INAUGURAL-DISSERTATION

zur
Erlangung der Doktorwürde
der
Naturwissenschaftlich-Mathematischen Gesamtfakultät
der Ruprecht-Karls-Universität
Heidelberg

vorgelegt von

Xavier Loizeau

aus Nantes, France

Tag der mündlichen Prüfung:

Oracle and minimax optimality of Bayesian and frequentist methods for linear statistical ill-posed inverse problems under L^2-loss

Betreuer: Jan Johannes
Claudia Schilings

Acknowledgments

I am grateful to the administration of the *Combined Faculty of Natural Sciences and Mathematics* of the Ruprecht-Karls-Universität Heidelberg.

Zusammenfassung

Abstract in german

Abstract

Abstract in English

Contents

Chapter 1

Introduction

1.1 Statistical ill-posed inverse problems

1.1.1 Statistical model

Consider a measurable observation space $(\mathbb{Y}, \mathcal{Y})$, a parameter space Ξ and a family of probability distribution on $(\mathbb{Y}, \mathcal{Y})$ indexed by Ξ, which we denote $(\mathbb{P}_{Y|f})_{f \in \Xi}$. That is to say, for any f in Ξ, $\mathbb{P}_{Y|f}$ is a probability measure.

Throughout this thesis, we are interested in the estimation of the parameter f, under the two paradigmata of frequentist and Bayesian statistics, specified hereafter, and the quantification of the quality of such estimations.
From a frequentist point of view, one specifies a measurable application from $\mathbb{Y}$ to Ξ, called estimator. Ideally, an estimator solely depends of the observations and needs no knowledge whatsoever about the parameter of interest in order to be implemented properly. Once this application specified the duty of the statistician is to study properties such as consistency, rate of convergence, and asymptotic distribution of the estimator. These properties are presented in more details in SECTION 1.2.

In the Bayesian paradigm, one defines a σ-algebra $\mathcal{B}$ on Ξ and a probability distribution $\mathbb{P}_{\boldsymbol{f}}$ on $(\Xi, \mathcal{B})$ called prior distribution which represents the prior knowledge about the parameter, for example gathered by experts or prior experiments. One is then interested in the posterior distribution, that is, the distribution of the parameter of interest given the observations. From a purely Bayesian point of view, being able to define a prior distribution on Ξ and compute (or approach) the posterior distribution is all that is needed as one does not assume existence of a true underlying parameter. However, from a frequentist Bayesian (also called pragmatic Bayesian) approach, described in details in SECTION 1.3, one assumes existence of a true parameter and wonders if the posterior distribution contracts around this true parameter.

We are particularly interested in a specific class of models where Ξ is a subset of a function (or sequence) space. More specifically, let $\mathbb{T}$ and $\mathbb{D}$ be a subsets of $\mathbb{R}$ ($\mathbb{T}$ might be referred to as time domain and $\mathbb{D}$ as intensity domain). Denote $\|\cdot\|_{\Xi}$ a norm on the space of functions $\mathbb{T} \to \mathbb{D}$. Then Ξ is the space of functions $\{f : \mathbb{T} \to \mathbb{D}, \|f\|_{\Xi} < \infty\}$.

We will assume moreover that there exist a measure $\mathbb{P}^\circ$ on $(\mathbb{Y}, \mathcal{Y})$ dominating the family $(\mathbb{P}_{Y|f})_{f\in\Xi}$ and we denote $L : (\Xi \times \mathbb{Y}, \mathcal{B} \otimes \mathcal{Y}) \to \big(\overline{\mathbb{R}}_+, \mathcal{B}(\mathbb{R})\big)$ the likelihood with respect to $\mathbb{P}^\circ$:

$$\frac{\mathbb{P}_{Y|f}}{\mathbb{P}^\circ}(f, y) = L(f, y).$$

1.1.2 (Compact operator)

We will give particular interest to inverse problems, a family of models where one wants to infer on a parameter f but the data we observe is generated through the distribution with parameter $T(f)$ where T is an operator from Ξ to itself.
Hence we have:

$$\begin{aligned} T : \Xi &\to \Xi \qquad ; \\ f &\to T(f) \\ Y &\sim \mathbb{P}_{Y|T(f)}. \end{aligned}$$

These models gathered interest for a long time because many of them have the particularity to be ill-posed in the sense of Hadamard (1902). That is to say, if we build an estimator $\widehat{T(f)}$ of $T(f)$ from the data Y and try to apply T^{-1} to this estimator in order to estimate f, one of the following problems might arise:

- non existence (the equation $T(x) = \widehat{T(f)}$ does not have a solution);
- non unicity (the equation $T(x) = \widehat{T(f)}$ has multiple solutions);
- non stability (the solutions to the equations $T(x) = \widehat{T(f)}$ and $T(x) = \widehat{T(f)} + \varepsilon$ are arbitrary far for ε arbitrary small with respect to $\|\cdot\|$).

Though Hadamard thought that inverse problems do not arise in practical situations and that problems of our realm only are of the well-posed kind; evolution of science proved him wrong and ill-posed problems now have many applications and the specific challenges they represent gather a lot of interest as we will show through two examples in SECTION 1.4 and SECTION 1.5.

More-Penrose inverse.
Self adjoint.
Orthogonal basis of eigen functions -> notation $(e_j)_{j\in\mathbb{F}}, (\lambda_j)_{j\in\mathbb{F}}$.
$\mathcal{F}$ the application $\Xi \to \Theta, f \mapsto \theta = (\langle f|e_j\rangle)_{j\in\mathbb{F}}$. $\mathcal{F}^{-1}$ the application $\Theta \to \Xi, \theta \mapsto f = (\int_{j\in\mathbb{F}} \theta_j e_j(x))_{x\in\mathbb{T}}$. Plancherel theorem. $\mathbb{P}_{Y|\theta}$, $\mathbb{P}_{\boldsymbol{\theta}}$, $L(\theta, y)$
Compactness, $\lambda_j \to 0$ -> inverse unbounded (condition 3 not check). $\mathcal{F}(T(f)) = \lambda\theta$

1.1.3 Ill-posed inverse problem with known operator

e_j and λ_j known

1.1.4 Ill-posed inverse problem with partially known operator

e_j known but only an estimator $\widehat{\lambda}_j$ of λ_j is available.

1.1.5 Popular regularisation methods

- spectral cut-off
- Tikhonov
- entropy minimisation
- ...?

1.2 Frequentist approach

1.2.1 Estimation

- (M/Z-estimation);
- projection; Introduce notation for subspaces of the shape $\mathbb{B}_{k,l}$ as well as operators $\Pi_{k,l}$
- kernel smoother...

1.2.2 Decision theory

As we have seen previously, for a given model, one could chose among a variety of estimators. This choice is in general not obvious and decision theory can be used to help in this process.

In order to apply decision theory in our context one needs to define some objects.

The loss function $l : (\{\mathbb{Y} \to \Theta\} \times \mathbb{Y} \times \Theta) \to \mathbb{R}_+$: this function represents the error made by using a certain estimator $\widehat{\theta}$ while estimating a true parameter θ° when the data at hand is Y.

A natural choice would be to consider a distance on Θ, say $d : \Theta \times \Theta \to \mathbb{R}_+$ and to define

$$\begin{aligned} l :& \{\mathbb{Y} \to \Theta\} \times \mathbb{Y} \times \Theta \to \mathbb{R}_+ \\ & (\widehat{\theta}, Y, \theta^\circ) \qquad\qquad \mapsto d(\widehat{\theta}(Y), \theta^\circ). \end{aligned}$$

As an example, in the functional parameter space case we describe in SECTION 1.1.1 one could consider the following objects.

Definition 1 Product $\cdot$ on Θ
We define the bi-linear operator

$$\begin{aligned} \cdot : \quad \Theta^2 &\to \Theta \\ (\theta, \theta') &\mapsto \theta \cdot \theta' := \left(j \mapsto \theta_j \theta'_j\right)_{j \in \mathbb{F}}. \end{aligned}$$

With this definition at hand we can define the inner product of Θ

Definition 2 Inner product $\langle \cdot | \cdot \rangle_{l^2}$ on Θ
We define the operator

$$\begin{aligned} \langle \cdot | \cdot \rangle_{l^2} : \quad \Theta^2 &\to \overline{\mathbb{C}} \\ (\theta, \theta') &\mapsto \int_{j \in \mathbb{F}} (\theta \cdot \overline{\theta'})_j dj. \end{aligned}$$

This leads to the natural l^2-norm

Definition 3 l^2-norm $\| \cdot \|_{l^2}$ on Θ
We define the norm

$$\begin{aligned} \| \cdot \|_{l^2} : \quad \Theta &\to \overline{\mathbb{R}_+} \\ \theta &\mapsto \sqrt{\langle \theta | \theta \rangle_{l^2}} = \left(\int_{j \in \mathbb{F}} |\theta_j|^2 dj \right)^{1/2}. \end{aligned}$$

It is common to consider the larger family of norms $\| \cdot \|_{l^p}$ for any number p in $[1, \infty]$ which however do not define an inner product space for $p \neq 2$ (they do not verify the parallelogram inequality):

Definition 4 l^p-norm $\| \cdot \|_{l^p}$ on Θ
We define the norm

$$\begin{aligned} \| \cdot \|_{l^p} : \Theta &\to \overline{\mathbb{R}_+}. \\ \theta &\mapsto \left(\int_{j \in \mathbb{F}} |\theta_j|^p \, dj \right)^{1/p} \end{aligned}$$

A last kind of norm which is of interest are the weighted norms. Using a weighted norm as loss function allows to give more interest to some specific features of the functions (high or low frequencies for example).

Definition 5 l^p_{u}-norm $\| \cdot \|_{l^p_{\mathrm{u}}}$ on Θ

Consider an element $[\mathfrak{u}]$ of Θ. We define the norm

$$\begin{aligned} \|\cdot\|_{l^p_{[\mathfrak{u}]}} :& \Theta \to \mathbb{R}_+. \\ \theta &\mapsto \left(\int_{j\in\mathbb{F}} |(\theta \cdot [\mathfrak{u}])_j|^p dj \right)^{1/p} \end{aligned}$$

In particular, if $[\mathfrak{u}]$ is the sequence constantly equal to 1, we find the definition of $\|\cdot\|_{l^p}$.

In order to apply decision theory, we have to assume that the objects we try to estimate belongs to the space where out loss function is finite, for which we give the following notations.

Definition 6 Spaces $\mathcal{L}^2, \mathcal{L}^p, \mathcal{L}^p_{[\mathfrak{u}]}$ of functions
We define the sets

$$\begin{aligned} \mathcal{L}^2 &:= \{\theta \in \Theta) : \|\theta\|_{l^2} < \infty\}\,; \\ \mathcal{L}^p_{\mathfrak{u}} &:= \{\theta \in \Theta : \|\theta\|_{l^p} < \infty\}\,; \\ \mathcal{L}^p_{[\mathfrak{u}]} &:= \left\{\theta \in \Theta : \|\theta\|_{l^p_{[\mathfrak{u}]}} < \infty\right\}. \end{aligned}$$

However, we stated in section 1.1.1 that our interest in the space Θ is justified by its duality with the function space Ξ. It is hence of interest to consider norms on Ξ and inquire their link with those on Θ before proceeding.

As with Θ we start by defining an inner product:

Definition 7 Scalar product $\langle\cdot|\cdot\rangle_{L^2}$ on Ξ
We define the scalar product

$$\begin{aligned} \langle\cdot|\cdot\rangle_{L^2} : \quad & \Xi \times \Xi \to \overline{\mathbb{R}}. \\ (f,g) &\mapsto \int_{\mathbb{T}} f(x)\overline{g(x)}dx \end{aligned}$$

We obtain with this scalar product the natural L^2 norm :

Definition 8 L^2-norm $\|\cdot\|_{L^2}$ on Ξ
We define the norm

$$\begin{aligned} \|\cdot\|_{L^2} :& \Xi \to \overline{\mathbb{R}_+}. \\ f &\mapsto \langle f|f\rangle_{L^2}^{1/2} = \left(\int_{\mathbb{T}} |f(x)|^2 dx \right)^{1/2} \end{aligned}$$

For statistical inference it is generally necessary to assume that the objects of interest have finite norm. We hence define the space $\mathbb{L}^2$:

Definition 9 Space $\mathbb{L}^2$ of functions
We define the set
$$\mathbb{L}^2 := \{f \in \Xi : \|f\|_{L^2} < \infty\}.$$

It is common to consider the larger family of norms $\|\cdot\|_{L^p}$ for any number p in $[1,\infty]$ which however do not define an inner product space :

Definition 10 L^p-norm $\|\cdot\|_{L^p}$ on Ξ
We define the norm
$$\begin{aligned}\|\cdot\|_{L^p} :\Xi &\to \overline{\mathbb{R}_+}.\\ f &\mapsto \left(\int_{\mathbb{T}} |f(x)|^p dx\right)^{1/p}\end{aligned}$$

Obviously one can define the associated spaces:

Definition 11 Space $\mathbb{L}^p$ of functions
We define the set
$$\mathbb{L}^p := \{f \in \Xi : \|f\|_{L^p} < \infty\}.$$

A last kind of norm which is of interest are the weighted norms. Using a weighted norm as loss function allows to give more interest to some specific features of the functions (high or low frequencies for example). To do so, we need to define the convolution operator on Ξ.

Definition 12 Product $\star$ on Ξ
We define the bi-linear operator
$$\begin{aligned}\star :\Xi^2 &\to \Xi\\ (f,g) &\mapsto \left(t \mapsto \int_{s\in\mathbb{T}} f(s)\cdot g(t-s)ds\right).\end{aligned}$$

With this definition at hand we obtain the following norm.

Definition 13 $L^p_{\mathfrak{u}}$-norm $\|\cdot\|_{L^p_{\mathfrak{u}}}$ on Ξ
Consider a distribution $\mathfrak{u}$ from $\mathbb{T}$ to $\mathbb{D}$. We define the norm
$$\begin{aligned}\|\cdot\|_{L^r_{\mathfrak{u}}} :\Xi &\to \overline{\mathbb{R}_+}.\\ f &\mapsto \left(\int_{\mathbb{T}} |(f\star\mathfrak{u})(x)|^p dx\right)^{1/p}\end{aligned}$$

In particular, if $\mathfrak{u}$ is the Dirac distribution in 0, we find the definition of $\|\cdot\|_{L^p}$.

We finally define the associated spaces:

DEFINITION 14 SPACE $\mathbb{L}^p_{\mathfrak{u}}$ OF FUNCTIONS
We define the set

$$\mathbb{L}^p_{\mathfrak{u}} := \left\{ f \in \Xi) : \|f\|_{L^p_{\mathfrak{u}}} < \infty \right\}.$$

We have, for any p in $[1, \infty]$ and f in Ξ.

$$\begin{aligned}
\|f\|^r_{\mathfrak{u}} &= \left(\int_{\mathbb{T}} |(f \star \mathfrak{u})(x)|^p dx \right)^{1/p} \\
&= \left(\int_{\mathbb{T}} \left| \int_{j \in \mathbb{F}} ([f] \cdot [\mathfrak{u}])(j) \cdot e_j(x) dj \right|^p dx \right)^{1/p} \\
&\leqslant \left(\int_{\mathbb{T}} \int_{j \in \mathbb{F}} ((|[f] \cdot [\mathfrak{u}])(j)| \cdot |e_j(x)|)^p \, dj \, dx \right)^{1/p} \\
&\leqslant \left(\int_{j \in \mathbb{F}} |([f] \cdot [\mathfrak{u}])(j)|^p \int_{[0,1[} |e_j(x)|^r dx \, dj \right)^{1/p} \\
&\leqslant \left(\int_{j \in \mathbb{F}} |([f] \cdot [\mathfrak{u}])(j)|^p dj \cdot 1 \right)^{1/p} \\
&\leqslant \|[f] \cdot [\mathfrak{u}]\|_{l^p} \\
&\leqslant \|[f]\|_{l^p_{[\mathfrak{u}]}}.
\end{aligned}$$

For the specific case of $p = 2$, the theorem of Plancherel holds and we have

$$\begin{aligned}
\|f\|^2_{\mathfrak{u}} &= \|f \star \mathfrak{u}\|^2 \\
&= \|[f] \cdot [\mathfrak{u}]\|_{l^2} \\
&= \|[f]\|_{l^2_{[\mathfrak{u}]}}.
\end{aligned}$$

We can hence conclude that applying decision theory to θ while using the l^2-norm naturally gives results on f. We cannot say as much concerning other types of norms.

We hence assume from now on that the parameter of interest has finite norm.

ASSUMPTION 1 The parameter of interest θ° is in $\mathcal{L}^2$.

The risk function $(\mathcal{R}_n : (\{\mathbb{Y} \to \Theta\} \times \Theta) \to \mathbb{R}_+)_{n \in \mathbb{N}}$ One can notice the the loss function defined previously depends on the observation and, as such, is a random object that cannot, in general be optimised over the choice of estimator.

A way to overcome this limitation is considering a so called risk function such as the expected loss function.

DEFINITION 15 EXPECTED LOSS RISK FUNCTION
We define the sequence of functions

$$\begin{aligned}\mathcal{R}_n :&(\mathbb{Y} \to \Theta) \times \Theta \to \mathbb{R}_+ \\ &(\widehat{\theta}, \theta^\circ) \qquad\qquad \mapsto \mathbb{E}^n_{\theta^\circ}\left[l(\widehat{\theta}, Y, \theta^\circ)\right].\end{aligned}$$

In particular, the mean square error, will be of particular interest throughout this thesis.

DEFINITION 16 MEAN SQUARE ERROR
We denote the risk function associated with the l_2-loss in the following fashion:

$$\begin{aligned}\Phi_n :&(\mathbb{Y} \to \Theta) \times \Theta \to \mathbb{R}_+ \\ &(\widehat{\theta}, \theta^\circ) \qquad\qquad \mapsto \mathbb{E}^n_{\theta^\circ}\left[\left\|\widehat{\theta}(Y) - \theta^\circ\right\|^2_{l_2}\right].\end{aligned}$$

EXAMPLE. PROJECTION ESTIMATOR
If one considers a projection estimator, as in SECTION 1.2.1, *one can carry the following computations out for any m subset of $\mathbb{F}$:*

$$\begin{aligned}\Phi_n(\overline{\theta}^m, \theta^\circ) =&\mathbb{E}^n_{\theta^\circ}\left[\|\overline{\theta}^m - \theta^\circ\|^2_{l^2}\right] \\ =&\mathbb{E}^n_{\theta^\circ}\left[\int\limits_{j\in\mathbb{F}} \left|\overline{\theta}^m_j - {\theta^\circ}_j\right|^2 dj\right] \\ =&\int\limits_m \mathbb{E}^n_{\theta^\circ}\left[\left|\overline{\theta}^m_j - {\theta^\circ}_j\right|^2\right] dj + \int\limits_{\mathbb{F}\setminus m} \left|{\theta^\circ}_j\right|^2 dj \\ =&\int\limits_m \left(\mathbb{V}^n_{\theta^\circ}\left[\overline{\theta}^m_j - {\theta^\circ}_j\right] + \left|\mathbb{E}^n_{\theta^\circ}\left[\overline{\theta}^m_j - {\theta^\circ}_j\right]\right|^2\right) dj + \int\limits_{\mathbb{F}\setminus m} \left|{\theta^\circ}_j\right|^2 dj \\ =&\int\limits_m \left(\mathbb{V}^n_{\theta^\circ}\left[\overline{\theta}^m_j\right] + \left|\mathbb{E}^n_{\theta^\circ}\left[\overline{\theta}^m_j\right] - {\theta^\circ}_j\right|^2\right) dj + \int\limits_{\mathbb{F}\setminus m} \left|{\theta^\circ}_j\right|^2 dj.\end{aligned}$$

We hence define for any m $V^n_m := \int\limits_m \left(\mathbb{V}^n_{\theta^\circ}\left[\overline{\theta}^m_j\right] + \left|\mathbb{E}^n_{\theta^\circ}\left[\overline{\theta}^m_j\right] - {\theta^\circ}_j\right|^2\right) dj$ and note that

$$\left[V^n_m \vee \mathfrak{b}^2_m\right] \leqslant \Phi_n(\overline{\theta}^m, \theta^\circ) \leqslant 2\left[V^n_m \vee \mathfrak{b}^2_m\right].$$

This fact is important as we are mostly interested in the convergence rate which is defined up to a constant and $\left[V^n_m \vee \mathfrak{b}^2_m\right]$ is easier to control by distinguishing the values *of m for which the bias dominates from those for which the variance is the leading factor.*

Definition 17 Mean square error for projection estimators
We denote the risk function associated with the l_2-loss for the projection estimator with threshold m by:

$$\begin{aligned}\Phi_n^m :&\Theta \to \mathbb{R}_+\\ &\theta^\circ \mapsto \left[V_m^n \vee \mathfrak{b}_m^2\right].\end{aligned}$$

The risk function hence allows us to quantify the performance of an estimator independently of the random observation.

Alternatively, one can consider the probability to exceed a certain loss.

Definition 18 Threshold overcome probability risk function
We define the sequence of functions

$$\begin{aligned}\mathfrak{R}_n :&(\mathbb{Y} \to \Theta) \times \Theta \times \mathbb{R}_+ \to \mathbb{R}_+\\ &(\widehat{\theta}, \theta^\circ, a \qquad \mapsto \mathbb{P}_{\theta^\circ}^n\left(l(\widehat{\theta}, Y, \theta^\circ) \geqslant a\right).\end{aligned}$$

In general, one is interested in the asymptotic behaviour of $\mathcal{R}$ or $\mathfrak{R}$ (and then replacing a by a sequence $(a_n)_{n\in\mathbb{N}}$) when n tends to infinity. In particular, for a given estimator $\widehat{\theta}$ and a fixed value θ° of the parameter of interest, the sequence $\mathcal{R}_n(\widehat{\theta}, \theta^\circ)$ is called convergence rate of $\widehat{\theta}$ at θ° and if $\mathfrak{R}_n(\widehat{\theta}, \theta^\circ, a_n)$ tends to 0 as n tends to infinity, a_n is called speed of convergence in probability of $\widehat{\theta}$ at θ°. If this sequence tends to zero, the estimator is called consistent.

While it is technically feasible to minimise the risk function over $\widehat{\theta}$ for each θ°, the result will be discountenancing as the minimisers will invariably be functions almost surely equal to θ° itself which brilliantly yields a loss function equal to 0, independently of the observation and hence a risk function equal to 0. Our goal being to estimate θ°, it is obvious that such an estimator is not at hand.

We are interested in this thesis in two formulations of optimality which allow to overcome this limitation.

Oracle optimality Introduce $\mathfrak{o}_n$, $\mathcal{O}_n$, $\mathfrak{o}_{\mathbb{P}}$, $\mathcal{O}_{\mathbb{P}}$, $\asymp$,... here

Consider $\mathcal{E}$, a family of estimators.

Definition 19 Oracle convergence rate
A sequence of functions $(\mathcal{R}_{\mathcal{E},n} : \Theta \to \mathbb{R}_+)_{n\in\mathbb{N}}$ is called oracle risk for the family of estimators $\mathcal{E}$ if there exist a constant C in $[1, \infty[$ such that, for any θ° in Θ, and all n, we have:

$$\mathcal{R}_{\mathcal{E},n}(\theta^\circ) \leqslant C \cdot \inf_{\widehat{\theta}\in\mathcal{E}} \mathcal{R}_n(\widehat{\theta}, \theta^\circ).$$

Definition 20 Exact oracle convergence rate, oracle optimal estimator

A sequence of functions $\mathcal{R}^{\circ}_{\mathcal{E},n} : \Theta \to \mathbb{R}_+$ is called exact oracle convergence rate for the family of estimators $\mathcal{E}$ if, in addition to being an oracle convergence rate, there exist an element $\widehat{\theta}$ of $\mathcal{E}$ such that, for any θ° in Θ and n in $\mathbb{N}$, we have:

$$\mathcal{R}^{\circ}_{\mathcal{E},n}(\theta^{\circ}) \geqslant C^{-1} \cdot \mathcal{R}_n(\widehat{\theta}, \theta^{\circ}).$$

An estimator such as $\widehat{\theta}$ is called oracle optimal.

We carry on with the projection estimators example.

Definition 21 Oracle optimal quadratic risk convergence rate for projection estimators

Define $m^{\circ}_n \in \underset{m \subset \mathbb{F}}{\arg\min} \{\Phi^m(\theta^{\circ})\}$, then the sequence $\Phi^{\circ}_n := \Phi^{m^{\circ}_n}_n$ is an exact oracle convergence rate and the projection estimator $\overline{\theta}^{m^{\circ}_n}$ is an oracle optimal estimator.

We see that, given a family of estimators, oracle optimality defines the best element of this family. However, this requires to restrict ourselves to a family of estimator.

Minimax optimality An alternative to oracle optimality is minimax optimality.

Definition 22 Maximal convergence rate

Considering a subset $\widetilde{\Theta}$ of Θ, and an estimator $\widetilde{\theta}$, we call "maximal convergence rate of $\widetilde{\theta}$ over $\widetilde{\Theta}$" the sequence indexed by n defined by

$$\mathcal{R}_{\widetilde{\Theta},n}(\widetilde{\theta}) := \sup_{\theta^{\circ} \in \widetilde{\Theta}} \mathcal{R}_n(\widetilde{\theta}, \theta^{\circ}).$$

Definition 23 Minimax optimal convergence rate

Considering a subset $\widetilde{\Theta}$ of Θ, a sequence $\mathcal{R}^{\star}_{\widetilde{\Theta},n}$ is called minimax convergence rate if there exist a constant C greater than 1 such that, for any n in $\mathbb{N}$

$$\mathcal{R}^{\star}_{\widetilde{\Theta},n} \leqslant C \cdot \inf_{\widetilde{\theta} \in \{\mathbb{Y} \to \Theta\}} \mathcal{R}_{\widetilde{\Theta},n}(\widetilde{\theta}).$$

Moreover, $\mathcal{R}^{\star}_{\widetilde{\Theta},n}$ is called minimax optimal convergence rate if there exists some estimator $\widehat{\theta}$ such that

$$\mathcal{R}^{\star}_{\widetilde{\Theta},n} \geqslant C^{-1} \cdot \mathcal{R}_{\widetilde{\Theta},n}(\widehat{\theta}).$$

An estimator such as $\widehat{\theta}$ is called minimax optimal.

In this definition, be aware that the infimum is taken over all possible estimator of θ°. An example of space which we use in this thesis as $\widetilde{\Theta}$ are Sobolev's ellipsoids.

Definition 24 Sobolev's ellipsoids

Given a constant r in $\mathbb{R}_+$, and a positive, decreasing sequence of numbers smaller than 1, $(\mathfrak{a}_j)_{j\in\mathbb{F}}$, we define the Sobolev's ellipsoid $\Theta(\mathfrak{a},r)$ by

$$\Theta(\mathfrak{a},r) := \{\theta \in \Theta : \|\theta\|_{\mathfrak{a}} \leqslant r\}.$$

Those ellipsoid are interesting as they can directly be related to classes of regularity for the counterpart space Ξ.

On those spaces, the projection estimators yield the following maximal convergence rate.

Definition 25 Maximal convergence rate of projection estimators over Sobolev's ellipsoids
Start by highlighting the following fact:

$$\mathfrak{b}_m^2 = \int_{\mathbb{F}\setminus m} |\theta_j|^2 dj = \int_{\mathbb{F}\setminus m} \frac{\mathfrak{a}_j}{\mathfrak{a}_j} |\theta_j|^2 dj \leqslant \max_{j\notin m}\{\mathfrak{a}_j\} \int_{\mathbb{F}\setminus m} \frac{|\theta_j|^2}{\mathfrak{a}_j} dj \leqslant r \max_{j\notin m}\{\mathfrak{a}_j\}.$$

Hence we can write:

$$\mathcal{R}_{\Theta(\mathfrak{a},r)}(\overline{\theta}^m) \leqslant \sup_{\theta^\circ\in\Theta(\mathfrak{a},r)} \Phi_n^m \leqslant \sup_{\theta^\circ\in\Theta(\mathfrak{a},r)} \left[V_m^n \vee \mathfrak{b}_m^2\right] \leqslant \sup_{\theta^\circ\in\Theta(\mathfrak{a},r)} \left[V_m^n \vee r \max_{j\notin m}\{\mathfrak{a}_j\}\right] \leqslant$$

$[V_m^n \vee \mathfrak{a}_m]$

COMPLETE HERE

The minimax optimal and the oracle optimal rates depend in the true parameter f and the operator T and, hence in their respective counterpart θ° and λ. Some typical behaviours of θ° and λ are often considered in order to compare upper bounds obtained in theorems with the optimal rates as the comparison is not always obvious.
In particular, throughout this thesis, we shall distinguish the following two cases for θ°, respectively called parametric and non-parametric which commonly lead to very different behaviour of the optimal rates:

(p) there exist a finite subset K of $\mathbb{F}$ such that, for any subset K' of K, $\mathfrak{b}_{K'}(\theta^\circ) > 0$ and $\mathfrak{b}_K(\theta^\circ) = 0$;

(np) for all finite subset K of $\mathbb{F}$, $\mathfrak{b}_K(\theta^\circ) > 0$.

Note that the Fourier series expansion of the function of interest f is, in case (p), *finite*, i.e., $f = \sum_{j\in K} \theta_j^\circ e_j$ for some finite subset K of $\mathbb{F}$ while in the opposite case (np), it is *infinite*, i.e., not finite.

Numerical discussion 1.2.1.
The upper bounds we give will be discussed in such "numerical discussions" where we consider the following typical behaviours of θ° and λ and give an equivalent to the upper bound in terms of an explicit function of n.
Regarding the operator eigen-values λ, we consider the following two cases, respectively called ordinary smooth and super-smooth:

(o) there exists a strictly positive real number a such that $\Lambda_m \sim m^{2a}$, then $m\overline{\Lambda}_m \sim m^{2a+1}$ and $\Lambda_{(m)} \sim m^{2a}$;

(s) there exists a strictly positive real number a such that $\Lambda_m \sim \exp(m^{2a})$, then $m\overline{\Lambda}_m \sim m^{-(1-2a)_+} \exp(m^{2a})$ and $\Lambda_{(m)} \sim \exp(m^{2a})$.

As for the parameter of interest θ°, we express the typical behaviours in terms of its tails i.e., $(\mathfrak{b}_m(\theta^\circ))_{m\in\mathbb{F}} = \|\Pi_m^\perp \theta^\circ\|_{\ell^2}$, we distinguish the cases **(p)** and **(np)**, and with **(np)** distinguish the super smooth and ordinary smooth for the parameter of interest.

(o) there exists a strictly positive real number p such that $\mathfrak{b}_m^2(\theta^\circ) \sim m^{-2p}$;

(s) there exists a strictly positive real number p such that $\mathfrak{b}_m^2(\theta^\circ) \sim \exp(-m^{2p})$.

We consider the following situations: in the cases **[p-o]** and **[p-s]** the parameter of interest has a finite representation **(p)** and the operator is either ordinary smooth **(o)** or super smooth **(s)**. In the cases **[o-o]** and **[o-s]** the parameter of interest is ordinary smooth **(o)** and the operator is either ordinary smooth **(o)** or super smooth **(s)**. Case **[s-o]** is the opposite of case **[o-s]**. □

While the names given here to the typical cases may seem arbitrary, we shall justify them through the examples treated in this thesis where the decaying rate of θ° and λ respectively can be interpreted in terms of function smoothness.
The particular interest for these different cases will also appear natural as the behaviour of the optimal rate will be considerably different in our examples; moreover, this phenomenon is observed in many statistical models, also outside of our field of interest.

1.2.3 Adaptivity

- penalised contrast
- Lepski
- ...

1.3 Bayesian approach

1.3.1 The Bayesian paradigm

- Bayes' theorem;
- prior distribution;
- posterior distribution (include conditions of existence);

1.3.2 Typical priors for non-parametric models

- Gaussian process prior

- Sieve priors (specific case)

$$\mathbb{P}^n_{\boldsymbol{\theta}}(\theta) = \exp\left[-\frac{1}{2}\sum_{|j|\leqslant m}|\theta_j|^2\right]\cdot\prod_{|j|>m}\delta_0(\theta_j)$$

- Chinese restaurant process
- Dirichlet process

1.3.3 The pragmatic Bayesian approach

- Consistence
- contraction rate
- exact contraction rate
- uniform contraction rate
- oracle optimality
- minimax optimality

1.3.4 Existing central results

- Goshal Van der Vaart
- Nickl

1.3.5 Iteration procedure, self informative limit and Bayes carrier

1.4 First examples of inverse problem: the inverse Gaussian sequence space model

Consider the Gaussian process $Y(x)$, defined on $[0,1[$ with constant volatility $\frac{1}{n}$ with n in $\mathbb{N}^\star$ and mean process $f\star g$ where f and g are functions from $[0,1[$ to $\mathbb{R}$. In short, we have $dY(x) = (f\star g)(x)dx + \frac{1}{n}dW(x)$ where W is the Brownian motion. We want to estimate f while observing a realisation of Y. We assume that g is known.
We denote θ and λ respectively the Fourier transforms of f and g respectively.
The likelihood with respect to the standard Brownian motion, noted $\mathbb{P}^\circ$, for this model can be written as follows (see Liptser and Shiryaev (2013))

$$\frac{d\,\mathbb{P}^n_{Y^n|f,g}}{d\,\mathbb{P}^\circ} \propto \exp\left[\int_{[0,1[}\frac{1}{\sqrt{n}}(f\star g)(x)dW(x) - \frac{1}{2}\left\|\frac{f\star g}{\sqrt{n}}\right\|^2\right].$$

We use the fact that the volatility of the process is constant and the properties of the Fourier transform to show that there exist a sequence of independent random variables with standard normal distribution such that the likelihood of the Fourier transform of the process is given by:

$$\frac{d\,\mathbb{P}^n_{Y^n|(\theta,\lambda)}}{d\,\mathbb{P}^\circ} \propto \exp\left[-\frac{1}{2}\sum_{j\in\mathbb{Z}}\frac{(\theta_j\lambda_j-\xi_j)^2}{\sqrt{n}}\right].$$

Therefore, the Fourier transform of the observed process follows a Gaussian process indexed by $\mathbb{Z}$, with mean $\theta\cdot\lambda$ and variance $\frac{1}{n}$.
Note that if the volatility was not constant, we would obtain

$$\frac{d\,\mathbb{P}^n_{Y^n|(\theta,\lambda)}}{d\,\mathbb{P}^\circ} \propto \exp\left[-\frac{1}{2}\sum_{j\in\mathbb{Z}}((\sigma\star(\theta\lambda))_j-\xi_j)^2\right].$$

The mean process would hence be $\sigma\star(\theta\cdot\lambda)$, which can be rewritten as an inverse problem with a non diagonal operator, more precisely a Toeplitz operator. We do not consider this case in this thesis.
Another motivation for this model is the heat equation. Heat equation + oracle rate for projection estimate + minimax rate (so all notations are introduced before moving on)

Numerical discussion 1.4.1.
Let us now illustrate the oracle rate $(\mathcal{R}^\circ_n(f,\Lambda))_{n\in\mathbb{N}}$ in those cases. Firstly, recall that for any error density g and associated $(\Lambda_m)_{m\in\mathbb{N}}$ in case **(p)** the oracle rate is parametric, that is $\mathcal{R}^\circ_n(f,\Lambda)\sim n^{-1}$. Thereby, in both, the case **[p-o]** and **[p-s]**, the oracle rate is parametric, i.e. setting $n_f := \frac{K\overline{\Lambda}_K}{\mathfrak{b}^2_{K-1}(f)}$, for all $n\geqslant n_f$ holds $\mathfrak{b}^2_{K-1}(f) > K\overline{\Lambda}_K n^{-1}$, and hence $m^\circ_n = K$ and $\mathcal{R}^\circ_n(f,\Lambda) = K\overline{\Lambda}_K n^{-1}\sim n^{-1}$, where

[p-o] $n_f\sim K^{2a+1}/\mathfrak{b}^2_{K-1}(f)$,

[p-s] $n_f\sim K^{-(1-2a)_+}\exp(K^{2a})/\mathfrak{b}^2_{K-1}(f)$.

On the other hand side, to illustrate **(np)**, where the oracle rate is non-parametric, more precisely, $\lim_{n\to\infty} n\mathcal{R}^\circ_n(f,\Lambda)=\infty$ we consider the cases

[o-o] $\mathcal{R}^\circ_n(f,\Lambda)\sim(m^\circ_n)^{-2p}\sim(m^\circ_n)^{2a+1}n^{-1}$, and hence, $m^\circ_n\sim n^{1/(2p+2a+1)}$ and $\mathcal{R}^\circ_n(f,\Lambda)\sim n^{-2p/(2p+2a+1)}$

[o-s] $\mathcal{R}^\circ_n(f,\Lambda)\sim(m^\circ_n)^{-2p}\sim(m^\circ_n)^{-(1-2a)_+}\exp((m^\circ_n)^{2a})n^{-1}$, and hence, $m^\circ_n\sim(\log n-\frac{2p-(1-2a)_+}{2a}\log\log n)^{1/(2a)}$ and $\mathcal{R}^\circ_n(f,\Lambda)\sim(\log n)^{-p/a}$.

[s-o] $\mathcal{R}^\circ_n(f,\Lambda)\sim\exp(-(m^\circ_n)^{2p})\sim(m^\circ_n)^{2a+1}n^{-1}$, and hence, $m^\circ_n\sim(\log n-\frac{2a+1}{2p}\log\log n)^{1/(2p)}$ and $\mathcal{R}^\circ_n(f,\Lambda)\sim(\log n)^{(2a+1)/(2p)}n^{-1}$.

1.5 Second example of inverse problem: circular density deconvolution

1.5.1 The model

The circular deconvolution model is defined as follows: let X and ε be circular random variables (that is to say, taking values in the unit circle, identified to the interval $[0,1[$)

with respective distributions $\mathbb{P}^X$ and $\mathbb{P}^\varepsilon$ and densities f^X and f^ε with respect to some common and known dominating measure μ on $([0,1[,\mathcal{A})$. We would hence write for any x in $[0,1[$, $f^X(x) = \frac{d\mathbb{P}^X}{d\mu}(x)$ for instance.

Definition 26 Modular addition
From now on we denote by $\square$ the modular addition on $[0,1[$. That is to say

$$\forall (x,y) \in [0,1[^2, \quad x\square y = x + y[1] = x + y - \lfloor x+y \rfloor.$$

The object of interest is f^X while we only observe identically distributed replications $Y^n = (Y_k)_{k\in[\![1,n]\!]}$ of the random variable Y, defined by $Y := X\square\varepsilon$. We note $\mathbb{P}^Y$ the distribution of the random variable Y and f^Y its density with respect to μ. One would notice that $\mathbb{P}^Y$ and f^Y are respectively given, for any A in $\mathcal{A}$ and y in $[0,1[$, by $\mathbb{P}^Y(A) = (\mathbb{P}^X\star\mathbb{P}^\varepsilon)(A) = \int\limits_{[0,1[}\int\limits_{[0,1[} \mathbb{1}_A(x\square s)d\mathbb{P}^X(x)d\mathbb{P}^\varepsilon(s)$ and $f^Y(y) = (f^X\star f^\varepsilon)(y) = \int\limits_0^1 f^X(y\square(-s))f^\varepsilon(s)d\mu(s)$. Indeed, for any μ-measurable and μ-almost surely bounded function g, we have

$$\begin{aligned}
\mathbb{E}[g(Y)] =&\mathbb{E}[g(X\square\varepsilon)]\\
=&\int_0^1\int_0^1 g(x\square s)d\mathbb{P}^X(x)d\mathbb{P}^\varepsilon(s)\\
=&\int_0^1\int_0^1 g(y)d\mathbb{P}^X(y\square(-s))d\mathbb{P}^\varepsilon(s)\\
=&\int_0^1 g(y)\int_0^1 d\mathbb{P}^\varepsilon(s)d\mathbb{P}^X(y\square(-s))\\
=&\int_0^1 g(y)\int_0^1 f^X(y\square(-s))f^\varepsilon(s)d\mu(s)d\mu(y);
\end{aligned}$$

one should note that the integrals above converge, according to the dominated convergence theorem.
We will thus note $\mathcal{D}_\mu([0,1[)$ the space of densities on $[0,1[$ with respect to μ. Moreover we write indifferently ${}^\star\cdot$ the unary operator which associates to a distribution itself convoluted with $\mathbb{P}^\varepsilon$ and the unary operator which associates to a density itself convoluted with f^ε. That is to say, given a probability measure $\mathbb{P}$ on $([0,1[,\mathcal{A})$, ${}^\star\mathbb{P}$ is such that, for any A in $\mathcal{A}$, ${}^\star\mathbb{P}_f(A) = (\mathbb{P}^\varepsilon\star\mathbb{P}_f)(A)$. And for any element f of $\mathcal{D}_\mu([0,1[)$, ${}^\star f$ is such that, for any x in $[0,1[$, ${}^\star f(x) = (f\star f^\varepsilon)(x)$. The model can therefore at first be written $([0,1[^n, {}^\star\mathbb{P}_f, f\in\mathcal{D}_\mu([0,1[))$, where $\mathbb{P}_f$ is the probability distribution with density f with respect to μ.
I think it should be possible to show that $\mathbb{P}^\varepsilon$ does not have to be continuous w.r.t μ and that $\mathbb{P}^Y$ would be anyway. Hence we do not need a density for $\mathbb{P}^\varepsilon$ and we can compute the Fourier transform of the distribution anyway.

Definition 27 Positive (semi-)definitiveness

A sequence/function $[f]$ from $\mathbb{Z}$ to $\mathbb{C}$ is positive (semi-)definite iff, for any finite subset $\{x_1, \ldots, x_n\}$, the Toeplitz matrix $A = (a_{i,j})_{(i,j)\in[\![1,n]\!]^2}$ with $a_{i,j}$ defined by $[f](x_i - x_j)$ is positive (semi-)definite.
In particular, this requires that $[f](x) = \overline{[f](-x)}$, $[f](0) > 0$ and for all x, $[f](x) \leqslant [f](0)$.

Then, by denoting $\mathcal{M}([0,1[)$ the set of all probability measures on $[0,1[$ and $\mathcal{S}^+(\mathbb{Z})$ the set of all positive definite, complex valued, functions $[f]$ on $\mathbb{Z}$ with $[f](0) = 1$, we define the Fourier transform.

Definition 28 Fourier transform of measures
We denote by $\mathcal{F}$ the Fourier transform operator on measures :

$$\begin{aligned}\mathcal{F}: \quad \mathcal{M}([0,1[) &\to \mathcal{S}^+(\mathbb{Z}) \\ \nu \qquad &\mapsto \left(j \mapsto \int_0^1 \exp\left[-2i\pi jx\right] d\nu(x) \right).\end{aligned}$$

Notation 1 Fourier basis functions
As we will operate in the frequency domain for most of the remaining note, it is convenient to use the following notation for the orthonormal basis used in Fourier transform :

$$\forall j \in \mathbb{Z}, \forall x \in [0,1[, \quad e_j(x) := \exp[-2i\pi jx].$$

Remark 1.5.1. *It is convenient to note that for any x and s in $[0,1[$ and j in $\mathbb{Z}$, we have $e_j[x\square s] = e_j[x]e_j[s]$, due to the periodicity of the complex exponential function.*

As we are interested in densities of probability distributions dominated by a common measure μ we define the Fourier transform with respect to μ.

Definition 29 Fourier transform of densities
We denote by $\mathcal{F}_\mu$ the Fourier transform operator of densities with respect to the measure μ :

$$\begin{aligned}\mathcal{F}_\mu: \quad \mathcal{D}_\mu([0,1[) &\to \mathcal{S}^+(\mathbb{Z}) \\ f \qquad &\mapsto \left(j \mapsto \int_0^1 e_j(x) f(x) d\mu(x) \right).\end{aligned}$$

Notation 2 Fourier transform of useful functions
From now on we adopt the following notations for the functions which will appear regularly :

$$\begin{aligned}\forall j \in \mathbb{Z}, \qquad & \theta_j^\circ := \mathcal{F}_\mu(f^X)(j); \\ & \lambda_j := \mathcal{F}_\mu(f^\varepsilon)(j); \\ \forall f \in \mathcal{D}_\mu([0,1[), \forall j \in \mathbb{Z}, \quad & [f](j) := \mathcal{F}_\mu(f)(j).\end{aligned}$$

Obviously, we have

$$\begin{aligned}
\forall j \in \mathbb{Z}, \mathcal{F}(f^Y)(j) =& \int_0^1 e_j(y)\mathbb{P}^Y(dy) \\
=& \int_0^1 \int_0^1 e_j(x \square s)\mathbb{P}^X(dx)\mathbb{P}^\varepsilon(ds) \\
=& \int_0^1 e_j(s) \int_0^1 e_j(x)\mathbb{P}^X(dx)\mathbb{P}^\varepsilon(ds) \\
=& \int_0^1 e_j(s)\mathbb{P}^\varepsilon(ds) \int_0^1 e_j(x)\mathbb{P}^X(dx) \\
=& \mathcal{F}(\mathbb{P}^\varepsilon)(j)\mathcal{F}(\mathbb{P}^X)(j) \\
=& \int_0^1 f^\varepsilon(s)e_j(s)d\mu(s) \int_0^1 e_j(x)f^X(x)\mu(dx) \\
=& \mathcal{F}_\mu(f^\varepsilon)(j)\mathcal{F}_\mu(f^X)(j) \\
=& \theta_j^\circ \lambda_j
\end{aligned}$$

so the Fourier transform, exchanges convolution with point-wise product.

The following theorem, which is a special case of Bochner's theorem, allows us to formulate an inverse for the Fourier transform.

Theorem 1.5.1. Herglotz's representation theorem
A function $[f]$ from $\mathbb{Z}$ to $\mathbb{C}$ with $[f](0) = 1$ is semi-definite positive iff there exist μ in $\mathcal{M}([0,1[)$ such that for all j in $\mathbb{Z}$, we have

$$[f](j) = \int_{[0,1[} \exp[-2i\pi jx]d\mu(x).$$

The properties of the set $\mathcal{S}^+(\mathbb{Z})$ can be interpreted as follow :

$$\begin{aligned}
\mathcal{F}(f)(j) =& \overline{\mathcal{F}(f^Y)(-j)} \quad && f \text{ is real valued;} \\
\mathcal{F}(f)(0) =& 1 && f \text{ integrates at } 1;
\end{aligned}$$

and $\mathcal{F}(f)$ positive semi-definitive implies the positivity of f.

The Fourier transform being bijective, one can safely write its inversion and we have, for any function $[f]$ in $\mathcal{S}^+$:

$$\forall A \in \mathcal{A}, \qquad \mathcal{F}^{-1}[f](A) = \int_A \sum_{j\in\mathbb{Z}} [f](j)e_j(x)dx;$$
$$\forall x \in [0,1[, \quad \mathcal{F}_\mu^{-1}[f](x) = \sum_{j\in\mathbb{Z}} [f](j)e_j(x).$$

However, in the most general case, the above mentioned series do not necessarily converge and one would need to consider the densities on our model as Schwartz distributions (see Billingsley (1986)). We avoid this difficulty by assuming the considered distributions dominated by the Lebesgue measure. We hence drop the μ index from now on (and, for example note $\mathcal{D}([0,1[)$ instead of $\mathcal{D}_\mu([0,1))$.
We will hence consider the model written in these terms : $\big([0,1[^n, \mathbb{P}_{[f]}, f \in \mathcal{S}^+(\mathbb{Z})\big)$; where $\mathbb{P}_{[f]}$ is the distribution which admits the density with respect to μ which Fourier transform is $[f]$.

1.5.2 Known noise density

Considering an iid. n-sample $Y_1, \dots, Y_n$ from h we denote by $\mathbb{E}^n_Y$ the expectation with respect to their joint distribution $\mathbb{P}^n_Y$. Given an estimator $\widetilde{f}$ of $f \in L^2$ based on the observations we measure its accuracy by a quadratic risk, that is, $\mathbb{E}^n_Y \|\widetilde{f} - f\|^2_{L^2}$. Keep in mind that throughout the paper we assume that $|\lambda_j| > 0$ holds for all $j \in \mathbb{Z}$. Considering $\Lambda = (\Lambda_j)_{j\in\mathbb{N}}$ with $\Lambda_j := |\lambda_j|^{-2}$ for $j \in \mathbb{N}$, we set $\Lambda_{(m)} = \max\{\Lambda_j, j \in [\![1,m]\!]\}$ and $\overline{\Lambda}_m = \frac{1}{m}\sum_{j=1}^m \Lambda_j$. Keeping for each $j \in \mathbb{Z}\backslash\{0\}$ in mind that $\lambda_0 = 1 \geqslant |\lambda_j| = |\lambda_{-j}|$ it follows $\Lambda_{|j|} = |\lambda_j|^{-2}$ and consequently, $\Lambda_{(m)} = \max\big\{|\lambda_j|^{-2}, j \in [\![-m,m]\!]\big\}$ and $\sum_{|j|\in[\![1,m]\!]} |\lambda_j|^{-2} = 2m\overline{\Lambda}_m$. Finally, since $(\lambda_j)_{j\in\mathbb{Z}}$ is a ℓ^2 sequence having only non-zero components bounded by one, i.e., $0 < |\lambda_j| \leqslant 1$, for all $j \in \mathbb{Z}$, it follows $\lim_{m\to\infty} \Lambda_{(m)} = \infty$ and for any diverging sequence $(m_n)_{n\in\mathbb{N}}$ of positive integers, i.e., $\lim_{n\to\infty} m_n = \infty :\Leftrightarrow \forall K > 0 : \exists n_o \in \mathbb{N} : \forall n \geqslant n_o : m_n \geqslant K$, holds $\lim_{m\to\infty} m_n\overline{\Lambda}_{m_n} = \infty$. Given $m \in \mathbb{N}$ letting $\mathfrak{b}^2_m(f) := \|\theta^\circ - \overline{[f]}m\|^2_{L^2}$ we observe that

$$\mathbb{E}^n_Y\|\widetilde{f} - f\|^2_{L^2} = \frac{1}{n}\sum_{|j|\in[\![1,m]\!]} \frac{1 - |\phi_j|^2}{|\lambda_j|^2} + \mathfrak{b}^2_m(f) = \frac{1}{n}\sum_{0<|j|\leqslant m}\{\frac{1}{|\lambda_j|^2} - |\theta^\circ_j|^2\} + \mathfrak{b}^2_m(f)$$
$$= \frac{2m}{n}\overline{\Phi}_m - \frac{1}{n}\sum_{0<|j|\leqslant m} |\theta^\circ_j|^2 + \mathfrak{b}^2_m(f)$$

and using $\mathfrak{b}^2_m(f) = \|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2}\mathfrak{b}^2_m(f)$ and $\sum_{0<|j|\leqslant m} |\theta^\circ_j|^2 = \|\theta^{\circ,m}\|^2_{L^2} - 1 = \|\Pi^\perp_{\mathbb{U}_0}\theta^{\circ,m}\|^2_{L^2}$ hence

$$\mathbb{E}^n_Y\|\widetilde{f} - f\|^2_{L^2} + \frac{1}{n}\|\Pi^\perp_{\mathbb{U}_0}\theta^{\circ,m}\|^2_{L^2} = \frac{2m}{n}\overline{\Phi}_m + \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_m(f). \tag{1.1}$$

Defining $\mathcal{R}^\circ_n[m,f,\Lambda] := [\mathfrak{b}^2_m(f) \vee m\overline{\Lambda}_m n^{-1}]$ for $m,n \in \mathbb{N}$ it follows immediately

$$[\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2} \wedge 2]\,\mathcal{R}^\circ_n[m,f,\Lambda] \leqslant \mathbb{E}^n_Y\|\widetilde{f}m - f\|^2_{L^2} + \frac{1}{n}\|\Pi^\perp_{\mathbb{U}_0}\theta^{\circ,m}\|^2_{L^2} \leqslant (\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2} + 2)\,\mathcal{R}^\circ_n[m,f,\Lambda]. \tag{1.2}$$

Let us select $m_n^\circ := \arg\min\{\mathcal{R}_n^\circ[m,f,\Lambda], m\in\mathbb{N}\}$ and set $\mathcal{R}_n^\circ(f,\Lambda) := \mathcal{R}_n^\circ[m_n^\circ, f,\Lambda]$ where $\arg\min_{m\in A}\{a_m\} := \min\{m : a_m \leqslant a_{m'}, \forall m'\in A\}$ for a sequence $(a_j)_{j\in\mathbb{N}}$ with minimal value in $A\subset\mathbb{N}$. We shall emphasise that $\mathcal{R}_n^\circ(f,\Lambda)\geqslant n^{-1}$ for all $n\in\mathbb{N}$, and $\mathcal{R}_n^\circ(f,\Lambda) = o(1)$ as $n\to\infty$. Observe that for all $\delta>0$ there exists $m_\delta\in\mathbb{N}$ and $n_\delta\in\mathbb{N}$ such that for all $n\geqslant n_\delta$ holds $\mathfrak{b}^2_{m_\delta}(\theta)\leqslant\delta$ and $m_\delta\overline{\Lambda}_{m_\delta}n^{-1}\leqslant\delta$, and whence $\mathcal{R}_n^\circ(f,\Lambda)\leqslant\mathcal{R}_n^\circ[m_\delta,f,\Lambda]\leqslant\delta$. Moreover, we have $m_n^\circ\leqslant n$. Indeed, by construction holds $\mathfrak{b}_n^2(f)\leqslant 1<(n+1)n^{-1}\leqslant(n+1)\overline{\Lambda}_{n+1}n^{-1}$, and hence $\mathcal{R}_n^\circ[n,f,\Lambda]<\mathcal{R}_n^\circ[m,f,\Lambda]$ for all $m\in[\![n+1,\infty[\![$ which in turn implies the claim $m_n^\circ\leqslant n$. Obviously, it follows thus $\mathcal{R}_n^\circ(f,\Lambda) = \min\{\mathcal{R}_n^\circ[m,f,\Lambda], m\in[\![1,n]\!]\}$ for all $n\in\mathbb{N}$. We shall use those elementary findings in the sequel without further reference. However, using the dimension $m_n^\circ\in[\![1,n]\!]$ it follows immediately

$$
\begin{aligned}
&[\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\wedge 2]\mathcal{R}_n^\circ(f,\Lambda) - \tfrac{1}{n}\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\leqslant\inf\{\mathbb{E}_Y^n\|\widetilde{f}m - f\|^2_{L^2}, m\in\mathbb{N}\}\\
&\leqslant\mathbb{E}_Y^n\|\widetilde{f}m_n^\circ - f\|^2_{L^2}\leqslant(\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}+2)\mathcal{R}_n^\circ(f,\Lambda) = \tfrac{(\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}+2)}{[\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\wedge 2]}[\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\wedge 2]\mathcal{R}_n^\circ(f,\Lambda)\\
&\leqslant(1+\tfrac{2}{\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}}\vee\tfrac{\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}}{2})\{\inf\{\mathbb{E}_Y^n\|\widetilde{f}m - f\|^2_{L^2}, m\in\mathbb{N}\}+\tfrac{1}{n}\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\}. \quad (1.3)
\end{aligned}
$$

Consequently, the rate $(\mathcal{R}_n^\circ(f,\Lambda))_{n\in\mathbb{N}}$, the dimension parameters $(m_n^\circ)_{n\in\mathbb{N}}$ and the OSE's $(\widetilde{f}m_n^\circ)_{n\in\mathbb{N}}$, respectively, is an oracle rate, an oracle dimension and oracle optimal (up to the constant $2(1+\frac{2}{\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}}\vee\frac{\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}}{2})$).

