minimax optimality.

Overview

§10 Minimax theory: a general approach

Suppose that the function of interest θ belongs to a class $\Theta \subseteq \mathbb{H}$. For each noise level $n \in \mathbb{N}$ let $\mathbb{P}^n_\Theta := (\mathbb{P}^n_\theta)_{\theta\in\Theta}$ denote a family of probability measures and let $\mathbb{E}^n_\theta$ be the expectation with respect to the measure $\mathbb{P}^n_\theta$ in $\mathbb{P}^n_\Theta$. Moreover, we assume that the probability measure associated with an observable quantity belongs to $\mathbb{P}^n_\Theta$.

§10|00.01 **GdSM (§01|03.06 continued).** Given $\ell_2 = \mathbb{L}_2(\mathbb{N}, 2^{\mathbb{N}}, \nu_{\mathbb{N}})$ consider a Gaussian direct sequence model (GdSM) as in §01|03.06. Here the observable stochastic process $\widehat{\theta}_\bullet = \theta_\bullet + n^{-1/2}\dot{\mathrm{B}}_\bullet \sim \mathrm{N}^n_\theta$ is a noisy version of $\theta_\bullet \in \Theta \subseteq \ell_2$ and $\dot{\mathrm{B}}_\bullet \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$. Consequently, $\widehat{\theta}_\bullet$ admits a N^n_θ-distribution belonging to the family $\mathrm{N}^n_\Theta := (\mathrm{N}^n_\theta)_{\theta_\bullet\in\Theta}$. Summarising the observations satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}^{\otimes\mathbb{N}}, \mathrm{N}^n_\Theta)$ where $\Theta \subseteq \ell_2$. □

Assume furthermore, that an estimator $\widetilde{\theta}$ of θ based on observable quantities is available which takes its values in $\mathbb{H}$ but does not necessarily belong to Θ. We shall measure the accuracy of any estimator $\widetilde{\theta}$ of θ by its distance $\mathfrak{d}_{\text{ist}}(\widetilde{\theta}, \theta)$ where $\mathfrak{d}_{\text{ist}}(\cdot,\cdot)$ is a certain semi metric to be specified below. Moreover, we call the quantity $\mathbb{E}^n_\theta\big(\mathfrak{d}^2_{\text{ist}}(\widetilde{\theta}, \theta)\big)$ risk of the estimator $\widetilde{\theta}$ of θ.

§10|00.02 **Definition.** Given an estimator $\widetilde{\theta}$ of a function of interest θ belonging to a class of solutions Θ based on observable quantities with probability measure $\mathbb{P}^n_\theta \in \mathbb{P}^n_\Theta$ we call

$$\mathcal{R}_n[\widetilde{\theta}\,|\,\Theta] := \sup\big\{\mathbb{E}^n_\theta\big(\mathfrak{d}^2_{\text{ist}}(\widetilde{\theta}, \theta)\big) : \theta \in \Theta\big\}$$

its *maximal risk* over Θ. □

§10|00.03 **Remark.** An advantage of taking a maximal risk instead of a risk is that the former does not depend on the unknown function θ. Imagine we would have taken a constant estimator, say $\widetilde{\theta} = h$, of θ. This would be the perfect estimator if by chance $\theta = h$, but in all other cases this estimator is likely to perform poorly. Therefore it is reasonable to consider the supremum over the whole class of possible functions in order to get consolidated findings. However, considering the maximal risk may be a very pessimistic point of view. □

§10|00.04 **Definition**. Consider a maximal risk $\mathcal{R}_n[\bullet\,|\,\Theta]$ over a family $\mathbb{P}^n_\Theta$ of probability measures. Let $\widehat{\theta}$ be an estimator of $\theta \in \Theta$, $\mathrm{C} \in \mathbb{R}_{>0}$ and for each $n \in \mathbb{N}$ let $\mathrm{R}^\star_n \in \mathbb{R}_{\geqslant 0}$ satisfy

(lower) $\mathrm{R}^\star_n$ is a *lower bound* up to the constant C^{-1} of the maximal risk over Θ, that is

$$\inf_{\widetilde{\theta}} \mathcal{R}_n[\widetilde{\theta}\,|\,\Theta] \geqslant \mathrm{C}^{-1}\,\mathrm{R}^\star_n$$

where the infimum is taken over all possible estimators of θ;

(upper) $\mathrm{R}^\star_n$ is an *upper bound* up to the constant C of the maximal risk over Θ, that is

$$\mathcal{R}_n[\widehat{\theta}\,|\,\Theta] \leqslant \mathrm{C}\,\mathrm{R}^\star_n$$

Then we call $\mathrm{R}^\star_n$ *minimax-bound* and the estimator $\widehat{\theta}$ *minimax-optimal* (up to the constant C). As a consequence, up to the constant C^2 the estimator $\widehat{\theta}$ attains the lower maximal risk bound that is, $\mathcal{R}_n[\widehat{\theta}\,|\,\Theta] \leqslant \mathrm{C}^2 \inf_{\widetilde{\theta}} \mathcal{R}_n[\widetilde{\theta}\,|\,\Theta]$. □

§10|00.05 **Remark**. We call a minimax-bound $(\mathrm{R}^\star_n)_{n\in\mathbb{N}}$ a *minimax-optimal rate* (of convergence) if in addition $\mathrm{R}^\star_n = \mathrm{o}(1)$ as $n \to \infty$. It is worth noting that a minimax-optimal rate is not unique since every other rate that is equivalent of order is also minimax-optimal. □

§10|00.06 **dSM (§01|03.05 continued)**. Given $\ell_2 = \mathbb{L}_2(\mathbb{N}, 2^{\mathbb{N}}, \nu_{\mathbb{N}})$ consider a Direct sequence model (dSM) as in §01|03.05. Here the observable stochastic process $\widehat{\theta}_\bullet = \theta_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ is a noisy version of $\theta_\bullet \in \Theta \subseteq \ell_2$ and $\dot{\varepsilon}_\bullet \sim \otimes_{j\in\mathbb{N}}\mathbb{P}^{\dot{\varepsilon}_j}$, where

(SM:ub) for $\sigma_\bullet \in \Sigma \subseteq \mathbb{R}^{\mathbb{N}}_{\geqslant 0} \cap \ell_\infty$ and $\mathrm{P}_{(0,1)} \in \mathscr{W}_2(\mathscr{B})$ we have $\mathbb{P}^{\dot{\varepsilon}_j} = \mathrm{P}_{(0,\sigma_j^2)}$ for all $j \in \mathbb{N}$,

Under **(SM:ub)** $\widehat{\theta}_\bullet$ admits a $\mathrm{P}^n_{\theta|\sigma}$-distribution belonging to the family $\mathrm{P}^n_{\Theta\times\Sigma} := (\mathrm{P}^n_{\theta|\sigma})_{\theta_\bullet\in\Theta,\sigma_\bullet\in\Sigma}$. Summarising the observations satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}^{\otimes\mathbb{N}}, \mathrm{P}^n_{\Theta\times\Sigma})$ where $\Theta \subseteq \ell_2$ and $\Sigma \subseteq \mathbb{R}^{\mathbb{N}}_{>0} \cap \ell_\infty$. □

More generally, given a class of solutions Θ, a class of nuissances parameters Ξ and a noise level $n \in \mathbb{N}$ let $\mathbb{P}^n_{\Theta\times\Xi} := (\mathbb{P}^n_{\theta|\xi})_{\theta\in\Theta,\xi\in\Xi}$ denote a family of probability measures. Moreover, we assume that the probability measure associated with an observable quantity belongs to $\mathbb{P}^n_{\Theta\times\Xi}$. Note that dismissing in Model §10|00.06 compared to Model §10|00.01 the assumption of a known sequence of variances $\sigma^2_\bullet$ the class of nuissances parameters Ξ equals Σ.

§10|00.07 **Definition**. Given an estimator $\widetilde{\theta}$ of a function of interest θ belonging to a class of solutions Θ based on observable quantities with probability measure $\mathbb{P}^n_{\theta|\xi} \in \mathbb{P}^n_{\Theta\times\Xi}$ we call

$$\mathcal{R}_n[\widetilde{\theta}\,|\,\Theta,\Xi] := \sup\left\{\mathbb{P}^n_{\theta|\xi}\big(\mathfrak{d}^2_{\mathrm{ist}}(\widetilde{\theta},\theta)\big) : \theta \in \Theta, \xi \in \Xi\right\}$$

its *maximal risk* over $\Theta \times \Xi$. □

§10|00.08 **diSM (§01|04.08 continued)**. Given $\ell_2 = \mathbb{L}_2(\mathbb{N}, 2^{\mathbb{N}}, \nu_{\mathbb{N}})$ and $\ell_\infty = \mathbb{L}_\infty(\mathbb{N}, 2^{\mathbb{N}}, \nu_{\mathbb{N}})$ consider a Diagonal inverse sequence model (diSM) as in §01|04.08 where $\mathfrak{s}_\bullet \in \ell_\infty$ is *known* in advance. Here the observable stochastic process $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ is a noisy version of $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \ell_2$ with $\theta_\bullet = \mathfrak{s}^\dagger_\bullet g_\bullet \in \Theta \subseteq \ell_2$ and $\dot{\varepsilon}_\bullet \sim \otimes_{j\in\mathbb{N}}\mathbb{P}^{\dot{\varepsilon}_j}$, where $\dot{\varepsilon}_\bullet$ satisfies **(SM:ub)** in Model §10|00.06 for $\Sigma \subseteq \mathbb{R}^{\mathbb{N}}_{\geqslant 0} \cap \ell_\infty$. Under **(SM:ub)** $\widehat{g}_\bullet$ admits a $\mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}$-distribution belonging to the family $\mathrm{P}^n_{\Theta\times\{\mathfrak{s}_\bullet\}\times\Sigma} := (\mathrm{P}^n_{\theta|\mathfrak{s}|\sigma})_{\theta_\bullet\in\Theta,\sigma_\bullet\in\Sigma}$. Summarising the observations satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}^{\otimes\mathbb{N}}, \mathrm{P}^n_{\Theta\times\{\mathfrak{s}_\bullet\}\times\Sigma})$ where $\Theta \subseteq \ell_2$ and $\Sigma \subseteq \mathbb{R}^{\mathbb{N}}_{>0} \cap \ell_\infty$. □

Given some transformation T defined on $\mathbb{H}$ let the probability measure associated with an observable quantity belong to a family of probability measures $\mathbb{P}^n_{\Theta\times\{\mathrm{T}\}\times\Xi} := (\mathbb{P}^n_{\theta|\mathrm{T}|\xi})_{\theta\in\Theta,\xi\in\Xi}$.

§10|00.09 **Definition**. Given an estimator $\widetilde{\theta}$ of a function of interest θ belonging to a class of solutions Θ based on an observable quantity with probability measure $\mathbb{P}^n_{g,\xi} \in \mathbb{P}^n_{\Theta\times\{T\}\times\Xi}$ we call

$$\mathcal{R}_n[\widetilde{\theta}\,|\,\Theta,\{T\},\Xi] := \sup\{\mathbb{E}^n_{\theta|T|\xi}\big(\mathfrak{d}^2_{\mathrm{ist}}(\widetilde{\theta},\theta)\big) : \theta\in\Theta, \xi\in\Xi\}$$

its *maximal risk* over $\Theta\times\{T\}\times\Xi$. □

§10|00.10 **diSM with noisy operator (§02|04.05 continued)**. Given $\ell_2 = \mathbb{L}_2(\mathbb{N},2^{\mathbb{N}},\nu_{\mathbb{N}})$ and $\ell_\infty = \mathbb{L}_\infty(\mathbb{N},2^{\mathbb{N}},\nu_{\mathbb{N}})$ consider a Diagonal inverse sequence model (diSM) with noisy operator as in §02|04.05 where $\mathfrak{s}_\bullet \in \mathcal{S} \subseteq \mathbb{R}^{\mathbb{N}}_{\setminus 0}\cap\ell_\infty$ is *not known* anymore. Here the observable stochastic process $\widehat{\mathfrak{s}}_\bullet = \mathfrak{s}_\bullet + k^{-1/2}\dot{\eta}_\bullet$ and $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet$ is a noisy version of $\mathfrak{s}_\bullet\in\mathcal{S}\subseteq\mathbb{R}^{\mathbb{N}}_{\setminus 0}\cap\ell_\infty$ and $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}})\subseteq\ell_2$ with $\theta_\bullet\in\Theta\subseteq\ell_2$, respectively, where $\dot{\varepsilon}_\bullet \sim \otimes_{j\in\mathbb{N}}\mathbb{P}^{\dot{\varepsilon}_j}$ and $\dot{\eta}_\bullet \sim \otimes_{j\in\mathbb{N}}\mathbb{P}^{\dot{\eta}_j}$ are *independent*. In addition, let $\dot{\varepsilon}_\bullet$ satisfy **(SM:ub)** in Model §10|00.06 for $\sigma_\bullet\in\Sigma\subseteq\mathbb{R}^{\mathbb{N}}_{\geqslant 0}\cap\ell_\infty$ and let $\dot{\eta}_\bullet$ fulfill

(SMnO:ub) for $\xi_\bullet\in\Xi\subseteq\mathbb{R}^{\mathbb{N}}_{\geqslant 0}\cap\ell_\infty$ we have $\mathbb{P}^{\dot{\eta}_j}\in\mathscr{W}_4(\mathscr{B})$ with $\xi^4_j = \mathbb{P}(\dot{\eta}^4_j)$ and $0 = \mathbb{P}(\dot{\eta}_j)$, $j\in\mathbb{N}$.

Under **(SM:ub)** $\widehat{g}_\bullet$ admits a $\mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}$-distribution belonging to the family $\mathrm{P}^n_{\Theta\times\mathcal{S}\times\Sigma} := (\mathrm{P}^n_{\theta|\mathfrak{s}|\sigma})_{\theta_\bullet\in\Theta,\mathfrak{s}_\bullet\in\mathcal{S},\sigma_\bullet\in\Sigma}$ and under **(SMnO:ub)** $\widehat{\mathfrak{s}}_\bullet$ admits a $\mathrm{P}^k_{\mathfrak{s}|\xi}$-distribution belonging to the family $\mathrm{P}^k_{\mathcal{S}\times\Xi} := (\mathrm{P}^k_{\mathfrak{s}|\xi})_{\mathfrak{s}_\bullet\in\mathcal{S},\xi_\bullet\in\Xi}$. Consequently, $(\widehat{g}_\bullet,\widehat{\mathfrak{s}}_\bullet)$ admits a joint $\mathrm{P}^{n\otimes k}_{\theta|\mathfrak{s}|\sigma|\xi} = \mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}\otimes\mathrm{P}^k_{\mathfrak{s}|\xi}$ distribution belonging to the family $\mathrm{P}^{n\otimes k}_{\Theta\times\mathcal{S}\times\Sigma\times\Xi} := (\mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}\otimes\mathrm{P}^k_{\mathfrak{s}|\xi})_{\theta_\bullet\in\Theta,\mathfrak{s}_\bullet\in\mathcal{S},\sigma\in\Sigma,\xi\in\Xi}$. Summarising the observations satisfy a statistical product experiment $\big(\mathbb{R}^{\mathbb{N}^2},\mathscr{B}^{\otimes\mathbb{N}^2},\mathrm{P}^{n\otimes k}_{\Theta\times\mathcal{S}\times\Sigma\times\Xi}\big)$ where $\Sigma,\Xi\subseteq\mathbb{R}^{\mathbb{N}}_{>0}\cap\ell_\infty$. $\mathcal{S}\subseteq\mathbb{R}^{\mathbb{N}}_{\setminus 0}\cap\ell_\infty$ and $\Theta\subseteq\ell_2$. □

Finally, given a class of solutions Θ, a class of operators $\mathbb{T}$, a class of nuissance parameters Ξ and noise levels $n,k\in\mathbb{N}$ let $\mathbb{P}^{n,k}_{\Theta\times\mathbb{T}\times\Xi} := (\mathbb{P}^{n,k}_{\theta|T|\xi})_{\theta\in\Theta,T\in\mathbb{T},\xi\in\Xi}$ denote a family of joint probability measures.