Interestingly, for any error density g and associated $(\Lambda_m)_{m\in\mathbb{N}}$ in case **(p)** the oracle rate is parametric, that is $\mathcal{R}_n^\circ(f,\Lambda)\sim n^{-1}$. More precisely, if $f = e_o$ then for each $m\in\mathbb{N}$, $\mathbb{E}_h^n\|\widetilde{f}m - f\|^2_{L^2} = 2m\overline{\Lambda}_m n^{-1}$, and hence $m_n^\circ = 1$ and $\mathcal{R}_n^\circ(f,\Lambda) = 2\overline{\Lambda}_1 n^{-1}\sim n^{-1}$. Otherwise if there is $K\in\mathbb{N}$ with $\mathfrak{b}_{K-1}(f)>0$ and $\mathfrak{b}_K(f)=0$, then setting $n_f := \frac{K\overline{\Lambda}_K}{\mathfrak{b}^2_{K-1}(f)}$, for all $n\geqslant n_f$ holds $\mathfrak{b}^2_{K-1}(f)>K\overline{\Lambda}_K n^{-1}$, and hence $m_n^\circ = K$ and $\mathcal{R}_n^\circ(f,\Lambda) = K\overline{\Lambda}_K n^{-1}\sim n^{-1}$. On the other hand side, in case **(np)** the oracle rate is non-parametric, more precisely, it holds $\lim_{n\to\infty} n\mathcal{R}_n^\circ(f,\Lambda)=\infty$. Indeed, since $\mathfrak{b}^2_{m_n^\circ}(f)\leqslant\mathcal{R}_n^\circ[m_n^\circ,f,\Lambda] = \mathcal{R}_n^\circ(f,\Lambda) = o(1)$ as $n\to\infty$ follows $m_n^\circ\to\infty$ and hence $m_n^\circ\overline{\Lambda}_{m_n^\circ}\to\infty$ which implies the claim because $n\mathcal{R}_n^\circ(f,\Lambda)\geqslant m_n^\circ\overline{\Lambda}_{m_n^\circ}$.

Numerical discussion 1.5.1.
Let us now illustrate the oracle rate $(\mathcal{R}_n^\circ(f,\Lambda))_{n\in\mathbb{N}}$ in those cases. Firstly, recall that for any error density g and associated $(\Lambda_m)_{m\in\mathbb{N}}$ in case **(p)** the oracle rate is parametric, that is $\mathcal{R}_n^\circ(f,\Lambda)\sim n^{-1}$. Thereby, in both, the case **[p-o]** and **[p-s]**, the oracle rate is parametric, i.e. setting $n_f := \frac{K\overline{\Lambda}_K}{\mathfrak{b}^2_{K-1}(f)}$, for all $n\geqslant n_f$ holds $\mathfrak{b}^2_{K-1}(f)>K\overline{\Lambda}_K n^{-1}$, and hence $m_n^\circ = K$ and $\mathcal{R}_n^\circ(f,\Lambda) = K\overline{\Lambda}_K n^{-1}\sim n^{-1}$, where

[p-o] $n_f\sim K^{2a+1}/\mathfrak{b}^2_{K-1}(f)$,

[p-s] $n_f\sim K^{-(1-2a)_+}\exp(K^{2a})/\mathfrak{b}^2_{K-1}(f)$.

On the other hand side, to illustrate **(np)**, where the oracle rate is non-parametric, more precisely, $\lim_{n\to\infty} n\mathcal{R}_n^\circ(f,\Lambda)=\infty$ we consider the cases

[o-o] $\mathcal{R}_n^\circ(f,\Lambda)\sim(m_n^\circ)^{-2p}\sim(m_n^\circ)^{2a+1}n^{-1}$, and hence, $m_n^\circ\sim n^{1/(2p+2a+1)}$ and $\mathcal{R}_n^\circ(f,\Lambda)\sim n^{-2p/(2p+2a+1)}$

[o-s] $\mathcal{R}_n^\circ(f,\Lambda)\sim(m_n^\circ)^{-2p}\sim(m_n^\circ)^{-(1-2a)_+}\exp((m_n^\circ)^{2a})n^{-1}$, and hence, $m_n^\circ\sim(\log n-\frac{2p-(1-2a)_+}{2a}\log\log n)^{1/(2a)}$ and $\mathcal{R}_n^\circ(f,\Lambda)\sim(\log n)^{-p/a}$.

[s-o] $\mathcal{R}_n^\circ(f,\Lambda)\sim\exp(-(m_n^\circ)^{2p})\sim(m_n^\circ)^{2a+1}n^{-1}$, and hence, $m_n^\circ\sim(\log n-\frac{2a+1}{2p}\log\log n)^{1/(2p)}$ and $\mathcal{R}_n^\circ(f,\Lambda)\sim(\log n)^{(2a+1)/(2p)}n^{-1}$. $\square$.

1.5.3 Unknown convolution density

Considering independent iid. n-sample $Y_1,\dots,Y_n$ from h and iid. q-sample $\varepsilon_1,\dots,\varepsilon_q$ from g we denote by $\mathbb{E}_{Y,\varepsilon}^{n,q}$, $\mathbb{E}_Y^n$ and $\mathbb{E}_\varepsilon^q$ the expectation with respect to their joint distribution $\mathbb{P}_{Y,\varepsilon}^{n,q}$, $\mathbb{P}_Y^n$ and $\mathbb{P}_\varepsilon^q$, respectively. Exploiting the independence assumption the risk of the OSE $\widehat{\theta^\circ}^m$ can be decomposed as follows

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ}^m-f\|_{L^2}^2=n^{-1}\sum_{|j|\in[\![1,m]\!]}\Lambda_j(1-|\phi_j|^2)\mathbb{E}_\varepsilon^q|\widehat{[g]}_j^+\lambda_j|^2+\|\Pi_{\mathbb{U}_0^\perp}f\|_{\mathbb{H}}^2\mathfrak{b}_m^2(f)$$
$$+\sum_{|j|\in[\![1,m]\!]}|\theta_j^\circ|^2\mathbb{E}_\varepsilon^q|\widehat{[g]}_j-\lambda_j|^2|\widehat{[g]}_j^+|^2+\sum_{|j|\in[\![1,m]\!]}\theta_j^{\circ 2}\mathbb{P}_\varepsilon^q(|\widehat{[g]}_j|^2<1/q).\quad(1.4)$$

Lemma 1.5.1. *There is a finite nurmerical constant $\mathcal{C}_4>0$ such that for all $j\in\mathbb{Z}$ hold $q^2\mathbb{E}_\varepsilon^q|\lambda_j-\widehat{[g]}_j|^4\leqslant\mathcal{C}_4$, (i) $\mathbb{E}_\varepsilon^q|\lambda_j\widehat{[g]}_j^+|^2\leqslant 4$; (ii) $\mathbb{P}_\varepsilon^q(|\widehat{[g]}_j^+|^2<1/q)\leqslant 4(1\wedge\Lambda_j/q)$, (iii) $\mathbb{E}_\varepsilon^q|\lambda_j-\widehat{[g]}_j|^2|\widehat{[g]}_j^+|^2\leqslant 4\mathcal{C}_4(1\wedge\Lambda_j/q)$.*

Proof of Lemma 1.5.1
Since $q\mathbb{E}_g^q|\lambda_j-\widehat{[g]}_j|^2=1-|\lambda_j|^2\leqslant 1$ we obtain (i) as follows

$$\mathbb{E}_g^q|\lambda_j\widehat{[g]}_j^+|^2\leqslant 2\mathbb{E}_g^q\{|\lambda_j-\widehat{[g]}_j|^2|\widehat{[g]}_j^+|^2+\mathbb{1}_{\{|\widehat{[g]}_j|^2\geqslant 1/q\}}\}\leqslant 2(q\mathbb{E}_g^q(\lambda_j-\widehat{[g]}_j)^2+1)\leqslant 4.$$

Consider (ii). Trivially, for any $j\in\mathbb{N}$ we have $\mathbb{P}_g^q(|\widehat{[g]}_j|^2<1/q)\leqslant 1$. If $1\leqslant 4/(q|\lambda_j|^2)=4\Lambda_j/q$, then obviously $\mathbb{P}_g^q(|\widehat{[g]}_j|^2<1/q)\leqslant\min(1,4\Lambda_j/q)$. Otherwise, we have $1/q<|\lambda_j|^2/4$ and hence using Tchebychev's inequaltiy,

$$\mathbb{P}_g^q(|\widehat{[g]}_j|^2<1/q)\leqslant\mathbb{P}_g^q(|\widehat{[g]}_j-\lambda_j|>|\lambda_j|/2)\leqslant 4\Lambda_j\mathbb{E}_g^q|\lambda_j-\widehat{[g]}_j|^2$$
$$\leqslant 4\Lambda_j/q=\min(1,4\Lambda_j/q)$$

where we have used again that $q\mathbb{E}_g^q|\lambda_j-\widehat{[g]}_j|^2\leqslant 1$. Combining both cases we obtain (ii). Consider (iii). there is a numerical constant $\mathcal{C}_4$ such that $q^2\mathbb{E}_g^q|\lambda_j-\widehat{[g]}_j|^4\leqslant\mathcal{C}_4$ due to

Theorem 2.10 of Petrov (1995), which in turn implies

$$\mathbb{E}_g^q|\lambda_j - \widehat{[g]}_j|^2|\widehat{[g]}_j^+|^2 \leqslant \mathbb{E}_g^q\Big\{|\lambda_j - \widehat{[g]}_j|^2|\widehat{[g]}_j^+|^2 2\Big[\frac{|\lambda_j - \widehat{[g]}_j|^2}{|\lambda_j|^2} + \frac{|\widehat{[g]}_j|^2}{|\lambda_j|^2}\Big]\Big\}$$
$$\leqslant \frac{2q\mathbb{E}_g^q|\lambda_j - \widehat{[g]}_j|^4}{|\lambda_j|^2} + \frac{2\mathbb{E}_g^q|\lambda_j - \widehat{[g]}_j|^2}{|\lambda_j|^2} \leqslant 4\mathcal{C}_4\Lambda_j/q.$$

Combining the last bound and $\mathbb{E}_g^q|\lambda_j - \widehat{[g]}_j|^2|\widehat{[g]}_j^+|^2 \leqslant q\mathbb{E}_g^q|\lambda_j - \widehat{[g]}_j|^2 \leqslant 1$ implies (iii), which completes the proof. □

Exploiting Lemma 1.5.1, $|\theta_j^\circ|^2 = |\theta_{-j}^\circ|^2$, $\Lambda_j = \Lambda_{-j}$ and from (1.4) follows

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ}^m - f\|_{L^2}^2 \leqslant 4n^{-1}\sum_{|j|\in[\![1,m]\!]}\Lambda_j + \|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2\mathfrak{b}_m^2(f)$$
$$+ 4\mathcal{C}_4\sum_{|j|\in[\![1,m]\!]}|\theta_j^\circ|^2[1\wedge\Lambda_j/q] + 4\sum_{|j|\in[\![1,m]\!]}|\theta_j^\circ|^2[1\wedge\Lambda_j/q]$$
$$\leqslant (\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2 + 8)\mathcal{R}_n^\circ[m,f,\Lambda] + 8(\mathcal{C}_4+1)\sum_{j\in\mathbb{N}}|\theta_j^\circ|^2[1\wedge\Lambda_j/q],$$

which in turn implies by setting $\mathcal{R}_q^\star(f,\Lambda) := \sum_{j\in\mathbb{N}}|\theta_j^\circ|^2[1\wedge\Lambda_j/q]$ and select again $m_n^\circ := \arg\min\{\mathcal{R}_n^\circ[m,f,\Lambda], m\in\mathbb{N}\}$ with $\mathcal{R}_n^\circ(f,\Lambda) := \mathcal{R}_n^\circ[m_n^\circ,f,\Lambda]$ that

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ}^{m_n^\circ} - f\|_{L^2}^2 \leqslant (\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2 + 8)\mathcal{R}_n^\circ(f,\Lambda) + 8(\mathcal{C}_4+1)\mathcal{R}_q^\star(f,\Lambda). \tag{1.5}$$

Reminder. *We note that $\mathcal{R}_n^\circ(f,\Lambda) \geqslant n^{-1}$ and $\mathcal{R}_q^\star(f,\Lambda) \geqslant q^{-1}\sum_{j\in\mathbb{N}}|\theta_j^\circ|^2 = \frac{1}{2}\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2 q^{-1}$, thereby whenever $f \neq e_0$ any additional term of order $n^{-1}+q^{-1}$ is negligible with respect to the rate $\mathcal{R}_n^\circ(f,\Lambda) + \mathcal{R}_q^\star(f,\Lambda)$, which we will use below without further reference. We shall emphasise that in case $n = q$ it holds*

$$\mathcal{R}_n^\star(f,\Lambda) = \sum_{j\in[\![1,m_n^\circ]\!]}|\theta_j^\circ|^2[1\wedge n^{-1}\Lambda_j] + \sum_{j>m_n^\circ}|\theta_j^\circ|^2[1\wedge n^{-1}\Lambda_j]$$
$$\leqslant \tfrac{1}{2}\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2 m_n^\circ\overline{\Lambda}_{m_n^\circ}/n + \tfrac{1}{2}\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2\mathfrak{b}_{m_n^\circ}^2 \leqslant \|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2\mathcal{R}_n^\circ[m_n^\circ,f,\Lambda] \tag{1.6}$$

which in turn implies $\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ}^m - f\|_{L^2}^2 \leqslant (8+[8\mathcal{C}_4+9]\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2)\mathcal{R}_n^\circ(f,\Lambda)$. In other words, the estimation of the unknown error density g is negligible whenever $n \leqslant q$. □

Let us again consider the two cases **(p)** and **(np)**. We note that in case **(p)** $\mathcal{R}_q^\star(f,\Lambda) \leqslant \|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2\Lambda_{(K)}q^{-1}$ and hence

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ}^{m_n^\circ} - f\|_{L^2}^2 \leqslant \mathcal{C}\{[1\vee\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2]\{K\overline{\Lambda}_K n^{-1} + \Lambda_{(K)}q^{-1}\} \tag{1.7}$$

for all $q\in\mathbb{N}$ and $n \geqslant n_f$. In other words the rate is parametric in both the ε-sample size q and the Y-sample size n. Thereby, the additional estimation of the error density is negligible whenever $q \geqslant n$. In the opposite case **(np)**, it is obviously of interest to characterise the minimal size q of the additional sample from ε needed to attain the same rate as in case of a known error density. Thus, in the next illustration we let the ε-sample size depend on the Y-sample size n as well.

Numerical discussion 1.5.2.
Consider as in Num. discussion 1.5.1 again the usual behaviours **[p-o]**, **[p-s]**, **[o-o]**, **[o-s]** and **[s-o]** for the sequences $(\mathfrak{b}_m(f))_{m\in\mathbb{N}}$ and $(\Lambda_m)_{m\in\mathbb{N}}$. Let $(q_n)_{n\in\mathbb{N}}$ be a sequence of positive integers.
Firstly, recall that due to (1.7) for any error density g and associated $(\Lambda_m)_{m\in\mathbb{N}}$ in case (p) the oracle rate is parametric, that is $\mathcal{R}^\circ_n(f,\Lambda) \sim n^{-1}$ and $\mathcal{R}^\star_q(f,\Lambda) \sim q^{-1}$. Thereby, in both, the case **[p-o]** and **[p-s]**, if $q_{\mathrm{p}} := \lim_{n\to\infty} nq_n^{-1}$ exists[footnote] then it follows that

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^\circ}^{m^\circ_n} - f\|^2_{L^2} = \begin{cases} O(n^{-1}), & \text{if } q_{\mathrm{p}} < \infty, \\ O(q_n^{-1}), & \text{otherwise .} \end{cases} \tag{1.8}$$

On the other hand side, to illustrate (np), where the oracle rate is non-parametric, more precisely, $\lim_{n\to\infty} n\mathcal{R}^\circ_n(f,\Lambda) = \infty$ we consider the cases

[o-o] $m^\circ_n \sim n^{1/(2p+2a+1)}$ and $\mathcal{R}^\circ_n(f,\Lambda) \sim n^{-2p/(2p+2a+1)}$ (cf. Num. discussion 1.5.1 **[o-o]**) and $\mathcal{R}^\star_q(f,\Lambda) \sim q^{-(p\wedge a)/a}$.
If $q_{\text{o-o}} := \lim_{n\to\infty} n^{2(p\vee a)/(2p+2a+1)} q_n^{-1}$ exists then it follows that as $n\to\infty$

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^\circ}^{m^\circ_n} - f\|^2_{L^2} = \begin{cases} O(n^{-2p/(2p+2a+1)}), & \text{if } q_{\text{o-o}} < \infty, \\ O(q_n^{-(p\wedge a)/a}), & \text{otherwise .} \end{cases} \tag{1.9}$$

[o-s] $m^\circ_n \sim (\log n)^{1/(2a)}$ and $\mathcal{R}^\circ_n(f,\Lambda) \sim (\log n)^{-p/a}$ (cf. Num. discussion 1.5.1 **[o-s]**) and $\mathcal{R}^\star_q(f,\Lambda) \sim (\log q)^{-p/a}$
If $q_{\text{o-s}} := \lim_{n\to\infty}(\log n)(\log q_n)^{-1}$ exists, then it follows that as $n\to\infty$

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^\circ}^{m^\circ_n} - f\|^2_{L^2} = \begin{cases} O\big((\log n)^{-p/a}\big), & \text{if } q_{\text{o-s}} < \infty, \\ O\big((\log q_n)^{-p/a}\big), & \text{otherwise .} \end{cases} \tag{1.10}$$

[s-o] $m^\circ_n \sim (\log n)^{1/(2p)}$ and $\mathcal{R}^\circ_n(f,\Lambda) \sim (\log n)^{(2a+1)/(2p)} n^{-1}$ (cf. Num. discussion 1.5.1 **[s-o]**) and $\mathcal{R}^\star_q(f,\Lambda) \sim q^{-1}$.
If $q_{\text{s-o}} := \lim_{n\to\infty} n(\log n)^{-(2a+1)/(2p)} q_n^{-1}$ exists, then it follows that as $n\to\infty$

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^\circ}^{m^\circ_n} - f\|^2_{L^2} = \begin{cases} O\big((\log n)^{(2a+1)/(2p)} n^{-1}\big), & \text{if } q_{\text{s-o}} < \infty, \\ O\big(q_n^{-1}\big), & \text{otherwise .} \end{cases} \tag{1.11}$$

The existence of the limits q_{p}, $q_{\text{o-o}}$, $q_{\text{o-s}}$, and $q_{\text{s-o}}$ is required only to exclude the case of oscillating sequences, which we are not interested in here. In this case, none of the two terms in the upper bound is asymptotically dominant, and the convergence rate is the alternating maximum of the two terms.

[footnote] The limit "∞" is authorized.

In the case **[p-o]** and **[p-s]**, whenever $n = O(q_n)$ we obtain the rate of known error density. However, this is true in the case **[o-o]**, whenever $n^{2(p\vee a)/(2p+2a+1)} = O(q_n)$, which is much less than $q_n = n$. This is even more visible in the case **[o-s]**, where the rate of known error density is attained even if $q_n = n^r$ for arbitrarily small $r > 0$. Moreover, we emphasise the influence of the parameter a as in Num. discussion 1.5.1 **(o)** or **(s)** that characterises the rate of decay of the Fourier coefficients of the error density g. Because a smaller value of a leads to faster rates of convergence, this parameter is often called *degree of ill-posedness*.

□

1.5.4 Empirical distribution and projection estimates

Within this framework, a natural approach for frequentist is, first, estimating of the Fourier coefficients through the empirical distribution and, secondly, reduce the dimension by projecting the sequence.
Consider the empirical distribution of the sample $\overline{\mathbb{P}}^n_{Y^n} = \frac{1}{n} \sum_{k\in[\![1,n]\!]} \delta_{Y_k^n}$. It is a sum of shifted Dirac distributions. Hence, the Fourier transform of this probability distribution can be written, for any j in $\mathbb{Z}$, as $\mathcal{F}(\overline{\mathbb{P}}^n_{Y^n})(j) = \frac{1}{n} \sum_{k\in[\![1,n]\!]} \mathcal{F}(\delta_{Y_k^n}) = \frac{1}{n} \sum_{k\in[\![1,n]\!]} \exp[-2i\pi Y_k^n j]\mathcal{F}(\delta_0) = \frac{1}{n} \sum_{k\in[\![1,n]\!]} \exp[-2i\pi Y_k^n j]$.

This Fourier transform is a random variable for which computation of the expected value and variance are rather straightforward :

$$
\begin{aligned}
\mathbb{E}_{\theta^\circ}[\mathcal{F}(\overline{\mathbb{P}}^n_{Y^n})(j)] =& \mathbb{E}_{\theta^\circ}\left[\frac{1}{n} \sum_{k\in[\![1,n]\!]} \exp[-2i\pi Y_k^n j]\right] \\
=& \mathbb{E}_{\theta^\circ}\left[\exp[-2i\pi Y_1^n j]\right] \\
=& \int_{[0,1[} \exp[-2i\pi y j] d\mathbb{P}^Y(y) \\
=& \theta_j^\circ \lambda_j; \\
\mathbb{V}_{\theta^\circ}[\mathcal{F}(\overline{\mathbb{P}}^n_{Y^n})(j)] =& \mathbb{V}_{\theta^\circ}\left[\frac{1}{n} \sum_{k\in[\![1,n]\!]} \exp[-2i\pi Y_k^n j]\right] \\
=& \frac{1}{n}\mathbb{V}_{\theta^\circ}\left[\exp[-2i\pi Y_1^n j]\right] \\
=& \frac{1}{n}\left(\mathbb{E}_{\theta^\circ}\left[|\exp[-2i\pi Y_1^n j]|^2\right] - |\mathbb{E}_{\theta^\circ}\left[\exp\left[-2i\pi Y_1^n j\right]\right]|^2\right) \\
=& \frac{1}{n}\left(1 - |\theta_j^\circ|^2|\lambda_j|^2\right).
\end{aligned}
$$

We hence obtain an unbiased estimator for θ_j°, $\overline{\theta}_j^n := \frac{1}{n\lambda_j} \sum_{k\in[\![1,n]\!]} \exp[-2i\pi Y_k^n j]$ with variance $\frac{1}{n|\lambda_j|^2}\left(1 - |\theta_j^\circ|^2|\lambda_j|^2\right)$. Note that the variance cancels if and only if $|\theta_j^\circ| = |\lambda_j| = 1$ which is the case for all j only for Dirac distributions.
In order to build an estimator for θ° from these estimators of its components, consider a

threshold m in $\mathbb{N}$ and the estimator $\overline{\theta}^{n,m} = \left(\overline{\theta}_j^n \cdot \mathbb{1}_{|j| \leqslant m}\right)_{j \in \mathbb{Z}}$. Then the risk (that is to say the expected loss) takes the form

$$\begin{aligned}
\mathcal{R}_{\mathfrak{u}}^2(\overline{\theta}^{n,m}|f^X) =& \mathbb{E}_{\theta^\circ}\left[\|\overline{\theta}^{n,m} - \theta^\circ\|_{l_{\mathfrak{u}}^2}^2\right] \\
=& \mathbb{E}_{\theta^\circ}\left[\sum_{j \in \mathbb{Z}} \left|\left(\left(\overline{\theta}^{n,m} - \theta^\circ\right) \cdot [\mathfrak{u}]\right)(j)\right|^2\right] \\
=& \sum_{|j| \leqslant m} \mathbb{E}_{\theta^\circ}\left[\left|\left(\left(\overline{\theta}^{n,m} - \theta^\circ\right) \cdot [\mathfrak{u}]\right)(j)\right|^2\right] + \sum_{|j| > m} |(\theta^\circ \cdot [\mathfrak{u}])(j)|^2 \\
=& \sum_{|j| \leqslant m} \left(\mathbb{V}_{\theta^\circ}\left[\left(\left(\overline{\theta}^{n,m} - \theta^\circ\right) \cdot [\mathfrak{u}]\right)(j)\right] + \left|\mathbb{E}_{\theta^\circ}\left[\left(\left(\overline{\theta}^{n,m} - \theta^\circ\right) \cdot [\mathfrak{u}]\right)(j)\right]\right|^2\right) + \mathfrak{b}_m^2 \\
=& \sum_{|j| \leqslant m} |[\mathfrak{u}](j)|^2 \mathbb{V}_{\theta^\circ}\left[\overline{\theta}_j^{n,m}\right] + \mathfrak{b}_m^2 \\
=& \sum_{|j| \leqslant m} \frac{|[\mathfrak{u}](j)|^2}{n|\lambda_j|^2}\left(1 - |\theta_j^\circ|^2|\lambda_j|^2\right) + \mathfrak{b}_m^2 \\
=& \sum_{|j| \leqslant m} \frac{|[\mathfrak{u}](j)|^2 \Lambda_j}{n} - \frac{|[\mathfrak{u}](j)|^2|\theta_j^\circ|^2}{n} + \mathfrak{b}_m^2.
\end{aligned}$$

Using the notations from previous chapter, we have $\phi_n^{m_n}(f^X, [\mathfrak{u}]) \leqslant \mathcal{R}_{\mathfrak{u}}^2(\overline{\theta}^{n,m}|f^X) + \sum_{|j| \leqslant m_n} \frac{|[\mathfrak{u}](j)|^2|\theta_j^\circ|^2}{n} \leqslant 2 \cdot \phi_n^{m_n}(f^X, [\mathfrak{u}])$

One could minimise this risk with respect to m, hence defining an oracle estimator, that is to say the best projection estimator for a specific value of θ°.

Chapter 2

Bayesian interpretation of penalised contrast model selection

In this chapter, we consider the family of Bayesian methods described as "Gaussian sieve priors" in SECTION 1.3.2 as well as an adaptive variant of these priors, the hierarchical sieve priors where the threshold parameter is a random variable with a specified prior distribution. We study their behaviour under two asymptotic, respectively described in SECTION 1.3.3 and SECTION 1.3.5.

In SECTION 2.1 we consider the self informative Bayes carrier of Gaussian sieve priors under continuity assumptions for the likelihood and show that its support is contained in the maximum likelihood set. Then, in SECTION 2.2 we show that the distribution of the hyper-parameter in the hierarchical prior contracts around the set of maximisers of a penalised contrast criterion. This section highlights a new link between Bayesian adaptive estimation and the frequentist penalised contrast model selection.

In SECTION 2.3, while considering the noise asymptotic, we line out two strategies of proof which allow to obtain contraction rates. The first relies on posterior moment bounding and which, up to our knowledge, is new; the second is specific to the hierarchical sieve prior and is similar to the one used in JOHANNES ET AL. (2016). In SECTION 2.4 we apply this strategies to the specific inverse Gaussian sequence space model. Doing so, we obtain exact contraction rate for the (iterated) Gaussian sieve prior using the first scheme of proof; and the iterated hierarchical prior using the second. This yields optimality for sieve priors with properly chosen threshold parameter; as well as for penalised contrast model selection; and for any iterated version of the hierarchical prior we consider. The most interesting point of this subsection is the novel way to show optimality of the penalised contrast model selection.

Finally, we conclude this chapter in SECTION 2.5 with a note about the shape of the posterior mean of the hierarchical prior, motivating the shape of the frequentist estimators we use in CHAPTER 3.

2.1 Iterated Gaussian sieve prior

We consider in this part a statistical model with a functional parameter space as described in SECTION 1.1.1. We adopt a sieve prior as described in SECTION 1.3.1 and first give interest to the asymptotic presented in SECTION 1.3.5.

We first remind the following notations. The parameter space Θ is a function space $\Theta = \{\theta : \mathbb{F} \to \mathbb{A}\}$; with $\mathbb{F}$ a subset of $\mathbb{R}$ and $\mathbb{A}$ a subset of $\mathbb{C}$.
To derive the self informative Bayes carrier we formulate the following hypothesis.

ASSUMPTION 2 COUNTABILITY ASSUMPTION
We assume that the set $\mathbb{F}$ is countable.

We equip Θ with the usual $\mathbb{L}^2$ norm $\|\cdot\|$, that is, for any θ in Θ, we have $\|\theta\|^2 = \sum_{j\in\mathbb{F}} |\theta_j|^2$ and consider the Borel sigma algebra $\mathcal{B}$ of the topology generated by this $\mathbb{L}^2$ norm.
On the other hand, our observation Y takes values in the space $(\mathbb{Y}, \mathcal{Y})$ with distribution in the family $(\mathbb{P}_{Y|\theta})_{\theta\in\Theta}$.
We assume the existence of a function $l : (\Theta, \mathcal{B}) \times (\mathbb{Y}, \mathcal{Y}) \to \mathbb{R}$ such that the likelihood with respect to some reference measures $\mathbb{P}^\circ$ is given by:

$$L(\boldsymbol{\theta}, y) \propto \exp\left[-l(\boldsymbol{\theta}, y)\right].$$

Then, the family of Gaussian sieve priors is indexed by a threshold parameter m in the set of subsets of $\mathbb{F}$, denoted $\mathcal{P}(\mathbb{F})$, and we denote by $\mathbb{P}_{\boldsymbol{\theta}^m}$ the element of this family with index m; moreover, we denote $\boldsymbol{\theta}^m$ a random variable following this distribution. There exists a reference measure $\mathbb{Q}^\circ$ such that the Gaussian sieve prior with threshold parameter m admits a density of the shape

$$\frac{d\,\mathbb{P}_{\boldsymbol{\theta}^m}}{d\,\mathbb{Q}^\circ}(\boldsymbol{\theta}) \propto \exp\left[-\frac{1}{2}\sum_{j\in m} |\boldsymbol{\theta}_j|^2\right] \cdot \prod_{j\notin m} \delta_0(\boldsymbol{\theta}_j).$$

If we denote by Θ_m the set $\{\theta \in \Theta : \forall j \notin m, \theta_j = 0\}$, Bayes' theorem gives the following shape for the iterated posterior distribution:

$$\begin{aligned}\frac{d\,\mathbb{P}^\eta_{\boldsymbol{\theta}^m|Y}}{d\,\mathbb{Q}^\circ}(\boldsymbol{\theta}, y) &= \frac{\exp\left[-\left(\frac{1}{2}\sum_{j\in m}|\boldsymbol{\theta}_j|^2 + \eta l(\boldsymbol{\theta}, y)\right)\right] \cdot \prod_{j\notin m}\delta_0(\boldsymbol{\theta}_j)}{\int_{\Theta_m} \exp\left[-\left(\frac{1}{2}\sum_{j\in m}|\mu_j|^2 + \eta l(\mu, y)\right)\right] d\mu} \\ &= \frac{\prod_{j\notin m}\delta_0(\boldsymbol{\theta}_j)}{\int_{\Theta_m} \exp\left[-\frac{1}{2}\sum_{j\in m}\left(|\mu_j|^2 - |\boldsymbol{\theta}_j|^2\right)\right] \exp\left[-\eta\left(l(\mu, y) - l(\boldsymbol{\theta}, y)\right)\right] d\mu}.\end{aligned}$$

The following assumption is also needed to obtain the self informative Bayes carrier.

Assumption 3 Continuous likelihood assumption
Assume that for any m in $\mathcal{P}(\mathbb{F})$ and y, $\Theta_m \to \mathbb{R}_+, \theta \mapsto l(\theta, y)$ is continuous.

The use of a threshold parameter brings us back to the study of a parametric model and the results from ref Bunke can be used to derive the self informative Bayes carrier.

Theorem 2.1.1. Self informative Bayes carrier for a sieve prior
Under Assumption 2 and Assumption 3 the support of the Bayesian carrier is contained in the set of minimisers of $\theta \mapsto l(\theta, y)$.

Proof of Theorem 2.1.1.
Let's remind that the definition of continuity gives us:

$$\forall \theta \in \Theta_m, \forall \varepsilon \in \mathbb{R}_+^\star, \exists \delta \in \mathbb{R}_+^\star : \forall \mu \in \Theta_m, \|\mu - \theta\| < \delta \Rightarrow |l(\mu, y) - l(\theta, y)| < \varepsilon.$$

Then, for any B in $\mathcal{B}$ such that $\inf_{\theta \in B} l(\theta, y) > \inf_{\mu \in \Theta_m} l(\mu, y)$, there exist δ in $\mathbb{R}_+^\star$ and a ball $\mathcal{E}$ of Θ_m of radius δ such that, $\sup_{\mu \in \mathcal{E}} l(\mu, y) < \inf_{\theta \in B} l(\theta, y)$ and hence $\sup_{\mu \in \mathcal{E}} l(\mu, y) - \inf_{\theta \in B} l(\theta, y) < 0$.
Hence we can write

$$\begin{aligned}
\mathbb{P}^{\eta}_{\boldsymbol{\theta}^m|Y}(B) =& \int_B \frac{\prod_{|j|>m} \delta_0(\boldsymbol{\theta}_j)}{\int_{\Theta_m} \exp\left[-\frac{1}{2}\sum_{|j|\leqslant m}(|\mu_j|^2 - |\boldsymbol{\theta}_j|^2)\right] \exp\left[-\eta\left(l(\mu, y) - l(\boldsymbol{\theta}, y)\right)\right] d\mu} d\theta \\
\leqslant& \int_B \frac{\prod_{|j|>m} \delta_0(\theta_j)}{\exp\left[-\eta\left(\sup_{\mu \in \mathcal{E}} l(\mu, y) - \inf_{\theta \in B} l(\theta, y)\right)\right] \int_{\mathcal{E}} \exp\left[-\frac{1}{2}\sum_{|j|\leqslant m}(|\mu_j|^2 - |\boldsymbol{\theta}_j|^2)\right] d\mu} d\theta \\
\leqslant& \frac{1}{\exp\left[-\eta\left(\sup_{\mu \in \mathcal{E}} l(\mu, y) - \inf_{\theta \in B} l(\theta, y)\right)\right]} \int_B \frac{\prod_{|j|>m} \delta_0(\theta_j) \exp\left[-\frac{1}{2}\sum_{|j|\leqslant m}|\boldsymbol{\theta}_j|^2\right]}{\int_{\mathcal{E}} \exp\left[-\frac{1}{2}\sum_{|j|\leqslant m}|\mu_j|^2\right] d\mu} d\theta \\
\to& 0.
\end{aligned}$$

□

We have hence showed that under the iteration asymptotic, the posterior distribution contracts itself on maximisers of the likelihood, constrained by $\theta_j = 0$ for any $|j| > m$.
Add remark with several maximisers
There is hence a clear link between this type of prior distribution and projection estimators. We will see that, while considering the noise asymptotic, the choice of the threshold is determinant for the quality of the estimation. The choice of the threshold for the projection estimators and for sieve priors should be led in a similar fashion, that is, balancing the bias (small value of the threshold) and the variance (high value of the threshold). As stated previously, the ideal choice of this parameter is however dependent on the parameter

of interest and hence not available. It is hence important to inquire adaptive methods for the selection of this parameter. Some methods for the frequentist estimation were outlined in the introduction such as the penalised contrast model selection. In the next section, we introduce the hierarchical sieve prior which consists in modelling the threshold parameter as a random variable. We will show that by selecting the prior distribution for this hyper-parameter properly, the iteration asymptotic gives a Bayesian interpretation to the penalised contrast model selection.

2.2 Adaptivity using a hierarchical prior

We denote $\mathbb{P}_{\boldsymbol{\theta}^M}$ a so called hierarchical prior distribution, described hereafter, and $\boldsymbol{\theta}^M$ a random variable following this prior. Define G a finite subset of $\mathbb{F}$ and $\mathrm{pen}: \mathcal{P}(G) \to \mathbb{R}_+$ a so-called penalty function. The threshold parameter noted m for the sieve prior described in the previous section is now a $\mathcal{P}(G)$-valued random variable denoted M. We note $\mathbb{P}_M$ the distribution of this parameter.

The density of $\mathbb{P}_M$ with respect to the counting measure (denoted $\#$) has the shape

$$\frac{d\,\mathbb{P}_M}{d\#}(m) \propto \exp[-\,\mathrm{pen}(m)]\mathbb{1}_{m\subset G}.$$

The dependance structure between the different quantities of the model is then the following:

$$\begin{aligned}\mathbb{P}_{\boldsymbol{\theta}^M|M=m} &= \mathbb{P}_{\boldsymbol{\theta}^m};\\ \mathbb{P}_{Y|\boldsymbol{\theta},M} &= \mathbb{P}_{Y|\boldsymbol{\theta}}\,.\end{aligned}$$

The following proposition is obtained by direct calculus.

Proposition 2.2.1. Iterated posterior distribution for the hyper parameter

Using the convention $\mathbb{P}^0_{\boldsymbol{\theta}^m|Y^n} = \mathbb{P}_{\boldsymbol{\theta}^m}$, *define for any* η *in* $\mathbb{N}^\star$, Y *in* $\mathbb{Y}$, *and* m *included in* G

$$\begin{aligned}\exp[\Upsilon^\eta(Y,m)] &:= \int_\Theta \frac{d\,\mathbb{P}_{Y|\boldsymbol{\theta}^M}}{d\,\mathbb{P}^\circ}(y,\theta)\,\frac{d\,\mathbb{P}^{\eta-1}_{\boldsymbol{\theta}^m|Y^n}}{d\mathbb{Q}^\circ}(m,\theta)d\mathbb{Q}^\circ(\theta)\\ &= \int_\Theta \exp\left[-\left(\frac{1}{2}\sum_{j\in m}|\theta_j|^2+\eta l(\theta,y)\right)\right]d\mathbb{Q}^\circ(\theta);\end{aligned}$$

The iterated posterior distribution of the threshold parameter is given, for any m *subset of*

G and y in $\mathbb{Y}$ by:

$$\begin{aligned}\mathbb{P}^{\eta}_{M|Y}(m,y) =& \frac{\exp\left[-\operatorname{pen}(m)+\eta\Upsilon^{\eta}(y,m)\right]}{\sum\limits_{j\subset G}\exp\left[-\operatorname{pen}(j)+\eta\Upsilon^{\eta}(y,j)\right]}\mathbb{1}_{m\subset G}\\ =& \frac{1}{\sum\limits_{j\subset G}\exp\left[\eta\left(\Upsilon^{\eta}(y,j)-\eta\Upsilon^{\eta}(y,m)\right)-\left(\operatorname{pen}(j)-\operatorname{pen}(m)\right)\right]}\mathbb{1}_{m\subset G}.\end{aligned}$$

Proof of Proposition 2.2.1

$$\begin{aligned}\frac{d\,\mathbb{P}_{M|Y}}{d\#}(m,y) \propto& \frac{d\,\mathbb{P}_{M,Y}}{d\#\,d\,\mathbb{P}^{\circ}}(m,y)\\ \propto& \int_{\Theta}\frac{d\,\mathbb{P}_{M,Y,\boldsymbol{\theta}^M}}{d\#\,d\,\mathbb{P}^{\circ}\ d\,\mathbb{Q}^{\circ}}(m,y,\theta)d\mathbb{Q}^{\circ}(\theta)\\ \propto& \int_{\Theta}\frac{d\,\mathbb{P}_{Y|M,\boldsymbol{\theta}^M}}{d\,\mathbb{P}^{\circ}}(m,y,\theta)\ \frac{d\,\mathbb{P}_{M,\boldsymbol{\theta}^M}}{d\#\,d\mathbb{Q}^{\circ}}(m,\theta)d\mathbb{Q}^{\circ}(\theta)\\ \propto& \int_{\Theta}\frac{d\,\mathbb{P}_{Y|\boldsymbol{\theta}^M}}{d\,\mathbb{P}^{\circ}}(y,\theta)\ \frac{d\,\mathbb{P}_{\boldsymbol{\theta}^M|M}}{d\mathbb{Q}^{\circ}}(m,\theta)\frac{d\,\mathbb{P}_{M}}{d\#}(m)d\mathbb{Q}^{\circ}(\theta)\\ \propto& \frac{d\,\mathbb{P}_{M}}{d\#}(m)\int_{\Theta}\frac{d\,\mathbb{P}_{Y|\boldsymbol{\theta}^M}}{d\,\mathbb{P}^{\circ}}(y,\theta)\ \frac{d\,\mathbb{P}_{\boldsymbol{\theta}^m}}{d\mathbb{Q}^{\circ}}(m,\theta)d\mathbb{Q}^{\circ}(\theta)\\ =& \frac{\frac{d\,\mathbb{P}_{M}}{d\#}(m)\int_{\Theta}\frac{d\,\mathbb{P}_{Y|\boldsymbol{\theta}^M}}{d\,\mathbb{P}^{\circ}}(y,\theta)\ \frac{d\,\mathbb{P}_{\boldsymbol{\theta}^m}}{d\mathbb{Q}^{\circ}}(m,\theta)d\mathbb{Q}^{\circ}(\theta)}{\sum\limits_{j\subset G}\frac{d\,\mathbb{P}_{M}}{d\#}(j)\int_{\Theta}\frac{d\,\mathbb{P}_{Y|\boldsymbol{\theta}^M}}{d\,\mathbb{P}^{\circ}}(y,\theta)\ \frac{d\,\mathbb{P}_{\boldsymbol{\theta}^m}}{d\mathbb{Q}^{\circ}}(j,\theta)d\mathbb{Q}^{\circ}(\theta)}\\ =& \frac{\exp[-\operatorname{pen}(m)]\int_{\Theta_m}\exp[-\frac{1}{2}(2l(y,\theta)+\sum\limits_{k\in m}|\theta_k|^2)]d\mathbb{Q}^{\circ}(\theta)}{\sum\limits_{j\subset G}\exp[-\operatorname{pen}(j)]\int_{\Theta_j}\exp[-\frac{1}{2}(2l(y,\theta)+\sum\limits_{k\in j}|\theta_k|^2)]d\mathbb{Q}^{\circ}(\theta)}.\end{aligned}$$

□

and we can deduce the self informative Bayes carrier.

Lemma 2.2.1. Self informative Bayes carrier of the hyper-parameter in a hierarchical sieve prior I

Note Υ the function

$$\begin{aligned}\Upsilon :& \mathbb{Y}\times\mathcal{P}(G)\to\mathbb{R}\\ &(y,m) \quad\mapsto \lim_{\eta\to\infty}\Upsilon^{\eta}(y,m).\end{aligned}$$

The support of the self informative Bayes carrier for the hyper-parameter M is

$$\arg\max_{m\subset G}\{\Upsilon(Y,m)\}.$$

Unfortunately, in many practical cases, the choice led by $\arg\max_{m\subset G}\{\Upsilon(y,m)\}$ is G itself and leads to inconsistent or suboptimal inference (as we will show later). However, if one allows the prior distribution to depend on η and to take the shape $\exp[-\eta\,\mathrm{pen}(m)]\mathbb{1}_{m\subset G}$, we obtain the following theorem.

Theorem 2.2.1. Self informative Bayes carrier of the hyper-parameter in a hierarchical sieve prior II
Using the modified prior which depends on η, the support of the self informative Bayes carrier for the hyper-parameter M is

$$\arg\max_{m\subset G}\{\Upsilon(Y,m)-\mathrm{pen}(Y,m)\}.$$

Proof of Theorem 2.2.1
For any finite set P of subsets of G such that $\max_{m\in P}\Upsilon(Y,m)-\mathrm{pen}(Y,m)<\max_{k\subset G}\Upsilon(Y,k)-\mathrm{pen}(Y,k)$, there exist a value of $\eta_\circ$ such that, for any η greater than $\eta_\circ$, $\max_{m\in P}\Upsilon^\eta(Y,m)-\mathrm{pen}(Y,m)<\max_{k\subset G}\Upsilon^\eta(Y,k)-\mathrm{pen}(Y,k)$ we can hence write

$$\begin{aligned}
\mathbb{P}^\eta_{M|Y}(P)=&\sum_{m\in P}\frac{1}{\sum_{j\subset G}\exp\left[\eta\left(\Upsilon^\eta(Y,j)-\Upsilon^\eta(Y,m)-(\mathrm{pen}(j)-\mathrm{pen}(m))\right)\right]}\mathbb{1}_{m\subset G}\\
\leqslant&\frac{\mathrm{Card}(P)}{\exp\left[\eta\left(\max_{j\subset G}\left(\Upsilon^\eta(Y,j)-\mathrm{pen}(j)\right)-\max_{m\in P}\left(\Upsilon^\eta(Y,m)-\mathrm{pen}(m)\right)\right)\right]}\mathbb{1}_{m\subset G}\\
\to&0.
\end{aligned}$$

□

Now that we determined the posterior distribution for the hyper-parameter, we can compute the posterior distribution for $\boldsymbol{\theta}^M$ itself.

Proposition 2.2.2. Iterated posterior distribution
The iterated posterior distribution is given by:

$$\begin{aligned}
\frac{d\mathbb{Q}^\eta_{\boldsymbol{\theta}^M|Y}}{d\,\mathbb{P}^\circ}(\theta,y)=&\sum_{m\subset G}\frac{d\,\mathbb{P}^\eta_{\boldsymbol{\theta}^m|Y}}{d\mathbb{Q}^\circ}(\theta,y,m)\frac{d\,\mathbb{P}^\eta_{M|Y}}{d\,\mathbb{P}^\circ}(m,y)\\
=&\sum_{m\subset G}\frac{\exp\left[-\left(\frac{1}{2}\sum_{j\in m}|\boldsymbol{\theta}_j|^2+\eta l(\boldsymbol{\theta},y)\right)\right]\cdot\prod_{j\notin m}\delta_0(\boldsymbol{\theta}_j)}{\int_{\Theta_m}\exp\left[-\left(\frac{1}{2}\sum_{j\in m}|\mu_j|^2+\eta l(\mu,y)\right)\right]d\mu}\frac{\exp[-\mathrm{pen}(m)+\eta\Upsilon^\eta(Y,m)]}{\sum_{j\subset G}\exp[-\mathrm{pen}(j)+\eta\Upsilon^\eta(Y,j)]}\mathbb{1}_{m\subset G}
\end{aligned}$$

Proof for Proposition 2.2.2

$$
\begin{aligned}
\frac{d\mathbb{Q}_{\boldsymbol{\theta}^M|Y}}{d\,\mathbb{P}^\circ}(\theta,y) &\propto \frac{d\,\mathbb{P}_{\boldsymbol{\theta}^M,Y}}{d\mathbb{Q}^\circ\,d\,\mathbb{P}^\circ}(\theta,y)\\
&\propto \sum_{m\subset J}\frac{d\,\mathbb{P}_{\boldsymbol{\theta}^M,Y,M}}{d\mathbb{Q}^\circ\,d\,\mathbb{P}^\circ\,d\,\mathbb{P}^\circ}(\theta,y,m)\\
&\propto \sum_{m\subset J}\frac{d\,\mathbb{P}_{\boldsymbol{\theta}^M|Y,M}}{d\mathbb{Q}^\circ}(\theta,y,m)\frac{d\,\mathbb{P}_{Y,M}}{d\,\mathbb{P}^\circ\,d\,\mathbb{P}^\circ}\\
&\propto \sum_{m\subset J}\frac{d\,\mathbb{P}_{\boldsymbol{\theta}^m|Y}}{d\mathbb{Q}^\circ}(\theta,y,m)\frac{d\,\mathbb{P}_{M|Y}}{d\,\mathbb{P}^\circ}\frac{d\,\mathbb{P}_{Y}}{d\,\mathbb{P}^\circ}(Y)\\
&= \sum_{m\subset J}\frac{d\,\mathbb{P}_{\boldsymbol{\theta}^m|Y}}{d\mathbb{Q}^\circ}(\theta,y,m)\frac{d\,\mathbb{P}_{M|Y}}{d\,\mathbb{P}^\circ}.
\end{aligned}
$$

□

And as a consequence, we can deduce the self informative Bayes carrier.

Theorem 2.2.2. Self informative carrier using a hierarchical sieve prior
Denote $\widehat{m} := \arg\max_{m\subset G}\{\Upsilon(Y,m) - \mathrm{pen}(m)\}$ *then the support of the self informative Bayes carrier is contained in* $\arg\max_{\theta\in\Theta_m, m\in\widehat{m}}\{-l(\theta,Y)\}$.

We have hence seen in these two first sections investigated the behaviour of the sieve prior and its hierarchical version under the iterative asymptotic and shown that under some mild assumptions, their self informative Bayes carriers correspond to some constrained maximum likelihood estimator and penalised contrast model selection version of it respectively. We should now investigate the behaviour of these (iterated) posteriorii under the noise asymptotic and define hypotheses under which they behave properly.

2.3 Proof strategies for contraction rates

In this section, we depict two proof strategies for contraction rates. They will be used in the next sections to compute contraction rates for sieve and hierarchical sieve priors respectively.
The first proof relies on moment bounding of the random variable $\|\boldsymbol{\theta} - \theta^\circ\|$. The second proof relies on the use of exponential concentration inequalities.

2.3.1 A moment control based method for contraction rate computation

In this section we outline a method to prove contraction rates which requires to bound properly some moments of the posterior distribution. We later use this method in the case of the inverse Gaussian sequence space with a sieve prior. Provided that bounds are available for the required moments, this method barely needs any other assumption on the model. Moreover, it appears that, in the example we display here, it leads to the same

rate as the frequentist optimal convergence rate without a logarithmic loss as it is often the case with popular methods.
A limitation is that moments of posterior distributions are not always explicitly available, in particular for non conjugate prior. A consequence is that we were not able to use this method for the deconvolution model nor for computation of contraction rate of the hierarchical prior.
However, we believe that the method could be generalised to wider cases, for example using convergence of distribution in Wasserstein distance implying convergence of moments.
A similar method to obtain lower bounds is described in annex. Unfortunately, it could not be used in any practical case here.

For all this section, Φ_n is the sequence which we want to prove to be a contraction rate; it is in general a function of θ° but we do not make this dependence appear in this section as it has no influence on the proof.