§10|00.11 **Definition**. Given an estimator $\widetilde{\theta}$ of a function of interest θ belonging to a class of solutions Θ based on observable quantities with joint probability measure $\mathbb{P}^{n,k}_{\theta|T|\xi}\in\mathbb{P}^{n,k}_{\Theta\times\mathbb{T}\times\Xi}$ we call

$$\mathcal{R}_{n,k}[\widetilde{\theta}\,|\,\Theta,\mathbb{T},\Xi] := \sup\big\{\mathbb{E}^{n,k}_{\theta|T|\xi}\big(\mathfrak{d}^2_{\mathrm{ist}}(\widetilde{\theta},\theta)\big) : \theta\in\Theta, T\in\mathbb{T},\xi\in\Xi\big\}$$

its *maximal risk* over $\Theta\times\mathbb{T}\times\Xi$. □

§10|00.12 **Remark**. Taking the supremum over the class of operators allows us to quantify the additional complexity due to the estimation of the operator. Moreover, if there exist an estimator $\widehat{\theta}$, a constant $\mathrm{C}\in\mathbb{R}_{>0}$ and for each $n,k\in\mathbb{N}$ there is $\mathrm{R}^\star_{n,k}\in\mathbb{R}_{\geqslant 0}$ such that

(lower) $\mathrm{R}^\star_{n,k}$ is a *lower bound* up to the constant C^{-1} of the maximal risk over $\Theta\times\mathbb{T}\times\Xi$, that is

$$\inf_{\widetilde{\theta}}\mathcal{R}_{n,k}[\widetilde{\theta}\,|\,\Theta,\mathbb{T},\Xi] \geqslant \mathrm{C}^{-1}\,\mathrm{R}^\star_{n,k}$$

where the infimum is taken over all possible estimators of θ;

(upper) $\mathrm{R}^\star_{n,k}$ is an *upper bound* up to the constant C of the maximal risk over $\Theta\times\mathbb{T}\times\Xi$, that is

$$\mathcal{R}_{n,k}[\widehat{\theta}\,|\,\Theta,\mathbb{T},\Xi] \leqslant \mathrm{C}\,\mathrm{R}^\star_{n,k},$$

then we call $\mathrm{R}^\star_{n,k}$ *minimax-bound* and the estimator $\widehat{\theta}$ *minimax-optimal* (up to the constant C). As a consequence, up to the constant C^2 the estimator $\widehat{\theta}$ attains the lower maximal risk bound that is, $\mathcal{R}_{n,k}[\widehat{\theta}\,|\,\Theta,\mathbb{T},\Xi] \leqslant \mathrm{C}^2\inf_{\widetilde{\theta}}\mathcal{R}_{n,k}[\widetilde{\theta}\,|\,\Theta,\mathbb{T},\Xi]$. Typically, we assume first that the nuissance parameter ξ is known *a priori*, and hence $\mathbb{P}^{n,k}_{\Theta,\mathbb{T},\{\xi\}}$ is a class of probability measures associated with the observable quantities. In this situation, we consider the maximal risk $\big\{\mathbb{E}^{n,k}_{\theta|T|\xi}\big(\mathfrak{d}^2_{\mathrm{ist}}(\widetilde{\theta},\theta)\big) : \theta\in\Theta, T\in\mathbb{T}\big\}$ and we seek a bound $\mathrm{R}^\star_{n,k}$ up to a constant which depends possibly on the nuissance parameter ξ. However, if the bound $\mathrm{R}^\star_{n,k}$ is a valid lower and upper bound up to a constant uniformly for all nuissance parameters $\xi\in\Xi$, then it is, obviously, also a bound of the maximal risk $\mathcal{R}_{n,k}[\widehat{\theta}\,|\,\Theta,\mathbb{T},\Xi]$. □

Considering the Hilbert space $\ell_2 = \mathbb{L}_2(\mathbb{N}, 2^{\mathbb{N}}, \nu_{\mathbb{N}})$ and a surjective partial isometry $U \in \mathbb{L}(\mathbb{H}, \ell)$, which is *fixed* and presumed to be *known* in advance, we study *statistical inverse problems* where observable quantities admit a probability measure $\mathbb{P}^{n,k}_{\theta|T|\xi} \in \mathbb{P}^{n,k}_{\Theta\times\mathbb{T}\times\Xi}$ for some class Θ, $\mathbb{T}$ and Ξ of solutions, operators and nuissance parameters, respectively. We consider the following global and local measures of accuracy.

§10|00.13 **Notation** (Reminder). For sequences $a_{\bullet}, b_{\bullet} \in (\mathbb{K})^{\mathbb{N}}$ taking its values in $\mathbb{K} \in \{\mathbb{R}, \mathbb{R}_{>0}, \mathbb{Z}, \dots\}$ we write $a_{\bullet} \in (\mathbb{K})^{\mathbb{N}}_{\nearrow}$ and $b_{\bullet} \in (\mathbb{K})^{\mathbb{N}}_{\searrow}$ if $a_{\bullet}$ and $b_{\bullet}$, respectively, is monotonically *non-decreasing* and *non-increasing*. If in addition $a_n \to \infty$ and $b_n \to 0$ as $n \to \infty$, then we write $a_{\bullet} \in (\mathbb{K})^{\mathbb{N}}_{\uparrow\infty}$ and $b_{\bullet} \in (\mathbb{K})^{\mathbb{N}}_{\downarrow 0}$ for short. For $w_{\bullet} \in \ell_\infty = \mathbb{L}_\infty(\nu_{\mathbb{N}})$ we set $w_{[0]} := 0$, $w_{[\bullet]} = (w_{[j]} := \|w_{\bullet}\mathbb{1}^{j}_{\bullet}\|_{\ell_\infty})_{j\in\mathbb{N}}$, $w_{(0)} := \|w_{\bullet}\|_{\ell_\infty}$, and $w_{(\bullet)} = (w_{(j)} := \|w_{\bullet}\mathbb{1}^{j|\perp}_{\bullet}\|_{\ell_\infty})_{j\in\mathbb{N}}$, where by construction $w_{[\bullet]} \in (\mathbb{R}_{\geqslant 0})^{\mathbb{N}}_{\nearrow}$ and $w_{(\bullet)} \in (\mathbb{R}_{\geqslant 0})^{\mathbb{N}}_{\searrow}$. □

§10|00.14 **Assumption** (Maximal global $\mathfrak{v}$-risk). Consider weights $\mathfrak{t}_{\bullet}, \mathfrak{a}_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ and $\mathfrak{v}_{\bullet} \in \mathbb{R}^{\mathbb{N}}_{>0}$ such that $(\mathfrak{a}\mathfrak{v})_{\bullet} = \mathfrak{a}_{\bullet}\mathfrak{v}_{\bullet} \in \ell_\infty$, and $(\mathfrak{a}\mathfrak{v})_{(\bullet)} \in (\mathbb{R}_{\geqslant 0})^{\mathbb{N}}_{\downarrow 0}$ and there exists $C_{(\mathfrak{a}\mathfrak{v})} \in (0, 1]$ such that for all $m \in \mathbb{N}$

$$(\mathfrak{a}\mathfrak{v})^2_{(m-1)} \geqslant \min\left\{(\mathfrak{a}\mathfrak{v})^2_j : j \in [\![m]\!]\right\} \geqslant C_{(\mathfrak{a}\mathfrak{v})}(\mathfrak{a}\mathfrak{v})^2_{(m-1)}$$

or in equal $1 \geqslant C_{(\mathfrak{a}\mathfrak{v})}\|(\mathfrak{a}\mathfrak{v})^{-2}_{\bullet}\mathbb{1}^m_{\bullet}\|_{\ell_\infty}(\mathfrak{a}\mathfrak{v})^2_{(m-1)}$. □

§10|00.15 **Reminder** (Maximal global $\mathfrak{v}$-risk). Under Assumption §10|00.14 we introduce $\ell_2^{\mathfrak{a}} = \mathrm{dom}(M_{\mathfrak{a}^{-1}}) = \ell_2\mathfrak{a}_{\bullet} = \ell_2(\mathfrak{a}^{-2}_{\bullet})$ endowed with $\|\cdot\|_{\mathfrak{a}^{-1}} := \|\cdot\|_{\ell_2(\mathfrak{a}^{-2}_{\bullet})}$ and the ellipsod $\ell_2^{\mathfrak{a},\mathrm{r}} := \left\{h_{\bullet} \in \ell_2^{\mathfrak{a}} : \|h_{\bullet}\|^2_{\mathfrak{a}^{-1}} \leqslant \mathrm{r}^2\right\} \subseteq \ell_2^{\mathfrak{a}}$, where the measures $\nu_{\mathbb{N}}$, $\mathfrak{v}^2_{\bullet}\nu_{\mathbb{N}}$ and $\mathfrak{a}^{-2}_{\bullet}\nu_{\mathbb{N}}$ dominate mutually each other. Under Assumption §10|00.14 we consider the following global measure of accuracy. Introduce $\ell_2(\mathfrak{v}^2_{\bullet}) = \mathbb{L}_2(\mathfrak{v}^2_{\bullet}\nu_{\mathbb{N}}) = \mathrm{dom}(M_{\mathfrak{v}}) = \ell_2\mathfrak{v}^{-1}_{\bullet} \subseteq \ell_2$ and $\|\cdot\|_{\mathfrak{v}} = \|M_{\mathfrak{v}}\cdot\|_{\ell_2}$, where $\ell_2^{\mathfrak{a},\mathrm{r}} \subseteq \ell_2(\mathfrak{v}^2_{\bullet})$ (Property §04|02.11). For $\theta_{\bullet} = U\theta \in \ell_2^{\mathfrak{a},\mathrm{r}}$ we call $\mathfrak{d}_{\mathrm{ist}}(\widetilde{\theta}_{\bullet}, \theta_{\bullet}) = \|\widetilde{\theta}_{\bullet} - \theta_{\bullet}\|_{\mathfrak{v}}$ *global $\mathfrak{v}$-error*, $\mathbb{P}^{n,k}_{\theta|T|\xi}\left(\|\widetilde{\theta}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}\right)$ *global $\mathfrak{v}$-risk* and

$$\mathcal{R}^{\mathfrak{v}}_{n,k}[\widetilde{\theta}_{\bullet} \mid \ell_2^{\mathfrak{a},\mathrm{r}}, \mathbb{T}, \Xi] := \sup\left\{\mathbb{P}^{n,k}_{\theta|T|\xi}\left(\|\widetilde{\theta}_{\bullet} - \theta_{\bullet}\|^2_{\mathfrak{v}}\right) : \theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}, T \in \mathbb{T}, \xi \in \Xi\right\}$$

maximal $\mathfrak{v}$-risk over $\ell_2^{\mathfrak{a},\mathrm{r}} \times \mathbb{T} \times \Xi$. Note that $(\mathfrak{a}\mathfrak{v})^2_{(\bullet)} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ by definition, hence $(\mathfrak{a}\mathfrak{v})^2_{(\bullet)} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$ is satisfied if and only if $(\mathfrak{a}\mathfrak{v})^2_{(m)} = o(1)$ as $m \to \infty$ (i.e. the maximal global approximation is consistent). Moreover if $(\mathfrak{a}\mathfrak{v})^2_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$ then we have trivially $(\mathfrak{a}\mathfrak{v})^2_{(\bullet)} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$ and $\|(\mathfrak{a}\mathfrak{v})^{-2}_{\bullet}\mathbb{1}^m_{\bullet}\|_{\ell_\infty} = (\mathfrak{a}\mathfrak{v})^{-2}_m = (\mathfrak{a}\mathfrak{v})^{-2}_{(m-1)}$ for all $m \in \mathbb{N}$, i.e. Assumption §10|00.14 is satisfied with $C_{(\mathfrak{a}\mathfrak{v})} = 1$. □

§10|00.16 **Assumption** (Maximal local ϕ-risk). Consider weights $\mathfrak{t}_{\bullet}, \mathfrak{a}_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ and $\phi \in \mathbb{R}^{\mathbb{N}}_{\setminus 0}$ such that $(\mathfrak{a}\phi)_{\bullet} := \mathfrak{a}_{\bullet}\phi_{\bullet} \in \ell_2$ and $(\mathfrak{a}\mathfrak{t})_{\bullet} := \mathfrak{a}_{\bullet}\mathfrak{t}_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$. □

§10|00.17 **Reminder** (Maximal local ϕ-risk). Under Assumption §10|00.16 introduce $\ell_2^{\mathfrak{a}} = \ell_2(\mathfrak{a}^{-2}_{\bullet})$ endowed with $\|\cdot\|_{\mathfrak{a}^{-1}} := \|\cdot\|_{\ell_2(\mathfrak{a}^{-2}_{\bullet})}$ and the ellipsod $\ell_2^{\mathfrak{a},\mathrm{r}} := \left\{h_{\bullet} \in \ell_2^{\mathfrak{a}} : \|h_{\bullet}\|^2_{\mathfrak{a}^{-1}} \leqslant \mathrm{r}^2\right\} \subseteq \ell_2^{\mathfrak{a}}$, where the measures $\nu_{\mathbb{N}}$, $|\phi_{\bullet}|\nu_{\mathbb{N}}$ and $\mathfrak{a}^{-2}_{\bullet}\nu_{\mathbb{N}}$ dominate mutually each other. Under Assumption §10|00.16 we consider the following local measure of accuracy. Under Assumption §10|00.16 introduce $\mathrm{dom}(\phi\nu_{\mathbb{N}}) := \left\{h_{\bullet} \in \ell_2 : \phi_{\bullet}h_{\bullet} \in \ell_1 = \mathbb{L}_1(\mathbb{N}, 2^{\mathbb{N}}, \nu_{\mathbb{N}})\right\}$ and the linear functional $\phi\nu_{\mathbb{N}} : \ell_2 \supseteq \mathrm{dom}(\phi\nu_{\mathbb{N}}) \to \mathbb{R}$ with $h_{\bullet} \mapsto \phi\nu_{\mathbb{N}}(h_{\bullet}) := \nu_{\mathbb{N}}(\phi_{\bullet}h_{\bullet})$ where $\ell_2^{\mathfrak{a},\mathrm{r}} \subseteq \mathrm{dom}(\phi\nu_{\mathbb{N}})$ (Property §04|02.23). For $\theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}$ we call $\mathfrak{d}_{\mathrm{ist}}(\widetilde{\theta}_{\bullet}, \theta_{\bullet}) = |\phi\nu_{\mathbb{N}}(\widetilde{\theta}_{\bullet} - \theta_{\bullet})|$ *local ϕ-error*, $\mathbb{P}^{n,k}_{\theta|T|\xi}\left(|\phi\nu_{\mathbb{N}}(\widetilde{\theta}_{\bullet} - \theta_{\bullet})|^2\right)$ *local ϕ-risk* and

$$\mathcal{R}^{\phi}_{n,k}[\widetilde{\theta}_{\bullet} \mid \ell_2^{\mathfrak{a},\mathrm{r}}, \mathbb{T}, \Xi] := \sup\left\{\mathbb{P}^{n,k}_{\theta|T|\xi}\left(|\phi\nu_{\mathbb{N}}(\widetilde{\theta}_{\bullet} - \theta_{\bullet})|^2\right) : \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}, T \in \mathbb{T}, \xi \in \Xi\right\}$$

maximal ϕ-risk over $\ell_2^{\mathfrak{a},\mathrm{r}} \times \mathbb{T} \times \Xi$. Imposing by Assumption §10|00.16 $\mathfrak{t}_{\bullet}, \mathfrak{a}_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ and hence $(\mathfrak{a}\mathfrak{t})^2_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\searrow}$ is rather weak. If in addition $\liminf_{j\to\infty} (\mathfrak{a}\mathfrak{t})^2_j \geqslant c \in \mathbb{R}_{>0}$ is satisfied, and hence

$(\mathfrak{a}\mathfrak{t})^2_{\bullet}, \mathfrak{a}^2_{\bullet}, \mathfrak{t}^2_{\bullet} \notin (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$, then $\mathfrak{a}^2_{\bullet} \notin (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$ and the assumption $(\mathfrak{a}\phi)_{\bullet} \in \ell_2$ implies $\phi_{\bullet} \in \ell_2$, which together with $\mathfrak{t}^2_{\bullet} \notin (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$ implies $(\phi/\mathfrak{t})_{\bullet} \in \ell_2$, and thus the rate $\mathrm{R}^{\star}_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi)$ is parametric (Illustration §07|01.78 or Illustration §08|01.72). Since we are interested in the case of a non-parametric rate, the additional assumption $(\mathfrak{a}\mathfrak{t})^2_{\bullet} \in (\mathbb{R}_{>0})^{\mathbb{N}}_{\downarrow 0}$ imposes a rather weak condition satisfied also in Illustration §07|01.78 or Illustration §08|01.72. □

§10|00.18 **Comment**. We formulate the results in terms of $\theta_{\bullet} = \mathrm{U}\theta \in \mathbb{J}$ rather than directly for $\theta \in \mathbb{H}$. Since U is known, considering the class $\mathbb{H}^{\mathfrak{a},\mathrm{r}} := \mathrm{U}^{\star}\mathbb{J}^{\mathfrak{a},\mathrm{r}} := \{\mathrm{U}^{\star}\theta_{\bullet} \colon \theta_{\bullet} \in \mathbb{J}^{\mathfrak{a},\mathrm{r}}\}$ we obtain immediately also bounds over $\mathbb{H}^{\mathfrak{a},\mathrm{r}}$ for the maximal global risk

$$\mathcal{R}^{\mathfrak{v}}_{n,k}[\widetilde{\theta}\,|\,\mathrm{U}^{\star}\mathbb{J}^{\mathfrak{a},\mathrm{r}}, \mathbb{T}, \Xi\,] := \sup\left\{\mathbb{E}^{n,k}_{\theta|\mathrm{T}|\xi}\left(\|\mathrm{U}(\widetilde{\theta}-\theta)\|^2_{\mathfrak{v}}\right) \colon \theta \in \mathbb{H}^{\mathfrak{a},\mathrm{r}}, \mathrm{T} \in \mathbb{T}, \xi \in \Xi\right\}$$

and maximal local risk

$$\mathcal{R}^{\phi}_{n,k}[\widetilde{\theta}\,|\,\mathrm{U}^{\star}\mathbb{J}^{\mathfrak{a},\mathrm{r}}, \mathbb{T}, \Xi\,] := \sup\left\{\mathbb{E}^{n,k}_{\theta|\mathrm{T}|\xi}\left(|\phi\nu_{\mathbb{N}}(\mathrm{U}(\widetilde{\theta}-\theta))|^2\right) \colon \theta \in \mathbb{H}^{\mathfrak{a},\mathrm{r}}, \mathrm{T} \in \mathbb{T}, \xi \in \Xi\right\}$$

which we do not explicitly state in the sequel. □

§11 Deriving a lower bound: a general reduction scheme

For a detailed discussion of several other strategies to derive lower bounds we refer the reader, for example, to the text book by Tsybakov [2009].

§11|00.01 **Definition**. Let $\mathbb{P}_0$ and $\mathbb{P}_1$ be two probability measures on a measurable space $(\mathcal{X}, \mathscr{X})$.

(a) The function

$$\mathrm{KL}(\mathbb{P}_0|\mathbb{P}_1) = \begin{cases} \mathbb{P}_0\left(\log\frac{\mathrm{d}\mathbb{P}_0}{\mathrm{d}\mathbb{P}_1}\right) = \int \log\left(\frac{\mathrm{d}\mathbb{P}_0}{\mathrm{d}\mathbb{P}_1}\right)d\mathbb{P}_0, & \text{if } \mathbb{P}_0 \ll \mathbb{P}_1, \\ +\infty, & \text{otherwise} \end{cases}$$

is called *Kullback-Leibler-divergence* of $\mathbb{P}_0$ with respect to $\mathbb{P}_1$.

Let $\mu \in \mathscr{M}_{\sigma}(\mathscr{X})$ be a $\mathbb{P}_0$ and $\mathbb{P}_1$ dominating σ-finite measure (e.g. $\mathbb{P}_0, \mathbb{P}_1 \ll \mu = \mathbb{P}_0 + \mathbb{P}_1$). We write $\mathrm{d}\mathbb{P}_0 := \mathrm{d}\mathbb{P}_0/\mathrm{d}\mu$ and $\mathrm{d}\mathbb{P}_1 := \mathrm{d}\mathbb{P}_1/\mathrm{d}\mu$ for short.

(b) The *Hellinger distance* between $\mathbb{P}_0$ and $\mathbb{P}_1$ is defined by

$$\mathrm{H}(\mathbb{P}_0, \mathbb{P}_1) := \left(\int |\sqrt{\mathrm{d}\mathbb{P}_0} - \sqrt{\mathrm{d}\mathbb{P}_1}|^2\right)^{1/2} := \|\sqrt{\mathrm{d}\mathbb{P}_0} - \sqrt{\mathrm{d}\mathbb{P}_1}\|_{\mathbb{L}_2(\mu)}$$

(c) and the *Hellinger affinity* is given by

$$\rho(\mathbb{P}_0, \mathbb{P}_1) := \int \sqrt{\mathrm{d}\mathbb{P}_0}\sqrt{\mathrm{d}\mathbb{P}_1} := \langle\sqrt{\mathrm{d}\mathbb{P}_0}, \sqrt{\mathrm{d}\mathbb{P}_1}\rangle_{\mathbb{L}_2(\mu)},$$

where both do not depend on the choice of the dominating measure μ. □

§11|00.02 **Remark**. The Kullback-Leibler-divergence satisfies $\mathrm{KL}(\mathbb{P}_0|\mathbb{P}_1) \geqslant 0$ as well as $\mathrm{KL}(\mathbb{P}_0|\mathbb{P}_1) = 0$ if and only if $\mathbb{P}_0 = \mathbb{P}_1$, but $\mathrm{KL}(\cdot|\cdot)$ is not symmetric. Moreover, for product measures holds $\mathrm{KL}(\mathbb{P}_{0,1} \otimes \mathbb{P}_{0,2}|\mathbb{P}_{1,1} \otimes \mathbb{P}_{1,2}) = \mathrm{KL}(\mathbb{P}_{0,1}|\mathbb{P}_{1,1}) + \mathrm{KL}(\mathbb{P}_{0,2}|\mathbb{P}_{1,2})$ and $\rho(\mathbb{P}_{0,1} \otimes \mathbb{P}_{0,2}, \mathbb{P}_{1,1} \otimes \mathbb{P}_{1,2}) = \rho(\mathbb{P}_{0,1}, \mathbb{P}_{1,1})\rho(\mathbb{P}_{0,2}, \mathbb{P}_{1,2})$. □

§11|00.03 **Lemma**. (i) $0 \leqslant \mathrm{H}^2(\mathbb{P}_0, \mathbb{P}_1) \leqslant 2$; (ii) $\rho(\mathbb{P}_0, \mathbb{P}_1) = 1 - \frac{1}{2}\mathrm{H}^2(\mathbb{P}_0, \mathbb{P}_1)$; *and* (iii) $\mathrm{H}^2(\mathbb{P}_0, \mathbb{P}_1) \leqslant \mathrm{KL}(\mathbb{P}_0|\mathbb{P}_1)$.

§11|00.04 **Proof** of Lemma §11|00.03. Given in the lecture course Statistik 2 (Lemma §13.12, p.54). □

§11|00.05 **Lemma**. *For $a_\bullet, b_\bullet \in \ell_\infty$ and $n \in \mathbb{N}$ we have* $\mathrm{KL}(\mathrm{N}^n_a|\mathrm{N}^n_b) = \frac{n}{2}\|a_\bullet - b_\bullet\|^2_{\ell_2}$.

§11|00.06 **Proof** of Lemma §11|00.05. Given in the lecture course Statistik 2 (Lemma §13.14, p.54). □

§11|00.07 **Assumption**. The distribution $\mathbb{P} \in \mathscr{W}(\mathscr{B})$ admits a Lebesgue-density $\mathbb{p} := \mathrm{d}\mathbb{P}/\mathrm{d}\lambda$ and there exist constants $\mathrm{C}_\circ, x_\circ \in \mathbb{R}_{>0}$ such that

$$\forall x \in [-x_\circ, x_\circ]: \quad \int \mathbb{p}(u)\log\Big(\frac{\mathbb{p}(u)}{\mathbb{p}(u-x)}\Big)\lambda(\mathrm{d}u) \leqslant \mathrm{C}_\circ x^2.$$

□

§11|00.08 **Lemma**. *Let $Y_\bullet \sim \mathbb{P}^{\otimes\mathbb{N}}$ where $\mathbb{P} \in \mathscr{W}(\mathscr{B})$ fulfills Assumption §11|00.07 with $\mathrm{C}_\circ, x_\circ \in \mathbb{R}_{>0}$. For $a_\bullet, b_\bullet, \sigma_\bullet \in \ell_\infty$ and $n \in \mathbb{N}$ consider $a_\bullet + n^{-1/2}\sigma_\bullet Y_\bullet \sim \mathbb{P}^n_{a|\sigma}$ and $b_\bullet + n^{-1/2}\sigma_\bullet Y_\bullet \sim \mathbb{P}^n_{b|\sigma}$. If $\|\sigma_\bullet^{-2}\|_{\ell_\infty} =: \mathbb{v}_\sigma \in \mathbb{R}_{>0}$ and $n^{1/2}\mathbb{v}_\sigma^{1/2}\|a_\bullet - b_\bullet\|_{\ell_\infty} \leqslant x_\circ$ then we have $\mathrm{KL}(\mathbb{P}^n_{a|\sigma}|\mathbb{P}^n_{b|\sigma}) \leqslant n\mathbb{v}_\sigma\mathrm{C}_\circ\|a_\bullet - b_\bullet\|^2_{\ell_2}$.*

§11|00.09 **Proof** of Lemma §11|00.08. Given in the lecture. □

§11|00.10 **Comment**. For $\sigma \in \mathbb{R}_{\setminus 0}$ the normal distribution $\mathrm{N}_{(0,\sigma^2)} \in \mathscr{W}(\mathscr{B})$ satisfy Assumption §11|00.07 with $\mathrm{C}_\circ = 1/(2\sigma^2)$ and $x_\circ = \infty$ (see Proof §11|00.06). □

§11|00.11 **Assumption**. The semi metric $\mathfrak{d}_{\mathrm{ist}}(\cdot,\cdot)$ is *symmetric* and satisfies the *triangular inequality*. Morever, for any estimator $\widetilde{\theta}$ and parameter θ_0 and θ_1 such that $\mathfrak{d}_{\mathrm{ist}}(\theta_0, \theta_1) \in \mathbb{R}_{>0}$ the quantities $\mathfrak{d}_{\mathrm{ist}}(\widetilde{\theta}, \theta_0)$ and $\mathfrak{d}_{\mathrm{ist}}(\widetilde{\theta}, \theta_1)$ are *measurable*. □

§11|00.12 **Lemma**. *Let $(\mathcal{X}, \mathscr{X})$ be a measurable space, let θ_0 and θ_1 be parameters with $\mathfrak{d}_{\mathrm{ist}}(\theta^0, \theta^1) \in \mathbb{R}_{>0}$, and let Assumption §11|00.11 be satisfied.*

(i) *If $\mathbb{P}_0, \mathbb{P}_1 \in \mathscr{W}(\mathscr{X})$ are probability measures then for any estimator $\widetilde{\theta}$ we have*

$$\mathbb{P}_0\big(\mathfrak{d}^2_{\mathrm{ist}}(\widetilde{\theta}, \theta^0)\big) + \mathbb{P}_1\big(\mathfrak{d}^2_{\mathrm{ist}}(\widetilde{\theta}, \theta^1)\big) \geqslant \frac{1}{2}\,\mathfrak{d}^2_{\mathrm{ist}}(\theta^0, \theta^1)\,\rho^2(\mathbb{P}_0, \mathbb{P}_1). \tag{11.01}$$

(ii) *If $\mathbb{P}_0, \mathbb{P}_1 \in \mathscr{W}(\mathscr{X})$ satisfy $\mathrm{H}(\mathbb{P}_0, \mathbb{P}_1) \leqslant 1$, then for any estimator $\widetilde{\theta}$ we have*

$$\mathbb{P}_0\big(\mathfrak{d}^2_{\mathrm{ist}}(\widetilde{\theta}, \theta^0)\big) + \mathbb{P}_1\big(\mathfrak{d}^2_{\mathrm{ist}}(\widetilde{\theta}, \theta^1)\big) \geqslant \frac{1}{8}\,\mathfrak{d}^2_{\mathrm{ist}}(\theta^0, \theta^1). \tag{11.02}$$

(iii) *For $n \in \mathbb{N}_{\geqslant 2}$ let $\mathbb{P}^n_0 := \otimes_{j\in[\![n]\!]}\mathbb{P}_{0|j} \in \mathscr{W}(\mathscr{X}^{\otimes n})$ and $\mathbb{P}^n_1 := \otimes_{j\in[\![n]\!]}\mathbb{P}_{1|j} \in \mathscr{W}(\mathscr{X}^{\otimes n})$ be product probability measures with marginals $\mathbb{P}_{0|j}, \mathbb{P}_{1|j} \in \mathscr{W}(\mathscr{X})$ fulfilling $\mathrm{H}(\mathbb{P}_{0|j}, \mathbb{P}_{1|j}) \leqslant 2n^{-1}$ for each $j \in [\![n]\!]$. Then for any estimator $\widetilde{\theta}$ we have*

$$\mathbb{P}^n_0\big(\mathfrak{d}^2_{\mathrm{ist}}(\widetilde{\theta}, \theta^0)\big) + \mathbb{P}^n_1\big(\mathfrak{d}^2_{\mathrm{ist}}(\widetilde{\theta}, \theta^1)\big) \geqslant \frac{1}{32}\,\mathfrak{d}^2_{\mathrm{ist}}(\theta^0, \theta^1). \tag{11.03}$$

§11|00.13 **Proof** of Lemma §11|00.12. Given in the lecture. □

§11|01 Lower bound based on two hypothesis

§11|01.01 **Lemma** (Lower bound based on two hypothesis). *Given a noise level $n \in \mathbb{N}$ let $\mathbb{P}^n_\Theta := (\mathbb{P}^n_\theta)_{\theta\in\Theta}$ be a family of probability measures. We measure the accuracy of an estimator $\widetilde{\theta}$ by its maximal risk*

$$\mathcal{R}_n[\widetilde{\theta}\,|\,\Theta] := \sup\big\{\mathbb{P}^n_\theta\big(\mathfrak{d}^2_{\mathrm{ist}}(\widetilde{\theta}, \theta)\big) : \theta \in \Theta\big\}.$$

(i) *If there are $\theta^0, \theta^1 \in \Theta$ with $\mathfrak{d}_{ist}(\theta^0, \theta^1) \in \mathbb{R}_{>0}$ and associated probability measures $\mathbb{P}^n_{\theta^0}$ and $\mathbb{P}^n_{\theta^1}$ such that Assumption §11|00.11 and $\mathrm{H}(\mathbb{P}^n_{\theta^0}, \mathbb{P}^n_{\theta^1}) \leqslant 1$ are satisfied then we have*

$$\inf_{\widetilde{\theta}} \mathcal{R}_n[\widetilde{\theta} \,|\, \Theta] \geqslant \frac{1}{16} \mathfrak{d}^2_{ist}(\theta^0, \theta^1) \qquad (11.04)$$

where the infimum is taken over all possible estimators.