Lemma 2.3.1. Upper bound for posterior expectation
Assume $\max\left\{\mathbb{E}^n_{\theta^\circ}\left[\mathbb{E}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]\right], \sqrt{\mathbb{V}^n_{\theta^\circ}\left[\mathbb{E}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]\right]}\right\} \in \mathcal{O}(\Phi_n)$. *Then, for any increasing unbounded sequence* c_n*, we have:*

$$\lim_{n\to\infty} \mathbb{P}^n_{\theta^\circ}\left(\mathbb{E}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right] \geqslant c_n\Phi_n\right) = 0.$$

Proof of Lemma 2.3.1
Define the sequence of random variables $\mathcal{S}_n := \frac{\mathbb{E}^n_{\boldsymbol{\theta}|Y^n}[\|\boldsymbol{\theta}-\theta^\circ\|]-\mathbb{E}^n_{\theta^\circ}\left[\mathbb{E}^n_{\boldsymbol{\theta}|Y^n}[\|\boldsymbol{\theta}-\theta^\circ\|]\right]}{\sqrt{\mathbb{V}^n_{\theta^\circ}\left[\mathbb{E}^n_{\boldsymbol{\theta}|Y^n}[\|\boldsymbol{\theta}-\theta^\circ\|]\right]}}$. This is a sequence of random variables with common expectation 0 and variance 1 and, as such, their distributions form a sequence of tight measures. Hence, for any increasing unbounded sequence c_n and $K_n := \mathbb{E}^n_{\theta^\circ}\left[\mathbb{E}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]\right] + c_n\sqrt{\mathbb{V}^n_{\theta^\circ}\left[\mathbb{E}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]\right]}$ we can write

$$\begin{aligned}\mathbb{P}^n_{\theta^\circ}\left(\mathbb{E}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right] \geqslant K_n\right) =& \mathbb{P}^n_{\theta^\circ}\left(S_n \geqslant \frac{K_n - \mathbb{E}^n_{\theta^\circ}\left[\mathbb{E}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]\right]}{\sqrt{\mathbb{V}^n_{\theta^\circ}\left[\mathbb{E}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]\right]}}\right)\\ =& \mathbb{P}^n_{\theta^\circ}\left(S_n \geqslant c_n\right)\end{aligned}$$

which tends to 0 as S_n is tight. □

Lemma 2.3.2. Upper bound for posterior variance
Assume $\max\left\{\mathbb{E}^n_{\theta^\circ}\left[\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]}\right], \sqrt{\mathbb{V}^n_{\theta^\circ}\left[\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]}\right]}\right\} \in \mathcal{O}(\Phi_n)$. *Then, for any increasing unbounded sequence* c_n*, we have:*

$$\lim_{n\to\infty} \mathbb{P}^n_{\theta^\circ}\left(\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]} \geqslant c_n\Phi_n\right) = 0.$$

Proof of Lemma 2.3.2

Define the sequence of random variables $\mathcal{S}_n := \frac{\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}[\|\boldsymbol{\theta}-\theta^\circ\|]}-\mathbb{E}^n_{\theta^\circ}\left[\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}[\|\boldsymbol{\theta}-\theta^\circ\|]}\right]}{\sqrt{\mathbb{V}^n_{\theta^\circ}\left[\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}[\|\boldsymbol{\theta}-\theta^\circ\|]}\right]}}$. This is a sequence of random variables with common expectation 0 and variance 1 and, as such, their distributions for a sequence of tight measures. Hence, for any increasing unbounded sequence c_n and $K_n := \mathbb{E}^n_{\theta^\circ}\left[\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]}\right] + c_n\sqrt{\mathbb{V}^n_{\theta^\circ}\left[\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]}\right]}$ we can write

$$\begin{aligned}\mathbb{P}^n_{\theta^\circ}\left(\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]} \geqslant K_n\right) =&\mathbb{P}^n_{\theta^\circ}\left(S_n \geqslant \frac{K_n - \mathbb{E}^n_{\theta^\circ}\left[\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]}\right]}{\sqrt{\mathbb{V}^n_{\theta^\circ}\left[\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]}\right]}}\right)\\ =&\mathbb{P}^n_{\theta^\circ}\left(S_n \geqslant c_n\right)\end{aligned}$$

which tends to 0 as S_n is tight. □

Theorem 2.3.1. Upper bound
Under the hypotheses of Lemma 2.3.1 *and* Lemma 2.3.2 *we have for any increasing unbounded sequence c_n*

$$\lim_{n\to\infty} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(\|\boldsymbol{\theta}-\theta^\circ\| > c_n\Phi_n\right)\right] = 0.$$

Proof of Theorem 2.3.1
Define the tight sequence of random variables $S_n := \frac{\|\boldsymbol{\theta}-\theta^\circ\|-\mathbb{E}^n_{\boldsymbol{\theta}|Y^n}[\|\boldsymbol{\theta}-\theta^\circ\|]}{\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}[\|\boldsymbol{\theta}-\theta^\circ\|]}}$. We consider the sequence of events $\Omega_n := \left\{\left\{\mathbb{E}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right] \geqslant c_n\Phi_n\right\} \cap \left\{\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]} \geqslant c_n\Phi_n\right\}\right\}$. We have $\mathbb{P}^n_{\theta^\circ}(\Omega_n) \leqslant \max\left(\mathbb{P}^n_{\theta^\circ}\left(\left\{\mathbb{E}_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right] \geqslant c_n\Phi_n\right\}\right), \mathbb{P}^n_{\theta^\circ}\left(\left\{\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]} \geqslant c_n\Phi_n\right\}\right)\right)$ which hence tends to 0. Hence, for $K_n := c_n\Phi_n(1+c_n)$, we can write

$$\begin{aligned}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(\|\boldsymbol{\theta}-\theta^\circ\| > K_n\right)\right] &= \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(S_n > \frac{K_n - \mathbb{E}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]}{\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]}}\right)\right]\\
&= \mathbb{E}^n_{\theta^\circ}\left[\mathbb{1}_{\Omega_n}\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(S_n > \frac{K_n - \mathbb{E}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]}{\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]}}\right)\right]\\
&\quad + \mathbb{E}^n_{\theta^\circ}\left[\mathbb{1}_{\Omega_n^c}\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(S_n > \frac{K_n - \mathbb{E}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]}{\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]}}\right)\right]\\
&\leqslant \mathbb{P}^n_{\theta^\circ}\left(\Omega_n\right) + \mathbb{P}^n_{\theta^\circ}\left(\Omega_n^c\right)\cdot\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(S_n > \frac{K_n - c_n\Phi_n}{c_n\Phi_n}\right)\right]\\
&\leqslant \mathbb{P}^n_{\theta^\circ}\left(\Omega_n\right) + \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^n_{\boldsymbol{\theta}|Y^n}\left(S_n > c_n\right)\right].\end{aligned}$$

We can conclude as S_n is a tight sequence, c_n tends to infinity and $\mathbb{P}^n_{\theta^\circ}\left(\Omega_n\right)$ tends to 0. □

2.3.2 An exponential concentration inequality based proof for contraction rates of hierarchical sieve priors

We give here the structure of the proof we use to prove the optimality of the (finally) iterated hierarchical sieve prior. This method takes advantage of the structure of the hierarchical prior and the additive form of the l^2 norm. It is similar to the one used in JOHANNES ET AL. (2016).

ASSUMPTION 4 NON ASYMPTOTIC LOADING OF SMALL SETS
There exist a sequence of sets $G_n^- \subset G_n^\circ$ and a sequence of real numbers $K_{A,n}$ such that the sequence of events $\mathcal{A}_{m,n} := \{\Upsilon^\eta(y, G_n^\circ \setminus m) < K_{A,n}\}$ verifies

$$\sum_{m \subset G_n^-} \exp\left[\eta\left(K_{A,n} - (\mathrm{pen}(m) - \mathrm{pen}(G_n^\circ))\right)\right] \in \mathfrak{o}_n(1)$$
$$\sum_{m \subset G_n^-} \mathbb{P}^n_{\theta^\circ}\left[\mathcal{A}^c_{m,n}\right] \in \mathfrak{o}_n(1)$$

ASSUMPTION 5 NON ASYMPTOTIC LOADING OF LARGE SETS
There exist a sequence of sets $G_n^+ \supset G_n^\circ$ and a sequence of real numbers $K_{B,n}$ such that the sequence of events $\mathcal{B}_{m,n} := \{\Upsilon^\eta(y, m \setminus G_n^\circ) < K_{B,n}\}$ verifies

$$\sum_{m \subset G_n^-} \exp\left[\eta\left(K_{B,n} - (\mathrm{pen}(m) - \mathrm{pen}(G_n^\circ))\right)\right] \in \mathfrak{o}_n(1)$$
$$\sum_{m \subset G_n^-} \mathbb{P}^n_{\theta^\circ}\left[\mathcal{B}^c_{m}\right] \in \mathfrak{o}_n(1)$$

ASSUMPTION 6 OPTIMAL CONTRACTION OF PROPER SIEVES
With the notations of ASSUMPTION 4 and ASSUMPTION 5 assume

$$\sum_{G_n^- \subset m \subset G_n^+} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}^m | Y^n}\left(\left\|\boldsymbol{\theta}^m - \theta_j^\circ\right\|^2_{l^2} > \Phi_n\right)\right] \in \mathfrak{o}_n(1)$$

Note that those assumptions are generally obtained using concentration inequalities such as the one displayed in APPENDIX A.

THEOREM 2.3.2. CONTRACTION RATE FOR ITERATED POSTERIOR OF HIERARCHICAL GAUSSIAN SIEVE PRIORS
Under ASSUMPTION 4, ASSUMPTION 5, *and* ASSUMPTION 4, *for any η in $[\![1, \infty[\![$ there exists a constant K such that*

$$\lim_{n\to\infty} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}^M | Y^n}\left(\left\|\boldsymbol{\theta}^M - \theta^\circ\right\|^2_{l^2} \geqslant K\Phi_n\right)\right] = 0.$$

PROOF OF THEOREM 2.3.2

First, notice that we have the following decomposition:

$$\begin{aligned}
\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}^M|Y^n}\left(\left\|\boldsymbol{\theta}^M-\theta^\circ\right\|^2_{l^2}>\Phi_n\right)\right] &= \mathbb{E}^n_{\theta^\circ}\left[\sum_{m\subset G_n}\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}^M|Y^n}\left(\left\{\left\|\boldsymbol{\theta}^M-\theta^\circ\right\|^2_{l^2}>\Phi_n\right\}\cap\{M=m\}\right)\right]\\
&= \sum_{m\subset G_n}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}^M|Y^n,M=m}\left(\left\{\left\|\boldsymbol{\theta}^M-\theta^\circ\right\|^2_{l^2}>\Phi_n\right\}\right)\cdot\mathbb{P}^{n,(\eta)}_{M|Y^n}(M=m)\right]\\
&= \sum_{m\subset G_n}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}^m|Y^n}\left(\left\{\|\boldsymbol{\theta}^m-\theta^\circ\|^2_{l^2}>\Phi_n\right\}\right)\cdot\mathbb{P}^{n,(\eta)}_{M|Y^n}(M=m)\right]
\end{aligned}$$

.

Then, for any three subsets G_n°, G_n^+ and G_n^- with $G_n^-\subset G_n^\circ\subset G_n^+\subset G_n$, we have

$$\begin{aligned}
\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}^M|Y^n}\left(\|\boldsymbol{\theta}-\theta^\circ\|^2_{l^2}>\Phi_n\right)\right] &\leqslant \sum_{m\subset G_n^-}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}(M=m)\right]+\sum_{m\supset G_n^+}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}(M=m)\right]\\
&\quad+\sum_{G_n^-\subset m\subset G_n^+}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}^m|Y^n}\left(\left\{\|\boldsymbol{\theta}^m-\theta^\circ\|^2_{l^2}>\Phi_n\right\}\right)\right]\\
&\leqslant \underbrace{\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(M\subset G_n^-\right)\right]}_{=:A}+\underbrace{\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(M\supset G_n^+\right)\right]}_{=:B}\\
&\quad+\sum_{G_n^-\subset m\subset G_n^+}\underbrace{\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}^m|Y^n}\left(\left\{\|\boldsymbol{\theta}^m-\theta^\circ\|^2_{l^2}>\Phi_n\right\}\right)\right]}_{=:C_m}
\end{aligned}$$

The goal is then to control the three sums using concentration inequalities.

We begin with A, where the conclusion is given by Assumption 5:

$$\begin{aligned}
A &= \sum_{m\subset G_n^-}\mathbb{E}^n_{\theta^\circ}\left[\frac{\exp\left[\eta\left(-\operatorname{pen}(m)+\Upsilon^\eta(y,m)\right)\right]}{\sum_{j\subset G}\exp\left[\eta\left(-\operatorname{pen}(j)+\Upsilon^\eta(y,j)\right)\right]}\mathbb{1}_{\mathcal{A}_m}\right]+\\
&\qquad\mathbb{E}^n_{\theta^\circ}\left[\frac{\exp\left[\eta\left(-\operatorname{pen}(m)+\Upsilon^\eta(y,m)\right)\right]}{\sum_{j\subset G}\exp\left[\eta\left(-\operatorname{pen}(j)+\Upsilon^\eta(y,j)\right)\right]}\mathbb{1}_{\mathcal{A}_m^c}\right]\\
&\leqslant \sum_{m\subset G_n^-}\mathbb{E}^n_{\theta^\circ}\left[\exp\left[\eta\left(\left(\Upsilon^\eta(y,G_n^\circ)-\Upsilon^\eta(y,m)\right)-\left(\operatorname{pen}(m)-\operatorname{pen}(G_n^\circ)\right)\right)\right]\mathbb{1}_{\mathcal{A}_m}\right]+\\
&\qquad\mathbb{P}^n_{\theta^\circ}\left[\mathcal{A}_m^c\right]\\
&\leqslant \sum_{m\subset G_n^-}\mathbb{E}^n_{\theta^\circ}\left[\exp\left[\eta\left(\Upsilon^\eta(y,G_n^\circ\setminus m)-\left(\operatorname{pen}(m)-\operatorname{pen}(G_n^\circ)\right)\right)\right]\mathbb{1}_{\mathcal{A}_m}\right]+\mathbb{P}^n_{\theta^\circ}\left[\mathcal{A}_m^c\right]\\
&\leqslant \sum_{m\subset G_n^-}\exp\left[\eta\left(K_{A,n}-\left(\operatorname{pen}(m)-\operatorname{pen}(G_n^\circ)\right)\right)\right]+\mathbb{P}^n_{\theta^\circ}\left[\mathcal{A}_m^c\right]\\
&\in\mathfrak{o}_n(1)
\end{aligned}$$

We process similarly for B, where the conclusion is given by ASSUMPTION 5:

$$\begin{aligned}
B = \sum_{m\supset G_n^+} &\mathbb{E}_{\theta^\circ}^n \left[\frac{\exp\left[\eta\left(-\operatorname{pen}(m) + \Upsilon^\eta(y,m)\right)\right]}{\sum\limits_{j\subset G} \exp\left[\eta\left(-\operatorname{pen}(j) + \Upsilon^\eta(y,j)\right)\right]} \mathbb{1}_{\mathcal{B}_m} \right] + \\
&\mathbb{E}_{\theta^\circ}^n \left[\frac{\exp\left[\eta\left(-\operatorname{pen}(m) + \Upsilon^\eta(y,m)\right)\right]}{\sum\limits_{j\subset G} \exp\left[\eta\left(-\operatorname{pen}(j) + \Upsilon^\eta(y,j)\right)\right]} \mathbb{1}_{\mathcal{B}_m^c} \right] \\
\leqslant \sum_{m\supset G_n^+} &\mathbb{E}_{\theta^\circ}^n \left[\exp\left[\eta\left(\left(\Upsilon^\eta(y,G_n^\circ) - \Upsilon^\eta(y,m)\right) - \left(\operatorname{pen}(m) - \operatorname{pen}(G_n^\circ)\right)\right)\right] \mathbb{1}_{\mathcal{B}_m}\right] + \\
&\mathbb{P}_{\theta^\circ}^n \left[\mathcal{B}_m^c\right] \\
\leqslant \sum_{m\supset G_n^+} &\mathbb{E}_{\theta^\circ}^n \left[\exp\left[\eta\left(-\Upsilon^\eta(y, m\setminus G_n^\circ) - \left(\operatorname{pen}(m) - \operatorname{pen}(G_n^\circ)\right)\right)\right] \mathbb{1}_{\mathcal{B}_m}\right] + \mathbb{P}_{\theta^\circ}^n \left[\mathcal{B}_m^c\right] \\
\leqslant \sum_{m\supset G_n^+} &\exp\left[\eta\left(K_{B,n} - \left(\operatorname{pen}(m) - \operatorname{pen}(G_n^\circ)\right)\right)\right] + \mathbb{P}_{\theta^\circ}^n \left[\mathcal{B}_m^c\right] \\
\in \mathfrak{o}_n(1)
\end{aligned}$$

Finally, C_m is directly controlled by ASSUMPTION 6.

□

2.3.3 An exponential concentration inequality based proof for contraction rates of self informative Bayes carrier of hierarchical sieve priors

In the previous section, we described the kind of proof used in JOHANNES ET AL. (2016) and argued that it can also be used with a finitely iterated posterior. We present here an adaptation of this scheme for the self informative Bayes carrier. The main subtlety lies in the fact that the hyper-parameter only loads extrema of a penalised contrast function.

ASSUMPTION 7 NON ASYMPTOTIC LOADING OF SMALL SETS, REVISITED
There exist a sequence of sets $G_n^- \subset G_n^\circ$ such that

$$\sum_{m\subset G_n^-} \mathbb{P}_{\theta^\circ}^n \left(-\Upsilon(G_n^\circ \setminus m, Y^n) < \operatorname{pen}(G_n^\circ - \operatorname{pen}(m))\right) \in \mathfrak{o}_n(1)$$

ASSUMPTION 8 NON ASYMPTOTIC LOADING OF LARGE SETS, REVISITED
There exist a sequence of sets $G_n^+ \supset G_n^\circ$ such that

$$\sum_{m\supset G_n^+} \mathbb{P}_{\theta^\circ}^n \left(\Upsilon(m \setminus G_n^\circ, Y^n) < \operatorname{pen}(G_n^\circ - \operatorname{pen}(m))\right) \in \mathfrak{o}_n(1)$$

Assumption 9 Optimal contraction of proper sieves, revisited
With the notations of Assumption 7 and Assumption 8 assume

$$\sum_{G_n^- \subset m \subset G_n^+} \mathbb{P}_{\theta^\circ}^n \left[\left\| \overline{\theta}^m - \theta_j^\circ \right\|_{l^2}^2 > \Phi_n \right] \in \mathfrak{o}_n(1)$$

Note that those assumptions are generally obtained using concentration inequalities such as the one displayed in appendix A.

Theorem 2.3.3. Contraction rate for iterated posterior of hierarchical Gaussian sieve priors
Under Assumption 7, Assumption 8, and Assumption 9, there exists a constant K such that

$$\lim_{n\to\infty} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}^M|Y^n}^{n,(\infty)} \left(\|\boldsymbol{\theta}^M - \theta^\circ\|_{l^2}^2 \geqslant K\Phi_n \right) \right] = 0.$$

Proof of Theorem 2.3.3
We start the proof in a similar fashion to Theorem 2.3.2:

$$\begin{aligned}
\mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}^M|Y^n}^{n,(\infty)} \left(\|\boldsymbol{\theta}^M - \theta^\circ\|_{l^2}^2 > \Phi_n \right) \right] &= \mathbb{E}_{\theta^\circ}^n \left[\sum_{m \subset G_n} \mathbb{P}_{\boldsymbol{\theta}^M|Y^n}^{n,(\infty)} \left(\left\{ \|\boldsymbol{\theta}^M - \theta^\circ\|_{l^2}^2 > \Phi_n \right\} \cap \{M = m\} \right) \right] \\
&= \sum_{m \subset G_n} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}^M|Y^n, M=m}^{n,(\infty)} \left(\left\{ \|\boldsymbol{\theta}^M - \theta^\circ\|_{l^2}^2 > \Phi_n \right\} \right) \cdot \mathbb{P}_{M|Y^n}^{n,(\infty)} (M = m) \right] \\
&= \sum_{m \subset G_n} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}^m|Y^n}^{n,(\infty)} \left(\left\{ \|\boldsymbol{\theta}^m - \theta^\circ\|_{l^2}^2 > \Phi_n \right\} \right) \cdot \mathbb{P}_{M|Y^n}^{n,(\infty)} (M = m) \right]
\end{aligned}$$

.

Then, for any three subsets G_n°, G_n^+ and G_n^- with $G_n^- \subset G_n^\circ \subset G_n^+ \subset G_n$, we have

$$\begin{aligned}
\mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}^M|Y^n}^{n,(\infty)} \left(\|\boldsymbol{\theta} - \theta^\circ\|_{l^2}^2 > \Phi_n \right) \right] &\leqslant \sum_{m \subset G_n^-} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{M|Y^n}^{n,(\infty)} (M = m) \right] + \sum_{m \supset G_n^+} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{M|Y^n}^{n,(\infty)} (M = m) \right] \\
&\quad + \sum_{G_n^- \subset m \subset G_n^+} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}^m|Y^n}^{n,(\infty)} \left(\left\{ \|\boldsymbol{\theta}^m - \theta^\circ\|_{l^2}^2 > \Phi_n \right\} \right) \right] \\
&\leqslant \underbrace{\mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{M|Y^n}^{n,(\eta)} \left(M \subset G_n^- \right) \right]}_{=:A} + \underbrace{\mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{M|Y^n}^{n,(\eta)} \left(M \supset G_n^+ \right) \right]}_{=:B} \\
&\quad + \sum_{G_n^- \subset m \subset G_n^+} \underbrace{\mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}^m|Y^n}^{n,(\eta)} \left(\left\{ \|\boldsymbol{\theta}^m - \theta^\circ\|_{l^2}^2 > \Phi_n \right\} \right) \right]}_{=:C_m}
\end{aligned}$$

The goal is then to control the three sums using concentration inequalities.

We begin with A, the conclusion is given by Assumption 7:

$$
\begin{aligned}
A &= \mathbb{P}^n_{\theta^\circ}\left[\forall l \supset G_n^-, \mathrm{pen}(\widehat{m}) + \Upsilon(\widehat{m}, Y) < \mathrm{pen}(l) + \Upsilon(l, Y)\right] \\
&\leqslant \mathbb{P}^n_{\theta^\circ}\left[\exists m \subset G_n^-, \mathrm{pen}(m) + \Upsilon(m, Y) < \mathrm{pen}(G_n^\circ) + \Upsilon(G_n^\circ, Y)\right] \\
&\leqslant \sum_{m \subset G_n^-} \mathbb{P}^n_{\theta^\circ}\left[\mathrm{pen}(m) + \Upsilon(m, Y) < \mathrm{pen}(G_n^\circ) + \Upsilon(G_n^\circ, Y)\right] \\
&\leqslant \sum_{m \subset G_n^-} \mathbb{P}^n_{\theta^\circ}\left[-\Upsilon(G_n^\circ \setminus m, Y) < \mathrm{pen}(G_n^\circ) - \mathrm{pen}(m)\right] \\
&\in \mathfrak{o}_n(1).
\end{aligned}
$$

We process similarly for B, the conclusion is given by Assumption 8:

$$
\begin{aligned}
B &= \mathbb{P}^n_{\theta^\circ}\left[\forall l \subset G_n^+, \mathrm{pen}(\widehat{m}) + \Upsilon(\widehat{m}, Y) < \mathrm{pen}(l) + \Upsilon(l, Y)\right] \\
&\leqslant \mathbb{P}^n_{\theta^\circ}\left[\exists m \supset G_n^+, \mathrm{pen}(m) + \Upsilon(m, Y) < \mathrm{pen}(G_n^\circ) + \Upsilon(G_n^\circ, Y)\right] \\
&\leqslant \sum_{m \supset G_n^+} \mathbb{P}^n_{\theta^\circ}\left[\mathrm{pen}(m) + \Upsilon(m, Y) < \mathrm{pen}(G_n^\circ) + \Upsilon(G_n^\circ, Y)\right] \\
&\leqslant \sum_{m \supset G_n^+} \mathbb{P}^n_{\theta^\circ}\left[\Upsilon(m \setminus G_n^\circ, Y) < \mathrm{pen}(G_n^\circ) - \mathrm{pen}(m)\right] \\
&\in \mathfrak{o}_n(1).
\end{aligned}
$$

Finally, C_m is directly controlled by Assumption 9.
□

2.4 First application example: the inverse Gaussian sequence space model

In this section, we consider the inverse Gaussian sequence space model and use the methodology described in section 2.3 to compute upper bounds of the Gaussian sieve priors described in section 2.1 when applied to this specific model. Doing so, we will notice that it gives us, for a very general case, the same speed as the convergence rate of projection estimators and that, by choosing properly the threshold parameter, we reach the oracle rate of convergence as well as the minimax optimal rate, **without a log-loss**.
Then, using a methodology similar to JOHANNES et al. (2016) we show that under some regularity conditions, the iterated hierarchical prior leads to optimal posterior contraction rate. As a consequence, we can conclude about the oracle and minimax optimality of the penalised contrast model selection estimator with a new strategy of proof.

As a reminder $\Phi_n^{m_n}(\theta^\circ)$, $\Phi_n^\circ(\theta^\circ)$, and $\Phi_n^\star(\Theta_{\mathfrak{a},r})$ respectively are the convergence rate for the projection estimator with threshold m_n at θ°, the oracle optimal convergence rate for the family of projection estimators at θ°, and the minimax optimal convergence rate over the

Sobolev ellipsoid $\Theta_{\mathfrak{a},r}$. In this model, they are given by:

$$\begin{aligned}
\Phi_n^{m_n}(\theta^\circ) &= \frac{m_n \overline{\Lambda}_{m_n}}{n} \vee \mathfrak{b}^2_{m_n^\circ};\\
\Phi_n^{\circ}(\theta^\circ) &= \frac{m_n^\circ \overline{\Lambda}_{m_n^\circ}}{n} \vee \mathfrak{b}^2_{m_n^\circ};\\
\Phi_n^{\star}(\Theta_{\mathfrak{a},r}) &= \frac{m_n^\star \overline{\Lambda}_{m_n^\star}}{n} \vee \mathfrak{a}_{m_n^\star}.
\end{aligned}$$

2.4.1 Self informative Bayes carrier for Gaussian sieve in iGSSM

We first consider the asymptotic $\eta \to \infty$ for the Gaussian sieve prior.

Theorem 2.4.1. Self informative Bayes carrier for Gaussian sieve in iGSSM
The self informative Bayes carrier has support in the set of constrained maximisers of the likelihood function. In this model, this set only contains one element which is the projection estimator, introduced in section 1.2.1*:* $\overline{\theta}^{m_n} = \left(\overline{\theta}_j^{m_n}\right)_{j\in\mathbb{N}} = \left(\frac{Y_j}{\lambda_j}\mathbb{1}_{j\leqslant m_n}\right)_{j\in\mathbb{N}}$.

Proof of Theorem 2.4.1
In this model, we have, $\mathbb{F} = \mathbb{N}$ which is countable by definition, hence Assumption 2 is verified.
In addition, for any θ in Θ_m, and y in $\mathbb{Y}$ we have

$$l(\theta, y) \propto -\frac{1}{2\sqrt{n}}\left(\sum_{j=1}^{m}(y_j - \lambda_j\theta_j)^2\right) + C;$$

which is continuous with respect to θ; therefore, Assumption 3 is verified.

We can hence apply Theorem 2.1.1 which proves that the support of the self informative Bayes carrier is contained in the set of maximisers of $l(\theta, y)$ which is obviously the singleton $\{(y_j/\lambda_j \mathbb{1}_{j\leqslant m})_{j\in\mathbb{N}}\}$. □

As an alternative, one could have noticed that the prior and likelihood are conjugated. Define for any j in $\mathbb{N}$ and η in $\mathbb{N}^\star$ the quantities

$$\widetilde{\theta}_j^{(\eta)} := \frac{n\eta Y_j^n \lambda_j}{1 + n\eta\lambda_j^2}; \quad \sigma_j^{(\eta)} := \frac{1}{1 + n\eta\lambda_j^2}.$$

Then, for any j in $\mathbb{N}$, the posterior distribution of $\boldsymbol{\theta}_j$ after η iterations is given by

$$\boldsymbol{\theta}_j | Y^{n,\eta} \sim \mathcal{N}(\widetilde{\theta}_j^{(\eta)}, \sigma_j^{(\eta)})\mathbb{1}_{j\leqslant m_n} + \delta_0(\boldsymbol{\theta}_j)\mathbb{1}_{j>m_n}.$$

Considering the respective limits of $\widetilde{\theta}_j^{(\eta)}$ and $\sigma_j^{(\eta)}$ as η tends to ∞ for any j in $\mathbb{N}$ coincides with our previous statement.

2.4.2 Contraction rate for Gaussian sieve in iGSSM

We now investigate the behaviour of the Gaussian sieve prior applied to iGSSM as n tends to ∞. In this context, it is interesting to let η and m depend on n; we hence note η_n and m_n.

Assumption 10 Reasonable choice of threshold sequence
Assume that m_n and η_n are chosen in such a way that ether

$$\sum_{j=1}^{m_n} \frac{\Lambda_j}{n} = \mathcal{O}(1);$$

or

$$\sum_{j=1}^{m_n} \frac{\Lambda_j^2 \left(\theta_j^\circ\right)^2}{n^2 \eta_n^2} \in \mathcal{O}\left(\sum_{j=1}^{m_n} \frac{\Lambda_j}{n}\right); \text{ and } \quad \sum_{j=1}^{m_n} \frac{\sqrt{\Lambda_j^3} \left|\theta_j^\circ\right|}{\sqrt{n^3}\eta_n} \in \mathcal{O}\left(\sum_{j=1}^{m_n} \frac{\Lambda_j}{n}\right).$$

Numerical discussion 2.4.1.
Discuss here this assumption depending on o-s p-np

□

Corollary 2.4.1. *Under* Assumption 10*, for any θ° in Θ and increasing, unbounded sequence c_n, we have*

$$\lim_{n\to\infty} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}^{m_n}|Y^n}\left(\|\theta^\circ - \boldsymbol{\theta}^{m_n}\|^2 \leqslant c_n \Phi_n^{m_n}\right)\right] = 1.$$

We will apply Theorem 2.3.1. We hence need to show

$$\begin{aligned}
\mathbb{E}^n_{\theta^\circ}\left[\mathbb{E}_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta} - \theta^\circ\|\right]\right] &\in \mathcal{O}(\Phi_n^{m_n});\\
\sqrt{\mathbb{V}^n_{\theta^\circ}\left[\mathbb{E}_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta} - \theta^\circ\|\right]\right]} &\in \mathcal{O}(\Phi_n^{m_n});\\
\mathbb{E}^n_{\theta^\circ}\left[\sqrt{\mathbb{V}_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta} - \theta^\circ\|\right]}\right] &\in \mathcal{O}(\Phi_n^{m_n});\\
\sqrt{\mathbb{V}^n_{\theta^\circ}\left[\sqrt{\mathbb{V}_{\boldsymbol{\theta}|Y^n}\left[\|\boldsymbol{\theta} - \theta^\circ\|\right]}\right]} &\in \mathcal{O}(\Phi_n^{m_n}).
\end{aligned}$$

We use the fact that $\|\boldsymbol{\theta} - \theta^\circ\| = \sum_{j\leqslant m_n} (\boldsymbol{\theta}_j - \theta^\circ)^2 + \mathfrak{b}^2_{m_n}$ and that we know that distribution of $\boldsymbol{\theta}_j$. This gives us the expectation and variance of the posterior distribution of $\|\boldsymbol{\theta} - \theta^\circ\|$. We use in addition $\frac{1}{1+\frac{\Lambda_j}{n\eta_n}} \leqslant 1$ to obtain upper bounds for these quantities.

$$\begin{aligned}\mathbb{E}^n_{\boldsymbol{\theta}^{m_n}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right] =&\sum_{j=1}^{m_n}\frac{\Lambda_j}{n\eta_n}\cdot\left(\frac{1}{\frac{\Lambda_j}{n\eta_n}+1}\right)\left(1+\frac{\left(-\theta^\circ_j+\eta_n\sqrt{n}\xi_j\lambda_j\right)^2}{\frac{\eta_n n}{\Lambda_j}\left(\frac{\Lambda_j}{\eta_n n}+1\right)}\right)+\mathfrak{b}^2_{m_n}\\ \leqslant&\sum_{j=1}^{m_n}\frac{\Lambda_j}{n\eta_n}+\sum_{j=1}^{m_n}\frac{\Lambda_j^2}{n^2\eta_n^2}\left(-\theta^\circ_j+\eta_n\sqrt{n}\xi_j\lambda_j\right)^2+\mathfrak{b}^2_{m_n};\end{aligned}$$

$$\begin{aligned}\mathbb{V}^n_{\boldsymbol{\theta}^{m_n}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right] =&2\sum_{j=1}^{m_n}\left(\frac{\Lambda_j}{n\eta_n}\cdot\frac{1}{\frac{\Lambda_j}{n\eta_n}+1}\right)^2\left(1+2\frac{\left(-\theta^\circ_j+\eta_n\sqrt{n}\xi_j\lambda_j\right)^2}{\frac{\eta_n n}{\Lambda_j}\left(\frac{\Lambda_j}{\eta_n n}+1\right)}\right)\\ \leqslant&2\sum_{j=1}^{m_n}\frac{\Lambda_j^2}{n^2\eta_n^2}+4\sum_{j=1}^{m_n}\frac{\Lambda_j^3}{n^3\eta_n^3}\left(-\theta^\circ_j+\eta_n\sqrt{n}\xi_j\lambda_j\right)^2.\end{aligned}$$

In addition, we use the sub-additivity of the square root to obtain this upper bound:

$$\begin{aligned}\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}^{m_n}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]} \leqslant&\sqrt{2}\sum_{j=1}^{m_n}\frac{\Lambda_j}{n\eta_n}+2\sum_{j=1}^{m_n}\sqrt{\frac{\Lambda_j^3}{n^3\eta_n^3}\left(-\theta^\circ_j+\eta_n\sqrt{n}\xi_j\lambda_j\right)^2}\\ \leqslant&\sqrt{2}\sum_{j=1}^{m_n}\frac{\Lambda_j}{n\eta_n}+2\sum_{j=1}^{m_n}\sqrt{\frac{\Lambda_j^3}{n^3\eta_n^3}}\left|-\theta^\circ_j+\eta_n\sqrt{n}\xi_j\lambda_j\right|.\end{aligned}$$

Using linearity of the expectation and the standard Gaussian distribution of ξ_j we have:

$$\begin{aligned}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{E}^n_{\boldsymbol{\theta}^{m_n}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]\right] \leqslant&\sum_{j=1}^{m_n}\frac{\Lambda_j}{n\eta_n}+\sum_{j=1}^{m_n}\frac{\Lambda_j}{n}\cdot\left[1+\frac{\Lambda_j}{n\eta_n^2}\left(\theta^\circ_j\right)^2\right]+\mathfrak{b}^2_{m_n}\\ \leqslant&\sum_{j=1}^{m_n}\frac{\Lambda_j}{n\eta_n}+\sum_{j=1}^{m_n}\frac{\Lambda_j}{n}+\sum_{j=1}^{m_n}\frac{\Lambda_j^2}{n^2\eta_n^2}\left(\theta^\circ_j\right)^2+\mathfrak{b}^2_{m_n}\end{aligned}$$

The same properties give us this bound:

$$\begin{aligned}\mathbb{V}^n_{\theta^\circ}\left[\mathbb{E}^n_{\boldsymbol{\theta}^{m_n}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]\right] \leqslant&2\sum_{j=1}^{m_n}\frac{\Lambda_j^2}{n^2}\left[1+4\frac{\Lambda_j}{\eta_n^2 n}\left(\theta^\circ_j\right)^2\right]\\ \leqslant&2\sum_{j=1}^{m_n}\frac{\Lambda_j^2}{n^2}+4\sum_{j=1}^{m_n}\frac{\Lambda_j^3}{\eta_n^2 n^3}\left(\theta^\circ_j\right)^2;\end{aligned}$$

And we use the sub-additivity of the square root in addition:

$$\begin{aligned}\sqrt{\mathbb{V}^n_{\theta^\circ}\left[\mathbb{E}^n_{\boldsymbol{\theta}^{m_n}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]\right]} \leqslant& \sqrt{2\sum_{j=1}^{m_n}\frac{\Lambda_j^2}{n^2}+4\frac{\Lambda_j^3}{\eta_n^2 n^3}\left(\theta_j^\circ\right)^2} \\ \leqslant& \sqrt{2}\sum_{j=1}^{m_n}\frac{\Lambda_j}{n}+2\sum_{j=1}^{m_n}\frac{\sqrt{\Lambda_j^3}}{\eta_n\sqrt{n^3}}\left|\theta_j^\circ\right|.\end{aligned}$$

To control the moments of the posterior variance, we use the properties of the folded Gaussian random variables:

$$\begin{aligned}\mathbb{E}^n_{\theta^\circ}\left[\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}^{m_n}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]}\right] \leqslant& \sqrt{2}\sum_{j=1}^{m_n}\frac{\Lambda_j}{n\eta_n}+2\sum_{j=1}^{m_n}\sqrt{\frac{\Lambda_j^3}{n^3\eta_n^3}}\,\mathbb{E}_{\theta^\circ}\left[\left|-\theta_j^\circ+\eta_n\sqrt{n}\xi_j\lambda_j\right|\right] \\ \leqslant& \sqrt{2}\sum_{j=1}^{m_n}\frac{\Lambda_j}{n\eta_n}+2\sum_{j=1}^{m_n}\sqrt{\frac{\Lambda_j^3}{n^3\eta_n^3}}\left(\eta_n\lambda_j\sqrt{\frac{2}{\pi}}\exp\left[-\frac{\left(\theta_j^\circ\right)^2}{2\eta_n^2\lambda_j^2}\right]\right. \\ & \left.+\theta_j^\circ\operatorname{erf}\left\{\frac{\theta_j^\circ}{\sqrt{2}\eta_n\lambda_j}\right\}\right) \\ \leqslant& \sqrt{2}\sum_{j=1}^{m_n}\frac{\Lambda_j}{n\eta_n}+2\sum_{j=1}^{m_n}\sqrt{\frac{2}{\pi\cdot n^3\eta_n}}\Lambda_j\exp\left[-\frac{\left(\theta_j^\circ\right)^2\Lambda_j}{2\eta_n^2}\right]+\sum_{j=1}^{m_n}\sqrt{\frac{\Lambda_j^3}{n^3\eta_n^3}}\theta_j^\circ;\end{aligned}$$

$$\begin{aligned}\mathbb{V}^n_{\theta^\circ}\left[\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}^{m_n}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]}\right] \leqslant& 2\sum_{j=1}^{m_n}\frac{\Lambda_j^3}{n^3\eta_n^3}\,\mathbb{V}^n_{\theta^\circ}\left[\left|-\theta_j^\circ+\eta_n\sqrt{n}\xi_j\lambda_j\right|\right] \\ \leqslant& 2\sum_{j=1}^{m_n}\frac{\Lambda_j^3}{n^3\eta_n^3}\cdot \\ & \left[\left(\theta_j^\circ\right)^2\left(1-\operatorname{erf}^2\left[\frac{\theta_j^\circ}{\sqrt{2}\eta_n\lambda_j}\right]\right)+\right. \\ & \eta_n^2\lambda_j^2\left(1-\frac{2}{\pi}\exp\left[-\frac{\left(\theta_j^\circ\right)^2}{\eta_n^2\lambda_j^2}\right]\right)- \\ & \left.\left(2\eta_n\lambda_j\theta_j^\circ\sqrt{\frac{2}{\pi}}\exp\left[\frac{-\left(\theta_j^\circ\right)^2}{2\eta_n^2\lambda_j^2}\right]\operatorname{erf}\left[\frac{\theta_j^\circ}{\sqrt{2}\eta_n\lambda_j^2}\right]\right)\right] \\ \leqslant& 2\sum_{j=1}^{m_n}\frac{\Lambda_j^3}{n^3\eta_n^3}\cdot\left[\left(\theta_j^\circ\right)^2+\frac{\eta_n^2}{\Lambda_j}\right];\end{aligned}$$

$$\sqrt{\mathbb{V}^n_{\theta^\circ}\left[\sqrt{\mathbb{V}^n_{\boldsymbol{\theta}^{m_n}|Y^n}\left[\|\boldsymbol{\theta}-\theta^\circ\|\right]}\right]}\leqslant\sqrt{2}\sum_{j=1}^{m_n}\sqrt{\frac{\Lambda_j^3\left(\theta_j^\circ\right)^2}{n^3\eta_n^3}}+\sum_{j=1}^{m_n}\sqrt{\frac{\Lambda_j^2}{n^3\eta_n}}.$$

Using Assumption 10, the leading term in each of these bounds is for the most of order $\Phi_n^{m_n}$ and hence, we can apply Theorem 2.3.1 which proves the statement. □

Notice that if one selects $m_n = m_n^\circ$ we obtain the oracle rate of convergence of projection estimators.

Corollary 2.4.2. *For any θ° in Θ and increasing, unbounded sequence c_n, we have*

$$\lim_{n\to\infty}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}^{m_n^\circ}|Y^n}\left(\left\|\theta^\circ-\boldsymbol{\theta}^{m_n^\circ}\right\|^2\leqslant c_n\Phi_n^\circ\right)\right]=1.$$

We have hence seen that Gaussian sieve priors contract around the true parameter at the same rate as the projection estimator with identical threshold parameter contract and that, in particular, the best Gaussian sieve prior contracts at the oracle convergence rate of the projection estimators.

2.4.3 Self informative Bayes carrier for hierarchical Gaussian sieve in iGSSM

In this subsection, we propose an analytical shape for a hierarchical Gaussian sieve prior to use in the context of an inverse Gaussian sequence space model.
We doubly justify this choice, first by showing that the self informative limit is a penalised contrast maximiser projection estimator and, in the next subsection, that this choice yields good contraction properties.

First remind:

$$\widetilde{\theta}_j^{(\eta)}=\frac{n\eta Y_j^n\lambda_j}{1+n\eta\lambda_j^2};\ \text{and}\qquad \sigma_j^{(\eta)}=\frac{1}{1+n\eta\lambda_j^2};$$

and define for any k in $\mathbb{N}$ the notations

$$\sigma^{k,(\eta)}:=(\sigma_j^{(\eta)}\mathbb{1}_{\{j\leqslant k\}})_{j\in\mathbb{N}};\ \text{and}\qquad \widetilde{\theta}^{k,(\eta)}:=(\widetilde{\theta}_j^{(\eta)}\mathbb{1}_{\{j\leqslant k\}})_{j\in\mathbb{N}}.$$

Then, we define $G_n := \max\{m\in[\![1,n]\!]:\Lambda_m/n\leqslant\Lambda_1\}$ (we will see in the next subsection that this choice is important in terms of contraction rate). To keep the analogy with section 2.2, G_n would have to be replaced by $[\![1,G_n]\!]$. Moreover, our prior only gives weight to the subsets of $[\![1,G_n]\!]$ with the shape $[\![1,m]\!]$ with m in $[\![1,G_n]\!]$.
We give the following shape to the prior for the threshold parameter

$$\mathbb{P}^n_M(M=m)=\frac{\exp\left[-3\frac{m}{2}\eta-\frac{\eta}{2}\sum_{j=1}^{m}\sigma_j\right]}{\sum_{k=1}^{G_n}\exp\left[-3\frac{m}{2}\eta-\frac{\eta}{2}\sum_{j=1}^{k}\sigma_j\right]}.$$

Using the notations of SECTION 2.2 (and keeping in mind the notation for weighted norms given in SECTION 1.2.2 in the context of Sobolev's ellipsoid, and the convention "$0/0 = 0$"), we have

$$\begin{aligned}\mathrm{pen}(m) \quad &= 3\frac{m}{2}\eta + \frac{\eta}{2}\sum_{j=1}^{m}\log\left(\sigma_j^{(\eta)}\right);\\ \Upsilon^{\eta}(Y,m) &= \sum_{j=1}^{m}\frac{n\left(Y_j^n\right)^2}{\frac{\Lambda_j}{n\eta}+1} + \frac{1}{2}\sum_{j=1}^{m}\log\left(\sigma_j^{(\eta)}\right).\end{aligned}$$

Which leads us to the iterated prior of the hyper-parameter:

$$\mathbb{P}^{n,(\eta)}_{M|Y^n}(m) = \frac{\exp\left[-\frac{\eta}{2}\left(3m - n\sum_{j=1}^{m}\frac{\left(Y_j^n\right)^2}{\frac{\Lambda_j}{n\eta}+1}\right)\right]}{\sum_{k=1}^{G_n}\exp\left[-\frac{\eta}{2}\left(3k - n\sum_{j=1}^{k}\frac{\left(Y_j^n\right)^2}{\frac{\Lambda_j}{n\eta}+1}\right)\right]}.$$

We can hence simplify our notation in the following way:

$$\mathrm{pen}(m) = 3m \text{ and } \Upsilon^{\eta}(Y,m) = \sum_{j=1}^{m}\frac{n\left(Y_j^n\right)^2}{\frac{\Lambda_j}{n\eta}+1}.$$

Let us remind that the iterated distribution for $\boldsymbol{\theta}^M|Y$ is given by

$$\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}^M|Y^n} = \sum_{m\in\mathbb{N}^\star}\mathbb{P}^{n,(\eta)}_{\boldsymbol{\theta}^m|Y^n}\cdot\mathbb{P}^{n,(\eta)}_{M|Y^n}(m).$$

Hence, according to THEOREM 2.2.2, the self informative limit for the hyper-parameter is

$$\widehat{m} := \operatorname*{arg\,min}_{m\in G} 3m - n\sum_{j=1}^{m}(Y_j^n)^2;$$

and the self informative Bayes limit for $\boldsymbol{\theta}^M$ is the associated projection estimator.

Note that for all distinct k and m in $[\![1, G_n]\!]$, we almost surely have $E(k) - E(m) \neq 0$ since $\Upsilon(k) - \Upsilon(m)$ is a random variable with absolutely continuous distribution with respect to Lebesgue measure and hence, $\mathbb{P}_{\theta^\circ}[\{\Upsilon(k) - \Upsilon(m) = \mathrm{pen}(k) - \mathrm{pen}(m)\}] = 0$.

2.4.4 Contraction rate for the hierarchical prior

In this subsection, we discuss the contraction rate of the hierarchical Gaussian iterated posterior distribution by applying the methodology described in SECTION 2.3.2.
The results are similar to the ones obtained in JOHANNES ET AL. (2016) but extended to the iterated posterior distribution, included in the case of "$\eta = \infty$", in such a way that it offers a novel proof for optimality of the penalised contrast maximiser projection estimator.

We start by stating the set of assumptions which allow us to state our results.

Assumption 11 Suppose that λ is monotonically and polynomially decreasing, that is, there exist c in $[1, \infty[$ and a in $\mathbb{R}_+$ such that

$$\forall j \in \mathbb{N}^\star, \quad \frac{1}{c} j^{-a} \leqslant \lambda_j \leqslant c j^{-a}.$$

This assumption assures that there exist a constant $L := L(\lambda)$ in $[1, \infty[$, independent of θ° such that for any sequence $(m_n)_{n \in \mathbb{N}^\star}$

$$\sup_{n \in \mathbb{N}^\star} \frac{m_n \Lambda_{m_n}}{n \Phi_n^{m_n}} \leqslant \sup_{n \in \mathbb{N}^\star} \Lambda_{m_n} / \overline{\Lambda}_{m_n} \leqslant L.$$

Assumption 12 Let θ° and λ be such that there exists n° in $\mathbb{N}^\star$

$$0 < \kappa^\circ := \kappa^\circ(\theta^\circ, \lambda) := \inf_{n \geqslant n^\circ} \left\{ (\Phi_n^\circ)^{-1} \left[\mathfrak{b}_{m_n^\circ} \wedge \frac{m_n^\circ \overline{\Lambda}_{m_n^\circ}}{n} \right] \right\} \leqslant 1$$

Assumption 13 Let $\mathfrak{a}$ and λ be sequences such that there exists $n^\star$ in $\mathbb{N}^\star$

$$0 < \kappa^\star := \kappa^\star(\mathfrak{a}, \lambda) := \inf_{n > n^\star} \left\{ (\Phi_n^\star)^{-1} \left[\mathfrak{a}_{m_n^\star} \wedge \frac{m_n^\star \overline{\Lambda}_{m_n^\star}}{n} \right] \right\} \leqslant 1.$$

Numerical discussion 2.4.2.

□

As we have seen previously with the sieve priors, the iteration procedure conserves the contraction rate.

Corollary 2.4.3. *Under Assumption 11 and Assumption 12, if, in addition* $\log(G_n)/m_n^\circ \to 0$ *as* $n \to \infty$ *then with* $D^\circ := D^\circ(\theta^\circ, \lambda) = \lceil 5L/\kappa^\circ \rceil$ *and* $K^\circ := 10(2 \vee \|\theta^\circ\|^2) L^2 (16 \vee D^\circ \Lambda_{D^\circ})$ *we have, for any* η $(1 \leqslant \eta < \infty)$*:*

$$\lim_{n \to \infty} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}^M | Y^n}^{n, (\eta)} \left((K^\circ)^{-1} \Phi_n^\circ \leqslant \|\theta^\circ - \boldsymbol{\theta}^M\|_{l^2}^2 \leqslant K^\circ \Phi_n^\circ \right) \right] = 1.$$

Corollary 2.4.4. *Under Assumption 11 and Assumption 13, if, in addition,* $\log(G_n)/m_n^\star \to 0$ *as* $n \to \infty$ *then, for any* η $(1 \leqslant \eta < \infty)$

- *for all* θ° *in* $\Theta_{\mathfrak{a}}(r)$*, with* $D^\star := D^\star(\mathfrak{a}, \lambda) = \lceil 5L/\kappa^\star \rceil$ *and* $K^\star := 16\,(2 \vee r)\, L^2\, (16 \vee D^\star \Lambda_{D^\star})\, (1 \vee r)$*, we have*
$$\lim_{n \to \infty} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}^M | Y^n}^{n, (\eta)} \left(\|\theta^\circ - \boldsymbol{\theta}^M\|^2 \leqslant K^\star \Phi_n^\star \right) \right] = 1;$$
- *for any monotonically increasing and unbounded sequence* K_n *holds*

$$\lim_{n \to \infty} \inf_{\theta^\circ \in \Theta_{\mathfrak{a}}(r)} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}^M | Y^n}^{n, (\eta)} \left(\|\theta^\circ - \boldsymbol{\theta}^M\|^2 \leqslant K_n \Phi_n^\star \right) \right] = 1.$$

More interestingly, we were able to adapt the proofs of these results to the self informative

Bayes carrier, which gives the two following results. The proofs are displayed in the annexes.

QUESTIONS FOR JAN: Should I put the Lemmata here and only put the proofs in the annex so it is easier to make the link with the method described earlier?

THEOREM 2.4.2. *Under ASSUMPTION 11, ASSUMPTION 12 and the condition that* $\limsup_{n\to\infty} \frac{\log(G_n)}{m_n^\circ}$, *define* $D^\circ := \left\lceil \frac{3}{\kappa^\circ} + 1 \right\rceil$ *and* $K^\circ := 16L \cdot [9 \vee D^\circ \Lambda_{D^\circ}]$*; then, we have for all* θ° *in* Θ,

$$\lim_{n\to\infty} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}^M|Y^n}^{n,(\infty)} \left((K^\circ)^{-1} \Phi_n^\circ \leqslant \|\boldsymbol{\theta}^M - \theta^\circ\|^2 \leqslant K^\circ \Phi_n^\circ \right) \right] = 1.$$

NUMERICAL DISCUSSION 2.4.3.

□

QUESTIONS FOR JAN: I actually have a doubt at the very end of the proof, could we go through it again?

THEOREM 2.4.3. *Under ASSUMPTION 11, ASSUMPTION 13 and the condition that* $\limsup_{n\to\infty} \frac{\log(G_n)}{m_n^\star}$, *define* $D^\star := \left\lceil \frac{3(1\vee r)}{\kappa^\star L} + 1 \right\rceil$ *and* $K^\star := 6(1 \vee r)(9L \vee D^\star \Lambda_{D^\star})$*; then, we have for all* θ° *in* $\Theta^{\mathfrak{a}}(r)$,

$$\lim_{n\to\infty} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}^M|Y^n}^{n,(\infty)} \left(\|\boldsymbol{\theta}^M - \theta^\circ\|^2 \leqslant K^\star \Phi_n^\star \right) \right] = 1,$$

and, for any increasing sequence K_n *such that* $\lim_{n\to\infty} K_n = \infty$,

$$\lim_{n\to\infty} \inf_{\theta^\circ \in \Theta^{\mathfrak{a}}(r)} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}^M|Y^n}^{n,(\infty)} \left(\|\boldsymbol{\theta}^M - \theta^\circ\|^2 \leqslant K_n \Phi_n^\star \right) \right] = 1.$$

NUMERICAL DISCUSSION 2.4.4.

□

We have hence showed that the self informative Bayes carrier contracts around the true parameter with the oracle optimal rate of sieve priors and with minimax optimal rate over Sobolev's ellipsoids. We will see in SECTION 3.2 that the self informative limit also converges with optimal rates.

2.5 On the shape of the posterior mean

We have hence seen that in a general case, considering the asymptotic iteration , the posterior distribution using a sieve prior contracts around the projection estimator and while using a hierarchical prior, the posterior contracts around some penalised contrast maximiser projection estimator.

It is also interesting to note that for any number of iteration η, the posterior mean can be written both as a shrinkage and as an aggregation estimator. Indeed, we have

$$\begin{aligned}
\mathbb{E}^{\eta}_{\boldsymbol{\theta}^{M}|Y^n}\left[\boldsymbol{\theta}^{M}\right] &= \mathbb{E}^{\eta}_{\boldsymbol{\theta}^{M}|Y^n}\left[\sum_{m\in G}\boldsymbol{\theta}^{M}\mathbb{1}_{M=m}\right] \\
&= \sum_{m\in G}\mathbb{E}^{\eta}_{\boldsymbol{\theta}^{M}|Y^n}\left[\boldsymbol{\theta}^{M}\mathbb{1}_{M=m}\right] \\
&= \sum_{m\in G}\mathbb{P}^{\eta}_{M|Y^n}(m=M)\,\mathbb{E}^{\eta}_{\boldsymbol{\theta}^{m}|Y^n}\left[\boldsymbol{\theta}^{m}\right];
\end{aligned}$$

and we see here the aggregation form of this estimator.
On the other hand, if we write the expectation of the components individually, we obtain:

$$\begin{aligned}
\mathbb{E}^{\eta}_{\boldsymbol{\theta}^{M}|Y^n}\left[\boldsymbol{\theta}^{M}_{j}\right] &= \mathbb{E}^{\eta}_{\boldsymbol{\theta}^{M}|Y^n}\left[\boldsymbol{\theta}^{m}_{j}\mathbb{1}_{M\geqslant j}\right] \\
&= \mathbb{P}^{\eta}_{M|Y^n}(M\geqslant j)\,\mathbb{E}^{\eta}_{\boldsymbol{\theta}^{m}|Y^n}\left[\boldsymbol{\theta}^{m}_{j}\right];
\end{aligned}$$

where we see the shrinkage property.

Aggregation estimates gathered a lot of interest, see for example Rigollet and Tsybakov (2007). While considering such estimators, the goal is to reach the convergence rate of the best estimator contributing to the aggregation.

In the next chapter, we hence investigate the properties of this estimator both in inverse Gaussian sequence space model and circular density deconvolution.

Chapter 3

Minimax and oracle optimal adaptive aggregation

We inquire in this chapter the properties of aggregation estimators as introduced in section 2.5. We introduce first a skim of proof for oracle and minimax optimality of this kind of estimator before applying it to the inverse Gaussian sequence space and the circular deconvolution models respectively introduced in section 1.4 and section 1.5, including in presence of dependance and partially known operator.

3.1 Strategy of proof for optimality of aggregation estimator

As stated in section 2.5, the posterior mean obtained while using a hierarchical sieve prior has the shape

$$
\begin{aligned}
\omega_m^{(\eta)} &:= \mathbb{P}_{M|Y}^{\eta}(M = m) \\
&= \frac{\exp\left[-\eta\left(\mathrm{pen}(m) + \Upsilon(y, m)\right)\right]}{\sum_{j \subset G} \exp\left[-\eta\left(\mathrm{pen}(j) + \Upsilon(y, j)\right)\right]} \mathbb{1}_{m \subset G} \\
&= \frac{\mathbb{1}_{m \subset G}}{\sum_{j \subset G} \exp\left[-\eta\left\{\left(\Upsilon(y, j) - \Upsilon(y, m)\right) + \left(\mathrm{pen}(j) - \mathrm{pen}(m)\right)\right\}\right]}; \\
\widetilde{\theta}_j^{(\eta)} &= \mathbb{E}_{\boldsymbol{\theta}^j|Y^n}^{(\eta)}\left[\boldsymbol{\theta}^j\right]_j; \\
\widehat{\theta}^{\eta} &:= \mathbb{E}_{\boldsymbol{\theta}^M|Y^n}^{\eta}\left[\boldsymbol{\theta}^M\right] \\
&= \sum_{m \in G} \omega_m^{(\eta)} \, \mathbb{E}_{\boldsymbol{\theta}^m|Y^n}^{\eta}\left[\boldsymbol{\theta}^m\right] \\
&= \left(\sum_{m : j \in m} \omega_m^{(\eta)} \widetilde{\theta}_j^{(\eta)}\right)_{j \in \mathbb{F}}.
\end{aligned}
$$

Hence, we see that, for each model, by providing a sequence of estimators $\left(\widetilde{\theta}_j^{(\eta)}\right)_{j \in \mathbb{F}}$; and two functions pen and Υ; we define a family of adaptive estimators indexed by η.

We are here interested in the convergence properties of estimators of this kind, in particular their potential oracle and minimax optimality.

We have the following shape for the risk:

$$\begin{aligned}
\mathbb{E}^n_{\theta^\circ}\left[\left\|\widehat{\theta}^\eta - \theta^\circ\right\|^2\right] &= \mathbb{E}^n_{\theta^\circ}\left[\sum_{j\in\mathbb{F}}\left|\widehat{\theta}^\eta - \theta^\circ\right|^2\right] \\
&= \mathbb{E}^n_{\theta^\circ}\left[\sum_{j\in\mathbb{F}}\left|\omega_j\widetilde{\theta}^{(\eta)}_j - \theta^\circ_j\right|^2\right] \\
&= \mathbb{E}^n_{\theta^\circ}\left[\sum_{j\in G_n}\left|\omega_j\widetilde{\theta}^{(\eta)}_j - \theta^\circ_j\right|^2\right] + \sum_{j\notin G_n}|\theta^\circ_j|^2 \\
&= \mathbb{E}^n_{\theta^\circ}\left[\sum_{j\in G_n}\left|\omega_j\left(\widetilde{\theta}^{(\eta)}_j - \theta^\circ_j\right) - (1-\omega_j)\theta^\circ_j\right|^2\right] + \sum_{j\notin G_n}|\theta^\circ_j|^2 \\
&\leqslant \underbrace{\sum_{j\in G_n}\mathbb{E}^n_{\theta^\circ}\left[\left|\omega_j\left(\widetilde{\theta}^{(\eta)}_j - \theta^\circ_j\right)\right|^2\right]}_{=:A} + \underbrace{\sum_{j\in G_n}\mathbb{E}^n_{\theta^\circ}\left[\left|(1-\omega_j)\theta^\circ_j\right|^2\right]}_{=:B} + \underbrace{\sum_{j\notin G_n}|\theta^\circ_j|^2}_{=:C}
\end{aligned}$$

$$A = \sum_{j\in G_n}\mathbb{E}^n_{\theta^\circ}\left[\omega_j^2\left|\widetilde{\theta}^{(\eta)}_j - \theta^\circ_j\right|^2\right]$$

3.2 Inverse Gaussian sequence space model

This result states that the self informative Bayes carrier contracts with oracle optimal rate of the sieve priors under our set of assumptions.

Theorem 3.2.1. *Consider $\overline{\theta}^{\widehat{m}}$ the frequentist estimator given by the self-informative limit. Under Assumption 11, Assumption 12 and the condition that* $\limsup_{n\to\infty}\frac{\log\left(\frac{G_n^2}{\Phi_n^\circ}\right)}{m_n^\circ} \leqslant \frac{5}{9L}$, *we have*

$$\exists C^\circ \in \mathbb{R}^\star_+ : \forall \theta^\circ \in \Theta, \quad \mathbb{E}^n_{\theta^\circ}\left[\|\overline{\theta}^{\widehat{m}} - \theta^\circ\|^2\right] \leqslant C^\circ \Phi^\circ_n.$$

We obtain here optimality results both for the self informative limit and self informative Bayes carrier.

Theorem 3.2.2. *Consider $\overline{\theta}^{\widehat{m}}$ the frequentist estimator given by the self-informative limit. Then, under Assumption 11, Assumption 13 and the condition that* $\limsup_{n\to\infty}\frac{\log\left(\frac{G_n^2}{\Phi_n^\star}\right)}{m_n^\star} < \frac{5}{9L}$, *we have*

$$\exists C^\star \in \mathbb{R}^\star_+ : \quad \sup_{\theta^\circ\in\Theta}\mathbb{E}^n_{\theta^\circ}\left[\|\overline{\theta}^{\widehat{m}} - \theta^\circ\|^2\right] \leqslant C^\star \Psi^\star_n.$$

3.3 Inverse Gaussian sequence space model with partially known operator

3.4 Circular deconvolution with independent data and known noise density

As stated in CHAPTER 2, the non-conjugated nature of the hierarchical Gaussian sieve in the context of circular deconvolution does not allow to compute the posterior mean analytically. However, in this part we mimic the form of this posterior mean and construct an estimator from this idea.
We start by reminding the definition of the projection estimators, which we will use to surrogate the posterior mean of sieve priors, which appear in the structure of the posterior mean of hierarchical sieves.

REMINDER. PROJECTION ESTIMATORS
We recall the notation for the projection estimators, for any m in $\mathbb{Z}$, we have

$$\begin{aligned}\overline{\theta}_m &:= \mathbb{1}_{m=0} + \mathbb{1}_{m\neq 0}\frac{1}{n}\sum_{p=1}^{n}\frac{e_m(Y_p)}{\lambda_m};\\ \left(\overline{\theta}_j^m\right)_{j\in\mathbb{Z}} &:= \left(\mathbb{1}_{|j|\leqslant m}\overline{\theta}_j\right)_{j\in\mathbb{Z}}.\end{aligned}$$

As in the posterior mean of hierarchical sieves, we define a weight sequence, corresponding to the posterior distribution of the threshold parameter.

DEFINITION 30 WEIGHT SEQUENCE
Let be the following quantities:

$$\begin{aligned}\kappa &:= \frac{23}{2};\\ \psi_m^\Lambda &:= \frac{\log\left(m\Lambda_m \vee (m+2)\right)^2}{\log\left(m+2\right)^2};\\ \Delta_m^\Lambda &:= m\Lambda_m\psi_m^\Lambda;\\ \mathrm{pen}(m) &:= \frac{9}{2}\cdot 12\cdot\kappa\cdot\Delta_m^\Lambda;\\ \Upsilon(Y,m) &:= n\left\|\overline{\theta}^m\right\|_{l^2}^2.\end{aligned}$$

Then, for any couple of natural integers n and η, we define the distribution $\mathbb{P}^{n,(\eta)}_{M|Y^n}$, dominated by the counting measure on $\mathbb{N}^\star$ such that, for any m in $[\![1,n]\!]$

$$\mathbb{P}^{n,(\eta)}_{M|Y^n}(m) := \frac{\exp\left[\eta\left(-\mathrm{pen}(m)+\Upsilon(m,Y^n)\right)\right]}{\sum_{k=1}^{n}\exp\left[\eta\left(-\mathrm{pen}(k)+\Upsilon(k,Y^n)\right)\right]}.$$

With those definitions at hand, we are able to define an estimator that reproduces the structure of the posterior mean of iterated hierarchical sieves.

Definition 31 Aggregation/shrinkage estimator
Using the notations we just introduced, we define, for any strictly positive integer η the shrinkage/aggregation estimator $\widehat{\theta}^{(\eta)}$ such that, for any j in $\mathbb{Z}$

$$\begin{aligned}\widehat{\theta}_j^{(\eta)} &:= \mathbb{P}^{n,(\eta)}_{M|Y^n}(\llbracket |j|, n\rrbracket)\overline{\theta}_j;\\ \widehat{\theta}^{\eta} &:= \sum_{j=1}^{n} \mathbb{P}^{n,(\eta)}_{M|Y^n}(j)\overline{\theta}^j.\end{aligned}$$

As previously, one can notice that, as η tends to infinity, this estimator converges to the penalised contrast maximiser projection estimator with penalty function pen and contrast Υ.
Using the method described in section 3.1, we are able to show that, for any θ°, the sequence defined hereafter is a convergence rate.

Definition 32 Convergence rate
Let be the sequences:

$$m_n^\dagger := \arg\min_{m\in\mathbb{N}}\left\{\left[\mathfrak{b}_m^2(\theta^\circ)\mathfrak{b}_0^{-2}(\theta^\circ) \vee 2\frac{m\Lambda_{(m)}}{n}\psi_n\right]\right\};$$

and

$$\Phi_n^\dagger := \left[\mathfrak{b}^2_{m_n^\dagger}(\theta^\circ)\mathfrak{b}_0^{-2}(\theta^\circ) \vee 2\frac{m_n^\dagger\Lambda_{(m_n^\dagger)}}{n}\psi_n\right].$$

Definition 33 Accelerated convergence rate
Let be the sequences:

$$m_n^\circ := \arg\min_{m\in\mathbb{N}}\left\{\left[\mathfrak{b}_m^2(\theta^\circ)\mathfrak{b}_0^{-2}(\theta^\circ) \vee 2\frac{m\overline{\Lambda}_m}{n}\psi_n\right]\right\};$$

and

$$\Phi_n^\circ := \left[\mathfrak{b}^2_{m_n^\circ}(\theta^\circ)\mathfrak{b}_0^{-2}(\theta^\circ) \vee 2\frac{m_n^\circ\overline{\Lambda}_{m_n^\circ}}{n}\psi_n\right].$$

Assumption 14 Let θ° have a finite series expansion as definied in **(p)**, that is, either (a) $\theta^\circ = (\mathbb{1}_{j=0})_{j\in\mathbb{Z}}$, i.e., $\mathfrak{b}_0(\theta^\circ) = \|\Pi_{\mathbb{U}_0}^\perp \theta^\circ\|_{\ell^2}^2 = 0$ or (b) there is $K \in \mathbb{N}$ with $1 \geqslant \mathfrak{b}_{K-1}(\theta^\circ) > 0$ and $\mathfrak{b}_K(\theta^\circ) = 0$. In case (a) set $n_{\theta^\circ,\Lambda} := \lceil 15(\frac{300}{\sqrt{\kappa}})^4\rceil$ while in case (b) given $K_\phi := K \vee 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ and $c_{\theta^\circ} := \frac{2\|\Pi_{\mathbb{U}_0}^\perp \theta^\circ\|_{\ell^2}^2 + 484\kappa}{\|\Pi_{\mathbb{U}_0}^\perp \theta^\circ\|_{\ell^2}^2 \mathfrak{b}_{K-1}^2(\theta^\circ)}$ let there $n_{\theta^\circ,\Lambda} \in \mathbb{N}$ be with $n_{\theta^\circ,\Lambda} > \lceil c_{\theta^\circ}\Delta^\Lambda_{K_\phi} \vee 15(\frac{300}{\sqrt{\kappa}})^4\rceil$ such that $m_n^\bullet := \max\{m \in \llbracket K, n\rrbracket : c_{\theta^\circ}\Delta_m^\Lambda < n\}$ where the defining set contains K_ϕ and thus it is not empty, satisfies $\psi^\Lambda_{m_n^\bullet} m_n^\bullet \geqslant K_\phi(\log n)$ for all $n \geqslant n_{\theta^\circ,\Lambda}$.

Remark 3.4.1. *Keep in mind that* $(\mathfrak{b}_m(\theta^\circ))_{m\in\mathbb{N}} \subset [0,1]$ *is monotonically non increasing with* $\mathfrak{b}_1(\theta^\circ) \leqslant 1$ *and* $\lim_{m\to\infty}\mathfrak{b}_m(\theta^\circ) = 0$. *Thereby, in case* (b) *of* Assumption 14 *holds*

$\|\Pi_{\mathbb{U}_0}^{\perp}\theta^\circ\|_{\ell^2}^2 > 0$ *and* $\frac{2\|\Pi_{\mathbb{U}_0}^{\perp}\theta^\circ\|_{\ell^2}^2+484\kappa}{\|\Pi_{\mathbb{U}_0}^{\perp}\theta^\circ\|_{\ell^2}^2\mathfrak{b}_{K-1}^2(\theta^\circ)} \geqslant 1$. *We shall stress that in* Lemma J.2.1 *and* Lemma J.2.2 *in the* appendix G *we derive upper bounds for the partially data-driven aggregated OSE featuring a deterioration of the upper bound which, due to* Assumption 14 *is avoided in the next assertion.*

Numerical discussion 3.4.1.
Let us illustrate Assumption 14 considering as in Num. discussion 1.5.1 the commonly studied behaviours **(o)** and **(s)** for the sequence $(\Lambda_j)_{j\in\mathbb{N}}$.

- **(o)** Let $\Lambda_m \sim m^{2a}$, $a > 0$, then we have $\psi_m^\Lambda \sim 1$, $\Lambda_{(m)} \sim \overline{\Lambda}_m \sim m^{2a}$, $\Delta_m^\Lambda = \psi_m^\Lambda m\Lambda_{(m)} \sim m^{2a+1}$ and hence $1 \sim \Delta_{m_n^\bullet}^\Lambda n^{-1} \sim (m_n^\bullet)^{2a+1}n^{-1}$ implies $m_n^\bullet \sim n^{1/(2a+1)}$ and $m_n^\bullet\psi_{m_n^\bullet}^\Lambda \sim n^{1/(2a+1)}$.
- **(s)** Let $\Lambda_m \sim \exp(m^{2a})$, $a > 0$, then we have $\psi_m^\Lambda \sim (m^{2a})^2$, $\Delta_m^\Lambda = m\psi_m^\Lambda\Lambda_{(m)} \sim m^{1+4a}\exp(m^{2a})$ and hence $n \sim \Delta_{m_n^\bullet}^\Lambda \sim (m_n^\bullet)^{1+4a}\exp((m_n^\bullet)^{2a})$ implies $m_n^\bullet \sim (\log n - \frac{1+4a}{2a}\log\log n)^{1/(2a)}$ and $m_n^\bullet\psi_{m_n^\bullet}^\Lambda \sim (\log n)^{2+1/(2a)}$.

Clearly, in both cases **(o)** and **(s)**, there is $n_{\theta^\circ,\Lambda} \in \mathbb{N}$ such that $\psi_{m_n^\bullet}^\Lambda m_n^\bullet \geqslant K_\phi(\log n)$ for all $n \geqslant n_{\theta^\circ,\Lambda}$ holds true.

More precisely, we obtain the following theorem, for which the proof is given in appendix G.

Theorem 3.4.1. *Let θ° have a finite series expansion as defined in* (p)*. Under* Assumption 14 *there is a finite numerical constant $\mathcal{C}$ such that for all $n \in \mathbb{N}$,*

$$\mathbb{E}_{\theta^\circ}^n\left[\|\widetilde{f} - f\|_{\ell^2}^2\right] \leqslant \mathcal{C}\{\Delta_{n_{\theta^\circ,\Lambda}}^\Lambda + \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 n_{f,\Lambda} + \|\phi\|_{\ell^1}^2\}n^{-1}. \tag{3.1}$$

Numerical discussion 3.4.2.
Let us illustrate Theorem 3.4.1 considering as in Num. discussion 3.4.1 the behaviours **(o)** and **(s)** for the sequence $(\Lambda_j)_{j\in\mathbb{N}}$. Keeping in mind that as shown in Num. discussion 3.4.1 there is $n_{f,\Lambda} \in \mathbb{N}$ such that $\psi_{m_n^\bullet}^\Lambda m_n^\bullet \geqslant K_h(\log n)$ for all $n \geqslant n_{f,\Lambda}$ holds true, due to Lemma J.2.2 there is a constant $\mathcal{C}_{f,g}$ depending only on the densities f and g such that $\mathbb{E}_Y^n\|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}_{f,g}n^{-1}$ for all $n \in \mathbb{N}$. Comparing the last result with the oracle rate derived in **[p-o]** and **[s-o]** in Num. discussion 1.5.1 we conclude, that $\widetilde{f}$ is optimal in an oracle sense in both cases **[p-o]** and **[s-o]**.

Assumption 15 Let f have an infinite series expansion as definied in (np), that is, $1 \geqslant \mathfrak{b}_m(f) > 0$ for all $m \in \mathbb{N}$. Given $m_h := 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ and $\widetilde{m}_h = \min\{m \in \mathbb{N} : \mathfrak{b}_{m_h}(f) > \mathfrak{b}_m(f)\}$ there is $n_{f,\Lambda} \in \mathbb{N}$ with $n_{f,\Lambda} \geqslant \lceil \frac{\Delta_{\widetilde{m}_h}^\Lambda}{\mathfrak{b}_{\widetilde{m}_h}^2(f)} \vee 15(\frac{300}{\sqrt{\kappa}})^4\rceil$ such that either (a) $\psi_{m_n^\diamond}^\Lambda m_n^\diamond \geqslant m_h|\log\mathcal{R}_n^\diamond(f,\Lambda)|$ for all $n \geqslant n_{f,\Lambda}$ or (b) $m_n^\diamond \leqslant m_h|\log\mathcal{R}_n^\diamond(f,\Lambda)|$ for all $n \geqslant n_{f,\Lambda}$. We set in case (a) $m_n^\bullet := m_n^\diamond$ and in case (b) $m_n^\bullet := m_h|\log\mathcal{R}_n^\diamond(f,\Lambda)|$.

Remark 3.4.2. *Considering $m_h := 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ and $\widetilde{m}_h = \min\{m \in \mathbb{N} : \mathfrak{b}_{m_h}(f) > \mathfrak{b}_m(f)\}$ as defined in* Assumption 15 *the defining set is not empty since $\mathfrak{b}_m(f) > 0$ for all $m \in \mathbb{N}$*

and $\lim_{m\to\infty}\mathfrak{b}_m(f) = 0$. *Moreover, it holds* $\widetilde{m}_h > m_h$ *due to the the monotonicity of* $\mathfrak{b}_m(f)$. *Noting that* $\lceil \frac{\Delta^\Lambda_{\widetilde{m}_h}}{\mathfrak{b}^2_{\widetilde{m}_h}(f)} \vee 15(\frac{300}{\sqrt{\kappa}})^4\rceil \geqslant \Delta^\Lambda_{\widetilde{m}_h} \geqslant \widetilde{m}_h$ *by construction, for all* $n \geqslant n_{f,\Lambda}$ *as in* Assumption 15 *holds* $\mathcal{R}^\circ_n[m_h, f, \Lambda] \geqslant \mathfrak{b}^2_{m_h}(f) > \mathfrak{b}^2_{\widetilde{m}_h}(f) = \mathcal{R}^\circ_n[\widetilde{m}_h, f, \Lambda]$ *and hence, for all* $n \geqslant n_{f,\Lambda}$ *we have* $m^\diamond_n > m_h$. *We use these preliminary findings in the proof of* Theorem 3.4.2.

Numerical discussion 3.4.3.
Let us illustrate Assumption 15 considering as in Num. discussion 1.5.1 usual behaviour **[o-o]**, **[s-o]** and **[o-s]** for the sequences $(\mathfrak{b}_m(f))_{m\in\mathbb{N}}$ and $(\Lambda_m)_{m\in\mathbb{N}}$:

- **[o-o]** Since $\mathfrak{b}^2_m(f) \sim m^{-2p}$ and $\Delta^\Lambda_m \sim m^{2a+1}$ (cf. Num. discussion 3.4.1 **(o)**) follows $\mathcal{R}^\diamond_n(f,\Lambda) \sim (m^\diamond_n)^{-2p} \sim \Delta^\Lambda_{m^\diamond_n} n^{-1} \sim (m^\diamond_n)^{2a+1}n^{-1}$ which implies $m^\diamond_n \sim n^{1/(2p+2a+1)}$, $\psi^\Lambda_{m^\diamond_n} m^\diamond_n \sim n^{1/(2p+2a+1)}$, $\mathcal{R}^\diamond_n(f,\Lambda) \sim n^{-2p/(2p+2a+1)}$ and $|\log \mathcal{R}^\diamond_n(f,\Lambda)| \sim (\log n)$.
- **[o-s]** Since $\mathfrak{b}^2_m(f) \sim m^{-2p}$ and $\Delta^\Lambda_m \sim m^{1+4a}\exp(m^{2a})$ (cf. Num. discussion 3.4.1 **(s)**) follows $\mathcal{R}^\diamond_n(f,\Lambda) \sim (m^\diamond_n)^{-2p} \sim \Delta^\Lambda_{m^\diamond_n} n^{-1} \sim (m^\diamond_n)^{1+4a}\exp((m^\diamond_n)^{2a})$ which implies $m^\diamond_n \sim (\log n)^{1/(2a)}$, $\psi^\Lambda_{m^\diamond_n} m^\diamond_n \sim (\log n)^{2+1/(2a)}$, $\mathcal{R}^\diamond_n(f,\Lambda) \sim (\log n)^{-p/a}$ and $|\log \mathcal{R}^\diamond_n(f,\Lambda)| \sim (\log\log n)$.
- **[s-o]** Since $\mathfrak{b}^2_m(f) \sim \exp(-m^{2p})$ and $\Delta^\Lambda_m \sim m^{2a+1}$ (cf. Num. discussion 3.4.1 **(o)**) follows $\mathcal{R}^\diamond_n(f,\Lambda) \sim \exp(-(m^\diamond_n)^{2p}) \sim \Delta^\Lambda_{m^\diamond_n} n^{-1} \sim (m^\diamond_n)^{2a+1}n^{-1}$ which implies $m^\diamond_n \sim (\log n)^{1/(2p)}$, $\psi^\Lambda_{m^\diamond_n} m^\diamond_n \sim (\log n)^{1/(2p)}$, $\mathcal{R}^\diamond_n(f,\Lambda) \sim (\log n)^{(2a+1)/(2p)}n^{-1}$ and $|\log \mathcal{R}^\diamond_n(f,\Lambda)| \sim (\log n)$.

Clearly, there is $n_{f,\Lambda} \in \mathbb{N}$ such that for all $n \geqslant n_{f,\Lambda}$ in the cases **[o-o]** and **[o-s]** $\psi^\Lambda_{m^\diamond_n} m^\diamond_n \geqslant m_h|\log \mathcal{R}^\diamond_n(f,\Lambda)|$, i.e., Assumption 15 (a) holds, while in case **[s-o]** $m^\diamond_n \leqslant m_h|\log \mathcal{R}^\diamond_n(f,\Lambda)|$ for $p \geqslant 1/2$, i.e., Assumption 15 (b) holds, and $\psi^\Lambda_{m^\diamond_n} m^\diamond_n \geqslant m_h|\log \mathcal{R}^\diamond_n(f,\Lambda)|$ for $p < 1/2$, i.e., Assumption 15 (a) holds.

Theorem 3.4.2. *Let be the constants* $K := \frac{\sqrt{2}-1}{21\sqrt{2}}$, *and* $C_{\lambda,\theta^\circ} \geqslant \sum_{j=1}^{\infty} \exp\left[-\eta\frac{\psi_n m\overline{\Lambda}_m}{2}\right]$ *then, for any* n *and* η *integers greater than* 1, *we have*

$$\mathbb{E}^n_{\theta^\circ}\left[\|\widehat{\theta}^{(\eta)} - \theta^\circ\|^2_{l^2}\right] \leqslant 174\mathfrak{b}^2_0\Phi^\dagger_n + \frac{1}{n}32C_{\lambda,\theta}\exp\left[-K\left(\frac{\psi_n 2m^\circ_n}{\|\theta^\circ\|^2_{l^2}\|\lambda\|^2_{l^2}} \wedge \sqrt{n\psi_n}\right) + \log(\psi_n) + 2\log\left(\Lambda_n \vee n\right)\right];$$

assuming $m^\dagger_n$ *tends to* ∞, *this gives*

$$\mathbb{E}^n_{\theta^\circ}\left[\|\widehat{\theta}^{(\eta)} - \theta^\circ\|^2_{l^2}\right] \in \mathcal{O}_n\left(\Phi^\dagger_n\right).$$

Comparison with the oracle rate of projection estimators reveals that in many cases, we obtain an oracle optimal estimator.

Numerical discussion 3.4.4.
Assume $m^\dagger_n$ tends to infinity and let be two positive real numbers p and a.

If $\mathfrak{b}_m^2 \asymp_{m\to\infty} m^{-2p}$ and $\Lambda_m \asymp_{m\to\infty} m^{2a}$, then we have $\psi_n \asymp_{n\to\infty} 4a$; $m_n^\dagger \asymp_{n\to\infty} n^{\frac{1}{2a+2p+1}}$; and $\Phi_n^\dagger \asymp_{n\to\infty} n^{-\frac{2p}{2a+2p+1}}$ and we have

$$\begin{aligned}\mathbb{E}_{\theta^\circ}^n\left[\left\|\widehat{\theta}^{(\eta)}-\theta^\circ\right\|_{l^2}^2\right] \in &\mathcal{O}_n\left(n^{-\frac{2p}{2a+2p+1}}+\right.\\ &\left.\frac{1}{n}C_{\lambda,\theta^\circ}\exp\left[-K\left(\frac{8an^{\frac{1}{2a+2p+1}}}{\|\lambda\|_{l^2}\|\theta^\circ\|_{l_2}}\wedge\sqrt{4an}\right)+4a\log(n)+2n^{2a}\right]\right)\\ \in&\mathcal{O}_n(n^{-\frac{2p}{2a+2p+1}})\end{aligned}$$

On the other hand, if $\Lambda_m \asymp_{m\to\infty} \exp\left[m^{2a}\right]$, then we have $\psi_n \asymp_{n\to\infty} \frac{n^{4a}}{\log(n)^2}$; $m_n^\dagger \asymp_{n\to\infty} \log(n)^{\frac{1}{2a}}$; and $\Phi_n^\dagger \asymp_{n\to\infty} \log(n)^{-\frac{p}{a}}$. Which leads to

$$\begin{aligned}\mathbb{E}_{\theta^\circ}^n\left[\left\|\widehat{\theta}^{(\eta)}-\theta^\circ\right\|_{l^2}^2\right] \in &\mathcal{O}_n\left(\log(n)^{-\frac{p}{a}}+\right.\\ &\left.\frac{1}{n}C_{\lambda,\theta^\circ}\exp\left[-K\left(\frac{n^{4a}2\log(n)^{\frac{1-4a}{2a}}}{\|\lambda\|_{l^2}\|\theta^\circ\|_{l_2}}\wedge\frac{n^{2a+\frac{1}{2}}}{\log(n)}\right)+4a\log(n)+2n^{2a}\right]\right)\\ \in&\mathcal{O}_n(\log(n)^{-\frac{p}{a}}).\end{aligned}$$

Alternatively, under stronger assumptions, we obtain the following theorem, for which the proof is given in appendix I.

Theorem 3.4.3. *Let be the constants $K := \frac{\sqrt{2}-1}{21\sqrt{2}}$, and $C_{\lambda,\theta^\circ} \geqslant \sum_{j=1}^{\infty}\exp\left[-\eta\frac{\psi_n m\overline{\Lambda}_m}{2}\right]$ then, for any n greater than 1 and η greater than $118K$, we have*

$$\begin{aligned}\mathbb{E}_{\theta^\circ}^n\left[\|\widehat{\theta}^{(\eta)}-\theta^\circ\|_{l^2}^2\right] \leqslant &174\mathfrak{b}_0^2\Phi_n^\circ+\\ &\frac{1}{n}32C_{\lambda,\theta}\exp\left[-K\left(\frac{\psi_n 2m_n^\circ\overline{\Lambda}_{m_n^\circ}}{\Lambda_{(m_n^\circ)}\|\theta^\circ\|_{l^2}^2\|\lambda\|_{l^2}^2}\wedge\sqrt{n\psi_n}\right)+\log(\psi_n)+2\log\left(\Lambda_{(n)}\vee n\right)\right];\end{aligned}$$

assuming m_n° tends to ∞, this gives

$$\mathbb{E}_{\theta^\circ}^n\left[\|\widehat{\theta}^{(\eta)}-\theta^\circ\|_{l^2}^2\right] \in \mathcal{O}_n\left(\Phi_n^\dagger\right).$$

3.5 Circular deconvolution with beta mixing data and known noise density

Considering the positive results obtained in the previous section, we are now interested in generalising those results to the situation where our observations are not a sequence of independent identically distributed variables anymore but may suffer from dependence.

Assumption 16 Strictly stationary, absolutely regular process
We assume in this section that the process of observations $(Y_p)_{p\in\mathbb{Z}}$ is strictly stationary and absolutely regular as described in appendix B.

Assumption 17 Rich space
We assume in this section that the process of observations $(Y_p)_{p\in\mathbb{Z}}$ is strictly stationary and absolutely regular as described in appendix B.

Assumption 18 Assume that, for any integer p, the joint distribution $\mathbb{P}_{Y_0,Y_p|\theta^\circ}$ of Y_0 and Y_p admits a density f_{Y_0,Y_p} which is square integrable.
Let $\|f_{Y_0,Y_p}\|^2_{L^2} := \int_0^1\int_0^1 |f_{Y_0,Y_p}(x,y)|^2 dx\,dy < \infty$ with a slight abuse of notations. If we denote further by $h\otimes g : [0,1]^2 \to \mathbb{R}$ the bivariate function $[h\otimes g](x,y) := h(x)g(y)$ then let assume $\gamma_f := \sup_{p\geqslant 1} \|f_{Y_0^n,Y_p^n|\theta^\circ} - f_{Y_0^n|\theta^\circ} \otimes f_{Y_0^n|\theta^\circ}\|_{L^2} < \infty$.
Assume in addition $\sum_{p=1}^{\infty} \beta(Y_0,Y_p) < \infty$ and $\gamma := \sup_{\theta^\circ\in\Theta(\mathfrak{a},r)} \gamma_\theta < \infty$

As in the posterior mean of hierarchical sieves, we define a weight sequence, corresponding to the posterior distribution of the threshold parameter.

Definition 34 Weight sequence

With those definitions at hand, we are able to define an estimator that reproduces the structure of the posterior mean of iterated hierarchical sieves.

Definition 35 Aggregation/shrinkage estimator
Using the notations we just introduced, we define, for any strictly positive integer η the shrinkage/aggregation estimator $\widehat{\theta}^{(\eta)}$ such that, for any j in $\mathbb{Z}$

$$\begin{aligned}\widehat{\theta}^{(\eta)}_j &:= \mathbb{P}^{n,(\eta)}_{M|Y^n}([\![|j|,n]\!])\overline{\theta}_j;\\ \widehat{\theta}^{\eta} &:= \sum_{j=1}^{n} \mathbb{P}^{n,(\eta)}_{M|Y^n}(j)\overline{\theta}^j.\end{aligned}$$

As previously, one can notice that, as η tends to infinity, this estimator converges to the penalised contrast maximiser projection estimator with penalty function pen and contrast Υ.

Using the method described in section 3.1, we are able to show that, for any θ°, the sequence defined hereafter is a convergence rate.

Definition 36 Convergence rate

More precisely, we obtain the following theorem, for which the proof is given in appendix J.

Theorem 3.5.1.

Comparison with the oracle rate of projection estimators reveals that in many cases, we obtain an oracle optimal estimator.

Numerical discussion 3.5.1.

3.6 Circular deconvolution with independent data and partially known noise density

Finally, in this section, we consider the case of a partially known operator as described in SECTION 1.1.4.

As a consequence, we need here to simultaneously estimate the distribution $\mathbb{P}_\varepsilon$ of the noise random variable ε and the density of interest f^X. We now assume that we have at hand two independent samples. The first is an i.i.d. sample from $\mathbb{P}^\varepsilon$, denoted $\varepsilon^q = (\varepsilon_r)_{r \in [\![1,q]\!]}$; the second is, as in the known operator case, a sample from the convolved distribution, assumed here to be i.i.d., denoted $Y^n = (Y_p)_{p \in [\![1,n]\!]}$.

We want to adapt our aggregation estimator shape to this case. In this perspective let us define an estimators for λ^{-1}.

DEFINITION 37 THRESHOLDED ESTIMATOR
For any m in $\mathbb{Z}$ define, with the convention "$0/0 = 0$"

$$\begin{aligned}\widehat{\lambda}_m &:= \frac{1}{q}\sum_{r=1}^{q} e_m(\varepsilon_r);\\ \widehat{\lambda}_m^+ &:= \frac{1}{\widehat{\lambda}_m}\mathbb{1}_{|\widehat{\lambda}_m|^2 > \frac{1}{q}}.\end{aligned}$$

Mimicking notations we used until here we also note, for any m in $\mathbb{N}$

$$\begin{aligned}\widehat{\Lambda}_m &:= |\widehat{\lambda}_m^+|^2\\ \widehat{\Lambda}_{(m)} &:= \max\left\{\widehat{\Lambda}_k, k \in [\![1, m]\!]\right\};\end{aligned}$$

With this definition, we define an alternative form for projection estimators which we aggregated in the two previous sections.

DEFINITION 38 THRESHOLDED PROJECTION ESTIMATORS
For any m in $\mathbb{Z}$, let be

$$\begin{aligned}\overline{\theta}_m^+ &:= \mathbb{1}_{m=0} + \mathbb{1}_{m\neq 0}\frac{1}{n}\sum_{p=1}^{n} e_m(Y_p)\widehat{\lambda}_m^+;\\ \left(\overline{\theta}_j^{m,+}\right)_{j\in\mathbb{Z}} &:= \left(\mathbb{1}_{|j|\leqslant m}\overline{\theta}_j^+\right)_{j\in\mathbb{Z}}.\end{aligned}$$

We give the following shape to the weight sequence.

DEFINITION 39 WEIGHT SEQUENCE

Let be the following quantities:

$$
\begin{aligned}
\kappa \quad & \geqslant 1; \\
\sqrt{\delta_m^{\widehat{\Lambda}}} \quad & := \frac{\log\left(k\widehat{\Lambda}_{(m)} \vee (m+2)\right)}{\log(m+2)} \\
\Delta_m^{\widehat{\Lambda}} \quad & := \delta_m^{\widehat{\Lambda}} m \widehat{\Lambda}_{(m)} \\
\operatorname{pen}(m) \quad & := \frac{9}{2} 12\kappa \Delta_m^{\widehat{\Lambda}}; \\
\Upsilon(Y, \varepsilon, m) & := n \left\| \overline{\theta}^{m,+} \right\|_{l^2}^2 .
\end{aligned}
$$

Then, for any couple of natural integers n and η, we define the distribution $\mathbb{P}^{n,(\eta)}_{M|Y^n,\varepsilon^q}$, dominated by the counting measure on $\mathbb{N}^\star$ such that, for any m in $[\![1, n]\!]$

$$
\mathbb{P}^{n,(\eta)}_{M|Y^n,\varepsilon^q}(m) := \frac{\exp\left[\eta\left(-\operatorname{pen}(m) + \Upsilon(Y^n, \varepsilon^q, m)\right)\right]}{\sum_{k=1}^{n} \exp\left[\eta\left(-\operatorname{pen}(k) + \Upsilon(Y^n, \varepsilon^q, k)\right)\right]}.
$$

With those definitions at hand, we are able to define an estimator that reproduces the structure of the posterior mean of iterated hierarchical sieves.

Definition 40 Aggregation/shrinkage estimator

Using the notations we just introduced, we define, for any strictly positive integer η the shrinkage/aggregation estimator $\widehat{\theta}^{(\eta)}$ such that, for any j in $\mathbb{Z}$

$$
\begin{aligned}
\widehat{\theta}_j^{(\eta)} & := \mathbb{P}^{n,(\eta)}_{M|Y^n,\varepsilon^q}([\![|j|, n]\!])\overline{\theta}_j^{+}; \\
\widehat{\theta}^{\eta} \quad & := \sum_{j=1}^{n} \mathbb{P}^{n,(\eta)}_{M|Y^n,\varepsilon^q}(j)\overline{\theta}^{j,+}.
\end{aligned}
$$

As previously, one can notice that, as η tends to infinity, this estimator converges to the penalised contrast maximiser projection estimator with penalty function pen and contrast Υ.

In addition, this time, it is important to note that this estimator does not depend on characteristics of λ nor θ° and is hence fully data driven and well designed for the context of a partially unknown operator problem.

Using the method described in section 3.1, we are able to show that, for any θ°, the sequence defined hereafter is a convergence rate.

Definition 41 Convergence rate

Let be the sequences:

$$
m_n^\dagger := \underset{m \in \mathbb{N}}{\arg\min} \left\{ \left[\mathfrak{b}_m^2(\theta^\circ)\mathfrak{b}_0^{-2}(\theta^\circ) \vee 2\frac{m\Lambda_{(m)}}{n}\psi_n \right] \right\};
$$

and

$$\Phi_n^\dagger := \left\lceil \mathfrak{b}^2_{m_n^\dagger}(\theta^\circ)\mathfrak{b}_0^{-2}(\theta^\circ) \vee 2\frac{m_n^\dagger \Lambda_{(m_n^\dagger)}}{n}\psi_n \right\rceil.$$

ASSUMPTION 19 Let f have a finite series expansion as defined in **(p)**, that is, either (a) $f = e_0$, i.e., $\mathfrak{b}_0(f) = \|\Pi_{\mathbb{U}_0}^\perp f\|^2_{L^2} = 0$ or (b) there is $K \in \mathbb{N}$ with $1 \geqslant \mathfrak{b}_{[K-1]}(f) > 0$ and $\mathfrak{b}_K(f) = 0$. In case (a) set $K_h := \lceil 15\frac{300^4}{\kappa^2} \vee 3\frac{800^2}{\kappa^2}\rceil$ while in case (b) given $K_h := K \vee 3\frac{800^2\|\phi\|^2_{\ell^1}}{\kappa^2}$ and $c_f := \frac{2\|\Pi_{\mathbb{U}_0}^\perp f\|^2_{L^2}+7576\kappa}{\|\Pi_{\mathbb{U}_0}^\perp f\|^2_{L^2}\mathfrak{b}^2_{[K-1]}(f)}$ let there $n_{f,\Lambda}, q_{f,\Lambda} \in \mathbb{N}$ be with $n_{f,\Lambda} > \lceil c_f \Delta^\Lambda_{K_h} \vee 15\frac{300^4}{\kappa^2}\rceil$ and $q_{f,\Lambda} > \lceil 289\log(K_h+2)\psi^\Lambda_{K_h}\Lambda_{(K_h)}\rceil$ such that $m_n^\bullet := \max\{m \in [\![K, n]\!] : c_f\,\Delta^\Lambda_m < n\}$ and $m_q^\bullet := \max\{m \in [\![K_h, q]\!] : 289\log(m+2)\psi^\Lambda_m\Lambda_{(m)} \leqslant q\}$ where the defining sets contain K_h and thus they are not empty, satisfies $\psi^\Lambda_{m_n^\bullet} m_n^\bullet \geqslant K_h(\log n)$ for all $n \geqslant n_{f,\Lambda}$ and $\psi^\Lambda_{m_q^\bullet} m_q^\bullet \geqslant K_h(\log q)$ for all $q \geqslant q_{f,\Lambda}$, respectively.

NUMERICAL DISCUSSION 3.6.1.
Let us illustrate ASSUMPTION 19 considering as in NUM. DISCUSSION 1.5.1 the commonly studied behaviours **(o)** and **(s)** for the sequence $(\Lambda_j)_{j\in\mathbb{N}}$.

(o) Let $\Lambda_m \sim m^{2a}$, $a > 0$, then $m_n^\bullet \psi^\Lambda_{m_n^\bullet} \sim n^{1/(2a+1)}$ (cf. NUM. DISCUSSION 3.4.1 **(o)**), while $q \sim (\log m_q^\bullet)\psi^\Lambda_{m_n^\bullet}\Lambda_{(m_q^\bullet)} \sim (\log m_q^\bullet)(m_q^\bullet)^{2a}$ implies $m_q^\bullet \sim (q/\log q)^{1/(2a)}$ and $m_q^\bullet \psi^\Lambda_{m_q^\bullet} \sim (q/\log q)^{1/(2a)}$.

(s) Let $\Lambda_m \sim \exp(m^{2a})$, $a > 0$, then $m_n^\bullet \psi^\Lambda_{m_n^\bullet} \sim (\log n)^{2+1/(2a)}$ (cf. NUM. DISCUSSION 3.4.1 **(s)**), while $q \sim (\log m_q^\bullet)\psi^\Lambda_{m_q^\bullet}\Lambda_{(m_q^\bullet)} \sim (\log m_q^\bullet)(m_q^\bullet)^{4a}\exp((m_q^\bullet)^{2a})$ implies $m_q^\bullet \sim (\log q - \frac{1+4a}{2a}\log\log q - \frac{1}{2a}\log\log\log q)^{1/(2a)}$ and $m_q^\bullet \psi^\Lambda_{m_q^\bullet} \sim (\log q)^{2+1/(2a)}$.

Clearly, in both cases **(o)** and **(s)**, there are $n_{f,\Lambda}, q_{f,\Lambda} \in \mathbb{N}$ such that $\psi^\Lambda_{m_n^\bullet} m_n^\bullet \geqslant K_h(\log n)$ for all $n \geqslant n_{f,\Lambda}$ and $\psi^\Lambda_{m_q^\bullet} m_q^\bullet \geqslant K_h(\log q)$ for all $q \geqslant q_{f,\Lambda}$ hold true.

THEOREM 3.6.1. *Let f have a finite series expansion as defined in (p). Under ASSUMPTION 19 there is a finite numerical constant $\mathcal{C}$ such that for all $n, q \in \mathbb{N}$*

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^\circ} - f\|^2_{L^2} \leqslant \mathcal{C}(1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2})(\Delta^\Lambda_{n_{f,\Lambda}} n^{-1} + K_h\Lambda^2_{(K_h)}q^{-1} + q_{f,\Lambda}q^{-1}) \tag{3.2}$$

Comparison with the oracle rate of projection estimators reveals that in many cases, we obtain an oracle optimal estimator.

NUMERICAL DISCUSSION 3.6.2.
Let us illustrate THEOREM 3.6.1 considering as in NUM. DISCUSSION 3.6.1 the behaviours **(o)** and **(s)** for the sequence $(\Lambda_j)_{j\in\mathbb{N}}$. Keeping in mind that as shown in NUM. DISCUSSION 3.6.1 there are $n_{f,\Lambda}, q_{f,\Lambda} \in \mathbb{N}$ such that $\psi^\Lambda_{m_n^\bullet} m_n^\bullet \geqslant K_h(\log n)$ for all $n \geqslant n_{f,\Lambda}$ and $\psi^\Lambda_{m_q^\bullet} m_q^\bullet \geqslant K_h(\log q)$ for all $q \geqslant q_{f,\Lambda}$ hold true, due to THEOREM 3.6.1 there is a constant $\mathcal{C}_{f,g}$ depending only on the densities f and g such that $\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^\circ} - f\|^2_{L^2} \leqslant \mathcal{C}_{f,g}(n^{-1} + q^{-1})$ for all $n, q \in \mathbb{N}$. Comparing the last result with the oracle rates derived in NUM. DISCUSSION 1.5.2 we conclude, that $\widehat{\theta^\circ}$ is optimal in an oracle sense in both, the case **[p-o]** and **[s-o]**.

ASSUMPTION 20 Let f have an infinite series expansion as definied in **(np)**, that is, $1 \geqslant \mathfrak{b}_m(f) > 0$ for all $m \in \mathbb{N}$. Given $m_h := 3 * 800^2 \|\phi\|_{\ell^1}^2 \kappa^{-2}$ and $\widetilde{m}_h = \min\{m \in \mathbb{N} : \mathfrak{b}_{m_h}(f) > \mathfrak{b}_m(f)\}$ there are $n_{f,\Lambda}, q_{f,\Lambda} \in \mathbb{N}$ with $n_{f,\Lambda} \geqslant \Delta^\Lambda_{\widetilde{m}_h} \mathfrak{b}^{-2}_{\widetilde{m}_h}(f) \vee 15 * 300^4 \kappa^{-2}$ and $q_{f,\Lambda} \geqslant 289 \log(m_h + 2) \psi^\Lambda_{m_h} \Lambda_{(m_h)}$ such that (a) for all $q \geqslant q_{f,\Lambda}$, $m_q^\bullet := \max\{m \in [\![m_h, q]\!] : 289 \log(m+2) \psi^\Lambda_m \Lambda_{(m)} \leqslant q\}$, where the defining set containing m_h is not empty, satisfies $\psi^\Lambda_{m_q^\bullet} m_q^\diamond \geqslant m_h |\log \mathcal{R}^\star_q(f,\Lambda)|$ and either (b) $\psi^\Lambda_{m_n^\diamond} m_n^\diamond \geqslant m_h |\log \mathcal{R}^\diamond_n(f,\Lambda)|$ for all $n \geqslant n_{f,\Lambda}$ or (c) $m_n^\diamond \leqslant m_h |\log \mathcal{R}^\diamond_n(f,\Lambda)|$ for all $n \geqslant n_{f,\Lambda}$. We set $m_n^\bullet := \lceil m_h |\log \mathcal{R}^\diamond_n(f,\Lambda)| \rceil \wedge n$ for $n < n_{f,\Lambda}$ and $m_n^\bullet := \lceil m_h |\log \mathcal{R}^\diamond_n(f,\Lambda)| \rceil \vee m_n^\diamond$ for $n \geqslant n_{f,\Lambda}$, and in addition $m_q^\bullet := m_n^\bullet$ for $q < q_{f,\Lambda}$, where consequently in case (b) $m_n^\bullet = m_h |\log \mathcal{R}^\diamond_n(f,\Lambda)|$ for $n < n_{f,\Lambda}$, $m_n^\bullet = m_n^\diamond$ for $n \geqslant n_{f,\Lambda}$ and in case (c) $m_n^\bullet = m_h |\log \mathcal{R}^\diamond_n(f,\Lambda)|$ for all $n \in \mathbb{N}$.

NUMERICAL DISCUSSION 3.6.3.
Let us illustrate ASSUMPTION 20 considering as in NUM. DISCUSSION 1.5.1 usual behaviour **[o-o]**, **[s-o]** and **[o-s]** for the sequences $(\mathfrak{b}_m(f))_{m \in \mathbb{N}}$ and $(\Lambda_m)_{m \in \mathbb{N}}$:

- **[o-o]** $\psi^\Lambda_{m_n^\diamond} m_n^\diamond \sim n^{1/(2p+2a+1)}$ and $|\log \mathcal{R}^\diamond_n(f,\Lambda)| \sim (\log n)$ (cf. NUM. DISCUSSION 3.4.3 **[o-o]**) while $m_q^\bullet \psi^\Lambda_{m_q^\bullet} \sim (q/\log q)^{1/(2a)}$ (cf. NUM. DISCUSSION 3.6.1 **(o)**) and $|\log \mathcal{R}^\star_q(f,\Lambda)| \sim (\log q)$ (cf. NUM. DISCUSSION 1.5.2 **[o-o]**)
- **[o-s]** $\psi^\Lambda_{m_n^\diamond} m_n^\diamond \sim (\log n)^{2+1/(2a)}$ and $|\log \mathcal{R}^\diamond_n(f,\Lambda)| \sim (\log \log n)$ (cf. NUM. DISCUSSION 3.4.3 **[o-s]**) while $m_q^\bullet \psi^\Lambda_{m_q^\bullet} \sim (\log q)^{2+1/(2a)}$ (cf. NUM. DISCUSSION 3.6.1 **(s)**) and $|\log \mathcal{R}^\star_q(f,\Lambda)| \sim (\log \log q)$ (cf. NUM. DISCUSSION 1.5.2 **[o-s]**)
- **[s-o]** $\psi^\Lambda_{m_n^\diamond} m_n^\diamond \sim (\log n)^{1/(2p)}$ and $|\log \mathcal{R}^\diamond_n(f,\Lambda)| \sim (\log n)$ (cf. NUM. DISCUSSION 3.4.3 **[s-o]**) while $m_q^\bullet \psi^\Lambda_{m_q^\bullet} \sim (q/\log q)^{1/(2a)}$ (cf. NUM. DISCUSSION 3.6.1 **(o)**) and $|\log \mathcal{R}^\star_q(f,\Lambda)| \sim (\log q)$ (cf. NUM. DISCUSSION 1.5.2 **[s-o]**)

Clearly, there is $q_{f,\Lambda} \in \mathbb{N}$ such that for all $q \geqslant q_{f,\Lambda}$ in the three cases **[o-o]**, **[o-s]** and **[s-o]** $\psi^\Lambda_{m_q^\bullet} m_q^\bullet \geqslant m_h |\log \mathcal{R}^\star_q(f,\Lambda)|$, i.e., ASSUMPTION 20 (a) holds. On the other hand side, there is $n_{f,\Lambda} \in \mathbb{N}$ such that for all $n \geqslant n_{f,\Lambda}$ in the cases **[o-o]** and **[o-s]** $\psi^\Lambda_{m_n^\diamond} m_n^\diamond \geqslant m_h |\log \mathcal{R}^\diamond_n(f,\Lambda)|$, i.e., ASSUMPTION 20 (b) holds, while in case **[s-o]** $m_n^\diamond \leqslant m_h |\log \mathcal{R}^\diamond_n(f,\Lambda)|$ for $p \geqslant 1/2$, i.e., ASSUMPTION 20 (c) holds, and $\psi^\Lambda_{m_n^\diamond} m_n^\diamond \geqslant m_h |\log \mathcal{R}^\diamond_n(f,\Lambda)|$ for $p < 1/2$, i.e., ASSUMPTION 20 (b) holds.

THEOREM 3.6.2. *Let f have an infinite series expansion as definied in (np). Under ASSUMPTION 20 there is a finite numerical constant $\mathcal{C}$ such that for all $n, q \in \mathbb{N}$*

$$\mathbb{E}^{n,q}_{Y,\varepsilon} \|\widehat{\theta}^\circ - f\|^2_{L^2} \leqslant \mathcal{C}\big\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}]\, \mathcal{R}^\diamond_n[m_n^\bullet \wedge m_q^\bullet, f, \Lambda] + \mathcal{R}^\star_q(f,\Lambda) \\ + [1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^\Lambda_{n_{f,\Lambda}} n^{-1} + q_{f,\Lambda} q^{-1}) + \|\phi\|^2_{\ell^1} n^{-1} + q^2_{f,\Lambda} q^{-1}\big\}. \quad (3.3)$$

Comparison with the oracle rate of projection estimators reveals that in many cases, we obtain an oracle optimal estimator.

COROLLARY 3.6.1. *Under the assumptions of THEOREM 3.6.2 for $n \in \mathbb{N}$ let $q_n := q(n) \in \mathbb{N}$ such that $m_n^\diamond \leqslant m_{q_n}^\bullet$. If in addition $\lim_{n\to\infty} \psi^\Lambda_{m_n^\diamond} m_n^\diamond |\log \mathcal{R}^\diamond_n(f,\Lambda)|^{-1} = \infty$, then there is a finite constant $\mathcal{C}_{f,g}$ depending on the densities f and g such that*

$$\mathbb{E}^{n,q}_{Y,\varepsilon} \|\widehat{\theta}^\circ - f\|^2_{L^2} \leqslant \mathcal{C}_{f,g}(\mathcal{R}^\diamond_n(f,\Lambda) + \mathcal{R}^\star_{q_n}(f,\Lambda)) \text{ for all } n \in \mathbb{N}.$$

Numerical discussion 3.6.4.
Let us illustrate Theorem 3.6.2 considering as in Num. discussion 3.6.3 usual behaviour **[o-o]**, **[o-s]** and **[s-o]** for the sequences $(\mathfrak{b}_m(f))_{m\in\mathbb{N}}$ and $(\Lambda_m)_{m\in\mathbb{N}}$. In light of Num. discussion 3.6.3, we apply Theorem 3.6.2, where we need only check Assumption 20. The rates then follow by an evaluation of the upper bound. Let $(q_n)_{n\in\mathbb{N}}$ be a sequence of positive integers and suppose that the limits $q_{\text{o-o}}$, $q_{\text{o-s}}$, and $q_{\text{s-o}}$ defined in Num. discussion 1.5.2 exists in the respective cases.