(ii) *Let $n \in \mathbb{N}_{\geqslant 2}$ and for each $\theta \in \Theta$ let $\mathbb{P}^n_\theta = \otimes_{j \in [\![n]\!]} \mathbb{P}_\theta$ be a product probability measure with identically $\mathbb{P}_\theta$-distributed marginals. If there are $\theta^0, \theta^1 \in \Theta$ with $\mathfrak{d}_{ist}(\theta^0, \theta^1) \in \mathbb{R}_{>0}$ and associated marginal probability measures $\mathbb{P}_{\theta^0}$ and $\mathbb{P}_{\theta^1}$ such that Assumption §11|00.11 and $\mathrm{H}(\mathbb{P}_{\theta^0}, \mathbb{P}_{\theta^1}) \leqslant 2n^{-1}$ are satisfied then we have*

$$\inf_{\widetilde{\theta}} \mathcal{R}_n[\widetilde{\theta} \,|\, \Theta] \geqslant \frac{1}{64} \mathfrak{d}^2_{ist}(\theta^0, \theta^1) \qquad (11.05)$$

where the infimum is taken over all possible estimators.

§11|01.02 **Proof** of Lemma §11|01.01. Given in the lecture. □

§11|01.03 **Remark** (Lower bound for a local ϕ-risk). Assuming the bounded Hellinger distance as for example in Lemma §11|01.01, Le Cam's general method (see Le Cam [1973]) and Pinsker's inequality allow us to derive a lower bound for a local ϕ-risk as in Reminder §10|00.17. However, in this special setting a lower bound can be obtained elementarily from Lemma §11|01.01, which in case (i) for any estimator $\widetilde{\theta}$ states

$$\mathcal{R}^\phi_n[\widetilde{\theta} \,|\, \Theta] := \sup\left\{ \mathbb{P}^n_\theta\left(|\phi\nu_{\mathbb{N}}(\widetilde{\theta} - \theta)|^2\right) : \theta \in \Theta \right\} \geqslant \frac{1}{16} |\phi\nu_{\mathbb{N}}(\theta^0 - \theta^1)|^2.$$

If we consider furthermore candidates $\theta^0 := \theta^*$ and $\theta^1 = -\theta^*$ for some $\theta^* \in \Theta$ such that $-\theta^* \in \Theta$, then trivially $|\phi\nu_{\mathbb{N}}(\theta^0 - \theta^1)|^2 = 4|\phi\nu_{\mathbb{N}}(\theta^*)|^2$ which in turn under the conditions of Lemma §11|01.01 (i) implies

$$\inf_{\widetilde{\theta}} \mathcal{R}^\phi_n[\widetilde{\theta} \,|\, \Theta] \geqslant \frac{1}{4} |\phi\nu_{\mathbb{N}}(\theta^*)|^2. \qquad (11.06)$$

Similarly, under the conditions of Lemma §11|01.01 (ii) we get

$$\inf_{\widetilde{\theta}} \mathcal{R}^\phi_n[\widetilde{\theta} \,|\, \Theta] \geqslant \frac{1}{16} |\phi\nu_{\mathbb{N}}(\theta^*)|^2. \qquad (11.07)$$

Often a minimax-optimal lower bound can be found by constructing a candidate $\theta^* = \mathrm{U}\theta^* \in \Theta$ that has the largest possible $|\phi\nu_{\mathbb{N}}(\theta^*)|^2$-value but $\mathbb{P}^n_{\theta^*}$ and $\mathbb{P}^n_{-\theta^*}$ are still statistically indistinguishable in the sense that $\mathrm{H}(\mathbb{P}^n_{\theta^*}, \mathbb{P}^n_{-\theta^*}) \leqslant 1$ or $\mathrm{H}(\mathbb{P}_{\theta^*}, \mathbb{P}_{-\theta^*}) \leqslant 2n^{-1}$. □

§11|01.04 **Reminder** (Maximal local ϕ-risk in diSM (§10|00.08 continued)). In Subsection §07|01 we consider an orthogonal projection estimator (OPE) in a Diagonal inverse sequence model (diSM) as in Model §01|04.08 (summarised in Model §10|00.08). Here the observable noisy version $\widehat{g}$ satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}^{\otimes \mathbb{N}}, \mathrm{P}^n_{\Theta \times \{\mathfrak{s}\} \times \Sigma} := (\mathrm{P}^n_{\theta|\mathfrak{s}|\sigma})_{\theta \in \Theta, \sigma \in \Sigma})$ where $\mathfrak{s} \in \ell_\infty$ is *known*, $\Theta \subseteq \ell_2$ and $\Sigma \subseteq \mathbb{R}^{\mathbb{N}}_{>0} \cap \ell_\infty$. Under Assumption §07|01.64 (which is implied by Assumption §10|00.16) in Corollary §07|01.74 an upper bound for the maximal local ϕ-risk of an OPE is shown. More precisly, assuming a multiplication operator $\mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathbb{M}}(\mathbb{J})$ (compare Notation §01|04.01), which fulfills a link condition $\mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathrm{t,d}}$ for $\mathrm{d} \in \mathbb{R}_{\geqslant 1}$ (see Assumption §04|03.04), the performance of the OPE $\widehat{\theta}^m = \mathfrak{s}^\dagger \widehat{g} \mathbb{1}^m \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ with dimension $m \in \mathbb{N}$ is measured by its maximal local ϕ-risk over the ellipsoid $\Theta = \ell_2^{\mathfrak{a},\mathrm{r}}$ with $\mathrm{r} \in \mathbb{R}_{>0}$, that is

$$\mathcal{R}^\phi_n[\widehat{\theta}^m \,|\, \ell_2^{\mathfrak{a},\mathrm{r}}, \{\mathrm{M}_{\mathfrak{s}}\}, \{\sigma\}] := \sup\left\{ \mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m_n^*} - \theta)|^2) : \theta \in \ell_2^{\mathfrak{a},\mathrm{r}} \right\} \quad \forall n, m \in \mathbb{N}.$$

For $n, m \in \mathbb{N}$ setting (as in (07.20))

$$\mathrm{R}_n^m(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi) := \|\mathfrak{a}_\bullet \mathbb{1}^{m|\perp}\|_\phi^2 + n^{-1}\|\mathfrak{t}_\bullet^\dagger \mathbb{1}^m\|_\phi^2, \quad m_n^\star := \arg\min\{\mathrm{R}_n^m(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi) : m \in \mathbb{N}\}$$
$$\text{and} \quad \mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi) := \mathrm{R}_n^{m_n^\star}(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi) = \min\{\mathrm{R}_n^m(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi) : m \in \mathbb{N}\} \quad (11.08)$$

the OPE $\widehat{\theta}_\bullet^{m_n^\star} = \mathfrak{s}_\bullet^\dagger \widehat{g}_\bullet \mathbb{1}^{m_n^\star} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ with optimally choosen dimension $m_n^\star = m_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi)$ as in (11.08) fulfills

$$\mathcal{R}_n^\phi[\widehat{\theta}_\bullet^{m_n^\star} \,|\, \ell_2^{\mathfrak{a}_\bullet,\mathrm{r}}, \{\mathrm{M}_{\mathfrak{s}}\}, \{\sigma_\bullet\}] \leqslant (\mathbb{v}_\sigma \mathrm{d}^2 \vee \mathrm{r}^2)\, \mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi) \quad \forall n \in \mathbb{N} \quad (11.09)$$

with $\|\sigma_\bullet^2\|_{\ell_\infty} =: \mathbb{v}_\sigma \in \mathbb{R}_{>0}$. In the proof of the next proposition we make use of Lemma §08|01.61 which under Assumption §07|01.64 (implied by Assumption §11|00.07) states that $(\mathfrak{a}\mathfrak{t})_{m_n^\star}^2 > n^{-1} \geqslant (\mathfrak{a}\mathfrak{t})_{m_n^\star+1}^2 = (\mathfrak{a}\mathfrak{t})_{(m_n^\star)}^2$ for all $n \in \mathbb{N}$ with $(\mathfrak{a}\mathfrak{t})_2^2 > n^{-1}$, i.e. $n \in \mathbb{N}_{>(\mathfrak{a}\mathfrak{t})_2^{-2}}$. □

§11|01.05 **diSM (§10|00.08 continued).** Consider $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet \sim \mathrm{P}_{\theta|\mathfrak{s}|\sigma}^n$ as in Model §10|00.08, where $\dot{\varepsilon}_\bullet$ satisfies **(SM:ub)** with $\mathrm{P}_{(0,1)} \in \mathscr{W}(\mathscr{B})$ and $\sigma_\bullet \in \mathbb{R}_{\geqslant 0}^{\mathbb{N}} \cap \ell_\infty$. In addition
(SM:lb) $\mathrm{P}_{(0,1)} \in \mathscr{W}(\mathscr{B})$ fulfills Assumption §11|00.07 with $\mathrm{C}_{\dot{\varepsilon}}, x_{\dot{\varepsilon}} \in \mathbb{R}_{>0}$ and $\sigma_\bullet^{-2} \in \ell_\infty$. □

§11|01.06 **Corollary** (diSM §11|01.05 continued). *For $\mathfrak{s}_\bullet \in \ell_\infty$, $\theta_\bullet \in \ell_2$, hence $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \subseteq \ell_2$, consider $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet \sim \mathrm{P}_{\theta|\mathfrak{s}|\sigma}^n$ as in Model §11|01.05, where $\dot{\varepsilon}_\bullet$ fulfills* **(SM:ub)** *and* **(SM:lb)** *with $\mathrm{C}_{\dot{\varepsilon}}, x_{\dot{\varepsilon}} \in \mathbb{R}_{>0}$ $\|\sigma_\bullet^{-2}\|_{\ell_\infty} =: \mathbb{v}_\sigma \in \mathbb{R}_{>0}$. For each $\theta_\bullet^\star \in \ell_2$ with $2n^{1/2}\mathbb{v}_\sigma^{1/2}\|\mathfrak{s}_\bullet\theta_\bullet^\star\|_{\ell_\infty} \leqslant x_{\dot{\varepsilon}}$ setting $\theta_\bullet^0 := \theta_\bullet^\star$ and $\theta_\bullet^1 := -\theta_\bullet^\star$ the distributions $\mathrm{P}_{\theta^\tau|\mathfrak{s}|\sigma}^n \in \mathscr{W}(\mathscr{B}^{\otimes\mathbb{N}})$, $\tau \in \{0,1\}$, satisfy $\mathrm{H}^2(\mathrm{P}_{\theta^0|\mathfrak{s}|\sigma}^n, \mathrm{P}_{\theta^1|\mathfrak{s}|\sigma}^n) \leqslant 4n\mathbb{v}_\sigma\mathrm{C}_{\dot{\varepsilon}}\|\mathfrak{s}_\bullet\theta_\bullet^\star\|_{\ell_2}^2$.*

§11|01.07 **Proof** of Corollary §11|01.06. Given in the lecture. □

§11|01.08 **Proposition** (diSM §11|01.05 continued). *For $\mathfrak{s}_\bullet \in \ell_\infty$, $\theta_\bullet \in \ell_2$, hence $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \subseteq \ell_2$, consider $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet \sim \mathrm{P}_{\theta|\mathfrak{s}|\sigma}^n$ as in Model §11|01.05, where $\dot{\varepsilon}_\bullet$ fulfills* **(SM:ub)** *and* **(SM:lb)** *with $\mathrm{C}_{\dot{\varepsilon}}, x_{\dot{\varepsilon}} \in \mathbb{R}_{>0}$ and $\|\sigma_\bullet^{-2}\|_{\ell_\infty} =: \mathbb{v}_\sigma \in \mathbb{R}_{>0}$. Let Assumption §10|00.16 and in addition*

$$|\mathfrak{s}_\bullet| \leqslant \mathrm{d}\mathfrak{t}_\bullet \quad \nu_{\mathbb{N}}\text{-a.e.} \quad \text{for } \mathrm{d} \in \mathbb{R}_{\geqslant 1} \quad (11.10)$$

be satisfied. Then we have

$$\inf_{\widetilde{\theta}_\bullet} \mathcal{R}_n^\phi[\widetilde{\theta}_\bullet \,|\, \ell_2^{\mathfrak{a}_\bullet,\mathrm{r}}, \{\mathrm{M}_{\mathfrak{s}}\}, \{\sigma_\bullet\}] \geqslant \mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi) \times \tfrac{1}{16}\big(4\mathrm{r}^2 \wedge \mathbb{v}_\sigma^{-1}\mathrm{d}^{-2}(\mathrm{C}_{\dot{\varepsilon}}^{-1} \wedge x_{\dot{\varepsilon}}^2)\big) \quad \forall n \in \mathbb{N}_{>(\mathfrak{a}\mathfrak{t})_2^{-2}} \quad (11.11)$$

where the infimum is taken over all possible estimators.