[o-o] Since Assumption 20 (a) with $m_q^\bullet \sim (q/\log q)^{1/(2a)}$ and (b) with $m_n^\circ \sim n^{1/(2p+2a+1)}$ hold true (cf., respectively, Num. discussion 3.6.1 **(o)** and Num. discussion 3.4.3 **[o-o]**), due to Theorem 3.6.2 and Num. discussion 1.5.2 **[o-o]** there is a constant $\mathcal{C}_{f,g}$ depending on f and g such that

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}_{f,g}\{(m_n^\diamond \wedge m_q^\bullet)^{-2p} + q^{-(p\wedge a)/a}\}, \quad \forall\, n, q \in \mathbb{N}. \tag{3.4}$$

We consider two cases. Firstly, let $p > a$. If $q_{\text{o-o}} = \lim_{n\to\infty} n^{2p/(2p+2a+1)} q_n^{-1} < \infty$, then

$$\frac{m_n^\diamond}{m_q^\bullet} \sim \frac{n^{1/(2p+2a+1)}}{(q_n/\log q_n)^{1/(2a)}} = \frac{n^{1/(2p+2a+1)}}{(q_n)^{1/(2p)}} \frac{(\log q_n)^{1/(2a)}}{(q_n)^{1/(2a)-1/(2p)}} = o(1).$$

This means $m_n^\diamond \lesssim m_q^\bullet$ so the resulting upper bound is of order $(m_n^\diamond)^{-2p} + q_n^{-1} \lesssim (m_n^\diamond)^{-2p}$. Suppose now that $q_{\text{o-o}} = \infty$. If in addition $q_{\text{o-o}}^b = \lim_{n\to\infty} m_n^\diamond (m_q^\bullet)^{-1} < \infty$ then the first summand in the upper bound in (3.4) reduces to $(m_n^\diamond)^{-2p}$ and thus (keep $q_{\text{o-o}} = \infty$ in mind) the resulting upper bound is of order q_n^{-1}. Now consider $q_{\text{o-o}}^b = \infty$, then the upper bound in (3.4) is of order $(m_q^\bullet)^{-2p} + q_n^{-1} \lesssim q_n^{-1}$ because $p > a$. Combining both cases, we obtain in case $p > a$ that as $n \to \infty$

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 = \begin{cases} O(n^{-2p/(2p+2a+1)}), & \text{if } q_{\text{o-o}} < \infty, \\ O(q_n^{-1}), & \text{otherwise}\ . \end{cases}$$

Now assume $p \leqslant a$. First, suppose $q_{\text{o-o}}^b = \lim_{n\to\infty} m_n^\diamond (m_q^\bullet)^{-1} < \infty$ then the first summand in the upper bound in (3.4) reduces to $(m_n^\diamond)^{-2p}$ and moreover, it follows that $q_{\text{o-o}} < \infty$. Therefore, the resulting upper bound is of order $(m_n^\diamond)^{-2p}$. Now consider $q_{\text{o-o}}^b = \infty$, then the upper bound in (3.4) is of order $(q_n/\log q_n)^{-p/a} + q_n^{-p/a} \lesssim (q_n/\log q_n)^{-p/a}$. Combining both cases, we obtain in case $p \leqslant a$ that as $n \to \infty$

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 = \begin{cases} O(n^{-2p/(2p+2a+1)}), & \text{if } q_{\text{o-o}}^b < \infty, \\ O((q_n/\log q_n)^{-p/a}), & \text{otherwise}\ . \end{cases}$$

[o-s] Since Assumption 20 (a) with $m_q^\bullet \sim (\log q)^{1/(2a)}$ and (b) with $m_n^\circ \sim (\log n)^{1/(2a)}$ hold true (cf., respectively, Num. discussion 3.6.1 **(s)** and Num. discussion 3.4.3 **[o-s]**), due to Theorem 3.6.2 and Num. discussion 1.5.2 **[o-s]** there is a constant $\mathcal{C}_{f,g}$ depending on f and g such that

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}_{f,g}\{(\log n)^{-p/a} + (\log q)^{-p/a}\}, \quad \forall\, n, q \in \mathbb{N}. \tag{3.5}$$

Considering $q_{\text{o-s}} = \lim_{n\to\infty}(\log n)(\log q_n)^{-1}$ it follows that as $n \to \infty$

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^{\circ}} - f\|^2_{L^2} = \begin{cases} O\big((\log n)^{-p/a}\big), & \text{if } q_{\text{o-s}} < \infty, \\ O\big((\log q_n)^{-p/a}\big), & \text{otherwise .} \end{cases}$$

[s-o] Since ASSUMPTION 20 (a) with $m^{\bullet}_q \sim (q/\log q)^{1/(2a)}$, (c) with $m^{\bullet}_n \sim (\log n)$ for $p \geqslant 1/2$ and (b) with $m^{\bullet}_n \sim (\log n)^{1/(2p)}$ for $p < 1/2$ hold true (cf., respectively, NUM. DISCUSSION 3.6.1 **(o)** and NUM. DISCUSSION 3.4.3 **[s-o]**), due to THEOREM 3.6.2 and NUM. DISCUSSION 1.5.2 **[o-s]** there is a constant $\mathcal{C}_{f,g}$ depending on f and g such that

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^{\circ}} - f\|^2_{L^2} \leqslant \mathcal{C}_{f,g}\{\mathcal{R}^{\diamond}_n[m^{\bullet}_n \wedge (q/\log q)^{1/(2a)}, f, \Lambda] + q^{-1}\}, \quad \forall\, n, q \in \mathbb{N}. \tag{3.6}$$

Clearly, if $q^b_{\text{s-o}} = \lim_{n\to\infty} n(m^{\bullet}_n)^{-(2a+1)}q_n^{-1} < \infty$ then holds $m^{\bullet}_n = (\log n)^{1\vee 1/(2p)} \lesssim (q_n/\log q_n)^{1/(2a)}$ and hence $\mathcal{R}^{\diamond}_n[m^{\bullet}_n \wedge (q_n/\log q_n)^{1/(2a)}, f, \Lambda] + q_n^{-1} \lesssim (m^{\bullet}_n)^{2a+1}n^{-1}$. Suppose now that $q^b_{\text{s-o}} = \infty$, then

$$\begin{aligned} &\mathcal{R}^{\diamond}_n[m^{\bullet}_n \wedge (q_n/\log q_n)^{1/(2a)}, f, \Lambda] + q_n^{-1} \\ &\qquad\qquad \lesssim (m^{\bullet}_n)^{2a+1}n^{-1} \vee \mathcal{R}^{\diamond}_n[(q_n/\log q_n)^{1/(2a)}, f, \Lambda] + q_n^{-1} \\ &\qquad \lesssim \exp(-(q_n/\log q_n)^{p/a}) \vee n^{-1}(q_n/\log q_n)^{-(2a+1)/(2a)} + q_n^{-1} \lesssim q_n^{-1}. \end{aligned}$$

Consequently, it follows that as $n \to \infty$

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^{\circ}} - f\|^2_{L^2} = \begin{cases} O\big(n^{-1}(\log n)^{(2a+1)[1\vee 1/(2p)]}\big), & \text{if } q^b_{\text{o-s}} < \infty, \\ O\big(q_n^{-1}\big), & \text{otherwise .} \end{cases}$$

Comparing the last rates with the oracle rates derived in NUM. DISCUSSION 1.5.1 **[o-o]**, **[o-s]** and **[s-o]** we see that in case **[o-o]** with $p > a$, **[o-s]** and **[s-o]** with $p < 1/2$ $\widehat{\theta^{\circ}}$ attains the oracle rate, while in case **[o-o]** with $p \leqslant a$ and **[s-o]** with $p \geqslant 1/2$ the rate of the fully data-driven estimator $\widehat{\theta^{\circ}}$ features a detoriation by a logarithmic factor compared to the oracle rate.

Appendix A

Useful results

A.1 Lemmata for the gaussian sequence space model

Lemma A.1.1. *Let $\{X_j\}_{j\geqslant 1}$ be independent and normally distributed random variables with real mean α_j and standard deviation $\beta_j \geqslant 0$. For $m \in \mathbb{N}$, set $S_m := \sum_{j=1}^m X_j^2$ and consider $v_m \geqslant \sum_{j=1}^m \beta_j^2, t_m \geqslant \max_{1\leqslant j\leqslant m} \beta_j^2$ and $r_m \geqslant \sum_{j=1}^m \alpha_j^2$. Then for all $c \geqslant 0$, we have*

$$\sup_{m\geqslant 1} \exp\left[\frac{c(c\wedge 1)(v_m+2r_m)}{4t_m}\right] \mathbb{P}\left(S_m - \mathbb{E}[S_m] \leqslant -c(v_m+2r_m)\right) \leqslant 1;$$
$$\sup_{m\geqslant 1} \exp\left[\frac{c(c\wedge 1)(v_m+2r_m)}{4t_m}\right] \mathbb{P}\left(S_m - \mathbb{E}[S_m] \geqslant \frac{3c}{2}(v_m+2r_m)\right) \leqslant 1.$$

Lemma A.1.2. *Let $\{X_j\}_{j\geqslant 1}$ be independent and normally distributed random variables with real mean α_j and standard deviation $\beta_j \geqslant 0$. For $m \in \mathbb{N}$, set $S_m := \sum_{j=1}^m X_j^2$ and consider $v_m \geqslant \sum_{j=1}^m \beta_j^2, t_m \geqslant \max_{1\leqslant j\leqslant m} \beta_j^2$ and $r_m \geqslant \sum_{j=1}^m \alpha_j^2$. Then for all $c \geqslant 0$, we have*

$$\sup_{m\geqslant 1} (6t_m)^{-1} \exp\left[\frac{c(v_m+2r_m)}{4t_m}\right] \mathbb{E}\left[S_m - \mathbb{E}[S_m] - \frac{3}{2}c(v_m+2r_m)\right]_+ \leqslant 1$$

with $(a)_+ := (a \vee 0)$.

A.2 Lemmata for circular deconvolution

Lemma A.2.1. Talagrand's inequalities

Let $X_1, \dots, X_n$ be independent $\mathcal{X}$-valued random variables and let $\overline{\nu}_t = \frac{1}{n}\sum_{i=1}^n [\nu_t(X_i) -$

$\mathbb{E}(\nu_t(X_i))]$ *for* ν_t *belonging to a countable class* $\{\nu_t, t \in \mathcal{T}\}$ *of measurable functions. Then,*

$$\mathbb{E}\left[\left(\sup_{t\in\mathcal{T}}|\overline{\nu}_t|^2 - 6H^2\right)_+\right] \leqslant C\left[\frac{v}{n}\exp\left(\frac{-nH^2}{6v}\right) + \frac{h^2}{n^2}\exp\left(\frac{-KnH}{h}\right)\right];$$
$$\mathbb{P}\left(\sup_{t\in\mathcal{T}}|\overline{\nu}_t| \geqslant 2H + \gamma\right) \leqslant 3\exp\left[-Kn\left(\frac{\gamma^2}{v}\wedge\frac{\gamma}{h}\right)\right];$$

for any $\gamma > 0$*, with numerical constants* $K = (\sqrt{2}-1)/(21\sqrt{2})$ *and* $C > 0$ *and where*

$$\sup_{t\in\mathcal{T}}\sup_{x\in\mathcal{X}}|\nu_t(x)| \leqslant h, \quad \mathbb{E}[\sup_{t\in\mathcal{T}}|\overline{\nu}_t|] \leqslant H, \quad \sup_{t\in\mathcal{T}}\frac{1}{n}\sum_{i=1}^{n}\mathbb{V}[\nu_t(X_i)] \leqslant v.$$

Lemma A.2.2. *(Talagrand's inequalities) Let* $Z_1, \dots, Z_n$ *be independent* $\mathcal{Z}$*-valued random variables and let* $\overline{\nu_h} = n^{-1}\sum_{i=1}^{n}[\nu_h(Z_i) - \mathbb{E}(\nu_h(Z_i))]$ *for* ν_h *belonging to a countable class* $\{\nu_h, h \in \mathcal{H}\}$ *of measurable functions. Then,*

$$\mathbb{E}\big(\sup_{h\in\mathcal{H}}|\overline{\nu_h}|^2 - 6\Psi^2\big)_+ \leqslant C\left[\frac{\tau}{n}\exp\left(\frac{-n\Psi^2}{6\tau}\right) + \frac{\psi^2}{n^2}\exp\left(\frac{-Kn\Psi}{\psi}\right)\right] \tag{A.1}$$

$$\mathbb{P}\big(\sup_{h\in\mathcal{H}}|\overline{\nu_h}| \geqslant 2\Psi + \lambda\big) \leqslant 3\exp\big[-Kn\big(\frac{\lambda^2}{\tau}\wedge\frac{\lambda}{\psi}\big)\big] \leqslant 3\big[\exp\big(\frac{-Kn\lambda^2}{\tau}\big) + \exp\big(\frac{-Kn\lambda}{\psi}\big)\big] \tag{A.2}$$

for any $\lambda > 0$*, with numerical constants* $K = (\sqrt{2}-1)/(21\sqrt{2})$ *and* $C > 0$ *and where*

$$\sup_{h\in\mathcal{H}}\sup_{z\in\mathcal{Z}}|\nu_h(z)| \leqslant \psi, \qquad \mathbb{E}(\sup_{h\in\mathcal{H}}|\overline{\nu_h}|) \leqslant \Psi, \qquad \sup_{h\in\mathcal{H}}\frac{1}{n}\sum_{i=1}^{n}\mathbb{V}\mathrm{ar}(\nu_h(Z_i)) \leqslant \tau.$$

Remark A.2.1. *Consequently, setting* $\lambda = \sqrt{2}(\sqrt{3}-\sqrt{2})\Psi = \frac{(\sqrt{6}-\sqrt{4})(\sqrt{6}+\sqrt{4})}{(\sqrt{6}+\sqrt{4})}\Psi = \frac{\sqrt{2}}{(\sqrt{3}+\sqrt{2})}\Psi$*, and hence* $\sqrt{2}\sqrt{3}\Psi = \sqrt{2}\sqrt{2}\Psi + \sqrt{2}(\sqrt{3}-\sqrt{2})\Psi$ *we have*

$$\mathbb{E}\big(\sup_{h\in\mathcal{H}}|\overline{\nu_h}|^2 - 6\Psi^2\big)_+ \leqslant C\left[\frac{\tau}{n}\exp\left(\frac{-n\Psi^2}{6\tau}\right) + \frac{\psi^2}{n^2}\exp\left(\frac{-n\Psi}{100\psi}\right)\right] \tag{A.3}$$

$$\mathbb{P}\big(\sup_{h\in\mathcal{H}}|\overline{\nu_h}|^2 \geqslant 6\Psi^2\big) \leqslant 3\big[\exp\big(\frac{-n\Psi^2}{400\tau}\big) + \exp\big(\frac{-n\Psi}{200\psi}\big)\big] \tag{A.4}$$

where we used that $K\frac{2}{(\sqrt{3}+\sqrt{2})^2} = \frac{(\sqrt{2}-1)}{(21\sqrt{2})}\frac{2}{(\sqrt{3}+\sqrt{2})^2} = \frac{(2-\sqrt{2})}{21(\sqrt{3}+\sqrt{2})^2} \geqslant \frac{1}{400}$ *and* $K\frac{\sqrt{2}}{(\sqrt{3}+\sqrt{2})} = \frac{\sqrt{2}-1}{21(\sqrt{3}+\sqrt{2})} \geqslant \frac{1}{200}$ *and* $K \geqslant \frac{1}{100}$. □

Remark A.2.2. *Introduce further the unit ball* $\mathbb{B}_m := \big\{h \in \mathbb{U}_m : \|h\|_{L^2} \leqslant 1\big\}$ *contained in the linear subspace* $\mathbb{U}_m = \overline{\mathrm{lin}}\{e_j, |j| \in [\![1, m]\!]\}$*. Setting* $\nu_h(Y) = \sum_{|j|\in[\![1,m]\!]}\overline{[h]}_j g_j^{-1} e_j(-Y)$ *with* $\mathbb{E}_h^n\nu_h(Y) = \sum_{|j|\in[\![1,m]\!]}\overline{[h]}_j g_j^{-1}\phi_j$*, hence* $\overline{\nu_h} = \frac{1}{n}\sum_{i=1}^{n}\sum_{|j|\in[\![1,m]\!]}\overline{[h]}_j g_j^{-1}(e_j(-Y_i) - \phi_j)$

and keeping in mind that $\overline{[h]}_j = \langle e_j, h\rangle_{L^2}$ *we have*

$$\|\widetilde{f}m - \theta^{\circ,m}\|_{L^2}^2 = \sup_{h\in\mathbb{B}_m} |\langle \widetilde{f}m - \theta^{\circ,m}, h\rangle_{L^2}|^2 = \sup_{h\in\mathbb{B}_m} |\sum_{|j|\in[\![1,m]\!]} g_j^{-1}(\widehat{[h]}_j - \phi_j)\overline{[h]}_j|^2$$

$$= \sup_{h\in\mathbb{B}_m} |\sum_{|j|\in[\![1,m]\!]} g_j^{-1}\{\tfrac{1}{n}\sum_{i=1}^{n}(e_j(-Y_i) - \phi_j)\}\overline{[h]}_j|^2 = \sup_{h\in\mathbb{B}_m} |\overline{\nu_h}|^2.$$

The last identity provides the necessary argument to link the condition **??** *and Talagrand's inequality. Note that, the unit ball* $\mathbb{B}_m$ *is not a countable set of functions, however, it contains a countable dense subset, say* $\mathcal{H}$*, since* L^2 *is separable, and it is straightforward to see that* $\sup_{h\in\mathbb{B}_m} |\overline{\nu_h}|^2 = \sup_{h\in\mathcal{H}} |\overline{\nu_h}|^2$. □

Remark A.2.3. Special case

Using $\gamma = \sqrt{2}\,(\sqrt{3} - \sqrt{2})\,H$ *we have*

$$\mathbb{E}\left(\sup_{t\in\mathcal{T}} |\overline{\nu}_t|^2 - 6H^2\right)_+ \leqslant C\left[\frac{v}{n}\exp\left(\frac{-nH^2}{6v}\right) + \frac{h^2}{n^2}\exp\left(\frac{-nH}{100h}\right)\right];$$

$$\mathbb{P}\left(\sup_{t\in\mathcal{T}} |\overline{\nu}_t|^2 \geqslant 6H^2\right) \leqslant 3\left(\exp\left[-\frac{nH^2}{400v}\right] + \exp\left[-\frac{nH}{200h}\right]\right);$$

A.3 Commonly used inequalities

Lemma A.3.1. Young's inequality

Consider p, q, *and* r, *three real numbers greater than 1 such that* $\frac{1}{p} + \frac{1}{q} = 1 + \frac{1}{r}$; *as well as* x *and* y, *respectively in* $\mathcal{L}^p$ *and* $\mathcal{L}^q$. *Then,*

$$\|x \star y\|_r \leqslant \|f\|_p \cdot \|g\|_q.$$

Lemma A.3.2. Cauchy Schwarz inequality

Let be x *and* y *in an inner product space (e.g.* $\mathcal{L}^2$*), then we have*

$$|\langle x|y\rangle|^2 \leqslant \|x\|^2 \cdot \|y\|^2.$$

Lemma A.3.3. Hölder's inequality

Let p *and* q *be elements of* $[1,\infty]$ *such that* $\frac{1}{p} + \frac{1}{q} = 1$. *Then, for any measurable complex-valued functions* f *and* g,

$$\|hb\|_{L^1} = \|h\|_{L^p} \cdot \|b\|_{L^q}$$

Appendix B

Dependent data

We present in this annex definitions, and results which are used in SECTION 3.5 in order to compute the convergence rate of the adaptive aggregation estimator for strictly stationary, absolutely regular process.

Definition 42 Strict stationarity

Consider a sequence of random variables $(Y_p)_{p\in\mathbb{N}}$. We say that $(Y_p)_{p\in\mathbb{N}}$ is strictly stationary if, for any integer q, any finite family of integers $(p_r)_{r\in[\![1,q]\!]}$, and integer h, $(Y_{p_r})_{r\in[\![1,q]\!]}$ is identically distributed to $(Y_{p_r+h})_{r\in[\![1,q]\!]}$. In particular, for any integers p and q, Y_p and Y_q have same distribution.

Definition 43 β-mixing coefficients

Let $(\Omega,\mathcal{A},\mathbb{P})$ be a probability space and $\mathcal{U}$ and $\mathcal{V}$ be two sub σ-algebras of $\mathcal{A}$. Then, we define the β-mixing coefficient of $\mathcal{U}$ and $\mathcal{V}$:

$$\beta(\mathcal{U},\mathcal{V}) := \frac{1}{2}\sup_{(U_j)_{j\in I}(V_j)_{j\in J}}\left\{\sum\sum|\,\mathbb{P}(U_j\,\mathbb{P}(V_k) - \mathbb{P}(U_j\cap V_k)|\right\}$$

where the sup is taken over all possible finite partition of Ω which are respectively $\mathcal{U}$ and $\mathcal{V}$ measurable.

In addition for two random variables Y_1 and Y_2 we note $\sigma(Y_1)$ and $\sigma(Y_2)$ the σ-algebra they generate and $\beta\left(Y_1,Y_2\right) = \beta\left(\sigma(Y_1),\sigma(Y_2)\right)$.

Definition 44 Absolutely regular process

Consider a stochastic process $(Y_p)_{p\in\mathbb{Z}}$. Denote, for any p in $\mathbb{N}$, by $\mathcal{F}_p^- := \sigma\left((Y_q)_{q\leqslant p}\right)$ and $\mathcal{F}_p^+ := \sigma\left((Y_q)_{q\geqslant p}\right)$. The stochastic process $(Y_p)_{p\in\mathbb{Z}}$ is said to be absolutely regular if

$$\lim_{p\to\infty}\beta(\mathcal{F}_0^-,\mathcal{F}_p^+) = 0.$$

Assumption 21 Rich space

Assume that the universe is rich enough in the sense that there exist a sequence of random variables with uniform distribution on $[0,1]$ which is independent of $(Y_p)_{p\in\mathbb{N}}$.

As a consequence, there exist a sequence $(Y_p^\perp)_{p\in\mathbb{N}}$ satisfying the following properties. For any positive integer s and for any strictly positive integer q, define the sets $(I_{q,p}^e)_{p\in[\![1,s]\!]} := [\![2(q-1)s+1,(2q-1)s]\!]$ and $\left(I_{q,p}^o\right)_{p\in[\![1,s]\!]} := [\![(2q-1)s+1,2qs]\!]$.

Define for any q in $\mathbb{N}$ the vectors of random variables $E_q := (Y^n_{I^e_{q,p}})_{p\in[\![1,s]\!]}$; $O_q := (Y^n_{I^o_{q,p}})_{p\in[\![1,s]\!]}$; and their counterparts $E_q^\perp := (Y^{n,\perp}_{I^e_{q,p}})_{p\in[\![1,s]\!]}$ and $O_q^\perp := (Y^{n,\perp}_{I^o_{q,p}})_{p\in[\![1,s]\!]}$.
Then, $\left(Y_p^\perp\right)_{p\in\mathbb{N}}$ satisfies:

- for any integer q, $E_q^\perp$, E_q, $O_q^\perp$, and O_q are identically distributed;
- for any integer q, $\mathbb{P}^n_{\theta^\circ}\left(E_q \neq E_q^\perp\right) \leqslant \beta_s$ and $\mathbb{P}^n_{\theta^\circ}\left(O_q \neq O_q^\perp\right) \leqslant \beta_s$;
- $\left(E_q^\perp\right)_{q\in\mathbb{N}}$ are independent and identically distributed and $\left(O_q^\perp\right)_{q\in\mathbb{N}}$ as well.

Appendix C

Proof of Theorem 2.4.2

C.1 Intermediate results

Proposition C.1.1. *Under Assumption 11, we have, for all m in $[\![1, G_n]\!]$*

$$\mathbb{P}^n_{\theta^\circ}\left[\left\|\overline{\theta}^m - \theta^\circ\right\|^2 < \frac{1}{2}\left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n}\right]\right] \leqslant \exp\left[-\frac{m}{16L}\right],$$
$$\mathbb{P}^n_{\theta^\circ}\left[\left\|\overline{\theta}^m - \theta^\circ\right\|^2 > 4\left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n}\right]\right] \leqslant \exp\left[-\frac{m}{9L}\right].$$

Definition 45 Define the following quantities :

$$G_n^- := \min\left\{m \in [\![1, m_n^\circ]\!] : \quad \mathfrak{b}_m \leqslant 9L\Phi_n^\circ\right\},$$
$$G_n^+ := \max\left\{m \in [\![m_n^\circ, G_n]\!] : \frac{(m - m_n^\circ)}{n} \leqslant 3\Lambda_{m_n^\circ}^{-1}\Phi_n^\circ\right\}.$$

Proposition C.1.2. *Under Assumption 11, we have the following concentration inequalities for the threshold hyper parameter :*

$$\mathbb{P}^n_{\theta^\circ}\left[M > G_n^+\right] \leqslant \exp\left[-\frac{5m_n^\circ}{9L} + \log\left(G_n\right)\right],$$
$$\mathbb{P}^n_{\theta^\circ}\left[M < G_n^-\right] \leqslant \exp\left[-\frac{4m_n^\circ}{9} + \log\left(G_n\right)\right].$$

C.2 Detailed proofs

Proof of Proposition C.1.1

Let be m in $[\![1, G_n]\!]$ and note that

$$\|\overline{\theta}^m - \theta^\circ\|^2 = \sum_{j=1}^{m}\left(\frac{Y_j}{\lambda_j} - \theta_j^\circ\right)^2 + \mathfrak{b}_m.$$

We will use LEMMA A.1.1; therefor define

$$\begin{aligned}
& & S_m &:= \sum_{j=1}^{m} \left(\frac{Y_j}{\lambda_j} - \theta^\circ_j \right)^2 & , \\
& & \mu_m &:= \mathbb{E}^n_{\theta^\circ} \left[S_m \right] & = \frac{m \overline{\Lambda}_m}{n}, \\
& \forall j \in [\![1, m]\!] & \beta_j^2 &:= \mathbb{V}^n_{\theta^\circ} \left[\frac{Y_j}{\lambda_j} - \theta^\circ_j \right] & = \frac{\Lambda_j}{n}, \\
& \forall j \in [\![1, m]\!] & \alpha_j^2 &:= \mathbb{E}^n_{\theta^\circ} \left[\frac{Y_j}{\lambda_j} - \theta^\circ_j \right] & = 0, \\
& & v_m &:= \sum_{j=1}^{m} \beta_j^2 & = \frac{m \overline{\Lambda}_m}{n}, \\
& & t_m &:= \max_{j \in [\![1, m]\!]} \beta_j^2 & = \frac{\Lambda_m}{n}.
\end{aligned}$$

We then control the concentration of S_m, first from above, using the following facts:

- for any a and b in $\mathbb{R}_+$, we have $a \vee b \leqslant a + b$;
- LEMMA A.1.1;
- ASSUMPTION 11

and we obtain

$$\begin{aligned}
\mathbb{P}^n_{\theta^\circ} \left[S_m + \mathfrak{b}_m \leqslant \frac{1}{2} \left[\frac{m \overline{\Lambda}_m}{n} \vee \mathfrak{b}_m \right] \right] &\leqslant \mathbb{P}^n_{\theta^\circ} \left[S_m + \mathfrak{b}_m \leqslant \frac{1}{2} \left(\frac{m \overline{\Lambda}_m}{n} + \mathfrak{b}_m \right) \right] \\
&\leqslant \mathbb{P}^n_{\theta^\circ} \left[S_m + \mathfrak{b}_m \leqslant \frac{1}{2} \frac{m \overline{\Lambda}_m}{n} + \mathfrak{b}_m \right] \\
&\leqslant \mathbb{P}^n_{\theta^\circ} \left[S_m \leqslant \frac{1}{2} \frac{m \overline{\Lambda}_m}{n} \right] \\
&\leqslant \mathbb{P}^n_{\theta^\circ} \left[S_m - \mu_m \leqslant -\frac{1}{2} \frac{m \overline{\Lambda}_m}{n} \right] \\
&\leqslant \exp \left[- \frac{n m \overline{\Lambda}_m}{16 n \Lambda_m} \right] \\
&\leqslant \exp \left[- \frac{m}{16 L} \right].
\end{aligned}$$

Finally, we control the concentration of S_m from bellow using the facts

- for any a and b in $\mathbb{R}_+$, we have $a \vee b \geqslant \frac{1}{2}(a + b)$;
- LEMMA A.1.1;
- ASSUMPTION 11;

which gives

$$
\begin{aligned}
\mathbb{P}^n_{\theta^\circ}\left[S_m + \mathfrak{b}_m \geqslant 4\left[\frac{m\overline{\Lambda}_m}{n} \vee \mathfrak{b}_m\right]\right] \leqslant & \mathbb{P}^n_{\theta^\circ}\left[S_m + \mathfrak{b}_m \geqslant 2\left(\frac{m\overline{\Lambda}_m}{n} + \mathfrak{b}_m\right)\right] \\
\leqslant & \mathbb{P}^n_{\theta^\circ}\left[S_m + \mathfrak{b}_m \geqslant 2\frac{m\overline{\Lambda}_m}{n} + \mathfrak{b}_m\right] \\
\leqslant & \mathbb{P}^n_{\theta^\circ}\left[S_m \geqslant 2\frac{m\overline{\Lambda}_m}{n}\right] \\
\leqslant & \mathbb{P}^n_{\theta^\circ}\left[S_m - \mu_m \geqslant \frac{m\overline{\Lambda}_m}{n}\right] \\
\leqslant & \exp\left[-\frac{nm\overline{\Lambda}_m}{9n\Lambda_m}\right] \\
\leqslant & \exp\left[-\frac{m}{9L}\right].
\end{aligned}
$$

Proof for Proposition C.1.2

First, let's proof the first inequality. Use the fact that :

$$
\begin{aligned}
\mathbb{P}_{\theta^\circ}\left[G_n^+ < \widehat{m} \leqslant G_n\right] = & \mathbb{P}_{\theta^\circ}\left[\forall l \in [\![1, G_n^+]\!], \quad \frac{3\widehat{m}}{n} - \sum_{j=1}^{\widehat{m}} Y_j^2 < \frac{3l}{n} - \sum_{j=1}^{l} Y_j^2\right] \\
\leqslant & \mathbb{P}_{\theta^\circ}\left[\exists m \in [\![G_n^+ + 1, G_n]\!] : \quad \frac{3m}{n} - \sum_{j=1}^{m} Y_j^2 < \frac{3m_n^\circ}{n} - \sum_{j=1}^{m_n^\circ} Y_j^2\right] \\
\leqslant & \sum_{m=G_n^+ + 1}^{G_n} \mathbb{P}_{\theta^\circ}\left[\frac{3m}{n} - \sum_{j=1}^{m} Y_j^2 < \frac{3m_n^\circ}{n} - \sum_{j=1}^{m_n^\circ} Y_j^2\right] \\
\leqslant & \sum_{m=G_n^+ + 1}^{G_n} \mathbb{P}_{\theta^\circ}\left[0 < \frac{3\left(m_n^\circ - m\right)}{n} + \sum_{j=m_n^\circ + 1}^{m} Y_j^2\right]
\end{aligned}
$$

We will now use Lemma A.1.1. For this purpose, define then for all m in $[\![G_n^+ + 1, G_n]\!]$: $S_m := \sum_{j=m_n^\circ+1}^{m} Y_j^2$, we then have $\mu_m := \mathbb{E}^n_{\theta^\circ}\left[S_m\right] = \frac{m - m_n^\circ}{n} + \sum_{j=m_n^\circ+1}^{m}\left(\theta_j^\circ \lambda_j\right)^2$, $\alpha_j^2 := \mathbb{E}^n_{\theta^\circ}\left[Y_j\right]^2 = \left(\theta_j^\circ \lambda_j\right)^2$ and $\beta_j^2 := \mathbb{V}^n_{\theta^\circ}\left[Y_j\right] = \frac{1}{n}$.

Now, using that λ is monotonically decreasing and $\mathfrak{b}_{m_n^\circ} \leqslant \Phi_n^\circ$, we note

$$\begin{aligned}
\sum_{j=m_n^\circ+1}^{m} \alpha_j^2 &= \sum_{j=m_n^\circ+1}^{m} \left(\theta_j^\circ \lambda_j\right)^2 \\
&\leqslant \Lambda_{m_n^\circ}^{-1} \sum_{j=m_n^\circ+1}^{m} \left(\theta_j^\circ\right)^2 \\
&\leqslant \Lambda_{m_n^\circ}^{-1} \mathfrak{b}_{m_n^\circ} \\
&\leqslant \Lambda_{m_n^\circ}^{-1} \Phi_n^\circ && =: r_m \\
\sum_{j=m_n^\circ+1}^{m} \beta_j^2 &= \frac{m - m_n^\circ}{n} && =: v_m \\
\max_{j \in [\![m_n^\circ+1, m]\!]} \beta_j &= \frac{1}{n} && =: t_m
\end{aligned}$$

Hence, we have, for all m in $[\![G_n^+, G_n]\!]$

$$\begin{aligned}
\mathbb{P}_{\theta^\circ}^n \left[\sum_{j=m_n^\circ+1}^{m} Y_j^2 - 3\frac{m - m_n^\circ}{n} > 0 \right] &= \mathbb{P}_{\theta^\circ}^n \left[S_m - \frac{m - m_n^\circ}{n} > 2\frac{m - m_n^\circ}{n} \right] \\
&= \mathbb{P}_{\theta^\circ}^n \left[S_m - \frac{m - m_n^\circ}{n} - \sum_{j=m_n^\circ+1}^{m} \left(\theta_j^\circ \lambda_j\right)^2 > 2\frac{m - m_n^\circ}{n} - \sum_{j=m_n^\circ+1}^{m} \left(\theta_j^\circ \lambda_j\right) \right. \\
&\leqslant \mathbb{P}_{\theta^\circ}^n \left[S_m - \mu_m > 2\frac{m - m_n^\circ}{n} - \Lambda_{m_n^\circ}^{-1} \Phi_n^\circ \right].
\end{aligned}$$

Using the definition of G_n^+, we have $\frac{m - m_n^\circ}{n} > 3\Lambda_{m_n^\circ}^{-1} \Phi_n^\circ$.
Hence, we can write, using Assumption 11 and Lemma A.1.1 with $c = 2/3$:

$$\begin{aligned}
\mathbb{P}_{\theta^\circ}^n \left[\sum_{j=m_n^\circ}^{m} Y_j^2 - 3\frac{m - m_n^\circ}{n} > 0 \right] &\leqslant \mathbb{P}_{\theta^\circ}^n \left[S_m - \mu_m > \frac{m - m_n^\circ}{n} + 2\Lambda_{m_n^\circ}^{-1} \Phi_n^\circ \right] \\
&\leqslant \mathbb{P}_{\theta^\circ}^n \left[S_m - \mu_m > v_m + 2r_m \right] \\
&\leqslant \exp\left[-n \frac{\frac{m - m_n^\circ}{n} + 2\Lambda_{m_n^\circ}^{-1} \Phi_n^\circ}{9} \right] \\
&\leqslant \exp\left[-n \frac{5\Lambda_{m_n^\circ}^{-1} \Phi_n^\circ}{9} \right] \\
&\leqslant \exp\left[-\frac{5 m_n^\circ}{9L} \right].
\end{aligned}$$

Finally we can conclude that

$$\mathbb{P}_{\theta^\circ}^n \left[G_n^+ < \widehat{m} \leqslant G_n \right] \leqslant \exp\left[-\frac{5 m_n^\circ}{9L} + \log\left(G_n\right) \right].$$

We now prove the second inequality.

We begin by writing the same kind of inclusion of events as for the first inequality :

$$\begin{aligned}
\mathbb{P}^n_{\theta^\circ}\left[1 \leqslant \widehat{m} < G_n^-\right] =& \mathbb{P}^n_{\theta^\circ}\left[\forall m \in [\![G_n^-, G_n]\!], \quad 3\frac{\widehat{m}}{n} - \sum_{j=1}^{\widehat{m}} Y_j^2 < 3\frac{m}{n} - \sum_{j=1}^{m} Y_j^2\right] \\
\leqslant& \mathbb{P}_{\theta^\circ}\left[\exists m \in [\![1, G_n^- - 1]\!], \quad 3\frac{m}{n} - \sum_{j=1}^{m} Y_j^2 < 3\frac{m_n^\circ}{n} - \sum_{j=1}^{m_n^\circ} Y_j^2\right] \\
\leqslant& \sum_{m=1}^{G_n^-} \mathbb{P}^n_{\theta^\circ}\left[3\frac{m}{n} - \sum_{j=1}^{m} Y_j^2 < 3\frac{m_n^\circ}{n} - \sum_{j=1}^{m_n^\circ} Y_j^2\right] \\
\leqslant& \sum_{m=1}^{G_n^-} \mathbb{P}^n_{\theta^\circ}\left[\sum_{j=m+1}^{m_n^\circ} Y_j^2 < 3\frac{m_n^\circ - m}{n}\right].
\end{aligned}$$

The LEMMA A.1.1 steps in again. Define $S_m := \sum_{j=m+1}^{m_n^\circ} Y_j^2$ and we want to control the concentration of this sum, hence we take the following notations :

$$\begin{aligned}
\mu_m :=& \mathbb{E}_{\theta^\circ}\left[S_m\right] \\
=& \frac{m_n^\circ - m}{n} + \sum_{j=m+1}^{m_n^\circ} \left(\theta_j^\circ \lambda_j\right)^2 \\
r_m :=& \sum_{j=m+1}^{m_n^\circ} \left(\theta_j^\circ \lambda_j\right)^2 \\
v_m :=& \frac{m_n^\circ - m}{n} \\
t_m :=& \frac{1}{n}.
\end{aligned}$$

Hence, we have, using ASSUMPTION 11 and the definition of G_n^-

$$\begin{aligned}
\mathbb{P}^n_{\theta^\circ}\left[S_m < 3\frac{m_n^\circ - m}{n}\right] =& \mathbb{P}^n_{\theta^\circ}\left[S_m - \mu_m < 3\frac{m_n^\circ - m}{n} - \frac{m_n^\circ - m}{n} - \sum_{j=m+1}^{m_n^\circ}\left(\theta_j^\circ \lambda_j\right)^2\right]\\
=& \mathbb{P}^n_{\theta^\circ}\left[S_m - \mu_m < 3\frac{m_n^\circ - m}{n} - \frac{2}{3}\frac{m_n^\circ - m}{n} - \frac{1}{3}\sum_{j=m+1}^{m_n^\circ}\left(\theta_j^\circ \lambda_j\right)^2 - \frac{1}{3}\left[v_m + 2r_m\right]\right]\\
\leqslant& \mathbb{P}^n_{\theta^\circ}\left[S_m - \mu_m < -\frac{1}{3}\left[v_m + 2r_m\right] + \frac{7}{3}\frac{m_n^\circ}{n} - \frac{7}{3}\frac{m}{n} - \frac{1}{3}\Lambda_{m_n^\circ}^{-1}\sum_{j=m+1}^{m_n^\circ}\left(\theta_j^\circ\right)^2\right]\\
\leqslant& \mathbb{P}^n_{\theta^\circ}\left[S_m - \mu_m < -\frac{1}{3}\left[v_m + 2r_m\right] + \frac{7}{3}\frac{m_n^\circ}{n} + \frac{1}{3}\Lambda_{m_n^\circ}^{-1}\left(\mathfrak{b}_{m_n^\circ}^2 - \mathfrak{b}_m^2\right)\right]\\
\leqslant& \mathbb{P}^n_{\theta^\circ}\left[S_m - \mu_m < -\frac{1}{3}\left[v_m + 2r_m\right] + \frac{7L}{3}\Phi_n^\circ\Lambda_{m_n^\circ}^{-1} + \frac{1}{3}\Lambda_{m_n^\circ}^{-1}\left(\Phi_n^\circ - 9L\Phi_n^\circ\right)\right]\\
\leqslant& \mathbb{P}^n_{\theta^\circ}\left[S_m - \mu_m < -\frac{1}{3}\left[v_m + 2r_m\right] + \frac{\Phi_n^\circ\Lambda_{m_n^\circ}^{-1}}{3}\left(1 - 2L\right)\right]
\end{aligned}$$

we now use LEMMA A.1.1

$$\begin{aligned}
\mathbb{P}^n_{\theta^\circ}\left[S_m < 3\frac{m_n^\circ - m}{n}\right] \leqslant& \mathbb{P}^n_{\theta^\circ}\left[S_m - \mu_m < -\frac{1}{3}\left[v_m + 2r_m\right]\right]\\
\leqslant& \exp\left[-n\frac{\frac{m_n^\circ - m}{n} + 2\sum_{j=m+1}^{m_n^\circ}\left(\theta_j^\circ\lambda_j\right)^2}{36}\right]\\
\leqslant& \exp\left[-n\frac{\frac{m_n^\circ - m}{n} + 2\Lambda_{m_n^\circ}^{-1}\mathfrak{b}_m - 2\Lambda_{m_n^\circ}^{-1}\mathfrak{b}_{m_n^\circ}}{36}\right]\\
\leqslant& \exp\left[-n\frac{16L\Phi_n^\circ\Lambda_{m_n^\circ}^{-1}}{36}\right]\\
\leqslant& \exp\left[-\frac{4m_n^\circ}{9}\right]
\end{aligned}$$

Which in turn implies

$$\mathbb{P}^\circ_{\theta^\circ}\left[1 \leqslant \widehat{m} < G_n^-\right] \leqslant \exp\left[-\frac{4m_n^\circ}{9} + \log\left(G_n\right)\right].$$

PROOF OF THEOREM 2.4.2

By the the total probability formula, we have :

$$\begin{aligned}&\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y^n}\left(\left(K^\circ\right)^{-1}\Phi^\circ_n \leqslant \|\boldsymbol{\theta}^M-\theta^\circ\|^2 \leqslant K^\circ\Phi^\circ_n\right)\right]\\ &\quad = 1-\underbrace{\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y^n}\left(\left(K^\circ\right)^{-1}\Phi^\circ_n > \|\boldsymbol{\theta}^M-\theta^\circ\|^2\right)\right]}_{=:A}-\underbrace{\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y^n}\left(K^\circ\Phi^\circ_n < \|\boldsymbol{\theta}^M-\theta^\circ\|^2\right)\right]}_{B}.\end{aligned}$$

Hence, we will control A and B separately.

We first control A :

$$\begin{aligned}A &= \sum_{m=1}^{G_n}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y^n}\left(\left\{\left(K^\circ\right)^{-1}\Phi^\circ_n > \|\boldsymbol{\theta}^M-\theta^\circ\|^2\right\}\cap\{M=m\}\right)\right]\\ &\leqslant \sum_{m=1}^{G_n^- -1}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\infty)}_{M|Y^n}\left(\{M=m\}\right)\right]+\sum_{m=G_n^+ +1}^{G_n}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\infty)}_{M|Y}\left(\{M=m\}\right)\right]\\ &\qquad + \sum_{m=G_n^-}^{G_n^+}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y,M=m}\left(\left\{\left(K^\circ\right)^{-1}\Phi^\circ_n > \|\boldsymbol{\theta}^M-\theta^\circ\|^2\right\}\right)\right]\\ &\leqslant \underbrace{\sum_{m=1}^{G_n^- -1}\mathbb{P}^n_{\theta^\circ}\left[\{\widehat{m}=m\}\right]}_{=:A_1}+\underbrace{\sum_{m=G_n^+ +1}^{G_n}\mathbb{P}^n_{\theta^\circ}\left[\{\widehat{m}=m\}\right]}_{=:A_2}+\sum_{m=G_n^-}^{G_n^+}\underbrace{\mathbb{P}^n_{\theta^\circ}\left[\left\{\left(K^\circ\right)^{-1}\Phi^\circ_n > \left\|\overline{\theta}^m-\theta^\circ\right\|^2\right\}\right]}_{=:A_{3,m}}.\end{aligned}$$

While A_1 and A_2 can respectively be controlled thanks to Proposition C.1.2 by

$$\begin{aligned}A_1 &\leqslant \exp\left[-\frac{5m^\circ_n}{9L}+\log\left(G_n\right)\right]\\ A_2 &\leqslant \exp\left[-\frac{7m^\circ_n}{9}+\log\left(G_n\right)\right],\end{aligned}$$

we now have to control $\sum_{m=G_n^-}^{G_n^+} A_{3,m}$.

Thanks to Proposition C.1.1, we have that, for all m in $[\![1,G_n]\!]$:

$$\mathbb{P}^n_{\theta^\circ}\left[\left\{\left\|\overline{\theta}^m-\theta^\circ\right\|^2 < \frac{1}{2}\left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n}\right]\right\}\right] \leqslant \exp\left[-\frac{m}{16L}\right].$$

Moreover, by definition of Φ°_n, we have for all m in $[\![G_n^-,G_n^+]\!]$ that $\Phi^\circ_n \leqslant \left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n}\right]$, which implies, with $K^\circ \geqslant 16L$,

$$\begin{aligned}\mathbb{P}^n_{\theta^\circ}\left[\left\{\left\|\overline{\theta}^m-\theta^\circ\right\|^2 < \left(K^\circ\right)^{-1}\Phi^\circ_n\right\}\right] &\leqslant \mathbb{P}^n_{\theta^\circ}\left[\left\{\left\|\overline{\theta}^m-\theta^\circ\right\|^2 < \frac{1}{2}\left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n}\right]\right\}\right]\\ &\leqslant \exp\left[-\frac{m}{K^\circ}\right].\end{aligned}$$

This allows us to conclude that

$$\begin{aligned}\sum_{m=G_n^-}^{G_n^+} A_{3,m} &\leqslant \sum_{m=G_n^-}^{G_n^+} \exp\left[-\frac{m}{K^\circ}\right] \\ &\leqslant \sum_{m=G_n^-}^{\infty} \exp\left[-\frac{m}{K^\circ}\right] \\ &\leqslant \int_{m=G_n^- -1}^{\infty} \exp\left[-\frac{m}{K^\circ}\right] dm \\ &\leqslant \left[K^\circ \exp\left[-\frac{m}{K^\circ}\right]\right]_{G_n^- -1}^{\infty} dm \\ &\leqslant K^\circ \exp\left[-\frac{G_n^- - 1}{K^\circ}\right]\end{aligned}$$

We now control B :

$$\begin{aligned}B = &\sum_{m=1}^{G_n} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y^n} \left(\left\{K^\circ \Phi_n^\circ < \|\boldsymbol{\theta}^M - \theta^\circ\|^2\right\} \cap \{M = m\}\right)\right] \\ \leqslant &\sum_{m=1}^{G_n^- -1} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y^n} (\{M = m\})\right] + \sum_{m=G_n^+ +1}^{G_n} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}^{n,(\infty)}_{M|Y^n} (\{M = m\})\right] \\ &+ \sum_{m=G_n^-}^{G_n^+} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y,M=m} \left(\left\{K^\circ \Phi_n^\circ < \|\boldsymbol{\theta}^M - \theta^\circ\|^2\right\}\right)\right] \\ \leqslant &\underbrace{\sum_{m=1}^{G_n^- -1} \mathbb{P}_{\theta^\circ}^n [\{\widehat{m} = m\}]}_{=:B_1} + \underbrace{\sum_{m=G_n^+ +1}^{G_n} \mathbb{P}_{\theta^\circ}^n [\{\widehat{m} = m\}]}_{=:B_2} + \sum_{m=G_n^-}^{G_n^+} \underbrace{\mathbb{P}_{\theta^\circ}^n \left[\left\{K^\circ \Phi_n^\circ < \left\|\overline{\theta}^m - \theta^\circ\right\|^2\right\}\right]}_{=:B_{3,m}}.\end{aligned}$$

As previously, B_1 and B_2 are controlled in PROPOSITION C.1.2. Hence, we now control $\sum_{m=G_n^-}^{G_n^+} B_{3,m}$.

Using PROPOSITION C.1.1 again, we have that, for all m in $[\![1, G_n]\!]$:

$$\mathbb{P}_{\theta^\circ}^n \left[\left\{\|\overline{\theta}^m - \theta^\circ\|_{l_2}^2 > 4\left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n}\right]\right\}\right] \leqslant \exp\left[-\frac{m}{9L}\right].$$

In addition to that, thanks to the definitions of G_n^- and G_n^+, the monotonicity of $\mathfrak{b}_m$ and $\frac{m\overline{\Lambda}_m}{n}$, on the one hand we have, for all m in $[\![G_n^-, m_n^\circ]\!]$:

$$\begin{aligned}\frac{m\overline{\Lambda}_m}{n} &\leqslant \frac{m_n^\circ \overline{\Lambda}_{m_n^\circ}}{n} \leqslant \Phi_n^\circ, \\ \mathfrak{b}_m &\leqslant 9L\Phi_n^\circ;\end{aligned}$$

and on the other hand, thanks to Assumption 12, with $D^\circ := \lceil \frac{3}{\kappa^\circ} + 1 \rceil$, then we have for all m in $[\![m_n^\circ, G_n^+]\!]$:

$$\begin{aligned} m &\leqslant \frac{3\Phi_n^\circ n}{\Lambda_{m_n^\circ}} + m_n^\circ \leqslant \frac{3nm_n^\circ \overline{\Lambda}_{m_n^\circ}}{\Lambda_{m_n^\circ} n\kappa^\circ} + m_n^\circ \leqslant \left(\frac{3}{\kappa^\circ} + 1 \right) m_n^\circ \leqslant D^\circ m_n^\circ, \\ \mathfrak{b}_m &\leqslant \mathfrak{b}_{m_n^\circ} \leqslant \Phi_n^\circ. \end{aligned}$$

Using $m \leqslant D^\circ m_n^\circ$ and Assumption 11 we have

$$\overline{\Lambda}_m \leqslant \Lambda_m \leqslant \Lambda_{D^\circ m_n^\circ} \leqslant \Lambda_{D^\circ} \Lambda_{m_n^\circ} \leqslant \Lambda_{D^\circ} L \overline{\Lambda}_{m_n^\circ}.$$

So finally we have for all m in $[\![G_n^-, G_n^+]\!]$, with $K^\circ \geqslant 4L \cdot [9 \vee D^\circ \Lambda_{D^\circ}]$:

$$\begin{aligned} \mathfrak{b}_m &\leqslant 9L\Phi_n^\circ, \\ \frac{m\overline{\Lambda}_m}{n} &\leqslant D^\circ \Lambda_{D^\circ} L \cdot \frac{m_n^\circ \overline{\Lambda}_{m_n^\circ}}{n} \leqslant D^\circ \Lambda_{D^\circ} L \cdot \Phi_n^\circ, \\ \left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n} \right] &\leqslant L \left[9 \vee D^\circ \Lambda_{D^\circ} \right] \Phi_n^\circ, \\ 4 \left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n} \right] &\leqslant K^\circ \Phi_n^\circ. \end{aligned}$$

Which leads us to the upper bound:

$$\sum_{m=G_n^-}^{G_n^+} B_{3,m} \leqslant \sum_{m=G_n^-}^{G_n^+} \exp\left[-\frac{m}{9L} \right] \leqslant K^\circ \exp\left[-\frac{G_n^- - 1}{K^\circ} \right].$$

Finally, we can conclude :

$$\begin{aligned} &\mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}^M | Y^n}^{n,(\infty)} \left((K^\circ)^{-1} \Phi_n^\circ \leqslant \| \boldsymbol{\theta}^M - \theta^\circ \|^2 \leqslant K^\circ \Phi_n^\circ \right) \right] \\ &\quad \geqslant 1 - 2\exp\left[-\frac{5m_n^\circ}{9L} + \log\left(G_n\right) \right] - 2\exp\left[-\frac{7m_n^\circ}{9} + \log\left(G_n\right) \right] - 2K^\circ \exp\left[-\frac{G_n^- - 1}{K^\circ} \right]. \end{aligned}$$

Appendix D

Proof of THEOREM 2.4.3

D.1 Intermediate results

PROPOSITION D.1.1. *Under ASSUMPTION 11, we have, for all m in $[\![1, G_n]\!]$ and c greater than $\frac{3}{2}$,*

$$\mathbb{P}^n_{\theta^\circ}\left[\left\|\overline{\theta}^m - \theta^\circ\right\|^2 > 4c\left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n}\right]\right] \leqslant \exp\left[-\frac{cm}{6L}\right].$$

DEFINITION 46 Define the following quantities :

$$G_n^{\star-} := \min\left\{m \in [\![1, m_n^\star]\!] : \quad \mathfrak{b}_m \leqslant 9\,(1 \vee r)\, L\Phi_n^\star\right\},$$
$$G_n^{\star+} := \max\left\{m \in [\![m_n^\star, G_n]\!] : \frac{m - m_n^\star}{n} \leqslant 3\Lambda_{m_n^\star}^{-1}\,(1 \vee r)\,\Phi_n^\star\right\}.$$

PROPOSITION D.1.2. *Under ASSUMPTION 11, we have the following concentration inequalities for the threshold hyper parameter :*

$$\mathbb{P}^n_{\theta^\circ}\left[\widehat{m} > G_n^{\star+}\right] \leqslant \exp\left[-\frac{5\,(1 \vee r)\, m_n^\star}{9L} + \log\left(G_n\right)\right],$$
$$\mathbb{P}^n_{\theta^\circ}\left[\widehat{m} < G_n^{\star-}\right] \leqslant \exp\left[-\frac{7\,(1 \vee r)\, m_n^\star}{9} + \log\left(G_n\right)\right].$$

D.2 Detailed proofs

PROOF OF PROPOSITION D.1.1
Let be m in $[\![1, G_n]\!]$ and note that

$$\|\overline{\theta}^m - \theta^\circ\|^2 = \sum_{j=1}^m \left(\frac{Y_j}{\lambda_j} - \theta_j^\circ\right)^2 + \mathfrak{b}_m,$$

hence, we will use LEMMA A.1.1.

We then define

$$
\begin{aligned}
& S_m := \sum_{j=1}^{m}\left(\frac{Y_j}{\lambda_j}-\theta^\circ_j\right)^2, \\
& \mu_m := \mathbb{E}^n_{\theta^\circ}\left[S_m\right] = \frac{m\overline{\Lambda}_m}{n}, \\
\forall j \in [\![1,m]\!] \quad & \beta_j^2 := \mathbb{V}^n_{\theta^\circ}\left[\frac{Y_j}{\lambda_j}-\theta^\circ_j\right] = \frac{\Lambda_j}{n}, \\
\forall j \in [\![1,m]\!] \quad & \alpha_j^2 := \mathbb{E}^n_{\theta^\circ}\left[\frac{Y_j}{\lambda_j}-\theta^\circ_j\right] = 0, \\
& v_m := \sum_{j=1}^{m}\beta_j^2 = \frac{m\overline{\Lambda}_m}{n}, \\
& t_m := \max_{j\in[\![1,m]\!]}\beta_j^2 = \frac{\Lambda_m}{n}.
\end{aligned}
$$

We then control the concentration of S_m, define c a constant greater than $\frac{3}{2}$. With the following facts

- for any a and b in $\mathbb{R}_+$ we have $a \vee b \geqslant \frac{1}{2}(a+b)$
- $c > 1$
- Lemma A.1.1

$$
\begin{aligned}
\mathbb{P}^n_{\theta^\circ}\left[S_m + \mathfrak{b}_m \geqslant 4c\left(\frac{m\overline{\Lambda}_m}{n} \vee \mathfrak{b}_m\right)\right] & \leqslant \mathbb{P}^n_{\theta^\circ}\left[S_m + \mathfrak{b}_m \geqslant 2c\frac{m\overline{\Lambda}_m}{n} + \mathfrak{b}_m\right] \\
& \leqslant \mathbb{P}^n_{\theta^\circ}\left[S_m \geqslant 2c\frac{m\overline{\Lambda}_m}{n}\right] \\
& \leqslant \mathbb{P}^n_{\theta^\circ}\left[S_m - \mu_m \geqslant c\frac{m\overline{\Lambda}_m}{n}\right] \\
& \leqslant \exp\left[-\frac{cnm\overline{\Lambda}_m}{6n\Lambda_m}\right] \\
& \leqslant \exp\left[-\frac{cm}{6L}\right].
\end{aligned}
$$

Proof of Proposition D.1.2

First, let's proof the first inequality. Use the fact that :

$$
\begin{aligned}
&\mathbb{P}^n_{\theta^\circ}\left[G_n^{\star +} < \widehat{m} \leqslant G_n\right] \\
&= \mathbb{P}^n_{\theta^\circ}\left[\forall l \in [\![1, G_n^{\star +}]\!], \quad \frac{3\widehat{m}}{n} - \sum_{j=1}^{\widehat{m}} Y_j^2 < \frac{3l}{n} - \sum_{j=1}^{l} Y_j^2\right] \\
&\leqslant \mathbb{P}^n_{\theta^\circ}\left[\exists m \in [\![G_n^{\star +} + 1, G_n]\!] : \quad \frac{3m}{n} - \sum_{j=1}^{m} Y_j^2 < \frac{3m_n^\star}{n} - \sum_{j=1}^{m_n^\star} Y_j^2\right] \\
&\leqslant \sum_{m=G_n^{\star +}+1}^{G_n} \mathbb{P}^n_{\theta^\circ}\left[\frac{3m}{n} - \sum_{j=1}^{m} Y_j^2 < \frac{3m_n^\star}{n} - \sum_{j=1}^{m_n^\star} Y_j^2\right] \\
&\leqslant \sum_{m=G_n^{\star +}+1}^{G_n} \mathbb{P}^n_{\theta^\circ}\left[0 < 3\frac{m_n^\star - m}{n} + \sum_{j=m_n^\star+1}^{m} Y_j^2\right]
\end{aligned}
$$

We will now use Lemma A.1.1. For this purpose, define then for all m in $[\![G_n^{\star +} + 1, G_n]\!]$: $S_m := \sum_{j=m_n^\star+1}^{m} Y_j^2$, we then have $\mu_m := \mathbb{E}^n_{\theta^\circ}[S_m] = \frac{m - m_n^\star}{n} + \sum_{j=m_n^\star+1}^{m} \left(\theta_j^\circ \lambda_j\right)^2$, $\alpha_j^2 := \mathbb{E}^n_{\theta^\circ}[Y_j]^2 = \left(\theta_j^\circ \lambda_j\right)^2$ and $\beta_j^2 := \mathbb{V}^n_{\theta^\circ}[Y_j] = \frac{1}{n}$.

Now we note, using the definition of $\Theta^{\mathfrak{a}}(r)$

$$
\begin{aligned}
\sum_{j=m_n^\star+1}^{m} \alpha_j^2 &= \sum_{j=m_n^\star+1}^{m} \left(\theta_j^\circ \lambda_j\right)^2 \\
&\leqslant \Lambda_{m_n^\star}^{-1} \sum_{j=m_n^\star+1}^{m} \left(\theta_j^\circ\right)^2 \\
&\leqslant \Lambda_{m_n^\star}^{-1} \mathfrak{b}_{m_n^\star} \\
&\leqslant \Lambda_{m_n^\star}^{-1} (1 \vee r)\, \Phi_n^\star && =: r_m \\
\sum_{j=m_n^\star}^{m} \beta_j^2 &= \frac{m - m_n^\star}{n} && =: v_m \\
\max_{j \in [\![m_n^\star, m]\!]} \beta_j &= \frac{1}{n} && =: t_m.
\end{aligned}
$$

Hence, we have, for all m in $[\![G_n^{\star+}+1, G_n]\!]$

$$\begin{aligned}
&\mathbb{P}_{\theta^\circ}^n\left[\sum_{j=m_n^\star+1}^{m} Y_j^2 - 3\frac{m-m_n^\star}{n} > 0\right] \\
&= \mathbb{P}_{\theta^\circ}^n\left[S_m - \frac{m-m_n^\star}{n} > 2\frac{m-m_n^\star}{n}\right] \\
&= \mathbb{P}_{\theta^\circ}^n\left[S_m - \frac{m-m_n^\star}{n} - \sum_{j=m_n^\star+1}^{m}\left(\theta_j^\circ\lambda_j\right)^2 > 2\frac{m-m_n^\star}{n} - \sum_{j=m_n^\star+1}^{m}\left(\theta_j^\circ\lambda_j\right)^2\right] \\
&\leqslant \mathbb{P}_{\theta^\circ}^n\left[S_m - \mu_m > 2\frac{m-m_n^\star}{n} - \Lambda_{m_n^\star}^{-1}\mathfrak{b}_{m_n^\star}\right] \\
&\leqslant \mathbb{P}_{\theta^\circ}^n\left[S_m - \mu_m > 2\frac{m-m_n^\star}{n} - \Lambda_{m_n^\star}^{-1}(1\vee r)\,\Phi_n^\star\right].
\end{aligned}$$

Using the definition of $G_n^{\star+}$, we have $\frac{m-m_n^\star}{n} > 3\Lambda_{m_n^\star}^{-1}(1\vee r)\,\Phi_n^\star$.

Hence, we can write :

$$\begin{aligned}
&\mathbb{P}_{\theta^\circ}^n\left[\sum_{j=m_n^\star}^{m} Y_j^2 - 3\frac{m-m_n^\star}{n} > 0\right] \\
&\leqslant \mathbb{P}_{\theta^\circ}^n\left[S_m - \mu_m > \frac{m-m_n^\star}{n} + 2\Lambda_{m_n^\star}^{-1}(1\vee r)\,\Phi_n^\star\right] \\
&\leqslant \mathbb{P}_{\theta^\circ}^n\left[S_m - \mu_m > v_m + 2r_m\right] \\
&\leqslant \exp\left[-n\frac{\frac{m-m_n^\star}{n} + 2\Lambda_{m_n^\star}^{-1}(1\vee r)\,\Phi_n^\star}{9}\right] \\
&\leqslant \exp\left[-n\frac{5\Lambda_{m_n^\star}^{-1}(1\vee r)\,\Phi_n^\star}{9}\right] \\
&\leqslant \exp\left[-\frac{5(1\vee r)\,m_n^\star}{9L}\right].
\end{aligned}$$

Finally we can conclude that

$$\mathbb{P}_{\theta^\circ}^n\left[G_n^{\star+} < \widehat{m} \leqslant G_n\right] \leqslant \exp\left[-\frac{5(1\vee r)\,m_n^\star}{9L} + \log(G_n)\right].$$

We now prove the second inequality.

We begin by writing the same kind of inclusion of events as for the first inequality :

$$\begin{aligned}
\mathbb{P}^n_{\theta^\circ}\left[1 \leqslant \widehat{m} < G_n^{\star -}\right] =& \mathbb{P}^n_{\theta^\circ}\left[\forall m \in [\![G_n^-, G_n]\!], \quad 3\frac{\widehat{m}}{n} - \sum_{j=1}^{\widehat{m}} Y_j^2 < 3\frac{m}{n} - \sum_{j=1}^{m} Y_j^2\right] \\
\leqslant& \mathbb{P}^n_{\theta^\circ}\left[\exists m \in [\![1, G_n^{\star -} - 1]\!], \quad 3\frac{m}{n} - \sum_{j=1}^{m} Y_j^2 < 3\frac{m_n^\star}{n} - \sum_{j=1}^{m_n^\star} Y_j^2\right] \\
\leqslant& \sum_{m=1}^{G_n^{\star -}} \mathbb{P}^n_{\theta^\circ}\left[3\frac{m}{n} - \sum_{j=1}^{m} Y_j^2 < 3\frac{m_n^\star}{n} - \sum_{j=1}^{m_n^\star} Y_j^2\right] \\
\leqslant& \sum_{m=1}^{G_n^{\star -}} \mathbb{P}^n_{\theta^\circ}\left[\sum_{j=m+1}^{m_n^\star} Y_j^2 < 3\frac{m_n^\star - m}{n}\right].
\end{aligned}$$

The LEMMA A.1.1 steps in again, define $S_m := \sum_{j=m+1}^{m_n^\star} Y_j^2$ and we want to control the concentration of this sum, hence we take the following notations :

$$\begin{aligned}
\mu_m :=& \mathbb{E}^n_{\theta^\circ}\left[S_m\right] \\
=& \frac{m_n^\star - m}{n} + \sum_{j=m+1}^{m_n^\star} \left(\theta_j^\circ \lambda_j\right)^2 \\
r_m :=& \sum_{j=m+1}^{m_n^\star} \left(\theta_j^\circ \lambda_j\right)^2 \\
v_m :=& \frac{m_n^\star - m}{n} \\
t_m :=& \frac{1}{n}.
\end{aligned}$$

Hence, we have

$$\begin{aligned}
&\mathbb{P}^n_{\theta^\circ}\left[S_m < 3\frac{m^\star_n - m}{n}\right] \\
&\quad = \mathbb{P}^n_{\theta^\circ}\left[S_m - \mu_m < 3\frac{m^\star_n - m}{n} - \frac{m^\star_n - m}{n} - \sum_{j=m+1}^{m^\star_n}\left(\theta^\circ_j\lambda_j\right)^2\right] \\
&\quad = \mathbb{P}^n_{\theta^\circ}\left[S_m - \mu_m < 3\frac{m^\star_n - m}{n} - \frac{2}{3}\frac{m^\star_n - m}{n} - \frac{1}{3}\sum_{j=m+1}^{m^\star_n}\left(\theta^\circ_j\lambda_j\right)^2 - \frac{1}{3}\left[v_m + 2r_m\right]\right] \\
&\quad \leqslant \mathbb{P}^n_{\theta^\circ}\left[S_m - \mu_m < \frac{7}{3}\frac{m^\star_n - m}{n} - \frac{1}{3}\Lambda^{-1}_{m^\star_n}\sum_{j=m+1}^{m^\star_n}\left(\theta^\circ_j\right)^2 - \frac{1}{3}\left[v_m + 2r_m\right]\right] \\
&\quad \leqslant \mathbb{P}^n_{\theta^\circ}\left[S_m - \mu_m < -\frac{1}{3}\left[v_m + 2r_m\right] + \frac{7}{3}\frac{m^\star_n}{n} + \frac{1}{3}\Lambda^{-1}_{m^\star_n}\mathfrak{b}_{m^\star_n} - \frac{1}{3}\Lambda^{-1}_{m^\star_n}\mathfrak{b}_m\right] \\
&\quad \leqslant \mathbb{P}^n_{\theta^\circ}\left[S_m - \mu_m < -\frac{1}{3}\left[v_m + 2r_m\right] + 3\left(1 \vee r\right)\Phi^\star_n\overline{\Lambda}^{-1}_{m^\star_n} - \frac{1}{3}\Lambda^{-1}_{m^\star_n}\mathfrak{b}_m\right]
\end{aligned}$$

now, using the definition of G^-_n, we have $\mathfrak{b}_m > 9L\left(1 \vee r\right)\Phi^\star_n$ so

$$\begin{aligned}
&\mathbb{P}^n_{\theta^\circ}\left[S_m < 3\frac{m^\star_n - m}{n}\right] \\
&\quad \leqslant \mathbb{P}^n_{\theta^\circ}\left[S_m - \mu_m < -\frac{1}{3}\left[v_m + 2r_m\right]\right] \\
&\quad \leqslant \exp\left[-\frac{n}{36}\left(\frac{m^\star_n - m}{n} + 2\sum_{j=m+1}^{m^\star_n}\left(\theta^\circ_j\lambda_j\right)^2\right)\right] \\
&\quad \leqslant \exp\left[-n\frac{\frac{m^\star_n - m}{n} + 2\Lambda^{-1}_{m^\star_n}\mathfrak{b}_m - 2\Lambda^{-1}_{m^\star_n}\mathfrak{b}_{m^\star_n}}{36}\right] \\
&\quad \leqslant \exp\left[-n\frac{16L\left(1 \vee r\right)\Phi^\star_n\Lambda^{-1}_{m^\star_n}}{36}\right] \\
&\quad \leqslant \exp\left[-\frac{4\left(1 \vee r\right)m^\star_n}{9}\right]
\end{aligned}$$

Which in turn implies

$$\mathbb{P}^n_{\theta^\circ}\left[1 \leqslant \widehat{m} < G^{\star-}_n\right] \leqslant \exp\left[-\frac{4\left(1 \vee r\right)m^\star_n}{9} + \log\left(G_n\right)\right].$$

PROOF OF THEOREM 2.4.3

By the the total probability formula, we have :

$$\begin{aligned}&\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y^n}\left(\|\boldsymbol{\theta}^M-\theta^\circ\|^2\leqslant K^\star\Phi^\star_n\right)\right]\\&=1-\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y^n}\left(K^\star\Phi^\star_n<\|\boldsymbol{\theta}^M-\theta^\circ\|^2\right)\right].\end{aligned}$$

Hence, we will control $\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y^n}\left(K^\star\Phi^\star_n<\|\boldsymbol{\theta}^M-\theta^\circ\|^2\right)\right]$.

We can write :

$$\begin{aligned}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y^n}\left(K^\star\Phi^\star_n<\|\boldsymbol{\theta}-\theta^\circ\|^2\right)\right]&=\sum_{m=1}^{G_n}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y^n}\left(\left\{K^\star\Phi^\star_n<\|\boldsymbol{\theta}^M-\theta^\circ\|^2\right\}\cap\{M=m\}\right)\right]\\
&\leqslant\sum_{m=1}^{G_n^{\star-}-1}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\infty)}_{M|Y^n}(\{M=m\})\right]+\sum_{m=G_n^{\star+}+1}^{G_n}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y^n}(\{M=m\})\right]\\
&+\sum_{m=G_n^{\star-}}^{G_n^{\star+}}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y,M=m}\left(\left\{K^\star\Phi^\star_n<\|\boldsymbol{\theta}^M-\theta^\circ\|^2\right\}\right)\right]\\
&\leqslant\underbrace{\sum_{m=1}^{G_n^{\star-}-1}\mathbb{P}^n_{\theta^\circ}\left[\{\widehat{m}=m\}\right]}_{=:A}+\underbrace{\sum_{m=G_n^{\star+}+1}^{G_n}\mathbb{P}^n_{\theta^\circ}\left[\{\widehat{m}=m\}\right]}_{=:B}+\sum_{m=G_n^{\star-}}^{G_n^{\star+}}\underbrace{\mathbb{P}^n_{\theta^\circ}\left[\left\{K^\star\Phi^\star_n<\left\|\overline{\theta}^m-\theta^\circ\right\|^2\right\}\right]}_{=:C_m}.\end{aligned}$$

We control A and B using Proposition D.1.2.

Hence, we now control $\sum_{m=G_n^{\star-}}^{G_n^{\star+}}C_m$.

Using Assumption 13 we have

$$\left[a_{m_n^\star}\wedge\frac{m_n^\star\Lambda_{m_n^\star}}{n}\right]\leqslant\Phi^\star_n\leqslant\frac{\left[a_{m_n^\star}\wedge\frac{m_n^\star\Lambda_{m_n^\star}}{n}\right]}{\kappa^\star}.$$

Hence, for any m in $[\![m_n^\star,G_n^{\star+}]\!]$ we have, using the definition of $G_n^{\star+}$

$$\begin{aligned}m&\leqslant3\Lambda^{-1}_{m_n^\star}(1\vee r)\Phi^\star_n n+m_n^\star\\
&\leqslant\frac{3(1\vee r)n}{\Lambda_{m_n^\star}\kappa^\star}\left[\mathfrak{a}_{m_n^\star}\wedge\frac{m_n^\star\overline{\Lambda_{m_n^\star}}}{n}\right]+m_n^\star\\
&\leqslant3\frac{(1\vee r)}{\kappa^\star}\frac{m_n^\star\overline{\Lambda_{m_n^\star}}}{\Lambda_{m_n^\star}}+m_n^\star\\
&\leqslant\left(\frac{3(1\vee r)}{\kappa^\star L}+1\right)m_n^\star\\
&\leqslant D^\star m_n^\star;\end{aligned}$$

and

$$\begin{aligned}\overline{\Lambda_m} &\leqslant \Lambda_m \\ &\leqslant \Lambda_{D^\star m_n^\star} \\ &\leqslant \Lambda_{D^\star} \Lambda_{m_n^\star} \\ &\leqslant \Lambda_{D^\star} L \overline{\Lambda_{m_n^\star}};\end{aligned}$$

which give together

$$\frac{m\overline{\Lambda_m}}{n} \leqslant D^\star \Lambda_{D^\star} L \frac{m_n^\star \Lambda_{m_n^\star}}{n} \leqslant D^\star \Lambda_{D^\star} L \Phi_n^\star;$$

moreover, we have

$$\mathfrak{b}_m^2 \leqslant \mathfrak{b}_{m_n^\star}^2 \leqslant \Phi_n^\star$$

which leads to the conclusion

$$\left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n}\right] \leqslant D^\star \Lambda_{D^\star} (1 \vee r) \Phi_n^\star.$$

On the other hand, for and m in $[\![G_n^{\star-}, m_n^\star]\!]$, the definition of $G_n^{\star-}$ directly gives us

$$\left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n}\right] \leqslant 9L (1 \vee r) \Phi_n^\star.$$

Using Proposition D.1.1, we have that, for all m in $[\![G_n^{\star-}, G_n^{\star+}]\!]$ and c greater than $\frac{3}{2}$:

$$\mathbb{P}_{\theta^\circ}^n \left[\left\{\|\overline{\theta}^m - \theta^\circ\|^2 > 4c (9L \vee D^\star \Lambda_{D^\star}) (1 \vee r) \Phi_n^\star\right\}\right] \leqslant \exp\left[-\frac{cm}{6L}\right].$$

Hence, we set $K^\star := 6 (9L \vee D^\star \Lambda_{D^\star}) (1 \vee r)$, which leads us to the upper bound:

$$\sum_{m=G_n^{\star-}}^{G_n^{\star+}} C_m \leqslant \sum_{m=G_n^{\star-}}^{G_n^{\star+}} \exp\left[-\frac{3m}{8L}\right] \leqslant 4L \exp\left[-\frac{G_n^{\star-}}{4L}\right].$$

Finally, we can conclude :

$$\begin{aligned}&\mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{\boldsymbol{\theta}^M|Y^n}^{n,(\infty)} \left(\|\boldsymbol{\theta}^M - \theta^\circ\|^2 \leqslant K^\star \Phi_n^\star\right)\right] \\ &\quad \geqslant 1 - \exp\left[-\frac{5(1 \vee r) m_n^\star}{9L} + \log(G_n)\right] - \exp\left[-\frac{7(1 \vee r) m_n^\star}{9} + \log(G_n)\right] - 4L \exp\left[-\frac{G_n^{\star-}}{4L}\right].\end{aligned}$$

This proves the first part of the theorem for any θ° such that $G_n^{\star-}$ tends to infinity when n tends to ∞. In the opposite case, it means that there exist n° such that for all n larger than n°, $G_n^{\star-} = G_{n^\circ}^{\star-}$. This means that $n \mapsto \mathfrak{b}_{G_n^{\star-}}$ is constant function for n larger that n° but, by definition of $G_n^{\star-}$, we also have $\mathfrak{b}_{G_n^{\star-}} \leqslant 9 (1 \vee r) L \Phi_n^\star \to 0$ which leads to the conclusion that for all m greater than $G_{n^\circ}^{\star-}$, $\mathfrak{b}_m = 0$. Hence, for all m greater than $G_{n^\circ}^{\star-}$, we can write

$$\frac{K^\star \cdot \Phi_n^\star}{\left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n}\right]} = \frac{K^\star \Phi_n^\star}{\left[\frac{m\overline{\Lambda}_m}{n}\right]} \geqslant \frac{K^\star \frac{m_n^\star \overline{\Lambda}_{m_n^\star}}{n}}{\left[\frac{m\overline{\Lambda}_m}{n}\right]} \geqslant \frac{K^\star m_n^\star}{Lm} \geqslant 9D^\star \Lambda_{D^\star}(1 \vee r)\frac{m_n^\star}{m} \geqslant 1.$$

Hence, we set $c := \frac{9}{4}D^\star\Lambda_{D^\star}(1 \vee r)\frac{m_n^\star}{m}$ and can write in this case for all n larger than n° :

$$\begin{aligned}
\sum_{m=G_n^{\star-}}^{G_n^{\star+}} \mathbb{P}_{\theta^\circ}^n \left[K^\star\Phi_n^\star < \left\|\overline{\theta}^m - \theta^\circ\right\|^2\right] &\leqslant \sum_{m=G_n^{\star-}}^{G_n^{\star+}} \mathbb{P}_{\theta^\circ}^n \left[4c\left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n}\right] < \left\|\overline{\theta}^m - \theta^\circ\right\|^2\right] \\
&\leqslant \sum_{m=G_n^{\star-}}^{G_n^{\star+}} \exp\left[-\frac{cm}{6L}\right] \\
&\leqslant \sum_{m=G_n^{\star-}}^{G_n^{\star+}} \exp\left[-\frac{3D^\star\Lambda_{D^\star}(1\vee r)\, m_n^\star}{8L}\right] \\
&\leqslant \exp\left[-\frac{3D^\star\Lambda_{D^\star}(1\vee r)\, m_n^\star}{8L} + \log(G_n)\right].
\end{aligned}$$

We can hence conclude that

$$\begin{aligned}
\mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}^{n,(\infty)}_{\boldsymbol{\theta}^M|Y^n}\left(\|\boldsymbol{\theta}^M - \theta^\circ\|^2 \leqslant K^\star\Phi_n^\star\right)\right] &> 1 - \exp\left[-\frac{3D^\star\Lambda_{D^\star}(1\vee r)\, m_n^\star}{8L} + \log(G_n)\right] \\
&- \exp\left[-\frac{5m_n^\star}{9L} + \log(G_n)\right] - \exp\left[-\frac{7m_n^\star}{9} + \log(G_n)\right].
\end{aligned}$$

Hence, we have shown here that $\Phi_n^\star$ is an upper bound for the contraction rate under the quadratic risk. We will now use this to prove that it is also for the maximal risk.

Note that $K^\star\,\Phi_n^\star \geqslant 4\left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n}\right]$ for all m in $[\![G_n^{\star-}, G_n^{\star+}]\!]$. Hence, for any increasing function K_n such that $\lim_{n\to\infty} K_n = \infty$, we have

$$K_n\Phi_n^\star \geqslant 4\frac{K_n}{K^\star}\left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n}\right].$$

So, if we define $\tilde{n}^\circ$, the smallest integer such that $\frac{K_n}{K^\star} \geqslant 1$, we can apply Proposition D.1.1 and we have

$$\begin{aligned}
\sum_{m=G_n^{\star-}}^{G_n^{\star+}} \mathbb{P}_{\theta^\circ}^n \left[K_n\Phi_n^\star < \left\|\overline{\theta}^m - \theta^\circ\right\|^2\right] &\leqslant \sum_{m=G_n^{\star-}}^{G_n^{\star+}} \mathbb{P}_{\theta^\circ}^n \left[4\frac{K_n}{K^\star}\left[\mathfrak{b}_m \vee \frac{m\overline{\Lambda}_m}{n}\right] < \left\|\overline{\theta}^m - \theta^\circ\right\|^2\right] \\
&\leqslant \sum_{m=G_n^{\star-}}^{G_n^{\star+}} \exp\left[-\frac{4K_n m}{9K^\star L}\right] \\
&\leqslant \exp\left[-\frac{4K_n}{9K^\star L}\right].
\end{aligned}$$

We hence here have a uniform upper bound for the maximal risk which concludes the proof.

Appendix E

Proof for Theorem 3.2.1

E.1 Intermediate results

E.2 Detailed proofs

The L^2 risk can be written :

$$\mathbb{E}^n_{\theta^\circ}\left[\left\|\overline{\theta}^{\widehat{m}}-\theta^\circ\right\|^2\right]=\mathbb{E}^n_{\theta^\circ}\left[\sum_{j=1}^{G_n}\left(\overline{\theta}^{\widehat{m}}_j-\theta^\circ_j\right)^2\right]+\mathbb{E}^n_{\theta^\circ}\left[\sum_{j=G_n+1}^{\infty}\left(\theta^\circ_j\right)^2\right].$$

Together with

$$\forall j\in[\![1,G_n]\!],\quad \overline{\theta}^{\widehat{m}}_j-\theta^\circ_j=\left(\frac{Y_j}{\lambda_j}-\theta^\circ_j\right)\mathbb{1}_{\{\widehat{m}\in[\![j,G_n]\!]\}}+\theta^\circ_j\mathbb{1}_{\{\widehat{m}\in[\![1,j-1]\!]\}},$$

implies that

$$\mathbb{E}^n_{\theta^\circ}\left[\left\|\overline{\theta}^{\widehat{m}}-\theta^\circ\right\|^2\right]\leqslant\underbrace{\sum_{j=1}^{G_n}\mathbb{E}^n_{\theta^\circ}\left[\left(\frac{Y_j}{\lambda_j}-\theta^\circ_j\right)^2\mathbb{1}_{\{\widehat{m}\in[\![j,G_n]\!]\}}\right]}_{=:A}$$
$$+\underbrace{\sum_{j=1}^{G_n}\mathbb{E}^n_{\theta^\circ}\left[\left(\theta^\circ_j\right)^2\mathbb{1}_{\{\widehat{m}\in[\![1,j-1]\!]\}}\right]}_{=:B}+\underbrace{\sum_{j=G_n+1}^{\infty}\mathbb{E}^n_{\theta^\circ}\left[\left(\theta^\circ_j\right)^2\right]}_{=:C}.$$

We will now control each of the three parts of the sum using Lemma A.1.2 and Proposition C.1.1.

First, consider A and let be some positive real number p to be specified later.

Then we can write

$$
\begin{aligned}
A &\leqslant \sum_{j=1}^{G_n^+} \mathbb{E}_{\theta^\circ}^n \left[\left(\frac{Y_j}{\lambda_j} - \theta_j^\circ \right)^2 \right] + \sum_{j=G_n^+ +1}^{G_n} \mathbb{E}_{\theta^\circ}^n \left[\left(\frac{Y_j}{\lambda_j} - \theta_j^\circ \right)^2 \mathbb{1}_{\{\widehat{m} \in [\![j, G_n]\!]\}} \right] \\
&\leqslant \sum_{j=1}^{G_n^+} \mathbb{E}_{\theta^\circ}^n \left[\left(\frac{Y_j}{\lambda_j} - \theta_j^\circ \right)^2 \right] + \sum_{j=G_n^+ +1}^{G_n} \mathbb{E}_{\theta^\circ}^n \left[\left(\frac{Y_j}{\lambda_j} - \theta_j^\circ \right)^2 \mathbb{1}_{\{\widehat{m} \in [\![G_n^+ +1, G_n]\!]\}} \right] \\
&\leqslant \sum_{j=1}^{G_n^+} \mathbb{E}_{\theta^\circ}^n \left[\left(\frac{Y_j}{\lambda_j} - \theta_j^\circ \right)^2 \right] + \sum_{j=G_n^+ +1}^{G_n} \mathbb{E}_{\theta^\circ}^n \left[\left(\frac{Y_j}{\lambda_j} - \theta_j^\circ \right)^2 \mathbb{1}_{\{\widehat{m} \in [\![G_n^+ +1, G_n]\!]\}} \right] \\
&\quad - p \mathbb{E}_{\theta^\circ}^n \left[\mathbb{1}_{\{\widehat{m} \in [\![G_n^+ +1, G_n]\!]} \right] + p \mathbb{E}_{\theta^\circ}^n \left[\mathbb{1}_{\{\widehat{m} \in [\![G_n^+ +1, G_n]\!]} \right] \\
&\leqslant \sum_{j=1}^{G_n^+} \mathbb{E}_{\theta^\circ}^n \left[\left(\frac{Y_j}{\lambda_j} - \theta_j^\circ \right)^2 \right] + \mathbb{E}_{\theta^\circ}^n \left[\left(\sum_{j=G_n^+ +1}^{G_n} \left(\frac{Y_j}{\lambda_j} - \theta_j^\circ \right)^2 - p \right) \mathbb{1}_{\{\widehat{m} \in [\![G_n^+ +1, G_n]\!]\}} \right] \\
&\quad + p \mathbb{E}_{\theta^\circ}^n \left[\mathbb{1}_{\{\widehat{m} \in [\![G_n^+ +1, G_n]\!]} \right] \\
&\leqslant \underbrace{\sum_{j=1}^{G_n^+} \mathbb{E}_{\theta^\circ}^n \left[\left(\frac{Y_j}{\lambda_j} - \theta_j^\circ \right)^2 \right]}_{=:A_3} + \underbrace{\mathbb{E}_{\theta^\circ}^n \left[\left(\sum_{j=G_n^+ +1}^{G_n} \left(\frac{Y_j}{\lambda_j} - \theta_j^\circ \right)^2 - p \right)_+ \right]}_{=:A_1} + \underbrace{p \mathbb{E}_{\theta^\circ}^n \left[\mathbb{1}_{\{\widehat{m} \in [\![G_n^+ +1, G_n]\!]} \right]}_{=:A_2}.
\end{aligned}
$$

First, we will control A_1 using Lemma A.1.2.

The goal is to give p a value that is large enough to control this object but small enough so $p \cdot \mathbb{P}_{\theta^\circ}^n \left[G_n^+ < \widehat{m} \leqslant G_n \right]$ is still for the most Φ_n°.

Define $S_n := \sum_{j=G_n^+ +1}^{G_n} \left(\frac{Y_j}{\lambda_j} - \theta_j^\circ \right)^2$.

We have, for all j in $[\![G_n^+ + 1, G_n]\!]$,

$$
\frac{Y_j}{\lambda_j} - \theta_j^\circ \sim \mathcal{N}\left(0, \frac{\Lambda_j}{n} \right),
$$

so $\mathbb{E}_{\theta^\circ}^n \left[S_n \right] = \frac{1}{n} \sum_{j=G_n^+ +1}^{G_n} \Lambda_j$.

And define, using the definition of G_n and **??**

$$
\begin{aligned}
t_m &:= \Lambda_1 \geqslant \frac{\Lambda_{G_n}}{n} \geqslant \max_{j \in [\![G_n^+ +1, G_n]\!]} \frac{\Lambda_j}{n} \\
v_m &:= G_n \Lambda_1 \geqslant \frac{G_n \Lambda_{G_n}}{n} \geqslant \frac{1}{n} \sum_{j=G_n^+ +1}^{G_n} \Lambda_j.
\end{aligned}
$$

We can take $p = \mathbb{E}^n_{\theta^\circ}\left[S_n\right] + 3v_m$ which gives, using the definition of G_n^+ and $G_n > G_n^+$

$$\begin{aligned}
A_1 =& \mathbb{E}^n_{\theta^\circ}\left[\left(S_n - \mathbb{E}^n_{\theta^\circ}\left[S_n\right] - 3v_m\right)_+\right]\\
\leqslant& 6\Lambda_1 \exp\left[-\frac{G_n}{2}\right]\\
\leqslant& 6\Lambda_1 \exp\left[-\frac{nG_n}{2n}\right]\\
\leqslant& 6\Lambda_1 \exp\left[-n\frac{3\Lambda^{-1}_{m_n^\circ}\Phi_n^\circ}{2} - \frac{m_n^\circ}{2}\right]\\
A_1 \leqslant& 6\Lambda_1 \exp\left[-\frac{2m_n^\circ}{L}\right].
\end{aligned}$$

Thanks to Proposition C.1.1 it is easily shown that

$$A_2 < 4\Lambda_1 \exp\left[-\frac{5m_n^\circ}{9L} + 2\log\left(G_n\right)\right].$$

Finally, we control A_3. Using the definition of G_n^+ we have

$$\begin{aligned}
A_3 =& \sum_{j=1}^{G_n^+} \mathbb{E}^n_{\theta^\circ}\left[\left(\overline{\theta}_j - \theta_j^\circ\right)^2\right]\\
=& \sum_{j=1}^{G_n^+} \frac{\Lambda_j}{n}\\
=& \frac{1}{n} G_n^+ \overline{\Lambda}_{G_n^+}\\
A_3 \leqslant& \frac{\overline{\Lambda}_{G_n^+}}{\Lambda_{m_n^\circ}} 3\Phi_n^\circ.
\end{aligned}$$

Note that, using **??** and the definition of G_n^+, we have that $\frac{\overline{\Lambda}_{G_n^+}}{\Lambda_{m_n^\circ}}$ is bounded for n large enough; indeed with $D^\circ := \left\lceil \frac{3}{\kappa^\circ} + 1 \right\rceil$,

$$\begin{aligned}
&G_n^+ \leqslant \frac{3n\Phi_n^\circ}{\Lambda_{m_n^\circ}} + m_n^\circ \leqslant \frac{3nm_n^\circ \overline{\Lambda}_{m_n^\circ}}{\Lambda_{m_n^\circ} n\kappa^\circ} + m_n^\circ \leqslant \left(\frac{3}{\kappa^\circ} + 1\right) m_n^\circ \leqslant D^\circ m_n^\circ\\
&\Rightarrow \frac{\overline{\Lambda}_{G_n^+}}{\Lambda_{m_n^\circ}} \leqslant \frac{\overline{\Lambda}_{D^\circ \cdot m_n^\circ}}{\Lambda_{D^\circ \cdot m_n^\circ}} \cdot \frac{\Lambda_{D^\circ \cdot m_n^\circ}}{\Lambda_{m_n^\circ}} \leqslant \Lambda_{D^\circ}.
\end{aligned}$$

Hence, we have

$$A \leqslant 6\Lambda_1 \exp\left[-\frac{2m_n^\circ}{L}\right] + 4\Lambda_1 \exp\left[-\frac{5m_n^\circ}{9L} + 2\log\left(G_n\right)\right] + \Lambda_{D^\circ} 3\Phi_n^\circ.$$

Now we control B We use a decomposition similar to the one used for A :

$$\begin{aligned}
B \leqslant & \sum_{j=G_n^-+1}^{G_n} \left(\theta_j^\circ\right)^2 + \sum_{j=1}^{G_n^-} \mathbb{E}_{\theta^\circ}^n \left[\left(\theta_j^\circ\right)^2 \mathbb{1}_{\{\widehat{m}\in[\![1,j-1]\!]\}}\right] \\
\leqslant & \sum_{j=G_n^-+1}^{G_n} \left(\theta_j^\circ\right)^2 + \sum_{j=1}^{G_n^-} \mathbb{E}_{\theta^\circ}^n \left[\left(\theta_j^\circ\right)^2 \mathbb{1}_{\{\widehat{m}\in[\![1,G_n^--1]\!]\}}\right] \\
B \leqslant & \underbrace{\sum_{j=G_n^-+1}^{G_n} \left(\theta_j^\circ\right)^2}_{=:B_1} + \underbrace{\sum_{j=1}^{G_n^-} \left(\theta_j^\circ\right)^2 \mathbb{P}_{\theta^\circ}^n \left[1 \leqslant \widehat{m} < G_n^-\right]}_{=:B_2}.
\end{aligned}$$

First, notice that $B_1 + C = \mathfrak{b}_{G_n^-} \leqslant 9L\Phi_n^\circ$ by the definition of G_n^-.

To control B_2, we use the fact that θ° is square summable and the Proposition C.1.1 :

$$\begin{aligned}
B_2 = & \mathbb{P}_{\theta^\circ}^n \left[1 \leqslant \widehat{m} < G_n^-\right] \sum_{j=1}^{G_n^-} \left(\theta_j^\circ\right)^2 \\
\leqslant & \exp\left[-\frac{7m_n^\circ}{9} + \log\left(G_n\right)\right] \cdot \|\theta^\circ\|^2.
\end{aligned}$$

So we have

$$B + C \leqslant 9L\Phi_n^\circ + \|\theta^\circ\|^2 \cdot \exp\left[-\frac{7m_n^\circ}{9} + \log\left(G_n\right)\right].$$

Which leads to :

$$\begin{aligned}
\mathbb{E}_{\theta^\circ}^n \left[\left\|\overline{\theta}^{\widehat{m}} - \theta^\circ\right\|^2\right] \leqslant & 3\left(\Lambda_{D^\circ} + 3L\right)\Phi_n^\circ + \\
& \left(6\Lambda_1 \exp\left[-\frac{2m_n^\circ}{L} - \log\left(\Phi_n^\circ\right)\right] + 4\Lambda_1 \exp\left[-\frac{5m_n^\circ}{9L} + \log\left(\frac{G_n^2}{\Phi_n^\circ}\right)\right]\right. \\
& \left. + \|\theta^\circ\|^2 \cdot \exp\left[-\frac{7m_n^\circ}{9} + \log\left(\frac{G_n}{\Phi_n^\circ}\right)\right]\right)\Phi_n^\circ,
\end{aligned}$$

which proves that there exist C such that, for n large enough,

$$\mathbb{E}_{\theta^\circ}^n \left[\left\|\overline{\theta}^{\widehat{m}} - \theta^\circ\right\|^2\right] \leqslant C\Phi_n^\circ.$$

Appendix F

Proof for Theorem 3.2.2

F.1 Intermediate results

F.2 Detailed proofs

The L^2 risk can be written :

$$\mathbb{E}^n_{\theta^\circ}\left[\left\|\overline{\theta}^{\widehat{m}} - \theta^\circ\right\|^2\right] = \mathbb{E}^n_{\theta^\circ}\left[\sum_{j=1}^{G_n}\left(\overline{\theta}_j - \theta^\circ_j\right)^2\right] + \mathbb{E}^n_{\theta^\circ}\left[\sum_{j=G_n+1}^{\infty}\left(\theta^\circ_j\right)^2\right].$$

Together with

$$\forall j \in [\![1, G_n]\!], \quad \overline{\theta}^{\widehat{m}}_j - \theta^\circ_j = \left(\frac{Y_j}{\lambda_j} - \theta^\circ_j\right)\mathbb{1}_{\{\widehat{m}\in[\![j,G_n]\!]\}} + \theta^\circ_j\mathbb{1}_{\{\widehat{m}\in[\![1,j-1]\!]\}},$$

implies that

$$\begin{aligned}\mathbb{E}^n_{\theta^\circ}\left[\left\|\overline{\theta}^{\widehat{m}} - \theta^\circ\right\|^2\right] \leqslant & \underbrace{\sum_{j=1}^{G_n}\mathbb{E}^n_{\theta^\circ}\left[\left(\frac{Y_j}{\lambda_j} - \theta^\circ_j\right)^2\mathbb{1}_{\{\widehat{m}\in[\![j,G_n]\!]\}}\right]}_{=:A} \\ & + \underbrace{\sum_{j=1}^{G_n}\mathbb{E}^n_{\theta^\circ}\left[\left(\theta^\circ_j\right)^2\mathbb{1}_{\{\widehat{m}\in[\![1,j-1]\!]\}}\right]}_{=:B} + \underbrace{\sum_{j=G_n+1}^{\infty}\left(\theta^\circ_j\right)^2}_{=:C}.\end{aligned}$$

We will now control each of the three parts of the sum.

First, consider A and let be some positive real number p to be specified later.

Then we can write

$$
\begin{aligned}
A \leqslant& \sum_{j=1}^{G_n^{\star+}} \mathbb{E}_{\theta^\circ}^n\left[\left(\frac{Y_j}{\lambda_j}-\theta_j^\circ\right)^2\right] + \sum_{j=G_n^{\star+}+1}^{G_n} \mathbb{E}_{\theta^\circ}^n\left[\left(\frac{Y_j}{\lambda_j}-\theta_j^\circ\right)^2 \mathbb{1}_{\{\widehat{m}\in[\![j,G_n]\!]\}}\right]\\
\leqslant& \sum_{j=1}^{G_n^{\star+}} \mathbb{E}_{\theta^\circ}^n\left[\left(\frac{Y_j}{\lambda_j}-\theta_j^\circ\right)^2\right] + \sum_{j=G_n^{\star+}+1}^{G_n} \mathbb{E}_{\theta^\circ}^n\left[\left(\frac{Y_j}{\lambda_j}-\theta_j^\circ\right)^2 \mathbb{1}_{\{\widehat{m}\in[\![G_n^{\star+},G_n]\!]\}}\right]\\
\leqslant& \sum_{j=1}^{G_n^{\star+}} \mathbb{E}_{\theta^\circ}^n\left[\left(\frac{Y_j}{\lambda_j}-\theta_j^\circ\right)^2\right] + \sum_{j=G_n^{\star+}+1}^{G_n} \mathbb{E}_{\theta^\circ}^n\left[\left(\frac{Y_j}{\lambda_j}-\theta_j^\circ\right)^2 \mathbb{1}_{\{\widehat{m}\in[\![G_n^{\star+},G_n]\!]\}}\right]\\
&- p\mathbb{E}_{\theta^\circ}^n\left[\mathbb{1}_{\{\widehat{m}\in[\![G_n^{\star+},G_n]\!]}\right] + p\mathbb{E}_{\theta^\circ}^n\left[\mathbb{1}_{\{\widehat{m}\in[\![G_n^{\star+},G_n]\!]}\right]\\
\leqslant& \sum_{j=1}^{G_n^{\star+}} \mathbb{E}_{\theta^\circ}^n\left[\left(\frac{Y_j}{\lambda_j}-\theta_j^\circ\right)^2\right] + \mathbb{E}_{\theta^\circ}\left[\left(\sum_{j=G_n^{\star+}+1}^{G_n}\left(\frac{Y_j}{\lambda_j}-\theta_j^\circ\right)^2 - p\right)\mathbb{1}_{\{\widehat{m}\in[\![G_n^{\star+},G_n]\!]\}}\right]\\
&+ p\mathbb{E}_{\theta^\circ}^n\left[\mathbb{1}_{\{\widehat{m}\in[\![G_n^{\star+},G_n]\!]}\right]\\
\leqslant& \underbrace{\sum_{j=1}^{G_n^{\star+}} \mathbb{E}_{\theta^\circ}^n\left[\left(\frac{Y_j}{\lambda_j}-\theta_j^\circ\right)^2\right]}_{=:A_3} + \underbrace{\mathbb{E}_{\theta^\circ}^n\left[\left(\sum_{j=G_n^{\star+}+1}^{G_n}\left(\frac{Y_j}{\lambda_j}-\theta_j^\circ\right)^2 - p\right)_+\right]}_{=:A_1} + \underbrace{p\mathbb{E}_{\theta^\circ}^n\left[\mathbb{1}_{\{\widehat{m}\in[\![G_n^{\star+},G_n]\!]}\right]}_{=:A_2}.
\end{aligned}
$$

First, we will control A_1 using Lemma A.1.2.

The goal is to give p a value that is large enough to control this object but small enough so $p \cdot \mathbb{P}_{\theta^\circ}^n\left[G_n^{\star+} < \widehat{m} \leqslant G_n\right]$ is still for the most $\Psi_n^\star$.

Define $S_n := \sum_{j=G_n^{\star+}+1}^{G_n}\left(\frac{Y_j}{\lambda_j}-\theta_j^\circ\right)^2$.

We have, for all j in $[\![G_n^{\star+}+1, G_n]\!]$,

$$
\frac{Y_j}{\lambda_j}-\theta_j^\circ \sim \mathcal{N}\left(0, \frac{\Lambda_j}{n}\right),
$$

so $\mathbb{E}_{\theta^\circ}^n\left[S_n\right] = \frac{1}{n}\sum_{j=G_n^{\star+}+1}^{G_n}\Lambda_j$.

And define

$$
\begin{aligned}
t_m :=& \Lambda_1 &\geqslant \frac{\Lambda_{G_n}}{n} &\geqslant \max_{j\in[\![G_n^{\star+}+1,G_n]\!]}\frac{\Lambda_j}{n},\\
v_m :=& G_n\Lambda_1 &\geqslant \frac{G_n\Lambda_{G_n}}{n} &\geqslant \frac{1}{n}\sum_{j=G_n^{\star+}+1}^{G_n}\Lambda_j.
\end{aligned}
$$

We can take $p = \mathbb{E}^n_{\theta^\circ}[S_n] + 3v_m$ which gives

$$\begin{aligned}
A_1 =& \mathbb{E}^n_{\theta^\circ}\left[(S_n - \mathbb{E}^n_{\theta^\circ}[S_n] - 3v_m)_+\right] \\
\leqslant& 6\Lambda_1 \exp\left[-\frac{G_n}{2}\right] \\
\leqslant& 6\Lambda_1 \exp\left[-\frac{nG_n}{2n}\right] \\
\leqslant& 6\Lambda_1 \exp\left[-n\frac{3\Lambda^{-1}_{m^\star_n}(1\vee r)\Psi^\star_n}{2} - \frac{m^\star_n}{2}\right] \\
A_1 \leqslant& 6\Lambda_1 \exp\left[-\frac{2m^\star_n}{L}\right].
\end{aligned}$$

Thanks to Proposition C.1.2 it is easily shown that

$$A_2 \leqslant 4\Lambda_1 \exp\left[-\frac{5m^\star_n}{9L} + 2\log(G_n)\right].$$

Finally, we control A_3. Using the definition of $G^{\star+}_n$ we have

$$\begin{aligned}
A_3 =& \sum_{j=1}^{G^{\star+}_n} \mathbb{E}^n_{\theta^\circ}\left[\left(\overline{\theta}_j - \theta^\circ_j\right)^2\right] \\
=& \sum_{j=1}^{G^{\star+}_n} \frac{\Lambda_j}{n} \\
=& \frac{1}{n} G^{\star+}_n \overline{\Lambda}_{G^{\star+}_n} \\
A_3 \leqslant& \frac{\overline{\Lambda}_{G^{\star+}_n}}{\Lambda_{m^\star_n}} \Psi^\star_n \left(3(1\vee r) + \frac{\Lambda_{m^\star_n}}{\overline{\Lambda}_{m^\star_n}}\right).
\end{aligned}$$

Note that, using ?? and the definition of $G^{\star+}_n$, we have that $\frac{\overline{\Lambda}_{G^{\star+}_n}}{\Lambda_{m^\star_n}}$ is bounded for n large enough; indeed with $D^\star := \left\lceil \frac{3(1\vee r)}{\kappa^\star} + 1 \right\rceil$,

$$\begin{aligned}
G^{\star+}_n \leqslant& \frac{3(1\vee r)\Psi^\star_n n}{\Lambda_{m^\star_n}} + m^\star_n \leqslant \frac{3(1\vee r) m^\star_n \overline{\Lambda}_{m^\star_n} n}{n\Lambda_{m^\star_n}\kappa^\star} + m^\star_n \leqslant \left(\frac{3(1\vee r)}{\kappa^\star} + 1\right) m^\star_n \leqslant D^\star m^\star_n \\
\Rightarrow& \frac{\overline{\Lambda}_{G^{\star+}_n}}{\Lambda_{m^\star_n}} \leqslant \Lambda_{D^\star}.
\end{aligned}$$

Hence, we have

$$A \leqslant 6\Lambda_1 \exp\left[-\frac{2m^\star_n}{L}\right] + 4\Lambda_1 \exp\left[-\frac{5m^\star_n}{9L} + 2\log(G_n)\right] + 4\Lambda_{D^\star} L(1\vee r)\Psi^\star_n.$$

Now we control B. We use a similar decomposition to the one used for A :

$$
\begin{aligned}
B &\leqslant \sum_{j=G_n^{\star-}+1}^{G_n} \left(\theta_j^\circ\right)^2 + \sum_{j=1}^{G_n^{\star-}} \mathbb{E}_{\theta^\circ}^n \left[\left(\theta_j^\circ\right)^2 \mathbb{1}_{\{\widehat{m}\in[\![1,j-1]\!]\}}\right] \\
&\leqslant \sum_{j=G_n^{\star-}+1}^{G_n} \left(\theta_j^\circ\right)^2 + \sum_{j=1}^{G_n^{\star-}} \mathbb{E}_{\theta^\circ}^n \left[\left(\theta_j^\circ\right)^2 \mathbb{1}_{\left\{\widehat{m}\in[\![1,G_n^{\star-}-1]\!]\right\}}\right] \\
B &\leqslant \underbrace{\sum_{j=G_n^{\star-}+1}^{G_n} \left(\theta_j^\circ\right)^2}_{=:B_1} + \underbrace{\sum_{j=1}^{G_n^{\star-}} \left(\theta_j^\circ\right)^2 \mathbb{P}_{\theta^\circ}^n \left[1 \leqslant \widehat{m} < G_n^{\star-}\right]}_{=:B_2}.
\end{aligned}
$$

First, notice that $B_1 + C = \mathfrak{b}_{G_n^{\star-}} \leqslant 9L\,(1 \vee r)\,\Psi_n^\star$ by the definition of $G_n^{\star-}$.

To control B_2, we use the definition of the Sobolev ellipsoid and the Proposition C.1.2:

$$
\begin{aligned}
B_2 =& \mathbb{P}_{\theta^\circ}^n \left[1 \leqslant \widehat{m} < G_n^{\star-}\right] \sum_{j=1}^{G_n^{\star-}} \left(\theta_j^\circ\right)^2 \\
\leqslant& \exp\left[-\frac{7\,(1 \vee r)\,m_n^\star}{9} + \log\left(G_n\right)\right] \cdot \sum_{j=1}^{G_n^{\star-}} \frac{\mathfrak{a}_j}{\mathfrak{a}_j} \left(\theta_j^\circ\right)^2 \\
\leqslant& \exp\left[-\frac{7\,(1 \vee L^\circ)\,m_n^\star}{9} + \log\left(G_n\right)\right] \cdot \mathfrak{a}_1 \sum_{j=1}^{G_n^{\star-}} \frac{1}{\mathfrak{a}_j} \left(\theta_j^\circ\right)^2 \\
\leqslant& \exp\left[-\frac{7m_n^\star}{9} + \log\left(G_n\right)\right] \cdot \mathfrak{a}_1 r
\end{aligned}
$$

So we have

$$
B + C \leqslant 9L\,(1 \vee r)\,\Psi_n^\star + \mathfrak{a}_1 L^\circ \cdot \exp\left[-\frac{7m_n^\star}{9} + \log\left(G_n\right)\right].
$$

Which leads to :

$$
\begin{aligned}
\mathbb{E}_{\theta^\circ}^n \left[\left\|\overline{\theta}^{\widehat{m}} - \theta^\circ\right\|^2\right] \leqslant& L\,(1 \vee L^\circ)\,(4\Lambda_{D^\star} + 9)\,\Psi_n^\star + \\
& \left(6\Lambda_1 \exp\left[-\frac{2m_n^\star}{L} - \log\left(\Psi_n^\star\right)\right] + 4\Lambda_1 \exp\left[-\frac{5m_n^\star}{9L} + \log\left(\frac{G_n^2}{\Psi_n^\star}\right)\right] + \right. \\
& \left. \mathfrak{a}_1 r \cdot \exp\left[-\frac{7m_n^\star}{9} + \log\left(\frac{G_n}{\Psi_n^\star}\right)\right]\right) \Psi_n^\star,
\end{aligned}
$$

which proves that there exist K such that, for n large enough,

$$\mathbb{E}^n_{\theta^\circ}\left[\left\|\overline{\theta}^{\widehat{m}} - \theta^\circ\right\|^2\right] \leqslant C\Psi^\star_n.$$

Add remark stating that the bound is uniform over the ellipsoid hence rate for the maximal risk?

Appendix G

Proof for Theorem 3.4.1

G.1 Intermediate results

Lemma G.1.1. *Consider the aggregated OSE* $\widetilde{f} = \sum_{m=1}^{n} w_m \widetilde{f}m$ *with weights* $w_m \in [0,1]$, $m \in [\![1,n]\!]$, *satisfying* $\sum_{m=1}^{n} w_m = 1$. *For any* $m_- \in [\![1,n]\!]$ *and* $m_+ \in [\![1,n]\!]$ *holds*

$$\|\widetilde{f} - f\|_{L^2}^2 \leqslant 2\|\widetilde{f}m_+ - \theta^{\circ,m_+}\|_{L^2}^2 + 2\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_{m_-}^2(f)$$
$$+ 2\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathbb{P}_w([\![1,m_-[\![) + (24\kappa/n) \sum_{m\in]\!]m_+,n]\!]} \Delta_m^\Lambda w_m \mathbb{1}\{\|\widetilde{f}m - \theta^{\circ,m}\|_{L^2}^2 < 12\kappa\Delta_m^\Lambda/n\}$$
$$+2 \sum_{m\in]\!]m_+,n]\!]} (\|\widetilde{f}m - \theta^{\circ,m}\|_{L^2}^2 - \mathcal{C}_1 12\kappa\Delta_m^\Lambda/n)_+ + (24\mathcal{C}_1\kappa/n) \sum_{m\in]\!]m_+,n]\!]} \Delta_m^\Lambda \mathbb{1}\{\|\widetilde{f}m - \theta^{\circ,m}\|_{L^2}^2 \geqslant 12\kappa\Delta_m^\Lambda/n\}.$$

Lemma G.1.2. *If* $f = e_0$ *then there is a finite numerical constant* $\mathcal{C}$ *such that for all* $n \in \mathbb{N}$ *we have* $\mathbb{E}_Y^n\|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\Delta_{n_o}^\Lambda n^{-1}$ *with* $n_o := \lceil 15(\frac{300}{\sqrt{\kappa}})^4\rceil$.