§11|01.09 **Proof** of Proposition §11|01.08. Given in the lecture. □

§11|01.10 **Comment.** By combining the lower bound in Proposition §11|01.08 and the upper bound in Corollary §07|01.74 for the maximal local ϕ-risk of an OPE in a diSM §11|01.05 we have shown that $\mathrm{R}_n^\star(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \phi)$ is a minimax-rate and the OPE with optimally chosen dimension parameter is minimax-optimal (up to a constant). □

§11|01.11 **GdiSM (§01|04.09 continued).** Consider a Gaussian diagonal inverse sequence model (GdiSM) as in §01|04.09 where $\mathfrak{s}_\bullet \in \ell_\infty$ is *known* in advance. Here the observable stochastic process $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\mathrm{B}}_\bullet \sim \mathrm{N}_{\theta|\mathfrak{s}}^n$ is a noisy version of $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \ell_2$ with $\theta_\bullet = \mathfrak{s}_\bullet^\dagger g_\bullet \in \Theta \subseteq \ell_2$ and $\dot{\mathrm{B}}_\bullet \sim \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}$. Consequently, $\widehat{g}_\bullet$ admits a $\mathrm{N}_{\theta|\mathfrak{s}}^n$-distribution belonging to the family $\mathrm{N}_{\Theta\times\{\mathfrak{s}_\bullet\}}^n := (\mathrm{N}_{\theta|\mathfrak{s}}^n)_{\theta_\bullet\in\Theta}$. Summarising the observations satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}^{\otimes\mathbb{N}}, \mathrm{N}_{\Theta\times\{\mathfrak{s}_\bullet\}}^n)$ where $\Theta \subseteq \ell_2$. Under Assumption §07|01.64 (which is implied by Assumption §10|00.16) in Corollary §07|01.72 an upper bound for the maximal local ϕ-risk of an OPE is shown. More precisely, assuming a multiplication operator $\mathrm{M}_{\mathfrak{s}} \in \mathbb{L}(\mathbb{J})$ (compare Notation §01|04.01), which fulfills a link condition $\mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ for $\mathrm{d} \in \mathbb{R}_{\geqslant 1}$ (see Assumption §04|03.04), the performance of the OPE $\widehat{\theta}^m_{\bullet} = \mathfrak{s}^{\dagger}_{\bullet}\widehat{g}_{\bullet}\mathbb{1}^m_{\bullet} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ with dimension $m \in \mathbb{N}$ is measured by its maximal local ϕ-risk over the ellipsoid $\Theta = \ell_2^{\mathfrak{a},\mathrm{r}}$ with $\mathrm{r} \in \mathbb{R}_{>0}$, that is

$$\mathcal{R}^{\phi}_{n}[\,\widehat{\theta}^m_{\bullet} \,|\, \ell_2^{\mathfrak{a}_{\bullet},\mathrm{r}}, \{\mathrm{M}_{\mathfrak{s}}\}\,] := \sup\left\{\mathrm{N}^n_{\theta|\mathfrak{s}}(|\phi\nu_{\mathbb{N}}(\widehat{\theta}^{m^\star_n}_{\bullet} - \theta_{\bullet})|^2) \colon \theta_{\bullet} \in \ell_2^{\mathfrak{a},\mathrm{r}}\right\} \quad \forall n, m \in \mathbb{N}.$$

The OPE $\widehat{\theta}^{m^\star_n}_{\bullet} = \mathfrak{s}^{\dagger}_{\bullet}\widehat{g}_{\bullet}\mathbb{1}^{m^\star_n}_{\bullet} \in \mathrm{dom}(\phi\nu_{\mathbb{N}})$ with optimally choosen dimension $m^\star_n = m^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet})$ as in (11.08) fulfills $\mathcal{R}^{\phi}_{n}[\,\widehat{\theta}^{m^\star_n}_{\bullet} \,|\, \ell_2^{\mathfrak{a}_{\bullet},\mathrm{r}}, \{\mathrm{M}_{\mathfrak{s}}\}\,] \leqslant (\mathrm{d}^2 \vee \mathrm{r}^2)\, \mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet})$ for all $n \in \mathbb{N}$. □

§11|01.12 **Corollary** (GdiSM §11|01.11 continued). *For $\mathfrak{s}_{\bullet} \in \ell_\infty$, $\theta_{\bullet} \in \ell_2$, hence $g_{\bullet} = \mathfrak{s}_{\bullet}\theta_{\bullet} \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^2}) \subseteq \ell_2$, consider $\widehat{g}_{\bullet} = g_{\bullet} + n^{-1/2}\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}^n_{\theta|\mathfrak{s}}$ as in Model §11|01.11, where $\dot{\mathrm{B}}_{\bullet} \sim \mathrm{N}^{\otimes\mathbb{N}}_{(0,1)}$. Let Assumption §10|00.16 and in addition (11.10) be satisfied. Then we have*

$$\inf_{\widetilde{\theta}} \mathcal{R}^{\phi}_{n}[\,\widehat{\theta}^m_{\bullet} \,|\, \ell_2^{\mathfrak{a}_{\bullet},\mathrm{r}}, \{\mathrm{M}_{\mathfrak{s}}\}\,] \geqslant \mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet}) \times \tfrac{1}{8}\big(2\mathrm{r}^2 \wedge \mathrm{d}^{-2}\big) \quad \forall n \in \mathbb{N}_{>(\mathfrak{a}\mathfrak{t})_2^{-2}} \tag{11.12}$$

where the infimum is taken over all possible estimators.

§11|01.13 **Proof** of Corollary §11|01.12. Given in the lecture. □

§11|01.14 **Comment**. By combining the lower bound in Corollary §11|01.12 and the upper bound in Corollary §07|01.72 for the maximal local ϕ-risk of an OPE in a GdiSM §11|01.11 we have shown that $\mathrm{R}^\star_n(\mathfrak{a}_{\bullet}, \mathfrak{t}_{\bullet}, \phi_{\bullet})$ is a minimax-rate and the OPE with optimally chosen dimension parameter is minimax-optimal (up to a constant). □

§11|01.15 **Remark**. Let $\mathbb{P}^{n\otimes k}_{\Theta\times\Xi} = (\mathbb{P}^{n\otimes k}_{\theta|\xi})_{\theta\in\Theta,\xi\in\Xi}$ be a family of product measures $\mathbb{P}^{n\otimes k}_{\theta|\xi} = \mathbb{P}^n_{\theta|\xi} \otimes \mathbb{P}^k_{\xi}$ depending on a function of interest $\theta \in \Theta$, a nuissance parameter $\xi \in \Xi$ and noise levels $n, k \in \mathbb{N}$. The Lemma §11|01.01 allows us to bound from below the maximal risk for each nuissance parameter $\xi \in \Xi$ and noise level $n \in \mathbb{N}$. To be more precise, given a noise level $n \in \mathbb{N}$ for $\tau \in \{0, 1\}$ consider $\theta^\tau \in \Theta$ with associated product probability measure $\mathbb{P}^{n\otimes k}_{\theta^\tau|\xi} = \mathbb{P}^n_{\theta^\tau|\xi} \otimes \mathbb{P}^k_{\xi}$, then we have $\rho(\mathbb{P}^{n\otimes k}_{\theta^0|\xi}, \mathbb{P}^{n\otimes k}_{\theta^1|\xi}) = \rho(\mathbb{P}^n_{\theta^0|\xi} \otimes \mathbb{P}^k_{\xi}, \mathbb{P}^n_{\theta^1|\xi} \otimes \mathbb{P}^k_{\xi}) = \rho(\mathbb{P}^n_{\theta^0|\xi}, \mathbb{P}^n_{\theta^1|\xi})$ due to the independence. Consequently, if $\mathrm{H}(\mathbb{P}^n_{\theta^0|\xi}, \mathbb{P}^n_{\theta^1|\xi}) \leqslant 1$, then for any estimator $\widetilde{\theta}$ we obtain

$$\mathcal{R}_{n,k}[\,\widetilde{\theta} \,|\, \Theta, \{\xi\}\,] := \sup\left\{\mathbb{P}^{n\otimes k}_{\theta|\xi}\big(\mathfrak{d}^2_{\mathrm{ist}}(\widetilde{\theta}, \theta)\big) \colon \theta \in \Theta\right\} \geqslant \frac{1}{16}\mathfrak{d}^2_{\mathrm{ist}}(\theta^0, \theta^1)$$

due to Lemma §11|01.01. It is worth noting that we obtain the same lower bound when disposing of the family $\mathbb{P}^n_{\Theta\times\{\xi\}} = (\mathbb{P}^n_{\theta|\xi})_{\theta\in\Theta}$ only, in other words assuming the nuissance parameter $\xi \in \Xi$ is known in advance. □

§11|01.16 **Corollary** (Lower bound based on two hypothesis). *Let $\mathbb{P}^{n\otimes k}_{\Theta\times\Xi} = (\mathbb{P}^{n\otimes k}_{\theta|\xi})_{\theta\in\Theta,\xi\in\Xi}$ be a family of product measures $\mathbb{P}^{n\otimes k}_{\theta|\xi} = \mathbb{P}^n_{\theta|\xi} \otimes \mathbb{P}^k_{\xi}$ depending on a function of interest $\theta \in \Theta$, a nuissance parameter $\xi \in \Xi$ and noise levels $n, k \in \mathbb{N}$. If for each $\tau \in \{0, 1\}$ there are $\theta^\tau \in \Theta$ and a nuissance parameter $\xi^\tau \in \Xi$ with associated product probability measure $\mathbb{P}^{n\otimes k}_{\theta^\tau|\xi^\tau} = \mathbb{P}^n_{\theta^\tau|\xi^\tau} \otimes \mathbb{P}^k_{\xi^\tau}$ such that $\mathbb{P}^n_{\theta^0|\xi^0} = \mathbb{P}^n_{\theta^1|\xi^1}$ and in addition $\mathrm{H}(\mathbb{P}^k_{\xi^0}, \mathbb{P}^k_{\xi^1}) \leqslant 1$ then for any estimator $\widetilde{\theta}$ we have*

$$\mathcal{R}_{n,k}[\,\widetilde{\theta} \,|\, \Theta, \Xi\,] := \sup\left\{\mathbb{P}^{n\otimes k}_{\theta|\xi}\big(\mathfrak{d}^2_{\mathrm{ist}}(\widetilde{\theta}, \theta)\big) \colon \theta \in \Theta, \xi \in \Xi\right\} \geqslant \frac{1}{16}\mathfrak{d}^2_{\mathrm{ist}}(\theta^0, \theta^1). \tag{11.13}$$

§11|01.17 **Proof** of Corollary §11|01.16. Given in the lecture. □

§11|01.18 **Remark**. The last assertion allows us often to derive a lower bound depending on the classes Θ and Ξ and the noise level k but not on the noise level n. Roughly speaking this means that we cover the influence of the estimation of the nuissance parameter. Typically we combine this lower bound with the lower bound obtained in Lemma §11|01.01 where the nuissance parameter is assumed to be known in advance. □

§11|01.19 **Reminder** (Maximal global $\mathfrak{v}$-risk in diSM with noisy operator (§10|00.10 continued)). For $\ell_2 = \mathbb{L}_2(\mathbb{N}, 2^{\mathbb{N}}, \nu_{\mathbb{N}})$ in Subsection §07|02 we consider a thresholded orthogonal projection estimator (tOPE) in a Diagonal inverse sequence model (diSM) with noisy operator as in Model §02|04.05 (summarised in Model §10|00.10). Here the observable noisy versions $(\widehat{g}, \widehat{\mathfrak{s}})$ satisfy a statistical product experiment

$$\left(\mathbb{R}^{\mathbb{N}^2}, \mathscr{B}^{\otimes \mathbb{N}^2}, \mathrm{P}^{n\otimes k}_{\Theta\times\mathcal{S}\times\Sigma\times\Xi} := (\mathrm{P}^{n\otimes k}_{\theta|\mathfrak{s}|\sigma|\xi} := \mathrm{P}^n_{\theta|\mathfrak{s}|\sigma} \otimes \mathrm{P}^k_{\mathfrak{s}|\xi})_{\theta\in\Theta,\mathfrak{s}\in\mathcal{S},\sigma\in\Sigma,\xi\in\Xi}\right)$$

where $\Sigma, \Xi \subseteq \mathbb{R}^{\mathbb{N}}_{>0} \cap \ell_\infty$. $\mathcal{S} \subseteq \mathbb{R}^{\mathbb{N}}_{\setminus 0} \cap \ell_\infty$ and $\Theta \subseteq \ell_2$. Under Assumption §07|02.32 (which is implied by Assumption §10|00.14) in Corollary §07|02.39 an upper bound for the maximal global $\mathfrak{v}$-risk of a tOPE is shown. More precisely, assuming a multiplication operator $\mathrm{M}_{\mathfrak{s}} \in \mathbb{L}^{\mathrm{M}}(\ell_2)$ (compare Notation §02|04.02), which fulfills a link condition $\mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ for $\mathrm{d} \in \mathbb{R}_{\geqslant 1}$ (see Assumption §04|03.04), the performance of the tOPE $\widehat{\theta}^m := \widehat{\mathfrak{s}}^{(k)|\dagger}\widehat{g}^m = \widehat{\mathfrak{s}}^{\dagger}\mathbb{1}^{\{\widehat{\mathfrak{s}}^2 \geqslant k^{-1}\}}\widehat{g}\mathbb{1}^m \in \ell_2(\mathfrak{v}^2)$ (see Definition §07|02.04) with dimension $m \in \mathbb{N}$ is measured by its maximal global $\mathfrak{v}$-risk over the ellipsoid $\Theta = \ell_2^{\mathfrak{a},\mathrm{r}}$ with $\mathrm{r} \in \mathbb{R}_{>0}$ and the link condition $\mathcal{S} = \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ with $\mathrm{d} \in \mathbb{R}_{\geqslant 1}$, that is

$$\mathcal{R}^{\mathfrak{v}}_{n,k}[\widehat{\theta}^m \,|\, \ell_2^{\mathfrak{a},\mathrm{r}}, \mathbb{M}_{\mathfrak{t},\mathrm{d}}, \{\sigma\}, \{\xi\}] := \sup\left\{\mathrm{P}^{n\otimes k}_{\theta|\mathfrak{s}|\sigma|\xi}(\|\widehat{\theta}^m - \theta\|^2_{\mathfrak{v}}) \colon \theta \in \ell_2^{\mathfrak{a},\mathrm{r}}, \mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}\right\} \quad \forall n, k, m \in \mathbb{N}.$$

For $n, m \in \mathbb{N}$ setting (as in (07.37))

$$\mathrm{R}^m_n(\mathfrak{a}, \mathfrak{t}, \mathfrak{v}) := [(\mathfrak{a}\mathfrak{v})^2_{(m)} \vee n^{-1}\|\mathfrak{t}^{\dagger}\mathbb{1}^m\|^2_{\mathfrak{v}}], \quad m^\star_n := \arg\min\left\{\mathrm{R}^m_n(\mathfrak{a}, \mathfrak{t}, \mathfrak{v}) : m \in \mathbb{N}\right\}$$

$$\text{and} \quad \mathrm{R}^\star_n(\mathfrak{a}, \mathfrak{t}, \mathfrak{v}) := \mathrm{R}^{m^\star_n}_n(\mathfrak{a}, \mathfrak{t}, \mathfrak{v}) = \min\left\{\mathrm{R}^m_n(\mathfrak{a}, \mathfrak{t}, \mathfrak{v}) : m \in \mathbb{N}\right\} \quad (11.14)$$

the OPE $\widehat{\theta}^{m^\star_n} = \widehat{\mathfrak{s}}^{(k)|\dagger}\widehat{g}^{m^\star_n} \in \ell_2(\mathfrak{v}^2)$ with optimally choosen dimension $m^\star_n = m^\star_n(\mathfrak{a}, \mathfrak{t}, \mathfrak{v})$ as in (11.22) fulfills

$$\mathcal{R}^{\mathfrak{v}}_{n,k}[\widehat{\theta}^{m^\star_n} \,|\, \ell_2^{\mathfrak{a},\mathrm{r}}, \mathbb{M}_{\mathfrak{t},\mathrm{d}}, \{\sigma\}, \{\xi\}] \leqslant \mathrm{R}^\star_n(\mathfrak{a}, \mathfrak{t}, \mathfrak{v}) \vee \|(\mathfrak{a}\mathfrak{v})^2(1 \vee k\mathfrak{t}^2)^{-1}\|_{\ell_\infty}$$
$$\times (\mathrm{r}^2 + 4\mathrm{K}_\xi^2\mathrm{K}_\sigma^2\mathrm{d}^2 + 8\mathrm{K}_\xi^4\mathrm{r}^2\mathrm{d}^2) \quad \forall n, k \in \mathbb{N} \quad (11.15)$$

with $\mathrm{K}_\sigma := \|\sigma\|_{\ell_\infty} \vee 1 \in \mathbb{R}_{\geqslant 1}$ and $\mathrm{K}_\xi := \|\xi\|_{\ell_\infty} \vee 1 \in \mathbb{R}_{\geqslant 1}$. □

§11|01.20 **diSM with noisy operator** (§10|00.10 continued). Consider $(\widehat{g} = g + n^{-1/2}\dot{\varepsilon}, \widehat{\mathfrak{s}} = \mathfrak{s} + k^{-1/2}\dot{\eta}) \sim \mathrm{P}^{n\otimes k}_{\theta|\mathfrak{s}|\sigma|\xi} = \mathrm{P}^n_{\theta|\mathfrak{s}|\sigma} \otimes \mathrm{P}^k_{\mathfrak{s}|\xi}$ as in Model §10|00.10, where $\dot{\eta} \sim \otimes_{j\in\mathbb{N}}\mathbb{P}^{\dot{\eta}_j}$ fulfills **(SMnO:ub)** in Model §10|00.10 with $\xi \in \Xi \subseteq \mathbb{R}^{\mathbb{N}}_{\geqslant 0} \cap \ell_\infty$ and hence $\mathbb{v}_j^{\dot{\eta}} := \mathbb{P}(\dot{\eta}_j^2) \leqslant \xi_j^2$ for each $j \in \mathbb{N}$. In addition

(SMnO:lb) there exists $\mathrm{P}_{(0,1)} \in \mathscr{W}(\mathscr{B})$ fulfilling Assumption §11|00.07 with $\mathrm{C}_{\dot{\eta}}, x_{\dot{\eta}} \in \mathbb{R}_{>0}$ such that $\mathbb{P}^{\dot{\eta}_j} = \mathrm{P}_{(0,\mathbb{v}_j^{\dot{\eta}})}$ for each $j \in \mathbb{N}$ and $(\mathbb{v}^{\dot{\eta}})^{-1} \in \ell_\infty$. □

§11|01.21 **Corollary** (diSM with noisy operator §11|01.20 continued). *Consider* $\widehat{\mathfrak{s}} = \mathfrak{s} + k^{-1/2}\dot{\eta} \sim \mathrm{P}^k_{\mathfrak{s}|\xi}$ *as in Model §11|01.20, where* $\dot{\eta}$ *fulfills* **(SMnO:ub)** *and* **(SMnO:lb)** *with* $\mathrm{C}_{\dot{\eta}}, x_{\dot{\eta}} \in \mathbb{R}_{>0}$ *and* $\|(\mathbb{v}^{\dot{\eta}})^{-1}\|_{\ell_\infty} =: \mathbb{v}_{\dot{\eta}} \in \mathbb{R}_{>0}$. *For any* $\mathfrak{s}^0, \mathfrak{s}^1 \in \ell_\infty$ *with* $k^{1/2}\mathbb{v}_{\dot{\eta}}^{1/2}\|\mathfrak{s}^0 - \mathfrak{s}^1\|_{\ell_\infty} \leqslant x_{\dot{\eta}}$ *we have* $\mathrm{H}^2(\mathrm{P}^k_{\mathfrak{s}^0|\xi}, \mathrm{P}^k_{\mathfrak{s}^1|\xi}) \leqslant k\mathbb{v}_{\dot{\eta}}\mathrm{C}_{\dot{\eta}}\|\mathfrak{s}^0 - \mathfrak{s}^1\|^2_{\ell_2}$.

§11|01.22 **Proof** of Corollary §11|01.21. Given in the lecture. □

§11|01.23 **Proposition** (diSM with noisy operator §11|01.20 continued). *Consider* $(\widehat{g} = g + n^{-1/2}\dot{\varepsilon}, \widehat{\mathfrak{s}} = \mathfrak{s} + k^{-1/2}\dot{\eta}) \sim \mathrm{P}^{n\otimes k}_{\theta|\mathfrak{s}|\sigma|\xi} = \mathrm{P}^n_{\theta|\mathfrak{s}|\sigma} \otimes \mathrm{P}^k_{\mathfrak{s}|\xi}$ *as in Model §10|00.10, where* $\dot{\eta}$ *fulfills* **(SMnO:ub)** *and* **(SMnO:lb)** *with* $\mathrm{C}_{\dot{\eta}}, x_{\dot{\eta}} \in \mathbb{R}_{>0}$ *and* $\|(\mathbb{v}^{\dot{\eta}})^{-1}\|_{\ell_\infty} =: \mathbb{v}_{\dot{\eta}} \in \mathbb{R}_{>0}$. *If Assumption §10|00.16 is satisfied then we have*

$$\inf_{\widetilde{\theta}} \mathcal{R}^{\mathfrak{v}}_{n,k}[\widetilde{\theta} \,|\, \ell_2^{\mathfrak{a},\mathrm{r}}, \mathbb{M}_{\mathfrak{t},\mathrm{d}}, \{\sigma\}, \{\xi\}] \geqslant \|(\mathfrak{a}\mathfrak{v})^2(1 \vee k\mathfrak{t}^2)^{-1}\|_{\ell_\infty}$$
$$\times \frac{\mathrm{r}^2}{16\mathrm{d}^2}(\mathbb{v}_{\dot{\eta}}^{-1}\mathrm{C}_{\dot{\eta}}^{-1} \wedge \mathbb{v}_{\dot{\eta}}^{-1}x_{\dot{\eta}}^2 \wedge 4(1 - \mathrm{d}^{-1})^2) \quad \forall n, k \in \mathbb{N} \quad (11.16)$$

where the infimum is taken over all possible estimators.

§11|01.24 **Proof** of Corollary §11|01.26. Given in the lecture. □

§11|01.25 **GdiSM with noisy operator** (§02|04.06 continued). Consider a Gaussian diagonal inverse sequence model (GdiSM) with noisy operator as in §02|04.06 where $\mathfrak{s}_\bullet \in \ell_\infty$ is *not known* anymore. Here the observable process $\widehat{\mathfrak{s}}_\bullet = \mathfrak{s}_\bullet + k^{-1/2}\dot{W}_\bullet \sim N^k_{\mathfrak{s}}$ and $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{B}_\bullet \sim N^n_{\theta|\mathfrak{s}}$ is a noisy version of $\mathfrak{s}_\bullet \in \mathcal{S} \subseteq \mathbb{R}^{\mathbb{N}}_{\setminus 0} \cap \ell_\infty$ and $g_\bullet = \mathfrak{s}_\bullet \theta_\bullet \in \mathrm{dom}(M_{\mathfrak{s}}) \subseteq \ell_2$ with $\theta_\bullet \in \Theta \subseteq \ell_2$, respectively, where $\dot{B}_\bullet \sim N^{\otimes \mathbb{N}}_{(0,1)}$ and $\dot{W}_\bullet \sim N^{\otimes \mathbb{N}}_{(0,1)}$ are *independent*. Consequently, $(\widehat{g}_\bullet, \widehat{\mathfrak{s}}_\bullet)$ admits a joint $N^{n\otimes k}_{\theta|\mathfrak{s}} = N^n_{\theta|\mathfrak{s}} \otimes N^k_{\mathfrak{s}}$ distribution belonging to the family $N^{n\otimes k}_{\Theta\times\mathcal{S}} := (N^n_{\theta|\mathfrak{s}} \otimes N^k_{\mathfrak{s}})_{\theta_\bullet\in\Theta,\mathfrak{s}_\bullet\in\mathcal{S}}$. Summarising the observations satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}^2}, \mathscr{B}^{\otimes\mathbb{N}^2}, N^{n\otimes k}_{\Theta\times\mathcal{S}})$ where $\Theta \subseteq \ell_2$ and $\mathcal{S} \subseteq \mathbb{R}^{\mathbb{N}}_{\setminus 0} \cap \ell_\infty$. Under Assumption §07|02.32 (which is implied by Assumption §10|00.14) in Corollary §07|02.37 an upper bound for the maximal global $\mathfrak{v}$-risk of a tOPE is shown. More precisely, the performance of the tOPE $\widehat{\theta}^m_\bullet = \widehat{\mathfrak{s}}^{(k)|\dagger}_\bullet \widehat{g}^m_\bullet \in \ell_2(\mathfrak{v}^2_\bullet)$ with dimension $m \in \mathbb{N}$ is measured by its maximal global $\mathfrak{v}$-risk over the ellipsoid $\Theta = \ell_2^{\mathfrak{a},\mathrm{r}}$ with $\mathrm{r} \in \mathbb{R}_{>0}$ and the link condition $\mathbb{M}_{\mathfrak{t},\mathrm{d}}$ with $\mathrm{d} \in \mathbb{R}_{\geqslant 1}$, that is

$$\mathcal{R}^{\mathfrak{v}}_{n,k}[\widehat{\theta}^m_\bullet \,|\, \ell_2^{\mathfrak{a},\mathrm{r}}, \mathbb{M}_{\mathfrak{t},\mathrm{d}}] := \sup\left\{ N^{n\otimes k}_{\theta|\mathfrak{s}}(\|\widehat{\theta}^m_\bullet - \theta_\bullet\|^2_{\mathfrak{v}}) \colon \theta_\bullet \in \ell_2^{\mathfrak{a},\mathrm{r}}, M_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}} \right\} \quad \forall n,k,m \in \mathbb{N}.$$

The tOPE $\widehat{\theta}^{m^\star_n}_\bullet = \widehat{\mathfrak{s}}^{(k)|\dagger}_\bullet \widehat{g}^{m^\star_n}_\bullet \in \ell_2(\mathfrak{v}^2_\bullet)$ with optimally choosen dimension $m^\star_n = m^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet)$ as in (11.22) fulfills $\mathcal{R}^{\mathfrak{v}}_{n,k}[\widehat{\theta}^{m^\star_n}_\bullet \,|\, \ell_2^{\mathfrak{a},\mathrm{r}}, \mathbb{M}_{\mathfrak{t},\mathrm{d}}] \leqslant \mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) \vee \|(\mathfrak{a}\mathfrak{v})^2_\bullet (1 \vee k\mathfrak{t}^2_\bullet)^{-1}\|_{\ell_\infty} \times (\mathrm{r}^2 + 4\mathrm{d}^2 + 12\mathrm{r}^2\mathrm{d}^2)$ for all $n, k \in \mathbb{N}$. □

§11|01.26 **Corollary** (GdiSM with noisy operator §11|01.25 continued). *Consider* $(\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{B}_\bullet, \widehat{\mathfrak{s}}_\bullet = \mathfrak{s}_\bullet + k^{-1/2}\dot{W}_\bullet) \sim N^{n\otimes k}_{\theta|\mathfrak{s}} = N^n_{\theta|\mathfrak{s}} \otimes N^k_{\mathfrak{s}}$ *as in Model §11|01.25, where* $\dot{B}_\bullet \sim N^{\otimes\mathbb{N}}_{(0,1)}$ *and* $\dot{W}_\bullet \sim N^{\otimes\mathbb{N}}_{(0,1)}$ *are independent. Let Assumption §10|00.14 be satisfied. Then we have*

$$\inf_{\widetilde{\theta}_\bullet} \mathcal{R}^{\mathfrak{v}}_{n,k}[\widetilde{\theta}_\bullet \,|\, \ell_2^{\mathfrak{a},\mathrm{r}}, \mathbb{M}_{\mathfrak{t},\mathrm{d}}] \geqslant \|(\mathfrak{a}\mathfrak{v})^2_\bullet (1 \vee k\mathfrak{t}^2_\bullet)^{-1}\|_{\ell_\infty} \times \frac{\mathrm{r}^2}{32\mathrm{d}^2}(1 \wedge 8(1-\mathrm{d}^{-1})^2) \quad \forall n,k \in \mathbb{N} \quad (11.17)$$

where the infimum is taken over all possible estimators.

§11|01.27 **Proof** of Corollary §11|01.26. Given in the lecture. □

§11|02 Lower bound based on m hypothesis

§11|02.01 **Notation**. For $m \in \mathbb{N}$ set $\mathcal{T}_m := \{-1, 1\}^m$ and for each $\tau := (\tau_j)_{j\in[\![m]\!]} \in \mathcal{T}_m$ and $j \in [\![m]\!]$ introduce $\tau^{(j)} \in \mathcal{T}_m$ given by $\tau^{(j)}_j := -\tau_j$ and $\tau^{(j)}_l := \tau_l$ for $l \in [\![m]\!] \setminus \{j\}$. □

§11|02.02 **Lemma** (Assouad's cube technique). *Given a noise level* $n \in \mathbb{N}$ *let* $\mathbb{P}^n_\Theta := (\mathbb{P}^n_\theta)_{\theta\in\Theta}$ *be a family of probability measures. Suppose there exist* $m \in \mathbb{N}$ *and distances* $\mathfrak{d}^{(j)}_{\mathrm{ist}}(\cdot,\cdot)$, $j \in [\![m]\!]$ *such that* $\mathfrak{d}^2_{\mathrm{ist}}(\cdot,\cdot) \geqslant \sum_{j\in[\![m]\!]} |\mathfrak{d}^{(j)}_{\mathrm{ist}}(\cdot,\cdot)|^2$. *We measure the accuracy of an estimator* $\widetilde{\theta}$ *by its maximal risk*

$$\mathcal{R}_n[\widetilde{\theta} \,|\, \Theta] := \sup\left\{ \mathbb{P}^n_\theta\left(\mathfrak{d}^2_{\mathrm{ist}}(\widetilde{\theta}, \theta)\right) \colon \theta \in \Theta \right\}.$$

(i) *If there exists* $\{\theta^\tau \colon \tau \in \mathcal{T}_m\} \subseteq \Theta$ *such that for all* $\tau \in \mathcal{T}_m$ *and* $j \in [\![m]\!]$ *we have Assumption §11|00.11 and* $\mathrm{H}(\mathbb{P}^n_{\theta^\tau}, \mathbb{P}^n_{\theta^{\tau^{(j)}}}) \leqslant 1$ *then we obtain*

$$\inf_{\widetilde{\theta}} \mathcal{R}_n[\widetilde{\theta} \,|\, \Theta] \geqslant 2^{-m} \sum_{\tau\in\mathcal{T}_m} \frac{1}{16} \sum_{j\in[\![m]\!]} |\mathfrak{d}^{(j)}_{\mathrm{ist}}(\theta^\tau, \theta^{\tau^{(j)}})|^2 \quad (11.18)$$

where the infimum is taken over all possible estimators.