Lemma G.1.3. *Assume there is* $K \in \mathbb{N}$ *with* $1 \geqslant \mathfrak{b}_{[K-1]}(f) > 0$ *and* $\mathfrak{b}_K(f) = 0$. *Set* $K_h := K \vee 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$, $c_f := \frac{2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 + 484\kappa}{\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 \mathfrak{b}_{[K-1]}^2(f)}$ *and* $n_{f,\Lambda} = \lceil c_f \Delta_{K_h}^\Lambda \vee 15(\frac{300}{\sqrt{\kappa}})^4\rceil$.
If $n \leqslant n_{f,\Lambda}$ *then let* $m_n^\bullet := K_h(\log n)$, *and otherwise if* $n > n_{f,\Lambda}$ *then let* $m_n^\bullet := \max\{m \in [\![K,n]\!] : c_f\, \Delta_m^\Lambda < n\}$ *where the defining set contains* K_h *and thus it is not empty. There is a finite numerical constant* $\mathcal{C}$ *such that for all* $n \in \mathbb{N}$ *holds*

$$\mathbb{E}_Y^n\|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\{\Delta_{n_{f,\Lambda}}^\Lambda + \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 n_{f,\Lambda} + \|\phi\|_{\ell^1}^2\}n^{-1} + 6\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\{\exp\big(\tfrac{-\kappa\psi_{m_n^\bullet}^\Lambda m_n^\bullet}{400\|\phi\|_{\ell^1}}\big) - \tfrac{1}{n}\}. \tag{G.1}$$

If there is $\widetilde{n}_{f,\Lambda} \in \mathbb{N}$ *such that for all* $n \geqslant \widetilde{n}_{f,\Lambda}$ *in addition* $\psi_{m_n^\bullet}^\Lambda m_n^\bullet \geqslant K_h(\log n)$ *holds true then*

$$\mathbb{E}_Y^n\|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\{\Delta_{[n_{f,\Lambda}\vee\widetilde{n}_{f,\Lambda}]}^\Lambda + \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 [n_{f,\Lambda} \vee \widetilde{n}_{f,\Lambda}] + \|\phi\|_{\ell^1}^2\}n^{-1}. \tag{G.2}$$

Lemma G.1.4. *Given* $m_+^\diamond, m_-^\diamond \in [\![1,n]\!]$ *let* m_+ *and* m_- *as in* (G.10). *Let* $\Delta_m^\Lambda = \psi_m^\Lambda m\Lambda_{(m)}$, $\sqrt{\psi_m^\Lambda} = \frac{\log(m\Lambda_{(m)}\vee(m+2))}{\log(m+2)} \geqslant 1$, $6\kappa \geqslant 2\log(3e)$ *and* $\kappa \geqslant 1$. *If* $\mathcal{R}_n^\diamond[m,f,\Lambda] = [\mathfrak{b}_m^2(f) \vee \Delta_m^\Lambda/n]$ *for any* $m \in [\![1,n]\!]$, *then*

(i) $\mathbb{P}_{\widetilde{w}}(\llbracket 1, m_- \llbracket) \leqslant [m_- - 1] \exp\big(-\kappa\kappa n \mathcal{R}_n^\diamond(m_-^\diamond, f, \Lambda) - \frac{\kappa \|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2}{4} n \mathfrak{b}_{[m_- - 1]}^2(f)\big)$
$+ \mathbb{1}\{\|\widetilde{f}m_-^\diamond - \theta^{\circ, m_-^\diamond}\|_{L^2}^2 \geqslant 12\kappa\Delta_{m_-^\diamond}^\Lambda / n\}$;

(ii) $\sum_{m=m_++1}^n \Delta_m^\Lambda \widetilde{w}_m \mathbb{1}\{\|\widetilde{f}m - \theta^{\circ,m}\|_{L^2}^2 < 12\kappa\Delta_m^\Lambda/n\} \leqslant \frac{1}{36\kappa^2\kappa^2} + \frac{1}{3\kappa\kappa}$.

Proposition G.1.1. *Let $\kappa \geqslant 1$, $\kappa \geqslant 1$, $\Delta_m^\Lambda = \psi_m^\Lambda m \Lambda_{(m)}$ with $\sqrt{\psi_m^\Lambda} = \frac{\log(m\Lambda_{(m)} \vee (m+2))}{\log(m+2)} \geqslant 1$, and $\mathcal{R}_n^\diamond[m, f, \Lambda] = [\mathfrak{b}_m^2(f) \vee \Delta_m^\Lambda n^{-1}]$ for $m \in \mathbb{N}$. Given $m_+^\diamond, m_-^\diamond \in \llbracket 1, n \rrbracket$ let m_- as in (G.10). If $m_+^\diamond \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ then there is an universal numerical constant $\mathcal{C}$ such that for all $n \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ holds*

$$\begin{aligned}\mathbb{E}_Y^n \|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\big\{[1 \vee \|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2]\mathcal{R}_n^\diamond(m_+^\diamond, f, \Lambda) + \|\phi\|_{\ell^1}^2 n^{-1}\big\} \\ + 2\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_{m_-}^2(f) + 6\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 \exp\big(\frac{-\kappa\psi_{m_-^\diamond}^\Lambda m_-^\diamond}{400\|\phi\|_{\ell^1}}\big) \\ + 2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 [m_- - 1]\exp\big(-\kappa\kappa n \mathcal{R}_n^\diamond(m_-^\diamond, f, \Lambda) - \frac{\kappa\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2}{4} n \mathfrak{b}_{[m_- - 1]}^2(f)\big).\end{aligned}$$

Lemma G.1.5. *Let $\overline{\Lambda}_m = \frac{1}{m}\sum_{j \in \llbracket 1, m \rrbracket} \Lambda_j$, $\Lambda_{(m)} = \max\{\Lambda_j, j \in \llbracket 1, m \rrbracket\}$ and $\psi_m^\Lambda \geqslant 1$. Let $\kappa \geqslant 1$ and $\Delta_m^\Lambda = \psi_m^\Lambda m \Lambda_{(m)}$ with $\psi_m^\Lambda \geqslant 1$ then there is a numerical constant $\mathcal{C}$ such that for all $n \in \mathbb{N}$ and $m \in \llbracket 1, n \rrbracket$ holds*

(i) $\mathbb{E}_Y^n\big(\|\widetilde{f}m - \theta^{\circ,m}\|_{L^2}^2 - 12\kappa\Delta_m^\Lambda/n\big)_+$
$$\leqslant \mathcal{C}\Big[\frac{\|\phi\|_{\ell^1}\,\Lambda_{(m)}}{n}\exp\big(\frac{-\kappa\psi_m^\Lambda m}{6\|\phi\|_{\ell^1}}\big) + \frac{4m^2\Lambda_{(m)}^2}{n^2}\exp\big(\frac{-\sqrt{n\kappa\psi_m^\Lambda}}{100}\big)\Big]$$

(ii) $\mathbb{P}_Y^n\big(\|\widetilde{f}m - \theta^{\circ,m}\|_{L^2}^2 \geqslant 12\kappa\Delta_m^\Lambda/n\big) \leqslant 3\Big[\exp\big(\frac{-\kappa\psi_m^\Lambda m}{400\|\phi\|_{\ell^1}}\big) + \exp\big(\frac{-\sqrt{n\kappa\psi_m^\Lambda}}{100}\big)\Big]$

Lemma G.1.6. *Consider random weight $\widetilde{w}_m$, $m \in \llbracket 1, n \rrbracket$ as in (30). For any $l \in \llbracket 1, n \rrbracket$ keeping $\mathcal{R}_n^\diamond(l, f, \Lambda) = [\mathfrak{b}_l^2(f) \vee \Delta_l^\Lambda n^{-1}]$ in mind*

(i) *for all $k \in \llbracket 1, l \llbracket$ we have*
$\widetilde{w}_m \mathbb{1}\{\|\widetilde{f}l - \theta^{\circ,l}\|_{L^2}^2 < 12\kappa\Delta_l^\Lambda/n\} \leqslant \exp\big(\kappa n\big\{-\frac{\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2}{2}\mathfrak{b}_m^2(f) + [120\kappa + \frac{\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2}{2}]\mathcal{R}_n^\diamond[l, f, \Lambda]\big\}\big)$

(ii) *for all $m \in \rrbracket l, n \rrbracket$ we have*
$\widetilde{w}_m \mathbb{1}\{\|\widetilde{f}m - \theta^{\circ,m}\|_{\mathbb{H}}^2 < 12\kappa\Delta_m^\Lambda/n\} \leqslant \exp\big(\kappa n\big\{-12\kappa\Delta_m^\Lambda/n + [\frac{3}{2}\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 + 54\kappa]\mathcal{R}_n^\diamond[l, f, \Lambda]\big\}\big)$.

Corollary G.1.1. *Given $n \in \mathbb{N}$ and $\theta^\circ, \breve{f} \in L^2$ consider the families of orthogonal projections $\{\breve{f}^m = \Pi_{\mathbb{U}_m}\breve{f}, m \in \llbracket 1, n \rrbracket\}$ and $\{\theta^{\circ,m} = \Pi_{\mathbb{U}_m} f, m \in \llbracket 1, n \rrbracket\}$. If $\|\Pi_{\mathbb{U}_m}^\perp f\|_{L^2}^2 = \|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 \mathfrak{b}_m^2(f)$ for all $m \in \llbracket 1, n \rrbracket$, then for any $l \in \llbracket 1, n \rrbracket$ holds*

(i) $\|\theta^{\circ,k}\|_{L^2}^2 - \|\breve{f}^l\|_{L^2}^2 \leqslant \frac{11}{2}\|\breve{f}^l - \theta^{\circ,l}\|_{L^2}^2 - \frac{1}{2}\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\{\mathfrak{b}_k^2(f) - \mathfrak{b}_l^2(f)\}$, *for all $k \in \llbracket 1, l \llbracket$;*

(ii) $\|\breve{f}^k\|_{L^2}^2 - \|\breve{f}^l\|_{L^2}^2 \leqslant \frac{7}{2}\|\breve{f}^k - \theta^{\circ,k}\|_{L^2}^2 + \frac{3}{2}\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\{\mathfrak{b}_l^2(f) - \mathfrak{b}_k^2(f)\}$, *for all $k \in \rrbracket l, n \rrbracket$.*

Lemma G.1.7. *Let $\Delta_m^\Lambda = \psi_m^\Lambda m \Lambda_{(m)}$ with $\sqrt{\psi_m^\Lambda} = \frac{\log(m\Lambda_{(m)} \vee (m+2))}{\log(m+2)} \geqslant 1$, then for all $n \in \mathbb{N}$ and $m_+ \in \llbracket 1, n \llbracket$*

(i) *if* $m_+ \geqslant 3(\frac{12\|\phi\|_{\ell^1}}{\kappa})^2$ *and* $n \geqslant 15(\frac{200}{\sqrt{\kappa}})^4$ *(alternatively* $\sqrt{\kappa}/200 \geqslant \sqrt{3}$*) then* $\sum_{m=1+m_+}^{n} \mathbb{E}_Y^n\big(\|\widetilde{f}m - \theta^{\circ,m}\|_{L^2}^2 - 12\kappa\Delta_m^\Lambda/n\big)_+ \leqslant \mathcal{C}\big[\frac{12\|\phi\|_{\ell^1}^2}{\kappa} + 4\big]n^{-1}$

(ii) *if* $m_+ \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ *and* $n \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ *(alternatively* $\sqrt{\kappa}/300 \geqslant \sqrt{3}$*) then* $\sum_{m=1+m_+}^{n} \Delta_m^\Lambda \mathbb{P}_Y^n\big(\|\widetilde{f}m - \theta^{\circ,m}\|_{L^2}^2 \geqslant 12\kappa\Delta_m^\Lambda/n\big) \leqslant 3\big[(\frac{400\|\phi\|_{\ell^1}}{\kappa})^2 + \frac{800\|\phi\|_{\ell^1}}{\kappa} + 1\big]$.

G.2 Detailed proofs

Proof of Lemma G.1.1.
We start the proof with the observation that $\overline{[\widetilde{f}]}_j - \overline{[f]}_j = \widetilde{f}_{-j} - \theta^\circ_{-j}$ for all $j \in \mathbb{Z}$ and

$$\widetilde{f}_j - \theta^\circ_j = g_j^{-1}(\widehat{[h]}_j - \phi_j)\mathbb{P}_w(\llbracket j, n\rrbracket) - \theta^\circ_j\mathbb{P}_w(\llbracket 1, j\llbracket) \text{ for all } j \in \llbracket 1, n\rrbracket,$$
$$\widetilde{f}_0 - \theta^\circ_0 = 0 \text{ and } \widetilde{f}_j - \theta^\circ_j = -\theta^\circ_j \text{ for all } j > n.$$

Consequently, we have

$$\begin{aligned}\|\widetilde{f} - f\|_{L^2}^2 &= 2\sum_{j\in\llbracket 1,n\rrbracket} |g_j^{-1}(\widehat{[h]}_j - \phi_j)\mathbb{P}_w(\llbracket j, n\rrbracket) - \theta^\circ_j\mathbb{P}_w(\llbracket 1, j\llbracket)|^2 + 2\sum_{j>n}|\theta^\circ_j|^2 \\ &\leqslant \sum_{j\in\llbracket 1,n\rrbracket} 4\{\Lambda_j|\widehat{[h]}_j - \phi_j|^2\mathbb{P}_w(\llbracket j, n\rrbracket)\} + \sum_{j\in\llbracket 1,n\rrbracket} 4|\theta^\circ_j|^2\mathbb{P}_w(\llbracket 1, j\llbracket) + 2\sum_{j>n}|\theta^\circ_j|^2, \quad \text{(G.3)}\end{aligned}$$

where we consider the first r.h.s and the two other r.h.s. terms separatly. Consider the first r.hs. term in (G.3). We split the sum into two parts which we bound separately. Precisely,

$$
\begin{aligned}
&2\sum_{j\in[\![1,n]\!]}\Lambda_j|\widehat{[h]}_j-\phi_j|^2\mathbb{P}_w([\![j,n]\!])\\
&\qquad\leqslant 2\sum_{j\in[\![1,m_+]\!]}\Lambda_j|\widehat{[h]}_j-\phi_j|^2+2\sum_{j\in]\!]m_+,n]\!]}\Lambda_j|\widehat{[h]}_j-\phi_j|^2\sum_{l\in[\![j,n]\!]}w_l\\
&\qquad= 2\sum_{j\in[\![1,m_+]\!]}\Lambda_j|\widehat{[h]}_j-\phi_j|^2+\sum_{l\in]\!]m_+,n]\!]}w_l\,2\sum_{j\in]\!]m_+,l]\!]}\Lambda_j|\widehat{[h]}_j-\phi_j|^2\\
&\qquad\leqslant \|\widetilde{f}m_+-\theta^{\circ,m_+}\|^2_{L^2}+\sum_{l\in]\!]m_+,n]\!]}w_l\|\widetilde{f}l-\theta^{\circ,l}\|^2_{L^2}\\
&\qquad\leqslant \|\widetilde{f}m_+-\theta^{\circ,m_+}\|^2_{L^2}+\sum_{l\in]\!]m_+,n]\!]}w_l\|\widetilde{f}l-\theta^{\circ,l}\|^2_{L^2}\mathbb{1}\{\|\widetilde{f}l-\theta^{\circ,l}\|^2_{L^2}\geqslant 12\kappa\Delta_l^\Lambda/n\}\\
&\qquad\qquad+(12\kappa/n)\sum_{l\in]\!]m_+,n]\!]}\Delta_l^\Lambda w_l\mathbb{1}\{\|\widetilde{f}l-\theta^{\circ,l}\|^2_{L^2}<12\kappa\Delta_l^\Lambda/n\}\\
&= \|\widetilde{f}m_+-\theta^{\circ,m_+}\|^2_{L^2}+\sum_{l\in]\!]m_+,n]\!]}w_l\Big(\|\widetilde{f}l-\theta^{\circ,l}\|^2_{L^2}-\mathcal{C}_1 12\kappa\Delta_l^\Lambda/n\Big)\mathbb{1}\{\|\widetilde{f}l-\theta^{\circ,l}\|^2_{L^2}\geqslant 12\kappa\Delta_l^\Lambda/n\}\\
&+(12\mathcal{C}_1\kappa/n)\sum_{l\in]\!]m_+,n]\!]}w_l\Delta_l^\Lambda\mathbb{1}\{\|\widetilde{f}l-\theta^{\circ,l}\|^2_{L^2}\geqslant 12\kappa\Delta_l^\Lambda/n\}+(12\kappa/n)\sum_{l\in]\!]m_+,n]\!]}\Delta_l^\Lambda w_l\mathbb{1}\{\|\widetilde{f}l-\theta^{\circ,l}\|^2_{L^2}<12\kappa\Delta_l^\Lambda/n\}\\
&\qquad\leqslant \|\widetilde{f}m_+-\theta^{\circ,m_+}\|^2_{L^2}+\sum_{l\in]\!]m_+,n]\!]}(\|\widetilde{f}l-\theta^{\circ,l}\|^2_{L^2}-\mathcal{C}_1 12\kappa\Delta_l^\Lambda/n)_+\\
&+(12\mathcal{C}_1\kappa/n)\sum_{l\in]\!]m_+,n]\!]}\Delta_l^\Lambda\mathbb{1}\{\|\widetilde{f}l-\theta^{\circ,l}\|^2_{L^2}\geqslant 12\kappa\Delta_l^\Lambda/n\}+(12\kappa/n)\sum_{l\in]\!]m_+,n]\!]}\Delta_l^\Lambda w_l\mathbb{1}\{\|\widetilde{f}l-\theta^{\circ,l}\|^2_{L^2}<12\kappa\Delta_l^\Lambda/n\}
\end{aligned}
\tag{G.4}
$$

Consider the second and third r.hs. term in (G.3). Splitting the first sum into two parts we obtain

$$
\begin{aligned}
&2\sum_{j\in[\![1,n]\!]}|\theta_j^\circ|^2\mathbb{P}_w([\![1,j[\![)+2\sum_{j>n}|\theta_j^\circ|^2\\
&\qquad\leqslant 2\sum_{j\in[\![1,m_-]\!]}|\theta_j^\circ|^2\mathbb{P}_w([\![1,j[\![)+2\sum_{j\in]\!]m_-,n]\!]}|\theta_j^{\circ 2}+2\sum_{j>n}|\theta_j^\circ|^2\\
&\qquad\qquad\leqslant \|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\{\mathbb{P}_w([\![1,m_-[\![)+\mathfrak{b}^2_{m_-}(f)\}
\end{aligned}
\tag{G.5}
$$

Combining (G.3) and the upper bounds (G.4) and (G.5) we obtain the assertion, which completes the proof. □

Proof of Lemma J.2.1.

Let $n_o := \lceil 15(\frac{300}{\sqrt{\kappa}})^4\rceil$. We destinguish for $n\in\mathbb{N}$ the following two cases cases, (a) $n\in[\![1,n_o[\![$ and (b) $n\geqslant n_o$.

Consider (a). We select $m_+=n\leqslant n_o$ and thus keeping in mind that $f=e_0$, and hence

$\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 = 0$ from Lemma G.1.1 follows for all $n \in [\![1, n_o[\![$

$$\mathbb{E}_Y^n \|\widetilde{f} - f\|_{L^2}^2 \leqslant 2\mathbb{E}_Y^n \|\widetilde{f}n - \theta^{\circ,n}\|_{L^2}^2 \leqslant 4n\overline{\Lambda}_n n^{-1} \leqslant 4n_o\overline{\Lambda}_{n_o} n^{-1} \leqslant 4\Delta_{n_o}^{\Lambda} n^{-1}. \tag{G.6}$$

Consider (b), i.e., $n \geqslant n_o$. We select $m_+^\diamond := n_o \in [\![1, n]\!]$. Note that $\|\phi\|_{\ell^1} = 1$ and hence, $m_+ \geqslant m_+^\diamond \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$. Therefore, for all $n \geqslant n_o \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ due to Proposition G.1.1 follows

$$\begin{aligned}\mathbb{E}_Y^n \|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\{[1 \vee \|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2]\mathcal{R}_n^\diamond(m_+^\diamond, f, \Lambda) + n^{-1}\}\\ + 2\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_{m_-}^2(f) + 6\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 \exp\big(\frac{-\kappa\psi_{m_-^\diamond}^{\Lambda} m_-^\diamond}{400\|\phi\|_{\ell^1}}\big)\\ + 2\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 [m_- - 1] \exp\big(- \kappa\kappa n\mathcal{R}_n^\diamond(m_-^\diamond, f, \Lambda) - \frac{\kappa\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2}{4} n\mathfrak{b}_{[m_- - 1]}^2(f)\big).\end{aligned}$$

Since $\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 = 0$, and thus $\mathcal{R}_n^\diamond(m_+^\diamond, f, \Lambda) = \Delta_{m_+^\diamond}^{\Lambda}/n = \Delta_{n_o}^{\Lambda}/n$, there is a numerical constant $\mathcal{C}$ such that $\mathbb{E}_Y^n \|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\Delta_{n_o}^{\Lambda} n^{-1}$ for all $n \geqslant n_o$. Combining the upper bounds for the two cases (a) and (b) we obtain the assertion which completes the proof. □

Proof of Lemma G.1.4.
Consider ***(i)***. From Lemma G.1.6 ***(i)*** with $l = m_-^\diamond$ follows for all $m < m_- \leqslant m_-^\diamond$, and hence $\mathfrak{b}_m \geqslant \mathfrak{b}_{m_- - 1}$ that

$$\begin{aligned}\widetilde{w}_m \mathbb{1}\{\|\widetilde{f}m_-^\diamond - \theta^{\circ,m_-^\diamond}\|_{L^2}^2 < 12\kappa\Delta_{m_-^\diamond}^{\Lambda}/n\} \leqslant \exp\big(\kappa n\{-\frac{\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2}{2}\mathfrak{b}_m^2(f) + [120\kappa + \frac{\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2}{2}]\mathcal{R}_n^\diamond(m_-^\diamond, f, \Lambda)\}\big)\\ = \exp\big(\kappa n\{\underbrace{-\frac{\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2}{4}\mathfrak{b}_m^2(f) + [121\kappa + \frac{\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2}{2}]\mathcal{R}_n^\diamond(m_-^\diamond, f, \Lambda)}_{\leqslant 0}\}\big)\\ \times \exp\big(\kappa n\{ - \kappa\mathcal{R}_n^\diamond(m_-^\diamond, f, \Lambda) - \frac{\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2}{4}\mathfrak{b}_m^2(f)\}\big)\\ \leqslant \exp\big(- \kappa n\kappa\mathcal{R}_n^\diamond(m_-^\diamond, f, \Lambda) - \kappa n\frac{\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2}{4}\mathfrak{b}_{m_- - 1}^2(f)\big)\end{aligned}$$

which in turn implies ***(i)***, that is,

$$\begin{aligned}\mathbb{P}_{\widetilde{w}}([\![1, m_-[\![) \leqslant \mathbb{P}_{\widetilde{w}}([\![1, m_-[\![)\mathbb{1}\{\|\widetilde{f}m_-^\diamond - \theta^{\circ,m_-^\diamond}\|_{L^2}^2 < 12\kappa\Delta_{m_-^\diamond}^{\Lambda}/n\} + \mathbb{1}\{\|\widetilde{f}m_-^\diamond - \theta^{\circ,m_-^\diamond}\|_{L^2}^2 \geqslant 12\kappa\Delta_{m_-^\diamond}^{\Lambda}/n\}\\ \leqslant [m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond(m_-^\diamond, f, \Lambda) - \kappa n\frac{\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2}{4}\mathfrak{b}_{[m_- - 1]}^2(f)\big) + \mathbb{1}\{\|\widetilde{f}m_-^\diamond - \theta^{\circ,m_-^\diamond}\|_{L^2}^2 \geqslant 12\kappa\Delta_{m_-^\diamond}^{\Lambda}/n\}\end{aligned}$$

Consider ***(ii)***. From Lemma G.1.6 ***(ii)*** with $l = m_+^\diamond$ follows for all $m > m_+ \geqslant m_+^\diamond$

$$\begin{aligned}\widetilde{w}_m \mathbb{1}\{\|\widetilde{f}m - \theta^{\circ,m}\|_{L^2}^2 < 12\kappa\Delta_m^{\Lambda}/n\} \leqslant \exp\big(\kappa n\{-12\kappa\Delta_m^{\Lambda}/n + [\tfrac{3}{2}\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 + 54\kappa]\mathcal{R}_n^\diamond(m_+^\diamond, f, \Lambda)\}\big)\\ = \exp\big(\kappa n\{ - \tfrac{1}{2} * 12\kappa\Delta_m^{\Lambda}/n \underbrace{- \tfrac{1}{2} * 12\kappa\Delta_m^{\Lambda}/n + [\tfrac{3}{2}\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 + 54\kappa]\mathcal{R}_n^\diamond(m_+^\diamond, f, \Lambda)}_{\leqslant 0}\}\big)\\ \leqslant \exp\big(\kappa n\{ - \tfrac{1}{2} * 12\kappa\Delta_m^{\Lambda}/n\}\big)\end{aligned}$$

which in turn with $\Delta^\Lambda_m = \psi^\Lambda_m m\Lambda_{(m)}$ implies

$$\sum_{m=m_++1}^{n} \Delta^\Lambda_m \widetilde{w}_m \mathbb{1}_{\{\|\widetilde{f}m - \theta^{\circ,m}\|^2_{L^2} \leqslant 12\kappa\Delta^\Lambda_m/n\}} \leqslant \sum_{k=m_++1}^{n} \Delta^\Lambda_m \exp\big(-\kappa 6\kappa \Delta^\Lambda_m\big)$$
$$= \sum_{k=m_++1}^{n} \psi^\Lambda_m m\Lambda_{(m)} \exp\big(-\kappa 6\kappa \psi^\Lambda_m m\Lambda_{(m)}\big) \quad (G.7)$$

Exploiting that $\sqrt{\psi^\Lambda_m} = \frac{\log(m\Lambda_{(m)} \vee (m+2))}{\log(m+2)} \geqslant 1$, $6\kappa \geqslant 2\log(3e)$ and $\kappa \geqslant 1$, then for all $k \in \mathbb{N}$ we have $6\kappa\kappa k - \log(k+2) \geqslant 1$, and hence by $a\exp(-ab) \leqslant \exp(-b)$ for $a, b \geqslant 1$, it follows

$$\psi^\Lambda_m m\Lambda_{(m)} \exp\big(-6\kappa\kappa\psi^\Lambda_m m\Lambda_{(m)}\big) \leqslant \psi^\Lambda_m \exp\big(-6\kappa\kappa\psi^\Lambda_m m\Lambda_{(m)} + \sqrt{\psi^\Lambda_m}\log(m+2)\big)$$
$$\leqslant \psi^\Lambda_m \exp\big(-\psi^\Lambda_m(6\kappa\kappa m - \log(m+2))\big) \leqslant \exp\big(-(6\kappa\kappa m - \log(m+2))\big)$$
$$= (m+2)\exp\big(-6\kappa\kappa m\big)$$

Since $\sum_{m\in\mathbb{N}} \mu m \exp(-\mu m) \leqslant \mu^{-1}$ und $\sum_{m\in\mathbb{N}} \mu \exp(-\mu m) \leqslant 1$ follows the assertion, that is

$$\sum_{k=m_++1}^{n} \psi^\Lambda_m m\Lambda_{(m)} \exp\big(-6\kappa\kappa\psi^\Lambda_m m\Lambda_{(m)}\big) \leqslant \sum_{k=m_++1}^{\infty} (m+2)\exp\big(-6\kappa\kappa m\big) \leqslant \tfrac{1}{36\kappa^2\kappa^2} + \tfrac{1}{3\kappa\kappa}$$

which completes the proof. □

PROOF OF LEMMA J.2.2.
Given $K \in \mathbb{N}$ with $1 \geqslant \mathfrak{b}_{[K-1]}(f) > 0$ and $\mathfrak{b}_m(f) = 0$ for all $m \geqslant K$ let $K_h := K \vee 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$, $c_f := \frac{2\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2} + 484\kappa}{\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2}\mathfrak{b}^2_{[K-1]}(f)}$ and $n_{f,\Lambda} = \lceil c_f \Delta^\Lambda_{K_h} \vee 15(\frac{300}{\sqrt{\kappa}})^4 \rceil$ we distinguish for $n \in \mathbb{N}$ the following two cases, (a) $n \in [\![1, n_{f,\Lambda}[\![$ and (b) $n > n_{f,\Lambda}$.

Firstly, consider (a), let $n \in [\![1, n_{f,\Lambda}[\![$, then setting $m_- = 1$ and $m_+ = n$ from Lemma G.1.1 follows

$$\mathbb{E}^n_Y \|\widetilde{f} - f\|^2_{L^2} \leqslant 2\mathbb{E}^n_Y \|\widetilde{f}n - \theta^{\circ,n}\|^2_{L^2} + 2\|\Pi_{\mathbb{U}^\perp_0} f\|^2_{L^2}\mathfrak{b}^2_1(f)$$
$$\leqslant 4n\overline{\Lambda}_n n^{-1} + 2\|\Pi_{\mathbb{U}^\perp_0} f\|^2_{L^2} \leqslant 4n_{f,\Lambda}\overline{\Lambda}_{n_{f,\Lambda}} n^{-1} + 2\|\Pi_{\mathbb{U}^\perp_0} f\|^2_{L^2} n_{f,\Lambda} n^{-1}$$
$$\leqslant (4\Delta^\Lambda_{n_{f,\Lambda}} + 2\|\Pi_{\mathbb{U}^\perp_0} f\|^2_{L^2} n_{f,\Lambda}) n^{-1}. \quad (G.8)$$

Secondly, consider (b), i.e., $n > n_{f,\Lambda}$. Setting $m^\diamond_+ := K_h \leqslant \Delta^\Lambda_{K_h} \leqslant n_{f,\Lambda}$, i.e., $m^\diamond_+ \in [\![1, n]\!]$ from $m^\diamond_+ = K_h \geqslant K$ follows $\mathfrak{b}_{m^\diamond_+}(f) = 0$ and hence $\mathcal{R}^\diamond_n(m^\diamond_+, f, \Lambda) = \Delta^\Lambda_{K_h} n^{-1}$. Keeping in mind that $m_+^\diamond \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ and $n \geqslant n_o \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ from Proposition G.1.1 follows

$$\begin{aligned}\mathbb{E}_Y^n\|\widetilde{f}-f\|_{L^2}^2 \leqslant \mathcal{C}\{[1\vee\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2]\Delta_{K_h}^{\Lambda}n^{-1}+\|\phi\|_{\ell^1}^2 n^{-1}\}\\ +2\|\Pi_{\mathbb{U}_0^{\perp}}f\|_{L^2}^2\mathfrak{b}_{m_-}^2(f)+6\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2\exp\big(\tfrac{-\kappa\psi_{m_-^\diamond}^{\Lambda}m_-^\diamond}{400\|\phi\|_{\ell^1}}\big)\\ +2\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2\,[m_- -1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond(m_-^\diamond,f,\Lambda)-\tfrac{\kappa\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2}{4}n\mathfrak{b}_{[m_- -1]}^2(f)\big).\quad (G.9)\end{aligned}$$

Since $n > n_{f,\Lambda} \geqslant c_f\Delta_{K_h}^{\Lambda}$ with $c_f = \frac{2\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2+484\kappa}{\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2\mathfrak{b}_{[K-1]}^2(f)}$ the defining set of $m_n^\bullet := \max\{m \in [\![K,n]\!] : n > c_{f,\Lambda}\Delta_m^{\Lambda}\}$ evenutally containing K_h is not empty. Consequently, $m_n^\bullet \geqslant K$ and $\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2\mathfrak{b}_{[K-1]}^2(f) > [2\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2+484\kappa]\Delta_{m_n^\bullet}^{\Lambda}/n = [2\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2+484\kappa]\mathcal{R}_n^\diamond(m_n^\bullet,f,\Lambda)$.

Keep in mind that for each $m\in\mathbb{N}$, $\|\Pi_{\mathbb{U}_m}^{\perp}f\|_{L^2}^2 = \|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2\mathfrak{b}_k^2(f)$ and introduce in addition $\mathcal{R}_n^\diamond[m,f,\Lambda] := [\mathfrak{b}_m^2(f)\vee\Delta_m^{\Lambda}n^{-1}]$. For any $m_+^\diamond, m_-^\diamond \in [\![1,n]\!]$ let us define

$$\begin{aligned} m_- &:= \min\Big\{k\in[\![1,m_-^\diamond]\!] : \|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2\mathfrak{b}_k^2(f) \leqslant [2\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2+484\kappa]\mathcal{R}_n^\diamond[m_-^\diamond,f,\Lambda]\Big\} \quad\text{and}\\ m_+ &:= \max\Big\{k\in[\![m_+^\diamond,n]\!] : 12\kappa\Delta_k^{\Lambda}/n \leqslant [3\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2+108\kappa]\mathcal{R}_n^\diamond[m_+^\diamond,f,\Lambda]\Big\} \quad (G.10)\end{aligned}$$

where the defining set obviously contains $m_-^\diamond$ and $m_+^\diamond$, respectively, and hence, they are not empty.

Therefore, setting $m_-^\diamond := m_n^\bullet$ the definition (G.10) implies $m_- = K$ and hence $\mathfrak{b}_{m_-}^2(f) = \mathfrak{b}_K^2(f) = 0$, $\mathfrak{b}_{[m_- -1]}^2(f) = \mathfrak{b}_{[K-1]}^2(f) > 0$. From (J.26) follows for all $n > n_{f,\Lambda}$ thus

$$\begin{aligned}\mathbb{E}_Y^n\|\widetilde{f}-f\|_{L^2}^2 \leqslant \mathcal{C}\{[1\vee\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2]\Delta_{K_h}^{\Lambda}n^{-1}+\|\phi\|_{\ell^1}^2 n^{-1}\}\\ +6\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2\exp\big(\tfrac{-\kappa\psi_{m_n^\bullet}^{\Lambda}m_n^\bullet}{400\|\phi\|_{\ell^1}}\big)\\ +2\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2\,m_n^\bullet\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond(m_-^\diamond,f,\Lambda)-\tfrac{\kappa\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2}{4}n\mathfrak{b}_{[K-1]}^2(f)\big)\\ \leqslant \mathcal{C}\{[1\vee\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2]\Delta_{K_h}^{\Lambda}n^{-1}+\|\phi\|_{\ell^1}^2 n^{-1}\}\\ +6\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2\exp\big(\tfrac{-\kappa\psi_{m_n^\bullet}^{\Lambda}m_n^\bullet}{400\|\phi\|_{\ell^1}}\big)\\ +2\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2\,[K-1]\underbrace{\exp\big(-\tfrac{\kappa\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2}{4}n\mathfrak{b}_{[m_- -1]}^2(f)\big)}_{\leqslant\frac{4}{\kappa\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2\mathfrak{b}_{[K-1]}^2(f)}n^{-1}\exp(-1)}\quad (G.11)\end{aligned}$$

Note that $\Delta_{K_h}^{\Lambda} \leqslant n_{f,\Lambda}$ and $\frac{8[K-1]}{e\kappa\mathfrak{b}_{[K-1]}^2(f)} \leqslant \frac{1}{\kappa}\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2 n_{f,\Lambda}$. Thereby, we obtain

$$\begin{aligned}\mathbb{E}_Y^n\|\widetilde{f}-f\|_{L^2}^2 \leqslant \mathcal{C}_2\{\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2 n_{f,\Lambda}+\|\phi\|_{\ell^1}^2\}n^{-1}\\ +6\|\Pi_{\mathbb{U}_0}^{\perp}f\|_{L^2}^2\{\exp\big(\tfrac{-\kappa\psi_{m_n^\bullet}^{\Lambda}m_n^\bullet}{400\|\phi\|_{\ell^1}}\big)-\tfrac{1}{n}\}\quad (G.12)\end{aligned}$$

for some finite numerical constant $\mathcal{C}_2$.
Combining the upper bounds (J.25) and (J.28) for the two cases (a) and (b) we obtain the assertion (J.23), that is, there is a finite numerical constant $\mathcal{C}$ such that for all $n \in \mathbb{N}$ holds

$$\mathbb{E}_Y^n \|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\{\Delta^\Lambda_{n_{f,\Lambda}} + \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 n_{f,\Lambda} + \|\phi\|_{\ell^1}^2\} n^{-1} \\ + 6\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 \{\exp\big(\tfrac{-\kappa\psi^\Lambda_{m_n^\bullet} m_n^\bullet}{400\|\phi\|_{\ell^1}}\big) - \tfrac{1}{n}\} \quad \text{(G.13)}$$

Assume finally, that there is in addition $\widetilde{n}_{f,\Lambda} \in \mathbb{N}$ such that $\psi^\Lambda_{m_n^\bullet} m_n^\bullet \geqslant K_h(\log n)$ for all $n \geqslant \widetilde{n}_{f,\Lambda}$. We shall use without further reference that then $\exp\big(\frac{-\kappa\psi^\Lambda_{m_n^\bullet} m_n^\bullet}{400\|\phi\|_{\ell^1}}\big) \leqslant n^{-1}$ for all $n \geqslant \widetilde{n}_{f,\Lambda}$ since $K_h \geqslant \frac{400\|\phi\|_{\ell^1}}{\kappa}$. Following line by line the proof of (J.29) using $\widetilde{n}_{f,\Lambda} \vee n_{f,\Lambda}$ rather than $n_{f,\Lambda}$ we obtain the assertion, that is, $\mathbb{E}_Y^n \|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\{\Delta^\Lambda_{[n_{f,\Lambda} \vee \widetilde{n}_{f,\Lambda}]} + \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 [n_{f,\Lambda} \vee \widetilde{n}_{f,\Lambda}] + \|\phi\|_{\ell^1}^2\} n^{-1}$, which completes the proof. □

Proof of Lemma G.1.5.
For $h \in \mathbb{B}_m$ setting $\nu_h(Y) = \sum_{|j| \in [\![1,m]\!]} \overline{[h]}_j \overline{g_j}^{-1} e_j(-Y)$ where $\mathbb{E}_Y^n \nu_h(Y) = \sum_{|j| \in [\![1,m]\!]} \overline{[h]}_j \overline{g_j}^{-1} \phi_j$ we observe (see Remark A.2.2) that $\|\widetilde{f}m - \theta^{\circ,m}\|_{L^2}^2 = \sup_{h \in \mathbb{B}_m} |\overline{\nu_h}|^2$. We intent to apply Lemma A.2.1. Therefore, we compute next quantities ψ, Ψ, and τ verifying the three inequalities required in Lemma A.2.1.

Consider ψ first:

$$\sup_{h \in \mathbb{B}_m} \sup_{y \in [0,1]} |\nu_h(y)|^2 = \sup_{y \in [0,1]} \sum_{|j| \in [\![1,m]\!]} |\lambda_j|^{-2} |e_j(y)|^2 = 2 \sum_{j \in [\![1,m]\!]} \Lambda_j = 2m\overline{\Lambda}_m \leqslant 2m\Lambda_{(m)} =: \psi^2.$$

Next, find Ψ. Notice that $\sup_{h \in \mathbb{B}_m} |\langle \widetilde{f}m - \theta^{\circ,m}, h\rangle_{L^2}|^2 = \sum_{|j| \in [\![1,m]\!]} \Lambda_{|j|} |\widehat{[h]}_j - \phi_j|^2$. As $\mathbb{E}_Y^n |\widehat{[h]}_j - \phi_j|^2 = \frac{1}{n}(1 - |\phi_j|^2) \leqslant \frac{1}{n}$, we define

$$\mathbb{E}_Y^n [\sup_{h \in \mathbb{B}_m} |\langle \widetilde{f}m - \theta^{\circ,m}\rangle_{L^2}^2] \leqslant 2m\overline{\Lambda}_m/n \leqslant \kappa\psi^\Lambda_m 2m\Lambda_{(m)}/n = 2\kappa\Delta^\Lambda_m/n =: \Psi^2.$$

Finally, consider τ. Given $h \in \mathbb{B}_m$ we observe with $\mathbb{E}[e_j(Y_1)e_{j'}(-Y_1)] = \phi_{j'-j}$ that

$$\mathbb{E}|\nu_h(Y_1)|^2 = \mathbb{E}_Y \left| \sum_{|j| \in [\![1,m]\!]} \overline{[h]}_j \overline{g_j}^{-1} e_j(-Y_1) \right|^2 = \sum_{|j|,|j'| \in [\![1,m]\!]} h_j \overline{[g]}_j^{-1} \mathbb{E}[e_j(Y_1)e_{j'}(-Y_1)] \overline{g_{j'}}^{-1} \overline{[h]}_{j'} \\ = \sum_{|j|,|j'| \in [\![1,m]\!]} h_j \overline{[g]}_j^{-1} \phi_{j'-j} \overline{g_{j'}}^{-1} \overline{[h]}_{j'} = \langle \mathcal{U}_k A \mathcal{U}_k\, h, h\rangle_{\ell^2}$$

defining the Hermitian and positive semi-definite matrix $A := (\overline{[g]}_j^{-1} \overline{g_{j'}}^{-1} \phi_{j'-j})_{j,j' \in \mathbb{Z}}$ and the mapping $\mathcal{U}_k : \mathbb{C}^\mathbb{Z} \to \mathbb{C}^\mathbb{Z}$ with $z \mapsto \mathcal{U}_k z = (z_l \mathbb{1}\{|l| \in [\![1,m]\!]\})_{l \in \mathbb{Z}}$. Obviously, $\mathcal{U}_k$ is an orthogonal projection from ℓ^2 onto the linear subspace spanned by all ℓ^2-sequences with support on the index-set $[\![-m,-1]\!] \cup [\![1,m]\!]$. Straightforward algebra shows $\sup_{h \in \mathbb{B}_m} \mathbb{V}\mathrm{ar}(\nu_h(Y_1)) \leqslant$

$\sup_{h\in\mathbb{B}_m}\langle\mathcal{U}_k A\mathcal{U}_k\, h,h\rangle_{\ell^2}$, hence

$$\sup_{h\in\mathbb{B}_m}\tfrac{1}{n}\sum_{i=1}^{n}\mathbb{V}\mathrm{ar}(\nu_h(Y_i)) \leqslant \sup_{h\in\mathbb{B}_m}\langle\mathcal{U}_k A\mathcal{U}_k h,h\rangle_{\ell^2} = \sup_{h\in\mathbb{B}_m}\|\mathcal{U}_k A\mathcal{U}_k h\|_{\ell^2} \leqslant \|\mathcal{U}_k A\mathcal{U}_k\|_s.$$

where $\|M\|_s := \sup_{\|x\|_{\ell^2}\leqslant 1}\|Mx\|_{\ell^2}$ denotes the spectral-norm of a linear map $M:\ell^2\to\ell^2$. For a sequence $z\in(\mathbb{C}\backslash\{0\})^{\mathbb{Z}}$ let ∇_z and ∇_z^{-1} be the multiplication operator given by $\nabla_z\, x := (z_jx_j)_{j\in\mathbb{Z}}$ and $\nabla_z^{-1} := \nabla_{z^{-1}}$, respectively. Clearly, we have $\mathcal{U}_k A\mathcal{U}_k = \mathcal{U}_k\,\nabla_\lambda^{-1}\,\mathcal{U}_k\mathcal{C}_\phi\mathcal{U}_k\,\nabla_{\overline{[g]}}^{-1}\,\mathcal{U}_k$, where $\mathcal{C}_\phi := (h_{j-j'})_{j,j'\in\mathbb{Z}}$. Consequently,

$$\begin{aligned}\sup_{h\in\mathbb{B}_m}\tfrac{1}{n}\sum_{i=1}^{n}\mathbb{V}\mathrm{ar}(\nu_h(Y_i)) &\leqslant \|\mathcal{U}_k\,\nabla_\lambda^{-1}\,\mathcal{U}_k\|_s\,\|\mathcal{C}_\phi\|_s\,\|\mathcal{U}_k\,\nabla_{\overline{[g]}}^{-1}\,\mathcal{U}_k\|_s\\ &= \|\mathcal{U}_k\,\nabla_\lambda^{-1}\,\mathcal{U}_k\|_s^2\,\|\mathcal{C}_\phi\|_s.\end{aligned}$$

We have that $\|\mathcal{U}_k\,\nabla_\lambda^{-1}\,\mathcal{U}_k\|_s^2 = \max\{\Lambda_j, j\in[\![1,m]\!]\} = \Lambda_{(m)}$. It remains to show the boundedness of $\|\mathcal{C}_\phi\|_s$. Keeping in mind that $(\mathcal{C}_\phi z)_k := \sum_{j\in\mathbb{Z}}\phi_{j-k}z_j$, $k\in\mathbb{Z}$, it is easily verified that $\|\mathcal{C}_\phi z\|_{\ell^2}^2 \leqslant \|\phi\|_{\ell^1}^2\|z\|_{\ell^2}^2$ and hence employing $\phi = \theta^\circ\lambda$ and the Cauchy-Schwarz inequality yields $\|\mathcal{C}_\phi\|_s^2 \leqslant \|\lambda\|_{\ell^2}^2\,\|\theta^\circ\|_{\ell^2}^2$ which implies

$$\sup_{h\in\mathbb{B}_m}\tfrac{1}{n}\sum_{i=1}^{n}\mathbb{V}\mathrm{ar}(\nu_h(Y_i)) \leqslant \|g\|_{L^2}\|f\|_{L^2}\,\Lambda_{(m)} =: \tau.$$

Alternatively, $\sup_{h\in\mathbb{B}_m}\tfrac{1}{n}\sum_{i=1}^{n}\mathbb{V}\mathrm{ar}(\nu_h(Y_i)) \leqslant \|\phi\|_{\ell^1}\Lambda_{(m)} =: \tau$. Note, if $f=e_0$, then $h=e_0$ and hence $\|\phi\|_{\ell^1}=1$, while $\|g\|_{L^2}\|f\|_{L^2}$ can become arbitrarly large depending on g. Replacing in Lemma A.2.1 (A.3) and (A.4) the quantities ψ,Ψ and τ together with $\Delta_m^\Lambda = \psi_m^\Lambda m\Lambda_{(m)}$ gives the assertion and completes the proof. □

Proof of Proposition G.1.1.

Given $m_+^\diamond, m_-^\diamond\in[\![1,n]\!]$ let m_+ and m_- as in (G.10). From Lemma G.1.1 together with Lemma G.1.4 follows

$$\begin{aligned}\mathbb{E}_Y^n\|\widetilde{f}-f\|_{L^2}^2 \leqslant 2\mathbb{E}_Y^n\|\widetilde{f}m_+ - \theta^{\circ,m_+}\|_{L^2}^2 + 2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\mathfrak{b}_{m_-}^2(f) + (\tfrac{2}{3\kappa\kappa^2}+\tfrac{8}{\kappa})n^{-1}\\ + 2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\,[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond(m_-^\diamond,f,\Lambda) - \tfrac{\kappa\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2}{4}n\mathfrak{b}_{[m_- -1]}^2(f)\big)\\ + 2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\mathbb{P}_Y^n\big(\|\widetilde{f}m_-^\diamond - \theta^{\circ,m_-^\diamond}\|_{L^2}^2 \geqslant 12\kappa\Delta_{m_-^\diamond}^\Lambda/n\big)\\ + 2\sum_{m\in]\!]m_+,n]\!]}\mathbb{E}_Y^n\big(\|\widetilde{f}m - \theta^{\circ,m}\|_{L^2}^2 - \mathcal{C}_1 12\kappa\Delta_m^\Lambda/n\big)_+\\ + (24\mathcal{C}_1\kappa/n)\sum_{m\in]\!]m_+,n]\!]}\Delta_m^\Lambda\mathbb{P}_Y^n\Big(\|\widetilde{f}m - \theta^{\circ,m}\|_{L^2}^2 \geqslant 12\kappa\Delta_m^\Lambda/n\Big)\end{aligned}$$

Since $m_+ \geqslant m_+^\diamond \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ and $n\geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ due to Lemma G.1.7 *(i)* and *(ii)* there is a finite numerical constant $\mathcal{C}$ such that

$$\begin{aligned}\mathbb{E}^n_Y\|\widetilde{f}-f\|^2_{L^2} \leqslant 2\mathbb{E}^n_Y\|\widetilde{f}m_+ - \theta^{\circ,m_+}\|^2_{L^2} + 2\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + (\tfrac{2}{3\kappa\kappa^2}+\tfrac{8}{\kappa})n^{-1}\\ + 2\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^\diamond_n(m^\diamond_-,f,\Lambda) - \tfrac{\kappa\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2}}{4}n\mathfrak{b}^2_{[m_- -1]}(f)\big)\\ + 2\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2}\mathbb{P}^n_Y\big(\|\widetilde{f}m^\diamond_- - \theta^{\circ,m^\diamond_-}\|^2_{L^2} \geqslant 12\kappa\Delta^\Lambda_{m^\diamond_-}/n\big)\\ + \mathcal{C}\big[\tfrac{24\|\phi\|^2_{\ell^1}}{\kappa}+8\big]n^{-1} + \mathcal{C}_1\big[(72*\tfrac{400^2\|\phi\|^2_{\ell^1}}{\kappa} + 72*800\|\phi\|_{\ell^1} + 72\kappa\big]n^{-1}\end{aligned}$$

and together with Lemma G.1.5 *(ii)* we obtain

$$\begin{aligned}\mathbb{E}^n_Y\|\widetilde{f}-f\|^2_{L^2} \leqslant 2\mathbb{E}^n_Y\|\widetilde{f}m_+ - \theta^{\circ,m_+}\|^2_{L^2} + 2\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + (\tfrac{2}{3\kappa\kappa^2}+\tfrac{8}{\kappa})n^{-1}\\ + 2\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^\diamond_n(m^\diamond_-,f,\Lambda) - \tfrac{\kappa\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2}}{4}n\mathfrak{b}^2_{[m_- -1]}(f)\big)\\ + 6\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2}\Big[\exp\big(\tfrac{-\kappa\psi^\Lambda_{m^\diamond_-}m^\diamond_-}{400\|\phi\|_{\ell^1}}\big) + \exp\big(\tfrac{-\sqrt{n\kappa\psi^\Lambda_{m^\diamond_-}}}{100}\big)\Big]\\ + \mathcal{C}\big[\tfrac{24\|\phi\|^2_{\ell^1}}{\kappa}+8\big]n^{-1} + \mathcal{C}_1\big[(72*\tfrac{400^2\|\phi\|^2_{\ell^1}}{\kappa} + 72*800\|\phi\|_{\ell^1} + 72\kappa\big]n^{-1}\end{aligned}$$

Moreover, for $n > n_{f,\Lambda} \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ holds $\sqrt{n} \geqslant \frac{300}{\sqrt{\kappa}}\log(n+2) \geqslant \frac{100}{\sqrt{\kappa}}\log(n+2)$ which in turn together with $\psi^\Lambda_{m^\diamond_-} \geqslant 1$ implies $n\exp\big(-\sqrt{n}\frac{\sqrt{\kappa\psi^\Lambda_{m^\diamond_-}}}{100}\big) \leqslant \exp\big(-\sqrt{n}\frac{\sqrt{\kappa}}{100} + \log(n+2)\big) \leqslant 1$, and thus

$$\begin{aligned}\mathbb{E}^n_Y\|\widetilde{f}-f\|^2_{L^2} \leqslant 2\mathbb{E}^n_Y\|\widetilde{f}m_+ - \theta^{\circ,m_+}\|^2_{L^2} + 2\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + (\tfrac{2}{3\kappa\kappa^2}+\tfrac{8}{\kappa})n^{-1}\\ + 2\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^\diamond_n(m^\diamond_-,f,\Lambda) - \tfrac{\kappa\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2}}{4}n\mathfrak{b}^2_{[m_- -1]}(f)\big)\\ + 6\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2}\Big[\exp\big(\tfrac{-\kappa\psi^\Lambda_{m^\diamond_-}m^\diamond_-}{400\|\phi\|_{\ell^1}}\big) + n^{-1}\Big]\\ + \mathcal{C}\big[\tfrac{24\|\phi\|^2_{\ell^1}}{\kappa}+8\big]n^{-1} + \mathcal{C}_1\big[(72*\tfrac{400^2\|\phi\|^2_{\ell^1}}{\kappa} + 72*800\|\phi\|_{\ell^1} + 72\kappa\big]n^{-1}\end{aligned}$$

Recalling that $\mathbb{E}^n_h\|\widetilde{f}m - \theta^{\circ,m}\|^2_{\mathbb{H}} \leqslant 2m\overline{\Lambda}_m/n \leqslant 2\Delta^\Lambda_m/n$ for all $m \in \mathbb{N}$. Taking into account the definition (G.10) of m_+ we obtain $\mathbb{E}^n_Y\|\widetilde{f}m_+ - \theta^{\circ,m_+}\|^2_{L^2} \leqslant 2\Delta^\Lambda_{m_+}/n \leqslant 2[\frac{3}{12\kappa}\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2} + \frac{108}{12}]\mathcal{R}^\diamond_n(m^\diamond_+,f,\Lambda)$ and hence

$$\begin{aligned}\mathbb{E}^n_Y\|\widetilde{f}-f\|^2_{L^2} \leqslant [\tfrac{1}{\kappa}\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2} + 36]\mathcal{R}^\diamond_n(m^\diamond_+,f,\Lambda) + 2\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + (\tfrac{2}{3\kappa\kappa^2}+\tfrac{8}{\kappa})n^{-1}\\ + 2\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^\diamond_n(m^\diamond_-,f,\Lambda) - \tfrac{\kappa\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2}}{4}n\mathfrak{b}^2_{[m_- -1]}(f)\big)\\ + 6\|\Pi^\perp_{\mathbb{U}_0} f\|^2_{L^2}\Big[\exp\big(\tfrac{-\kappa\psi^\Lambda_{m^\diamond_-}m^\diamond_-}{400\|\phi\|_{\ell^1}}\big) + n^{-1}\Big]\\ + \mathcal{C}\big[\tfrac{24\|\phi\|^2_{\ell^1}}{\kappa}+8\big]n^{-1} + \mathcal{C}_1\big[(72*\tfrac{400^2\|\phi\|^2_{\ell^1}}{\kappa} + 72*800\|\phi\|_{\ell^1} + 72\kappa\big]n^{-1}\end{aligned}$$

Recalling that $\mathcal{R}_n^\diamond[m,f,\Lambda] = [\mathfrak{b}_m^2(f) \vee \Delta_m^\Lambda\, n^{-1}] \geqslant n^{-1}$ there is an universal finite numerical constant $\mathcal{C}$ such that

$$\mathbb{E}_Y^n \|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\big\{[1 \vee \|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2]\mathcal{R}_n^\diamond(m_+^\diamond, f, \Lambda) + \|\phi\|_{\ell^1}^2 n^{-1}\big\}$$
$$+ 2\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_{m_-}^2(f) + 6\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 \exp\big(\frac{-\kappa\psi_{m_-^\diamond}^\Lambda m_-^\diamond}{400\|\phi\|_{\ell^1}}\big)$$
$$+ 2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\,[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond(m_-^\diamond, f, \Lambda) - \frac{\kappa\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2}{4} n\mathfrak{b}_{[m_- -1]}^2(f)\big)$$

which shows the assertion and completes the proof. □

Proof of Lemma G.1.6.