(ii) *Let $n \in \mathbb{N}_{\geqslant 2}$ and for each $\theta \in \Theta$ let $\mathbb{P}^n_\theta = \otimes_{j\in[\![n]\!]}\mathbb{P}_\theta$ be a product probability measure with identically $\mathbb{P}_\theta$-distributed marginals. If there exists $\{\theta^\tau : \tau \in \mathcal{T}_m\} \subseteq \Theta$ $\tau \in \mathcal{T}_m$ and $j \in [\![m]\!]$ we have Assumption §11|00.11 and the marginals satisfy $\mathrm{H}(\mathbb{P}_{\theta^\tau}, \mathbb{P}_{\theta^{\tau^{(j)}}}) \leqslant 2n^{-1}$ then we have*

$$\inf_{\widetilde{\theta}} \mathcal{R}_n[\widetilde{\theta} \,|\, \Theta] \geqslant 2^{-m} \sum_{\tau\in\mathcal{T}_m} \frac{1}{64} \sum_{j\in[\![m]\!]} |\mathfrak{d}^{(j)}_{\text{ist}}(\theta^\tau, \theta^{\tau^{(j)}})|^2 \tag{11.19}$$

where the infimum is taken over all possible estimators.

§11|02.03 **Proof** of Lemma §11|02.02. Given in the lecture. □

§11|02.04 **Remark** (Lower bound for a global $\mathfrak{v}$-risk). For $a_\bullet, b_\bullet \in \ell_2(\mathfrak{v}^2_\bullet)$ consider $\mathfrak{d}_{\text{ist}}(a_\bullet, b_\bullet) = \|a_\bullet - b_\bullet\|_{\mathfrak{v}}$. Evidently, for each $j \in \mathbb{N}$ setting $\mathfrak{d}^{(j)}_{\text{ist}}(a_\bullet, b_\bullet) := |\mathfrak{v}_j(a_j - b_j)|$ we have

$$\mathfrak{d}^2_{\text{ist}}(a_\bullet, b_\bullet) = \|a_\bullet - b_\bullet\|^2_{\mathfrak{v}} \geqslant \sum_{j\in[\![m]\!]} \mathfrak{v}^2_j |a_j - b_j|^2 = \sum_{j\in[\![m]\!]} |\mathfrak{d}^{(j)}_{\text{ist}}(a_\bullet, b_\bullet)|^2 \quad \forall m \in \mathbb{N}.$$

Consequently, a lower bound for a global $\mathfrak{v}$-risk can be obtained elementarily from Lemma §11|02.02, which in case (i) for any estimator $\widetilde{\theta}_\bullet$ states

$$\mathcal{R}^{\mathfrak{v}}_n[\widetilde{\theta}_\bullet \,|\, \Theta] := \sup\left\{\mathbb{P}^n_\theta\left(\|\widetilde{\theta}_\bullet - \theta_\bullet\|^2_{\mathfrak{v}}\right) : \theta \in \Theta\right\} \geqslant 2^{-m} \sum_{\tau\in\mathcal{T}_m} \frac{1}{16} \sum_{j\in[\![m]\!]} \mathfrak{v}^2_j |\theta^\tau_j - \theta^{\tau^{(j)}}_j|^2$$

If we consider furthermore candidates $\{\theta^\tau_\bullet := (\tau_j \theta^\star_j \mathbb{1}^m_j)_{j\in\mathbb{N}} : \tau \in \mathcal{T}_m\} \subseteq \Theta \subseteq \ell_2(\mathfrak{v}^2_\bullet)$ for some $\theta^\star_\bullet \in \ell_2(\mathfrak{v}^2_\bullet)$, then it is easily seen that $\sum_{j\in[\![m]\!]} \mathfrak{v}^2_j |\theta^\tau_j - \theta^{\tau^{(j)}}_j|^2 = 4\sum_{j\in[\![m]\!]} \mathfrak{v}^2_j |\theta^\star_j|^2 = 4\|\theta^\star_\bullet \mathbb{1}^m_\bullet\|^2_{\mathfrak{v}}$ which in turn under the conditions of Lemma §11|02.02 (i) implies

$$\inf_{\widetilde{\theta}} \mathcal{R}^{\mathfrak{v}}_n[\widetilde{\theta}_\bullet \,|\, \Theta] \geqslant 2^{-m} \sum_{\tau\in\mathcal{T}_m} \frac{1}{4} \|\theta^\star_\bullet \mathbb{1}^m_\bullet\|^2_{\mathfrak{v}} = \frac{1}{4} \|\theta^\star_\bullet \mathbb{1}^m_\bullet\|^2_{\mathfrak{v}}. \tag{11.20}$$

Similarly, under the conditions of Lemma §11|02.02 (ii) we get

$$\inf_{\widetilde{\theta}} \mathcal{R}^{\mathfrak{v}}_n[\widetilde{\theta}_\bullet \,|\, \Theta] \geqslant 2^{-m} \sum_{\tau\in\mathcal{T}_m} \frac{1}{16} \|\theta^\star_\bullet \mathbb{1}^m_\bullet\|^2_{\mathfrak{v}} = \frac{1}{16} \|\theta^\star_\bullet \mathbb{1}^m_\bullet\|^2_{\mathfrak{v}}. \tag{11.21}$$

Often a minimax-optimal lower bound can be found by choosing the parameter m and constructing a candidate $\theta^\star_\bullet = \mathrm{U}\theta^\star$ that have the largest possible $\|\theta^\star_\bullet \mathbb{1}^m_\bullet\|^2_{\mathfrak{v}}$-value although that the associated $\mathbb{P}^n_{\theta^\tau}$, $\tau \in \mathcal{T}_m$ are still statistically indistinguishable in the sense that $\mathrm{H}(\mathbb{P}^n_{\theta^\tau}, \mathbb{P}^n_{\theta^{\tau^{(j)}}}) \leqslant 1$ or $\mathrm{H}(\mathbb{P}_{\theta^\tau}, \mathbb{P}_{\theta^{\tau^{(j)}}}) \leqslant 2n^{-1}$ for all $j \in [\![m]\!]$ and $\tau \in \mathcal{T}_m$. □

§11|02.05 **Reminder** (Maximal global $\mathfrak{v}$-risk in diSM (§10|00.08 continued)). In Subsection §07|01 we consider an orthogonal projection estimator (OPE) in a Diagonal inverse sequence model (diSM) as in Model §01|04.08 (summarised in Model §10|00.08). Here the observable noisy version $\widehat{g}_\bullet$ satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}^{\otimes\mathbb{N}}, \mathrm{P}^n_{\Theta\times\{\mathfrak{s}_\bullet\}\times\Sigma} := (\mathrm{P}^n_{\theta|\mathfrak{s}|\sigma})_{\theta_\bullet\in\Theta, \sigma_\bullet\in\Sigma})$ where $\mathfrak{s}_\bullet \in \ell_\infty$ is *known*, $\Theta \subseteq \ell_2$ and $\Sigma \subseteq \mathbb{R}^{\mathbb{N}}_{>0} \cap \ell_\infty$. Under Assumption §07|01.30 (which is implied by Assumption §10|00.14) in Corollary §07|01.40 an upper bound for the maximal global $\mathfrak{v}$-risk of an OPE is shown. More precisely, assuming a multiplication operator $\mathrm{M}_{\mathfrak{s}} \in \mathbb{L}(\mathbb{J})$ (compare Notation §01|04.01), which fulfills a link condition $\mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathrm{t,d}}$ for $\mathrm{d} \in \mathbb{R}_{\geqslant 1}$ (see Assumption §04|03.04), the performance of the OPE $\widehat{\theta}^m_\bullet = \mathfrak{s}^\dagger_\bullet \widehat{g}_\bullet \mathbb{1}^m_\bullet \in \mathrm{dom}(\phi_{\mathcal{U}_{\mathbb{N}}})$ with dimension $m \in \mathbb{N}$ is measured by its maximal global $\mathfrak{v}$-risk over the ellipsoid $\Theta = \ell^{\mathfrak{a},\mathrm{r}}_2$ with $\mathrm{r} \in \mathbb{R}_{>0}$, that is

$$\mathcal{R}^{\mathfrak{v}}_n[\widehat{\theta}^m_\bullet \,|\, \ell^{\mathfrak{a},\mathrm{r}}_2, \{\mathrm{M}_{\mathfrak{s}}\}, \{\sigma_\bullet\}] := \sup\left\{\mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}(\|\widehat{\theta}^{m^\star_n}_\bullet - \theta_\bullet\|^2_{\mathfrak{v}}) : \theta_\bullet \in \ell^{\mathfrak{a},\mathrm{r}}_2\right\} \quad \forall n, m \in \mathbb{N}.$$

For $n, m \in \mathbb{N}$ setting (as in (07.07))

$$\mathrm{R}^m_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) := (\mathfrak{a}\mathfrak{v})^2_{(m)} \vee n^{-1}\|\mathfrak{t}_\bullet^{-1}\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}}, \quad m^\star_n := \arg\min\{\mathrm{R}^m_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) : m \in \mathbb{N}\}$$
$$\text{and} \quad \mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) := \mathrm{R}^{m^\star_n}_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) = \min\{\mathrm{R}^m_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) : m \in \mathbb{N}\} \quad (11.22)$$

the OPE $\widehat{\theta}^{m^\star_n}_\bullet = \mathfrak{s}^\dagger_\bullet \widehat{g}_\bullet \mathbb{1}^{m^\star_n}_\bullet \in \ell_2(\mathfrak{v}^2_\bullet)$ with optimally choosen dimension $m^\star_n = m^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet)$ as in (11.22) fulfills

$$\mathcal{R}^{\mathfrak{v}}_n[\widehat{\theta}^{m^\star_n}_\bullet \mid \ell_2^{\mathfrak{a}_\bullet,\mathrm{r}}, \{\mathrm{M}_{\mathfrak{s}}\}, \{\sigma\}] \leqslant (\mathbb{v}_\sigma \mathrm{d}^2 + \mathrm{r}^2)\, \mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet) \quad \forall n \in \mathbb{N} \quad (11.23)$$

with $\|\sigma^2_\bullet\|_{\ell_\infty} =: \mathbb{v}_\sigma \in \mathbb{R}_{>0}$. □

§11|02.06 **Lemma**. *Under Assumption §10|00.14 for $m^\star_n := m^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet)$ and $\mathrm{R}^\star_n := \mathrm{R}^{m^\star_n}_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet)$ as in (11.22) distinguish* case i) : $\mathrm{R}^\star_n = n^{-1}\|\mathfrak{t}_\bullet^{-1}\mathbb{1}^{m^\star_n}_\bullet\|^2_{\mathfrak{v}} > (\mathfrak{a}\mathfrak{v})^2_{(m^\star_n)}$ *and* case ii) : $\mathrm{R}^\star_n = (\mathfrak{a}\mathfrak{v})^2_{(m^\star_n)} \geqslant n^{-1}\|\mathfrak{t}_\bullet^{-1}\mathbb{1}^{m^\star_n}_\bullet\|^2_{\mathfrak{v}}$. *Then for all $n \in \mathbb{N}_{>(\mathfrak{v}/\mathfrak{t})^2_1(\mathfrak{a}\mathfrak{v})^{-2}_{(1)}}$, i.e. $(\mathfrak{a}\mathfrak{v})^2_{(1)} > n^{-1}(\mathfrak{v}/\mathfrak{t})^2_1$, in* case i) *we have $(\mathfrak{a}\mathfrak{v})^2_{(m^\star_n-1)} > n^{-1}\|\mathfrak{t}_\bullet^{-1}\mathbb{1}^{m^\star_n}_\bullet\|^2_{\mathfrak{v}}$ and in* case ii) *setting*

$$m^\diamond_n := \min\{m \in \mathbb{N}_{>m^\star_n} : n^{-1}\|\mathfrak{t}_\bullet^{-1}\mathbb{1}^m_\bullet\|^2_{\mathfrak{v}} \geqslant (\mathfrak{a}\mathfrak{v})^2_{(m)}\} \quad (11.24)$$

we obtain $(\mathfrak{a}\mathfrak{v})^2_{(m^\star_n)} = (\mathfrak{a}\mathfrak{v})^2_{(m^\diamond_n-1)} \leqslant n^{-1}\|\mathfrak{t}_\bullet^{-1}\mathbb{1}^{m^\diamond_n}_\bullet\|^2_{\mathfrak{v}}$.

§11|02.07 **Proof** of Lemma §11|02.06. Given in the lecture. □

§11|02.08 **Corollary** (diSM §11|01.05 continued). *For $\mathfrak{s}_\bullet \in \ell_\infty$, $\theta_\bullet \in \ell_2$, hence $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \subseteq \ell_2$, consider $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet \sim \mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}$ as in Model §11|01.05, where $\dot{\varepsilon}_\bullet$ fulfills* **(SM:ub)** *and* **(SM:lb)** *with $\mathrm{C}_{\dot{\varepsilon}}, x_{\dot{\varepsilon}} \in \mathbb{R}_{>0}$ $\|\sigma^{-2}_\bullet\|_{\ell_\infty} =: \mathbb{v}_\sigma \in \mathbb{R}_{>0}$. For each $\theta^\star_\bullet \in \ell_2$ with $2n^{1/2}\mathbb{v}_\sigma^{1/2}\|\mathfrak{s}_\bullet\theta^\star_\bullet\mathbb{1}^m_\bullet\|_{\ell_\infty} \leqslant x_{\dot{\varepsilon}}$ and for each $\tau \in \mathfrak{T}_m = \{-1, 1\}^m$ as in Notation §11|02.01 setting $\theta^\tau_\bullet := (\tau_j\theta^\star_j\mathbb{1}^m_j)_{j\in\mathbb{N}}$ the distribution $\mathrm{P}^n_{\theta^\tau|\mathfrak{s}|\sigma} \in \mathscr{W}(\mathscr{B}^{\otimes\mathbb{N}})$ satisfies $\mathrm{H}^2(\mathrm{P}^n_{\theta^\tau|\mathfrak{s}|\sigma}, \mathrm{P}^n_{\theta^{\tau^{(j)}}|\mathfrak{s}|\sigma}) \leqslant 4n\mathbb{v}_\sigma\mathrm{C}_{\dot{\varepsilon}}\|\mathfrak{s}_\bullet\theta^\star_\bullet\mathbb{1}^m_\bullet\|^2_{\ell_\infty}$ for all $j \in [\![m]\!]$.*

§11|02.09 **Proof** of Corollary §11|02.08. Given in the lecture. □

§11|02.10 **Proposition** (diSM §11|01.05 continued). *For $\mathfrak{s}_\bullet \in \ell_\infty$, $\theta_\bullet \in \ell_2$, hence $g_\bullet = \mathfrak{s}_\bullet\theta_\bullet \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \subseteq \ell_2$, consider $\widehat{g}_\bullet = g_\bullet + n^{-1/2}\dot{\varepsilon}_\bullet \sim \mathrm{P}^n_{\theta|\mathfrak{s}|\sigma}$ as in Model §11|01.05, where $\dot{\varepsilon}_\bullet$ fulfills* **(SM:ub)** *and* **(SM:lb)** *with $\mathrm{C}_{\dot{\varepsilon}}, x_{\dot{\varepsilon}} \in \mathbb{R}_{>0}$ and $\|\sigma^{-2}_\bullet\|_{\ell_\infty} =: \mathbb{v}_\sigma \in \mathbb{R}_{>0}$. Let Assumption §10|00.14 and in addition (11.10) for $\mathrm{d} \in \mathbb{R}_{\geqslant 1}$ be satisfied. Then we have*

$$\inf_{\widetilde{\theta}_\bullet} \mathcal{R}^{\mathfrak{v}}_n[\widetilde{\theta}_\bullet \mid \ell_2^{\mathfrak{a}_\bullet,\mathrm{r}}, \{\mathrm{M}_{\mathfrak{s}}\}, \{\sigma\}] \geqslant \mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet)$$
$$\times \tfrac{1}{16}\big(4\mathrm{C}_{(\mathfrak{a}\mathfrak{v})}\mathrm{r}^2 \wedge \mathbb{v}_\sigma^{-1}\mathrm{d}^{-2}(\mathrm{C}_{\dot{\varepsilon}}^{-1} \wedge x^2_{\dot{\varepsilon}})\big) \quad \forall n \in \mathbb{N}_{>(\mathfrak{v}/\mathfrak{t})^2_1(\mathfrak{a}\mathfrak{v})^{-2}_{(1)}} \quad (11.25)$$

where the infimum is taken over all possible estimators.

§11|02.11 **Proof** of Proposition §11|02.10. Given in the lecture. □

§11|02.12 **Comment**. By combining the lower bound in Proposition §11|02.10 and the upper bound in Corollary §07|01.40 for the maximal global $\mathfrak{v}$-risk of an OPE in a diSM §11|01.05 we have shown that $\mathrm{R}^\star_n(\mathfrak{a}_\bullet, \mathfrak{t}_\bullet, \mathfrak{v}_\bullet)$ is a minimax-rate and the OPE with optimally chosen dimension parameter is minimax-optimal (up to a constant). □

§11|02.13 **GdiSM** **(§11|01.11 continued)**. Recall that the observations satisfy a statistical product experiment $(\mathbb{R}^{\mathbb{N}}, \mathscr{B}^{\otimes\mathbb{N}}, \mathrm{N}^n_{\Theta\times\{\mathfrak{s}_\bullet\}})$ where $\Theta \subseteq \ell_2$. Under Assumption §07|01.30 (which is implied by Assumption §10|00.14) in Corollary §07|01.38 an upper bound for the maximal global $\mathfrak{v}$-risk of an OPE is shown. More precisely, assuming a multiplication operator $\mathrm{M}_{\mathfrak{s}} \in \mathbb{L}(\mathbb{J})$ (compare Notation §01|04.01), which fulfills a link condition $\mathrm{M}_{\mathfrak{s}} \in \mathbb{M}_{\mathfrak{t},\mathrm{d}}$ for $\mathrm{d} \in \mathbb{R}_{\geqslant 1}$ (see Assumption §04|03.04), the performance of the OPE $\widehat{\theta}^{m} = \mathfrak{s}^{\dagger}\widehat{g}\mathbb{1}^{m} \in \ell_2(\mathfrak{v}^2)$ with dimension $m \in \mathbb{N}$ is measured by its maximal global $\mathfrak{v}$-risk over the ellipsoid $\Theta = \ell_2^{\mathfrak{a},\mathrm{r}}$ with $\mathrm{r} \in \mathbb{R}_{>0}$, that is

$$\mathcal{R}_n^{\mathfrak{v}}[\widehat{\theta}^{m} \mid \ell_2^{\mathfrak{a},\mathrm{r}}, \{\mathrm{M}_{\mathfrak{s}}\}] := \sup\left\{\mathrm{N}_{\theta|\mathfrak{s}}^{n}(\|\widehat{\theta}^{m_n^\star} - \theta\|_{\mathfrak{v}}^2) : \theta \in \ell_2^{\mathfrak{a},\mathrm{r}}\right\} \quad \forall n, m \in \mathbb{N}.$$

The OPE $\widehat{\theta}^{m_n^\star} = \mathfrak{s}^{\dagger}\widehat{g}\mathbb{1}^{m_n^\star} \in \ell_2(\mathfrak{v}^2)$ with optimally choosen dimension $m_n^\star = m_n^\star(\mathfrak{a}, \mathfrak{t}, \mathfrak{v})$ as in (11.22) fulfills $\mathcal{R}_n^{\mathfrak{v}}[\widehat{\theta}^{m_n^\star} \mid \ell_2^{\mathfrak{a},\mathrm{r}}, \{\mathrm{M}_{\mathfrak{s}}\}] \leqslant (\mathrm{d}^2 + \mathrm{r}^2)\,\mathrm{R}_n^\star(\mathfrak{a}, \mathfrak{t}, \mathfrak{v})$ for all $n \in \mathbb{N}$. □

§11|02.14 **Corollary** (GdiSM §11|02.13 continued). *For $\mathfrak{s} \in \ell_\infty$, $\theta \in \ell_2$, hence $g = \mathfrak{s}\theta \in \mathrm{dom}(\mathrm{M}_{\mathfrak{s}^\dagger}) \subseteq \ell_2$, consider $\widehat{g} = g + n^{-1/2}\dot{\mathrm{B}} \sim \mathrm{N}_{\theta|\mathfrak{s}}^{n}$ as in Model §11|01.11, where $\dot{\mathrm{B}} \sim \mathrm{N}_{(0,1)}^{\otimes\mathbb{N}}$. Let Assumption §10|00.14 and in addition (??) be satisfied. Then we have*

$$\inf_{\widetilde{\theta}} \mathcal{R}_n^{\mathfrak{v}}[\widehat{\theta}^{m} \mid \ell_2^{\mathfrak{a},\mathrm{r}}, \{\mathrm{M}_{\mathfrak{s}}\}] \geqslant \mathrm{R}_n^\star(\mathfrak{a}, \mathfrak{t}, \mathfrak{v}) \times \tfrac{1}{8}\left(2\mathrm{r}^2 \wedge \mathrm{d}^{-2}\right) \quad \forall n \in \mathbb{N}_{>(\mathfrak{v}/\mathfrak{t})_1^2(\mathfrak{a}\mathfrak{v})_{(1)}^{-2}} \tag{11.26}$$

where the infimum is taken over all possible estimators.

§11|02.15 **Proof** of Corollary §11|02.14. Given in the lecture. □

§11|02.16 **Comment**. By combining the lower bound in Corollary §11|02.14 and the upper bound in Corollary §07|01.38 for the maximal global $\mathfrak{v}$-risk of an OPE in a GdiSM §11|02.13 we have shown that $\mathrm{R}_n^\star(\mathfrak{a}, \mathfrak{t}, \mathfrak{v})$ is a minimax-rate and the OPE with optimally chosen dimension parameter is minimax-optimal (up to a constant). □

§11|02.17 **Remark**. Let $\mathbb{P}_{\Theta\times\Xi}^{n\otimes k} = (\mathbb{P}_{\theta|\xi}^{n\otimes k})_{\theta\in\Theta,\xi\in\Xi}$ be a family of product measures $\mathbb{P}_{\theta|\xi}^{n\otimes k} = \mathbb{P}_{\theta|\xi}^{n}\otimes\mathbb{P}_{\xi}^{k}$ depending on a function of interest $\theta \in \Theta$, a nuissance parameter $\xi \in \Xi$ and noise levels $n, k \in \mathbb{N}$. Supposing there exist $m \in \mathbb{N}$ and distances $\mathfrak{d}_{\mathrm{ist}}^{(j)}(\cdot,\cdot)$, $j \in [\![m]\!]$ such that $\mathfrak{d}_{\mathrm{ist}}^{2}(\cdot,\cdot) \geqslant \sum_{j\in[\![m]\!]} |\mathfrak{d}_{\mathrm{ist}}^{(j)}(\cdot,\cdot)|^2$ the Lemma §11|02.02 allows us to bound from below the maximal risk for each nuissance parameter $\xi \in \Xi$ and noise level $n \in \mathbb{N}$. To be more precise, given noise levels $n, k \in \mathbb{N}$ for each $\tau \in \mathcal{T}_m$ consider $\theta^\tau \in \Theta$ with associated product probability measure $\mathbb{P}_{\theta^\tau|\xi}^{n\otimes k} = \mathbb{P}_{\theta^\tau|\xi}^{n} \otimes \mathbb{P}_{\xi}^{k}$, then for all $j \in [\![m]\!]$ we have $\rho(\mathbb{P}_{\theta^\tau|\xi}^{n\otimes k}, \mathbb{P}_{\theta^{\tau^{(j)}}|\xi}^{n\otimes k}) = \rho(\mathbb{P}_{\theta^\tau|\xi}^{n}\otimes\mathbb{P}_{\xi}^{k}, \mathbb{P}_{\theta^{\tau^{(j)}}|\xi}^{n}\otimes\mathbb{P}_{\xi}^{k}) = \rho(\mathbb{P}_{\theta^\tau|\xi}^{n}, \mathbb{P}_{\theta^{\tau^{(j)}}|\xi}^{n})$ due to the independence. Consequently, if $\mathrm{H}(\mathbb{P}_{\tau}^{n}, \mathbb{P}_{\tau^{(j)}}^{n}) \leqslant 1$ for all $\tau \in \mathcal{T}_m$ and $j \in [\![m]\!]$, then for any estimator $\widetilde{\theta}$ we obtain

$$\mathcal{R}_{n,k}[\widetilde{\theta} \mid \Theta, \{\xi\}] := \sup\left\{\mathbb{P}_{\theta|\xi}^{n\otimes k}\left(\mathfrak{d}_{\mathrm{ist}}^{2}(\widetilde{\theta}, \theta)\right) : \theta \in \Theta\right\} \geqslant 2^{-m} \sum_{\tau\in\mathcal{T}_m} \frac{1}{16} \sum_{j\in[\![m]\!]} |\mathfrak{d}_{\mathrm{ist}}^{(j)}(\theta^\tau, \theta^{\tau^{(j)}})|^2$$

due to Lemma §11|02.02. It is worth noting that we obtain the same lower bound when disposing of the family $\mathbb{P}_{\Theta\times\{\xi\}}^{n} = (\mathbb{P}_{\theta|\xi}^{n})_{\theta\in\Theta}$ only, in other words assuming the nuissance parameter $\xi \in \Xi$ is known in advance. □

§11|02.18 **Corollary** (Lower bound based on m hypothesis). *Let $\mathbb{P}_{\Theta\times\Xi}^{n\otimes k} = (\mathbb{P}_{\theta|\xi}^{n\otimes k})_{\theta\in\Theta,\xi\in\Xi}$ be a family of product measures $\mathbb{P}_{\theta|\xi}^{n,k} = \mathbb{P}_{\theta|\xi}^{n} \otimes \mathbb{P}_{\xi}^{k}$ depending on a function of interest $\theta \in \Theta$, a nuissance parameter $\xi \in \Xi$ and noise levels $n, k \in \mathbb{N}$. Suppose there exist distances $\mathfrak{d}_{\mathrm{ist}}^{(j)}(\cdot,\cdot)$, $j \in [\![m]\!]$ such that $\mathfrak{d}_{\mathrm{ist}}^{2}(\cdot,\cdot) \geqslant \sum_{j\in[\![m]\!]} |\mathfrak{d}_{\mathrm{ist}}^{(j)}(\cdot,\cdot)|^2$. If there exists $\left\{(\theta^\tau, \xi^\tau) : \tau \in \mathcal{T}_m\right\} \subseteq \Theta \times \Xi$ such that for all $\tau \in \mathcal{T}_m$ and $j \in [\![m]\!]$ Assumption §11|00.11, (C1) $\mathbb{P}_{\theta^\tau|\xi^\tau}^{n} = \mathbb{P}_{\theta^{\tau^{(j)}}|\xi^{\tau^{(j)}}}^{n}$ and (C2) $\mathrm{H}(\mathbb{P}_{\xi^\tau}^{k}, \mathbb{P}_{\xi^{\tau^{(j)}}}^{k}) \leqslant 1$ are fulfilled, then for any estimator $\widetilde{\theta}$ we have*

$$\mathcal{R}_{n,k}[\widetilde{\theta} \mid \Theta, \Xi] := \sup\left\{\mathbb{P}_{\theta|\xi}^{n\otimes k}\left(\mathfrak{d}_{\mathrm{ist}}^{2}(\widetilde{\theta}, \theta)\right) : \theta \in \Theta, \xi \in \Xi\right\} \geqslant 2^{-m} \sum_{\tau\in\mathcal{T}_m} \frac{1}{16} \sum_{j\in[\![m]\!]} |\mathfrak{d}_{\mathrm{ist}}^{(j)}(\theta^\tau, \theta^{\tau^{(j)}})|^2.$$

§11|02.19 **Proof** of Corollary §11|02.18. Given in the lecture. □

§11|02.20 **Remark**. The last assertion allows us often to derive a lower bound depending on the classes Θ and Ξ and the noise level k but not on the noise level n. Roughly speaking this means that we cover the influence of the estimation of the nuissance parameter. Typically we combine this lower bound with the lower bound obtained in Lemma §11|02.02 where the nuissance parameter is assumed to be known in advance. □

Bibliography

C. R. Baker. Joint measures and cross-covariance operators. *Transactions of the american mathematical society*, 186:273–289, 1973.

S. Efromovich and V. Koltchinskii. On inverse problems with unknown operators. *IEEE Transactions on Information Theory*, 47(7):2876–2894, 2001.

H. W. Engl, M. Hanke, and A. Neubauer. *Regularization of inverse problems.* Kluwer Academic, Dordrecht, 2000.

J. Hadamard. *Le Problème de Cauchy et les Équations aux Dérivées Partielles Linéaires Hyperboliques*. Paris, Hermann, 1932.

E. Heinz. Beiträge zur Störungstheorie der Spektralzerlegung. *Mathematische Annalen*, 123: 415–438, 1951.

R. Kress. *Linear integral equations*, volume 82 of *Applied Mathematical Sciences.* Springer, New York, NY, 2 edition, 1989.

L. Le Cam. Convergence of Estimates Under Dimensionality Restrictions. *Annals of Statistics*, 1(1):38–53, 1973.

A. B. Tsybakov. *Introduction to nonparametric estimation.* Springer Series in Statistics. Springer, New York, 2009.

D. Werner. *Funktionalanalysis*. Springer-Lehrbuch, 2011.

H. Witting. Mathematische Statistik I: Parametrische Verfahren bei festem Stichprobenumfang. Stuttgart: B. G. Teubner., 1985.

Index