Given $m, l \in [\![1,n]\!]$ and an event Ω_{ml} (to be specified below) it clearly follows

$$\widetilde{w}_m \mathbb{1}\Omega_{ml} = \frac{\exp(-\kappa n\{-\|\widetilde{f}m\|_{L^2}^2 + \frac{9}{2}12\kappa\Delta_m^\Lambda/n\})}{\sum_{s\in[\![1,n]\!]}\exp(-\kappa n\{-\|\widetilde{f}s\|_{L^2}^2 + \frac{9}{2}12\kappa\Delta_s^\Lambda/n\})}\mathbb{1}\Omega_{ml}$$
$$\leqslant \exp\big(\kappa n\big\{\|\widetilde{f}m\|_{L^2}^2 - \|\widetilde{f}l\|_{L^2}^2 + \tfrac{9}{2}(12\kappa\Delta_l^\Lambda/n - 12\kappa\Delta_m^\Lambda/n)\big\}\big)\mathbb{1}\Omega_{ml} \quad \text{(G.14)}$$

We distinguish the two cases $m < l$ and $m > l$. Consider first that $m < l$. From ***(i)*** in corollary G.1.1 (with $\breve{f} = \widetilde{f}n$) follows that

$$\widetilde{w}_m\mathbb{1}\Omega_{ml} \leqslant \exp\big(\kappa n\big\{\|\widetilde{f}m\|_{L^2}^2 - \|\widetilde{f}l\|_{L^2}^2 + \tfrac{9}{2}(12\kappa\Delta_l^\Lambda/n - 12\kappa\Delta_m^\Lambda/n)\big\}\big)\mathbb{1}\Omega_{ml}$$
$$\leqslant \exp\big(\kappa n\big\{\tfrac{11}{2}\|\widetilde{f}l - \theta^{\circ,l}\|_{L^2}^2 - \tfrac{1}{2}\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2(\mathfrak{b}_k^2(f) - \mathfrak{b}_l^2(f)) + \tfrac{9}{2}(12\kappa\Delta_l^\Lambda/n - 12\kappa\Delta_k^\Lambda/n)\big\}\big)\mathbb{1}\Omega_{kl}$$

If we define $\Omega_{kl} := \{\|\widetilde{f}l - \theta^{\circ,l}\|_{L^2}^2 < 12\kappa\Delta_l^\Lambda/n\}$ then the last bound implies

$$\widetilde{w}_k\mathbb{1}\{\|\widetilde{f}l - \theta^{\circ,l}\|_{\mathbb{H}}^2 < 12\kappa\Delta_l^\Lambda/n\}$$
$$\leqslant \exp\big(\kappa n\big\{\tfrac{11}{2}12\kappa\Delta_l^\Lambda/n - \tfrac{1}{2}\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2(\mathfrak{b}_k^2(f) - \mathfrak{b}_l^2(f)) + \tfrac{9}{2}(12\kappa\Delta_l^\Lambda/n - 12\kappa\Delta_k^\Lambda/n)\big\}\big)$$
$$= \exp\big(\kappa n\big\{10 * 12\kappa\Delta_l^\Lambda/n - \tfrac{1}{2}\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2(\mathfrak{b}_k^2(f) - \mathfrak{b}_l^2(f)) - \tfrac{9}{2}12\kappa\Delta_k^\Lambda/n\big\}\big)$$

and hence, by exploiting that $12\kappa\Delta_k^\Lambda/n \geqslant 0$ and $\mathcal{R}_n^\diamond(l,f,\Lambda) = [\mathfrak{b}_l^2(f) \vee \Delta_l^\Lambda n^{-1}]$ follows the assertion ***(i)***, that is

$$\widetilde{w}_k\mathbb{1}\{\|\widetilde{f}l - \theta^{\circ,l}\|_{L^2}^2 < 12\kappa\Delta_l^\Lambda/n\} \leqslant \exp\big(\kappa n\big\{-\frac{\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2}{2}\mathfrak{b}_k^2(f) + [10*12\kappa + \frac{\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2}{2}]\mathcal{R}_n^\diamond(l,f,\Lambda))\big\}\big).$$

Consider secondly that $k > l$. From ***(ii)*** in corollary G.1.1 (with $\breve{f} = \widetilde{f}n$) and (G.14) follows

$$\widetilde{w}_k\mathbb{1}\Omega_{lk} \leqslant \exp\big(\kappa n\big\{\|\widetilde{f}m\|_{L^2}^2 - \|\widetilde{f}l\|_{L^2}^2 + \tfrac{9}{2}(12\kappa\Delta_l^\Lambda/n - 12\kappa\Delta_m^\Lambda/n)\big\}\big)\mathbb{1}\Omega_{ml}$$
$$\leqslant \exp\big(\kappa n\big\{\tfrac{7}{2}\|\widetilde{f}k - \theta^{\circ,k}\|_{L^2}^2 + \tfrac{3}{2}\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2(\mathfrak{b}_l^2(f) - \mathfrak{b}_k^2(f)) + \tfrac{9}{2}(12\kappa\Delta_l^\Lambda/n - 12\kappa\Delta_k^\Lambda/n)\big\}\big)\mathbb{1}\Omega_{lk}$$

If we set $\Omega_{lk} := \{\|\widetilde{f}k - \theta^{\circ,k}\|_{L^2}^2 < 12\kappa\Delta_k^\Lambda/n\}$ then we clearly have

$$\widetilde{w}_k \mathbb{1}_{\{\|\widetilde{f}k - \theta^{\circ,k}\|_{L^2}^2 < 12\kappa\Delta_k^\Lambda/n\}} \\ \leqslant \exp\big(\kappa n\big\{\tfrac{7}{2}12\kappa\Delta_k^\Lambda/n + \tfrac{3}{2}\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2(\mathfrak{b}_l^2(f) - \mathfrak{b}_k^2(f)) + \tfrac{9}{2}(12\kappa\Delta_l^\Lambda/n - 12\kappa\Delta_k^\Lambda/n)\big\}\big)$$

and hence, by exploiting $\mathfrak{b}_k^2(f) \geqslant 0$ and $\mathcal{R}_n^\diamond(l, f, \Lambda) = [\mathfrak{b}_l^2(f) \vee \Delta_l^\Lambda n^{-1}]$ follows the assertion ***(ii)***, that is

$$\widetilde{w}_k \mathbb{1}_{\{\|\widetilde{f}k - \theta^{\circ,k}\|_{L^2}^2 < 12\kappa\Delta_k^\Lambda/n\}} \leqslant \exp\big(\kappa n\big\{-12\kappa\Delta_k^\Lambda n^{-1} + [\tfrac{3}{2}\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 + \tfrac{9}{2} * 12\kappa]\mathcal{R}_n^\diamond(l, f, \Lambda)\big\}\big).$$

which completes the proof. □

PROOF OF COROLLARY G.1.1.
Consider ***(i)***. For $l > k$ denote by $\Pi_{\mathbb{U}_{kl}} = \Pi_{\mathbb{U}_l} - \Pi_{\mathbb{U}_k} = \Pi_{\mathbb{U}_k}^\perp - \Pi_{\mathbb{U}_l}^\perp$ the orthogonal projection onto $\mathbb{U}_{kl} := \overline{lin}\{e_j, |j| \in \rrbracket k, l\rrbracket\}$ where $\|\Pi_{\mathbb{U}_{kl}} h\|_{L^2}^2 = \|\Pi_{\mathbb{U}_l} h\|_{L^2}^2 - \|\Pi_{\mathbb{U}_k} h\|_{L^2}^2 = \|\Pi_{\mathbb{U}_k}^\perp h\|_{L^2}^2 - \|\Pi_{\mathbb{U}_l}^\perp h\|_{L^2}^2$ for all $h \in L^2$. Let us define $\Upsilon(h) := \|h\|_{L^2}^2 - 2\langle h, \breve{f}\rangle_{L^2}$. For each $l \in \llbracket 1, n\rrbracket$ and for any $h \in \mathbb{U}_l$ we have $\langle h, \breve{f}\rangle_{L^2} = \langle h, \breve{f}^l\rangle_{L^2}$, which in turn implies $\Upsilon(h) = \|h\|_{L^2}^2 - 2\langle h, \breve{f}^l\rangle_{L^2} + \|\breve{f}^l\|_{L^2}^2 - \|\breve{f}^l\|_{L^2}^2 = \|h - \breve{f}^l\|_{L^2}^2 - \|\breve{f}^l\|_{L^2}^2$ and consequently, $\Upsilon(\breve{f}^l) = -\|\breve{f}^l\|_{L^2}^2$ for all $l \in \mathbb{N}$. Obviously, $\|\breve{f}^k\|_{L^2}^2 - \|\breve{f}^l\|_{L^2}^2 = \Upsilon(\breve{f}^l) - \Upsilon(\breve{f}^k)$, while $\|\Pi_{\mathbb{U}_{kl}} f\|_{L^2}^2 = \|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\{\mathfrak{b}_k^2(f) - \mathfrak{b}_l^2(f)\}$. Consequently, the claim ***(i)*** can equivalently be rewritten as

$$\Upsilon(\breve{f}^l) - \Upsilon(\breve{f}^k) \leqslant \tfrac{11}{2}\|\breve{f}^l - \theta^{\circ,l}\|_{L^2}^2 - \tfrac{1}{2}\|\Pi_{\mathbb{U}_{kl}} f\|_{L^2}^2. \tag{G.15}$$

Analogously, if $k > l$ then the claim ***(ii)*** can equivalently be rewritten as

$$\Upsilon(\breve{f}^l) - \Upsilon(\breve{f}^k) \leqslant \tfrac{7}{2}\|\breve{f}^k - \theta^{\circ,k}\|_{L^2}^2 + \tfrac{3}{2}\|\Pi_{\mathbb{U}_{lk}} f\|_{L^2}^2. \tag{G.16}$$

Proof of (G.15). For $x, y, z \in L^2$ we observe that

$$\begin{aligned}\Upsilon(x) - \Upsilon(y) &= \|x\|_{L^2}^2 - \|y\|_{L^2}^2 - 2\langle x - y, \breve{f}\rangle_{L^2} \\ &= \|x\|_{L^2}^2 - 2\langle x, z\rangle_{L^2} + \|z\|_{L^2}^2 - \|y\|_{L^2}^2 + 2\langle y, z\rangle_{L^2} - \|z\|_{L^2}^2 - 2\langle x - y, \breve{f} - z\rangle_{L^2} \\ &= \|x - z\|_{L^2}^2 - \|y - z\|_{L^2}^2 - 2\langle x - y, \breve{f} - z\rangle_{L^2}\end{aligned}$$

and in particular for $x_l = \Pi_{\mathbb{U}_l} x$ and $x_k = \Pi_{\mathbb{U}_k} x$ with $k < l$ where $x_l - x_k = \Pi_{\mathbb{U}_{kl}} x$ we have

$$\begin{aligned}\Upsilon(x_l) - \Upsilon(x_k) &= \|x_l - z\|_{L^2}^2 - \|x_k - z\|_{L^2}^2 - 2\langle x_l - x_k, \breve{f} - z\rangle_{L^2} \\ &= \|\Pi_{\mathbb{U}_l}(x - z)\|_{L^2}^2 - \|\Pi_{\mathbb{U}_k}(x - z)\|_{L^2}^2 + \|\Pi_{\mathbb{U}_l}^\perp z\|_{L^2}^2 - \|\Pi_{\mathbb{U}_k}^\perp z\|_{L^2}^2 - 2\langle \Pi_{\mathbb{U}_{kl}} x, \breve{f} - z\rangle_{L^2} \\ &= \|\Pi_{\mathbb{U}_{kl}}(x - z)\|_{L^2}^2 - \|\Pi_{\mathbb{U}_{kl}} z\|_{L^2}^2 - 2\langle \Pi_{\mathbb{U}_{kl}} x, \Pi_{\mathbb{U}_{kl}}(\breve{f} - z)\rangle_{L^2} \\ &= \|\Pi_{\mathbb{U}_{kl}}(x - z)\|_{L^2}^2 - \|\Pi_{\mathbb{U}_{kl}} z\|_{L^2}^2 - 2\|\Pi_{\mathbb{U}_{kl}} x\|_{L^2}\langle \tfrac{\Pi_{\mathbb{U}_{kl}} x}{\|\Pi_{\mathbb{U}_{kl}} x\|_{L^2}}, \Pi_{\mathbb{U}_{kl}}(\breve{f} - z)\rangle_{L^2}.\end{aligned} \tag{G.17}$$

Exploiting the elementary inequality $-2ab \leqslant \frac{1}{4}a^2 + 4b^2$ and setting $\mathbb{B}_{kl} := \{x \in \mathbb{U}_{kl} :$

$\|x\|_{L^2} = 1\}$ it follows

$$\Upsilon(x_l)-\Upsilon(x_k) \leqslant \|\Pi_{\mathbb{U}_{kl}}(x-z)\|_{L^2}^2-\|\Pi_{\mathbb{U}_{kl}} z\|_{L^2}^2+\tfrac{1}{4}\|\Pi_{\mathbb{U}_{kl}} x\|_{L^2}^2+4|\langle\tfrac{\Pi_{\mathbb{U}_{kl}} x}{\|\Pi_{\mathbb{U}_{kl}} x\|_{L^2}}, \Pi_{\mathbb{U}_{kl}}(\breve{f}-z)\rangle_{L^2}|^2$$
$$\leqslant \|\Pi_{\mathbb{U}_{kl}}(x-z)\|_{L^2}^2-\|\Pi_{\mathbb{U}_{kl}} z\|_{L^2}^2+\tfrac{1}{2}\|\Pi_{\mathbb{U}_{kl}}(x-z)\|_{L^2}^2+\tfrac{1}{2}\|\Pi_{\mathbb{U}_{kl}} z\|_{L^2}^2+4\sup_{y\in\mathbb{B}_{kl}}|\langle y, \Pi_{\mathbb{U}_{kl}}(\breve{f}-z)\rangle_{L^2}|^2$$
$$= \tfrac{3}{2}\|\Pi_{\mathbb{U}_{kl}}(x-z)\|_{L^2}^2 - \tfrac{1}{2}\|\Pi_{\mathbb{U}_{kl}} z\|_{L^2}^2 + 4\|\Pi_{\mathbb{U}_{kl}}(\breve{f}-z)\|_{L^2}^2.$$

Replacing x by $\breve{f}$ and z by f the last estimate implies (G.15), that is,

$$\Upsilon(\breve{f}^l) - \Upsilon(\breve{f}^k) \leqslant \tfrac{3}{2}\|\Pi_{\mathbb{U}_{kl}}(\breve{f}-f)\|_{L^2}^2 - \tfrac{1}{2}\|\Pi_{\mathbb{U}_{kl}} f\|_{L^2}^2 + 4\|\Pi_{\mathbb{U}_{kl}}(\breve{f}-f)\|_{L^2}^2$$
$$= \tfrac{11}{2}\|\Pi_{\mathbb{U}_{kl}}(\breve{f}-f)\|_{L^2}^2 - \tfrac{1}{2}\|\Pi_{\mathbb{U}_{kl}} f\|_{L^2}^2 \leqslant \tfrac{11}{2}\|\breve{f}^l - \theta^{\circ,l}\|_{L^2}^2 - \tfrac{1}{2}\|\Pi_{\mathbb{U}_{kl}} f\|_{L^2}^2$$

Proof of (G.16). In case of $k > l$ from (G.17) follows

$$\Upsilon(x_l) - \Upsilon(x_k) = -(\Upsilon(x_k) - \Upsilon(x_l))$$
$$= -\|\Pi_{\mathbb{U}_{lk}}(x-z)\|_{L^2}^2 + \|\Pi_{\mathbb{U}_{lk}} z\|_{L^2}^2 + 2\|\Pi_{\mathbb{U}_{lk}} x\|_{L^2}\langle\tfrac{\Pi_{\mathbb{U}_{lk}} x}{\|\Pi_{\mathbb{U}_{lk}} x\|_{L^2}}, \Pi_{\mathbb{U}_{lk}}(\breve{f}-z)\rangle_{L^2}$$

Exploiting again the elementary inequality $-2ab \leqslant \frac{1}{4}a^2 + 4b^2$ and keeping in mind that $\mathbb{B}_{lk} := \{x \in \mathbb{U}_{lk} : \|x\|_{L^2} = 1\}$ it follows

$$\Upsilon(x_l) - \Upsilon(x_k) \leqslant -\|\Pi_{\mathbb{U}_{lk}}(x-z)\|_{L^2}^2 + \|\Pi_{\mathbb{U}_{lk}} z\|_{L^2}^2 + \tfrac{1}{2}\|\Pi_{\mathbb{U}_{lk}}(x-z)\|_{L^2}^2$$
$$+ \tfrac{1}{2}\|\Pi_{\mathbb{U}_{lk}} z\|_{L^2}^2 + 4\sup_{y\in\mathbb{B}_{lk}}|\langle y, \Pi_{\mathbb{U}_{lk}}(\breve{f}-z)\rangle_{L^2}|^2$$
$$= -\tfrac{1}{2}\|\Pi_{\mathbb{U}_{lk}}(x-z)\|_{L^2}^2 + \tfrac{3}{2}\|\Pi_{\mathbb{U}_{lk}} z\|_{L^2}^2 + 4\|\Pi_{\mathbb{U}_{lk}}(\breve{f}-z)\|_{L^2}^2.$$

Replacing x by $\breve{f}^{\bullet}$ and z by f the last estimate implies (G.15), that is

$$\Upsilon(\breve{f}^l) - \Upsilon(\breve{f}^k) \leqslant -\tfrac{1}{2}\|\Pi_{\mathbb{U}_{lk}}(\breve{f}-f)\|_{L^2}^2 + \tfrac{3}{2}\|\Pi_{\mathbb{U}_{lk}} f\|_{L^2}^2 + 4\|\Pi_{\mathbb{U}_{lk}}(\breve{f}-f)\|_{L^2}^2$$
$$= \tfrac{7}{2}\|\Pi_{\mathbb{U}_{lk}}(\breve{f}-f)\|_{L^2}^2 + \tfrac{3}{2}\|\Pi_{\mathbb{U}_{lk}} f\|_{L^2}^2 \leqslant \tfrac{7}{2}\|\breve{f}^k - \theta^{\circ,k}\|_{L^2}^2 + \tfrac{3}{2}\|\Pi_{\mathbb{U}_{lk}} f\|_{L^2}^2,$$

which completes the proof. □

Proof of Lemma G.1.7.
Since $\psi_m^\Lambda \geqslant 1$ for $m \geqslant 3(\frac{12\|\phi\|_{\ell^1}}{\kappa})^2$ holds $\frac{\kappa\sqrt{\psi_m^\Lambda}m}{12\|\phi\|_{\ell^1}} - \log(m+2) \geqslant 0$ and hence

$$\Lambda_{(m)}\exp\big(\tfrac{-\kappa\psi_m^\Lambda m}{6\|\phi\|_{\ell^1}}\big) \leqslant \exp\big(\tfrac{-\kappa\psi_m^\Lambda m}{12\|\phi\|_{\ell^1}}\big)\exp\big(-\sqrt{\psi_m^\Lambda}[\tfrac{\kappa\sqrt{\psi_m^\Lambda}m}{12\|\phi\|_{\ell^1}} - \log(m+2)]\big)$$
$$\leqslant \exp\big(\tfrac{-\kappa\psi_m^\Lambda m}{12\|\phi\|_{\ell^1}}\big) \leqslant \exp\big(-\tfrac{\kappa}{12\|\phi\|_{\ell^1}}m\big)$$

consequently, if $m_+ \geqslant 3(\frac{12\|\phi\|_{\ell^1}}{\kappa})^2$ then

$$\sum_{m=1+m_+}^{n} \Lambda_{(m)} \exp\big(\tfrac{-\kappa\psi_m^\Lambda m}{6\|\phi\|_{\ell^1}}\big) \quad\leqslant\quad \sum_{m=1+m_+}^{n} \exp\big(\quad - \quad \tfrac{\kappa}{12\|\phi\|_{\ell^1}} m\big) \quad\leqslant\quad \frac{12\|\phi\|_{\ell^1}}{\kappa}$$

Using for all $n \geqslant 15(\frac{200}{\sqrt{\kappa}})^4$ holds $\sqrt{n} \geqslant \frac{200}{\sqrt{\kappa}} \log(n+2)$ (alternatively: $\sqrt{3x} \geqslant \log(x+2)$ for all $x \geqslant 1$ and $\sqrt{\kappa}/200 \geqslant \sqrt{3}$)

$$\frac{m^2\Lambda_{(m)}^2}{n} \exp\big(\tfrac{-\sqrt{n\kappa\psi_m^\Lambda}}{100}\big) \leqslant \tfrac{1}{n} \exp\big(-2\sqrt{\psi_m^\Lambda}[\tfrac{\sqrt{n\kappa}}{200} - \log(m+2)]\big)$$
$$\{\text{alternatively} \leqslant \tfrac{1}{n} \exp\big(-2\sqrt{\psi_m^\Lambda n} \quad \underbrace{[\tfrac{\sqrt{\kappa}}{200} - \sqrt{3}]}_{\geqslant 0 \text{ since } \sqrt{\kappa}/200 \geqslant \sqrt{3}} \quad\big)\}$$
$$\leqslant \tfrac{1}{n}$$

consequently,

$$\sum_{m=1+m_+}^{n} \frac{m^2\Lambda_{(m)}^2}{n} \exp\big(\tfrac{-\sqrt{n\kappa\psi_m^\Lambda}}{100}\big) \leqslant \sum_{m=1+m_+}^{n} \tfrac{1}{n} \leqslant 1$$

Combining the last two bounds and Lemma G.1.5 we obtain ***(i)***, that is

$$\sum_{m=1+m_+}^{n} \mathbb{E}_Y^n\big(\|\widetilde{f}m - \theta^{\circ,m}\|_{L^2}^2 - 12\kappa\Delta_m^\Lambda/n\big)_+$$
$$\leqslant \mathcal{C}\Big[\frac{\|\phi\|_{\ell^1}}{n} \sum_{m=1+m_+}^{n} \Lambda_{(m)} \exp\big(\tfrac{-\kappa\psi_m^\Lambda m}{6\|\phi\|_{\ell^1}}\big) + \tfrac{4}{n} \sum_{m=1+m_+}^{n} \frac{m^2\Lambda_{(m)}^2}{n} \exp\big(\tfrac{-\sqrt{n\kappa\psi_m^\Lambda}}{100}\big)\Big]$$
$$\leqslant \mathcal{C} n^{-1}\big[\frac{12\|\phi\|_{\ell^1}^2}{\kappa} + 4\big]$$

If $m \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ then $m \geqslant (\frac{800\|\phi\|_{\ell^1}}{\kappa}) \log(m+2)$ and hence $m - \frac{400\|\phi\|_{\ell^1}}{\kappa} \log(m+2) \geqslant \frac{400\|\phi\|_{\ell^1}}{\kappa} \log(m+2)$ or equivalently, $\frac{\kappa}{400\|\phi\|_{\ell^1}} m - \log(m+2) \geqslant \log(m+2) \geqslant 1$ and thus

$$m\psi_m^\Lambda \Lambda_{(m)} \exp\big(\tfrac{-\kappa\psi_m^\Lambda m}{400\|\phi\|_{\ell^1}}\big) \leqslant \psi_m^\Lambda \exp\big(-\psi_m^\Lambda\,[\tfrac{\kappa}{400\|\phi\|_{\ell^1}} m - \log(m+2)]\big)$$
$$\leqslant (m+2) \exp\big(-\tfrac{\kappa}{400\|\phi\|_{\ell^1}} m\big)$$

consequently, if $m_+ \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ exploiting $\sum_{m\in\mathbb{N}}(m+2)\exp(-\mu m) \leqslant \mu^{-2} + 2\mu^{-1}$ follows

$$\sum_{m=1+m_+}^{n} m\psi_m^\Lambda \Lambda_{(m)} \exp\big(\tfrac{-\kappa\psi_m^\Lambda m}{400\|\phi\|_{\ell^1}}\big) \leqslant \sum_{m=1+m_+}^{n} (k+2) \exp\big(-\tfrac{\kappa}{400\|\phi\|_{\ell^1}} m\big)$$
$$\leqslant (\frac{400\|\phi\|_{\ell^1}}{\kappa})^2 + \frac{800\|\phi\|_{\ell^1}}{\kappa}$$

Since $\log(m\Lambda_{(m)}) \leqslant \frac{1}{e}m\Lambda_{(m)}$ follows $\psi_m^\Lambda \leqslant m\Lambda_{(m)}$, and for all $n \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ holds $\sqrt{n} \geqslant \frac{300}{\sqrt{\kappa}}\log(n+2)$ (alternatively: using $\sqrt{3x} \geqslant \log(x+2)$ for all $x \geqslant 1$ and $\sqrt{\kappa}/300 \geqslant \sqrt{3}$)

$$\begin{aligned}
&m\psi_m^\Lambda\Lambda_{(m)}\exp\big(\tfrac{-\sqrt{n\kappa\psi_m^\Lambda}}{100}\big) \leqslant m^2\Lambda_{(m)}^2\exp\big(\tfrac{-\sqrt{n\kappa\psi_m^\Lambda}}{100}\big)\\
&\leqslant \tfrac{1}{n}\exp\big(-\sqrt{\psi_m^\Lambda}[\tfrac{\sqrt{n\kappa}}{100} - 2\log(m+2)] + \log(n+2)\big) \leqslant \tfrac{1}{n}\exp\big(-3\sqrt{\psi_m^\Lambda}[\tfrac{\sqrt{n\kappa}}{300} - \log(n+2)]\big)\\
&\qquad\{\text{alternatively} \leqslant \tfrac{1}{n}\exp\big(-3\sqrt{\psi_m^\Lambda n}\underbrace{[\tfrac{\sqrt{\kappa}}{300} - \sqrt{3}]}_{\geqslant 0 \text{ since } \sqrt{\kappa}/300 \geqslant \sqrt{3}}\big)\}\\
&\leqslant \tfrac{1}{n}
\end{aligned}$$

consequently,

$$\sum_{m=1+m_+}^{n} m\psi_m^\Lambda\Lambda_{(m)}\exp\big(\tfrac{-\sqrt{n\kappa\psi_m^\Lambda}}{100}\big) \leqslant \sum_{m=1+m_+}^{n}\tfrac{1}{n} \leqslant 1$$

Combining the last two bounds and Lemma G.1.5 we obtain ***(ii)***, that is

$$\begin{aligned}
&\sum_{m=1+m_+}^{n}\psi_m^\Lambda m\Lambda_{(m)}\mathbb{P}_Y^n\big(\|\widetilde{f}m - \theta^{\circ,m}\|_{L^2}^2 \geqslant 12\kappa\Delta_m^\Lambda/n\big)\\
&\leqslant 3\sum_{m=1+m_+}^{n}\psi_m^\Lambda m\Lambda_{(m)}\Big[\exp\big(\tfrac{-\kappa\psi_m^\Lambda m}{400\|\phi\|_{\ell^1}}\big) + \exp\big(\tfrac{-\sqrt{n\kappa\psi_m^\Lambda}}{100}\big)\Big]\\
&\leqslant 3\Big[\big(\frac{400\|\phi\|_{\ell^1}}{\kappa}\big)^2 + \frac{800\|\phi\|_{\ell^1}}{\kappa} + 1\Big]
\end{aligned}$$

which implies the result and completes the proof. □

PROOF OF THEOREM 3.4.1.

We distinguish the cases (a) and (b) of Assumption 14. Firstly, consider (a). Due to Lemma J.2.1 we have $\mathbb{E}_Y^n\|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\Delta_{n_{f,\Lambda}}^\Lambda n^{-1}$ since $n_{f,\Lambda} > \lceil 15(\frac{300}{\sqrt{\kappa}})^4\rceil$. Secondly, consider (b). Due to Lemma J.2.2 under Assumption 14 setting $m_n^\bullet := K_h(\log n)$ for $n \leqslant n_{f,\Lambda}$ and otherwise $m_n^\bullet := \max\{m \in [\![K, n]\!] : c_f\,\Delta_m^\Lambda < n\}$ for $n > n_{f,\Lambda}$, where the defining set contains K_h and thus it is not empty, there is a finite numerical constant $\mathcal{C}$ such that for all $n \in \mathbb{N}$ holds

$$\mathbb{E}_Y^n\|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\{\Delta_{n_{f,\Lambda}}^\Lambda + \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 n_{f,\Lambda} + \|\phi\|_{\ell^1}^2\}n^{-1} + 6\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\{\exp\big(\tfrac{-\kappa\psi_{m_n^\bullet}^\Lambda m_n^\bullet}{400\|\phi\|_{\ell^1}}\big) - \tfrac{1}{n}\}. \tag{G.18}$$

Since under Assumption 14 for all $n \geqslant n_{f,\Lambda}$ in addition $\psi_{m_n^\bullet}^\Lambda m_n^\bullet \geqslant K_h(\log n)$ holds true it follows

$$\mathbb{E}_Y^n\|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\{\Delta_{n_{f,\Lambda}}^\Lambda + \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 n_{f,\Lambda} + \|\phi\|_{\ell^1}^2\}n^{-1}. \tag{G.19}$$

Combining the upper bounds for both cases (a) and (b) we obtain the assertion (3.1), which completes the proof.

Appendix H

Proof for Theorem 3.4.2

H.1 Intermediate results

The proof of this theorem is mainly based on the one of ??. However, we need an altered version of ?? and ??.

Definition 47 Define the following quantities :

$$G_n^- := \min\left\{m \in [\![1, m_n^\circ]\!] : \quad \mathfrak{b}_m^2 \leqslant 244\mathfrak{b}_0^2\Phi_n^\circ\right\},$$
$$G_n^+ := \max\left\{m \in [\![m_n^\circ, n]\!] : \psi_n m \Lambda_m \leqslant \frac{59}{2} n\mathfrak{b}_0^2\Phi_n^\circ\right\}.$$

Proposition H.1.1. *For any n and η in $\mathbb{N}$, with $\eta \geqslant 118K$, we have*

$$\mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left([\![0, G_n^- - 1]\!]\right)\right] \leqslant 4m_n^\circ \exp\left[-K\left(\frac{2\psi_n m_n^\circ \overline{\Lambda}_{m_n^\circ}}{\Lambda_{m_n^\circ}\|\theta^\circ\|_{l^2}\|\lambda\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \tag{H.1}$$

$$\mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left([\![G_n^+ + 1, n]\!]\right)\right] \leqslant C_{\lambda,\theta^\circ} \exp\left[-K\left(2\frac{\psi_n m_n^\circ \overline{\Lambda}_{m_n^\circ}}{\Lambda_{(m_n^\circ)}\|\theta^\circ\|_{l^2}\|\lambda\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \tag{H.2}$$

Proposition H.1.2.

$$\sum_{0<|j|\leqslant n} \Lambda_j\, \mathbb{E}_{\theta^\circ}^n\left[\left(\lambda_j\overline{\theta}_j - \lambda_j\theta_j^\circ\right)^2 \mathbb{P}_{M|Y^n}^{n,(\eta)}\left([\![|j|, n]\!]\right)\right]$$
$$\leqslant 28\mathfrak{b}_0^2\Phi_n^\circ + \frac{1}{n}12C_{\lambda,\theta^\circ}\exp\left[-K\left(\frac{\psi_n 2m_n^\dagger}{\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right) + \log\left(\psi_n n\Lambda_n\right)\right] \tag{H.3}$$

$$\sum_{0<|j|\leqslant n} \left(\theta_j^\circ\right)^2 \mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left([\![0, j-1]\!]\right)\right] + \sum_{|j|>n}\left(\theta_j^\circ\right)^2$$
$$\leqslant 59\mathfrak{b}_0^2\Phi_n^\dagger + \frac{1}{n}4C_{\lambda,\theta^\circ}\exp\left[-K\left(\frac{\psi_n 2m_n^\circ}{\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right) + \log\left(nm_n^\dagger\right)\right] \tag{H.4}$$

H.2 Detailed proofs

Proof of Proposition I.1.1

Building upon **??**, consider m in $\mathbb{N}$ such that $m < m_n^\circ$. Set $\Delta^\star_{m,m_n^\circ} = \Lambda_{(m_n^\circ)}$ and $\delta^\star_{m,m_n^\circ} = m_n^\circ\overline{\Lambda}_{m_n^\circ}$, $\mathrm{pen}(k) = \kappa m\overline{\Lambda}_m\psi_n$, $\Phi_n^\circ = \left[\mathfrak{b}^2_{m_n^\circ}\mathfrak{b}_0^2 \vee 2\frac{m_n^\circ\overline{\Lambda}_{m_n^\circ}}{n}\psi_n\right]$, and finishing with $\kappa = \frac{47}{8}$, such that they verify the hypotheses of **??** and we obtain

$$\begin{aligned}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}(m)\right] \leqslant &\exp\left[\frac{\eta n}{2}\left(-\frac{3}{4}\mathfrak{b}^2_m(\theta^\circ) + \frac{63}{4}\frac{\psi_n m_n^\circ\overline{\Lambda_{m_n^\circ}}}{n} + \left(\frac{3}{4}\mathfrak{b}_0^2(\theta^\circ) + \kappa\right)\Phi_n^\circ\right)\right] + \\ &3\exp\left[-K\left(\frac{\psi_n 2m_n^\circ\overline{\Lambda_{m_n^\circ}}}{\Lambda_{(m_n^\circ)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \\ \leqslant &\exp\left[\frac{\eta n}{8}\left(-3\mathfrak{b}^2_m(\theta^\circ) + \frac{179}{2}\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\circ\right)\right] + \\ &3\exp\left[-K\left(\frac{\psi_n 2m_n^\circ\overline{\Lambda_{m_n^\circ}}}{\Lambda_{(m_n^\circ)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right].\end{aligned} \tag{H.5}$$

Building upon **??**, consider m in $\mathbb{N}$ such that $m > m_n^\circ$, for any k we set $\mathrm{pen}(k) = \kappa k\overline{\Lambda}_k\psi_n$, and $\kappa = \frac{47}{8}$, we obtain

$$\begin{aligned}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}(m)\right] \leqslant &\exp\left[\frac{n\eta}{2}\left(\left(2\kappa + \frac{5}{4}\mathfrak{b}_0^2(\theta^\circ)\right)\Phi_n^\circ + \psi_n\frac{m\overline{\Lambda}_m}{n}\left(\frac{39}{4} - 2\kappa\right) - \frac{5}{4}\mathfrak{b}^2_m(\theta^\circ)\right)\right] + \\ &3\exp\left[-K\left(\frac{\psi_n\delta^\star_{m,l}}{\Delta^\star_{m,l}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \\ \leqslant &\exp\left[\frac{n\eta}{2}\left(\frac{57}{8}\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\circ - 2\psi_n\frac{m\overline{\Lambda}_m}{n}\right)\right] + \\ &3\exp\left[-K\left(\frac{\psi_n 2m\overline{\Lambda}_m}{\Lambda_{(m)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right].\end{aligned} \tag{H.6}$$

We will now conclude using the definitions of G_n^- and G_n^+.

Consider EQ. (I.1). As for all $m < G_n^-$, we have $\mathfrak{b}^2_m(\theta^\circ) > 244\Phi_n^\circ\mathfrak{b}_0^2(\theta^\circ)$ and using EQ. (I.5)

$$\begin{aligned}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left([\![0, G_n^- - 1]\!]\right)\right] \leqslant &G_n^-\exp\left[-\frac{\mathfrak{b}_0^2(\theta^\circ)n\Phi_n^\circ}{2}\right] + 3G_n^-\exp\left[-K\left(\frac{\psi_n 2m_n^\circ\overline{\Lambda_{m_n^\circ}}}{\Lambda_{(m_n^\circ)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \\ \leqslant &4m_n^\circ\exp\left[-K\left(\frac{\psi_n 2m_n^\circ\overline{\Lambda_{m_n^\circ}}}{\Lambda_{(m_n^\circ)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right].\end{aligned}$$

Now for EQ. (I.2). As for all $m > G_n^+$, we have $\psi_n m\overline{\Lambda}_m > n\Phi_n^\circ\mathfrak{b}_0^2 \cdot \frac{59}{2}$, fixing $\eta > 59\cdot 2\cdot K > \frac{59\cdot 2\cdot K}{\Lambda_{(G_n^+)}\cdot\|\lambda\|\cdot\|\theta^\circ\|}$ and using EQ. (I.6)

$$
\begin{aligned}
\mathbb{E}^n_{\theta^\circ} & \left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket G^+_n+1,n\rrbracket\right)\right] \\
& \leqslant \sum_{G^+<m\leqslant n} \exp\left[\frac{n\eta}{2}\left(\frac{57}{8}\mathfrak{b}_0^2(\theta^\circ)\Phi^\circ_n - 2\psi_n\frac{m\overline{\Lambda}_m}{n}\right)\right] + \\
& \quad 3\sum_{G^+<m\leqslant n} \exp\left[-K\left(\frac{\psi_n 2m\overline{\Lambda_m}}{\Lambda_{(m)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \\
& \leqslant \sum_{G^+<m\leqslant n} \exp\left[\eta\left(-\frac{\mathfrak{b}_0^2(\theta^\circ)n\Phi^\circ_n}{2} - \frac{\psi_n m\overline{\Lambda}_m}{2}\right)\right] + \\
& \quad \exp\left[-K\left(\frac{\psi_n 2G^+_n\overline{\Lambda_{G^+_n}}}{\Lambda_{(G^+_n)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \cdot \sum_{G^+<m\leqslant n} \frac{\exp\left[-K\left(\frac{\psi_n 2m\overline{\Lambda_m}}{\Lambda_{(m)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right]}{\exp\left[-K\left(\frac{\psi_n 2G^+_n\overline{\Lambda_{G^+_n}}}{\Lambda_{(G^+_n)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right]} \\
& \leqslant \exp\left[-\frac{\eta\mathfrak{b}_0^2(\theta^\circ)n\Phi^\circ_n}{2}\right]\sum_{G^+<m\leqslant n}\exp\left[-\frac{\eta\psi_n m\overline{\Lambda}_m}{2}\right] + \\
& \quad \exp\left[-K\left(\frac{\psi_n 2G^+_n\overline{\Lambda_{G^+_n}}}{\Lambda_{(G^+_n)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \cdot \sum_{1<m\leqslant n} \exp\left[-K\left(\frac{\psi_n 2m\overline{\Lambda_m}}{\Lambda_{(m)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \\
& \leqslant C_{\lambda,\theta^\circ}\exp\left[-K\left(\frac{\psi_n 2G^+_n\overline{\Lambda_{G^+_n}}}{\Lambda_{(G^+_n)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right].
\end{aligned}
$$

Which completes the proof.

□

Proof of Proposition I.1.2

Begin with eq. (I.3).

We use this first decomposition in order to use eq. (I.2) and **??**

$$
\begin{aligned}
\sum_{0<|j|\leqslant n} & \Lambda_j\left(\lambda_j\overline{\theta}_j - \lambda_j\theta^\circ_j\right)^2 \mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket |j|,n\rrbracket\right) \\
& \leqslant \sum_{0<|j|\leqslant G^+_n}\Lambda_j\left(\lambda_j\overline{\theta}_j - \lambda_j\theta^\circ_j\right)^2 + \sum_{G^+_n<|j|\leqslant n}\Lambda_j\left(\lambda_j\overline{\theta}_j - \lambda_j\theta^\circ_j\right)^2\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket G^+_n+1,n\rrbracket\right) \\
& \leqslant \left\|\Pi_{G^+_n}\left(\theta^\circ-\overline{\theta}\right)\right\|^2_{l^2} + \left\|\Pi_{G^+_n,n}\left(\theta^\circ-\overline{\theta}\right)\right\|^2_{l^2}\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket G^+_n+1,n\rrbracket\right)
\end{aligned}
$$

Considering the result from **??**, and, hence, taking $\delta^\star_{G^+_n,n} \geqslant \sum_{G^+_n\leqslant|j|\leqslant n}\Lambda_j$, we obtain

$$\sum_{0<|j|\leqslant n} \Lambda_j \left(\lambda_j\overline{\theta}_j - \lambda_j\theta^\circ_j\right)^2 \mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket |j|, n\rrbracket\right)$$
$$\leqslant \left\|\Pi_{G_n^+}\left(\theta^\circ - \overline{\theta}\right)\right\|^2_{l^2} + \left(\sup_{t\in\mathbb{B}_{G_n^+,n}} \left|\langle t|\overline{\theta} - \theta^\circ\rangle_{l^2}\right|^2 - 6\frac{\psi_n\delta^\star_{G_n^+,n}}{n}\right)_+ +$$
$$6\frac{\psi_n\delta^\star_{G_n^+,n}}{n}\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket G_n^+ + 1, n\rrbracket\right)$$

We can hence use EQ. (I.2) and **??** to obtain

$$\sum_{0<|j|\leqslant n} \Lambda_j \,\mathbb{E}^n_{\theta^\circ}\left[\left(\lambda_j\overline{\theta}_j - \lambda_j\theta^\circ_j\right)^2 \mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket |j|, n\rrbracket\right)\right]$$
$$\leqslant \frac{2G_n^+\overline{\Lambda}_{G_n^+}}{n} +$$
$$C\left\{\frac{\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}\Delta^\star_{G_n^+,n}}{n}\exp\left[-\frac{\psi_n\delta^\star_{G_n^+,n}}{6\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}\Delta^\star_{G_n^+,n}}\right] + \frac{\delta^\star_{G_n^+,n}}{n^2}\exp\left[-K\sqrt{n\psi_n}\right]\right\} +$$
$$6\frac{\psi_n\delta^\star_{G_n^+,n}}{n}C_{\lambda,\theta^\circ}\exp\left[-K\left(2\frac{\psi_n m_n^\circ\overline{\Lambda}_{m_n^\circ}}{\Lambda_{(m_n^\circ)}\|\theta^\circ\|_{l^2}\|\lambda\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right].$$

Using the definition of G_n^+ and taking $\Delta^\star_{G_n^+,n} = \Lambda_{(n)}$ and $\delta^\star_{G_n^+,n} = 2n\overline{\Lambda}_n$ and the constraints $C_{\lambda,\theta^\circ} \geqslant C\left[\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2} \vee 1\right]$ we obtain

$$\sum_{0<|j|\leqslant n} \Lambda_j \,\mathbb{E}^n_{\theta^\circ}\left[\left(\lambda_j\overline{\theta}_j - \lambda_j\theta^\circ_j\right)^2 \mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket |j|, n\rrbracket\right)\right]$$
$$\leqslant 28\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\dagger +$$
$$\frac{1}{n}\left\{C\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}\Lambda_{(n)}\exp\left[-\frac{\psi_n 2n\overline{\Lambda}_n}{6\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}\Lambda_{(n)}}\right] + 2\overline{\Lambda}_n\exp\left[-K\sqrt{n\psi_n}\right]\right\} +$$
$$6\psi_n 2\overline{\Lambda}_n C_{\lambda,\theta^\circ}\exp\left[-K\left(2\frac{\psi_n m_n^\circ\overline{\Lambda}_{m_n^\circ}}{\Lambda_{(m_n^\circ)}\|\theta^\circ\|_{l^2}\|\lambda\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right]$$
$$\leqslant 28\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\dagger +$$
$$\frac{1}{n}C_{\lambda,\theta^\circ}12\psi_n n\Lambda_{(n)}C_{\lambda,\theta^\circ}\exp\left[-K\left(2\frac{\psi_n m_n^\circ\overline{\Lambda}_{m_n^\circ}}{\Lambda_{(m_n^\circ)}\|\theta^\circ\|_{l^2}\|\lambda\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right];$$

which proves the statement.

Consider now EQ. (I.4).

We split the sum in a similar manner:

$$\begin{aligned}
&\sum_{0<|j|\leqslant n} \left(\theta_j^\circ\right)^2 \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{M|Y^n}^{n,(\eta)} \left(\llbracket 0, j-1 \rrbracket\right)\right] + \sum_{|j|>n} \left(\theta_j^\circ\right)^2 \\
&\qquad \leqslant \sum_{j\in\llbracket 1, G_n^- \rrbracket} |\theta_j^\circ|^2 \, \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{M|Y^n}^{n,(\eta)} \left(\llbracket 1, |j|-1 \rrbracket\right)\right] + \sum_{|j|\in\llbracket G_n^-+1, n\rrbracket} |\theta_j^\circ|^2 + \sum_{|j|>n} |\theta_j^\circ|^2 \\
&\qquad \leqslant \|\theta^\circ\|_{l^2}^2 \, \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{M|Y^n}^{n,(\eta)} \left(\llbracket 0, G_n^- + 1 \rrbracket\right)\right] + \mathfrak{b}_{G_n^-}^2(\theta^\circ)
\end{aligned}$$

So we now use EQ. (I.1),

$$\begin{aligned}
&\sum_{0<|j|\leqslant n} \left(\theta_j^\circ\right)^2 \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{M|Y^n}^{n,(\eta)} \left(\llbracket 0, j-1 \rrbracket\right)\right] + \sum_{|j|>n} \left(\theta_j^\circ\right)^2 \\
&\qquad \leqslant \mathfrak{b}_{G_n^-}^2(\theta^\circ) + \frac{1}{n} \|\theta^\circ\|_{l^2}^2 \, 4 m_n^\circ \exp\left[-K \left(\frac{2\psi_n m_n^\circ \overline{\Lambda}_{m_n^\circ}}{\Lambda_{m_n^\circ} \|\theta^\circ\|_{l^2} \|\lambda\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right]
\end{aligned}$$

The proof is completed using the definition of G_n^-.
□

Appendix I

Proof for Theorem 3.4.3

I.1 Intermediate results

The proof of this theorem is mainly based on the one of ??. However, we need an altered version of ?? and ??.

Definition 48 Define the following quantities :

$$\begin{aligned} G_n^- &:= \min\left\{m \in [\![1, m_n^\circ]\!] : \quad \mathfrak{b}_m^2 \leqslant 244\mathfrak{b}_0^2\Phi_n^\circ\right\}, \\ G_n^+ &:= \max\left\{m \in [\![m_n^\circ, n]\!] : \psi_n m \Lambda_m \leqslant \frac{59}{2} n \mathfrak{b}_0^2 \Phi_n^\circ\right\}. \end{aligned}$$

Proposition I.1.1. *For any n and η in $\mathbb{N}$, with $\eta \geqslant 118K$, we have*

$$\mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left([\![0, G_n^- - 1]\!]\right)\right] \leqslant 4m_n^\circ \exp\left[-K\left(\frac{2\psi_n m_n^\circ \overline{\Lambda}_{m_n^\circ}}{\Lambda_{m_n^\circ}\|\theta^\circ\|_{l^2}\|\lambda\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \tag{I.1}$$

$$\mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left([\![G_n^+ + 1, n]\!]\right)\right] \leqslant C_{\lambda,\theta^\circ} \exp\left[-K\left(2\frac{\psi_n m_n^\circ \overline{\Lambda}_{m_n^\circ}}{\Lambda_{(m_n^\circ)}\|\theta^\circ\|_{l^2}\|\lambda\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \tag{I.2}$$

Proposition I.1.2.

$$\sum_{0<|j|\leqslant n} \Lambda_j \, \mathbb{E}_{\theta^\circ}^n\left[\left(\lambda_j\overline{\theta}_j - \lambda_j\theta_j^\circ\right)^2 \mathbb{P}_{M|Y^n}^{n,(\eta)}\left([\![|j|, n]\!]\right)\right]$$
$$\leqslant 28\mathfrak{b}_0^2\Phi_n^\circ + \frac{1}{n}12C_{\lambda,\theta^\circ}\exp\left[-K\left(\frac{\psi_n 2m_n^\dagger}{\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right) + \log\left(\psi_n n \Lambda_n\right)\right] \tag{I.3}$$

$$\sum_{0<|j|\leqslant n} \left(\theta_j^\circ\right)^2 \mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left([\![0, j-1]\!]\right)\right] + \sum_{|j|>n} \left(\theta_j^\circ\right)^2$$
$$\leqslant 59\mathfrak{b}_0^2\Phi_n^\dagger + \frac{1}{n}4C_{\lambda,\theta^\circ}\exp\left[-K\left(\frac{\psi_n 2m_n^\circ}{\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right) + \log\left(nm_n^\dagger\right)\right] \tag{I.4}$$

I.2 Detailed proofs

Proof of Proposition I.1.1

Building upon ??, consider m in $\mathbb{N}$ such that $m < m_n^\circ$. Set $\Delta^\star_{m,m_n^\circ} = \Lambda_{(m_n^\circ)}$ and $\delta^\star_{m,m_n^\circ} = m_n^\circ \overline{\Lambda}_{m_n^\circ}$, $\text{pen}(k) = \kappa m \overline{\Lambda}_m \psi_n$, $\Phi_n^\circ = \left[\mathfrak{b}^2_{m_n^\circ}\mathfrak{b}_0^2 \vee 2\frac{m_n^\circ \overline{\Lambda}_{m_n^\circ}}{n}\psi_n\right]$, and finishing with $\kappa = \frac{47}{8}$, such that they verify the hypotheses of ?? and we obtain

$$\begin{aligned}
\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}(m)\right] \leqslant & \exp\left[\frac{\eta n}{2}\left(-\frac{3}{4}\mathfrak{b}^2_m(\theta^\circ) + \frac{63}{4}\frac{\psi_n m_n^\circ \overline{\Lambda_{m_n^\circ}}}{n} + \left(\frac{3}{4}\mathfrak{b}_0^2(\theta^\circ) + \kappa\right)\Phi_n^\circ\right)\right] + \\
& 3\exp\left[-K\left(\frac{\psi_n 2 m_n^\circ \overline{\Lambda_{m_n^\circ}}}{\Lambda_{(m_n^\circ)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \\
\leqslant & \exp\left[\frac{\eta n}{8}\left(-3\mathfrak{b}^2_m(\theta^\circ) + \frac{179}{2}\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\circ\right)\right] + \\
& 3\exp\left[-K\left(\frac{\psi_n 2 m_n^\circ \overline{\Lambda_{m_n^\circ}}}{\Lambda_{(m_n^\circ)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right].
\end{aligned} \tag{I.5}$$

Building upon ??, consider m in $\mathbb{N}$ such that $m > m_n^\circ$, for any k we set $\text{pen}(k) = \kappa k \overline{\Lambda}_k \psi_n$, and $\kappa = \frac{47}{8}$, we obtain

$$\begin{aligned}
\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}(m)\right] \leqslant & \exp\left[\frac{n\eta}{2}\left(\left(2\kappa + \frac{5}{4}\mathfrak{b}_0^2(\theta^\circ)\right)\Phi_n^\circ + \psi_n\frac{m\overline{\Lambda}_m}{n}\left(\frac{39}{4} - 2\kappa\right) - \frac{5}{4}\mathfrak{b}^2_m(\theta^\circ)\right)\right] + \\
& 3\exp\left[-K\left(\frac{\psi_n \delta^\star_{m,l}}{\Delta^\star_{m,l}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \\
\leqslant & \exp\left[\frac{n\eta}{2}\left(\frac{57}{8}\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\circ - 2\psi_n\frac{m\overline{\Lambda}_m}{n}\right)\right] + \\
& 3\exp\left[-K\left(\frac{\psi_n 2 m \overline{\Lambda}_m}{\Lambda_{(m)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right].
\end{aligned} \tag{I.6}$$

We will now conclude using the definitions of G_n^- and G_n^+.

Consider EQ. (I.1). As for all $m < G_n^-$, we have $\mathfrak{b}^2_m(\theta^\circ) > 244\Phi_n^\circ\mathfrak{b}_0^2(\theta^\circ)$ and using EQ. (I.5)

$$\begin{aligned}
\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket 0, G_n^- - 1\rrbracket\right)\right] \leqslant & G_n^-\exp\left[-\frac{\mathfrak{b}_0^2(\theta^\circ) n \Phi_n^\circ}{2}\right] + 3G_n^-\exp\left[-K\left(\frac{\psi_n 2 m_n^\circ \overline{\Lambda_{m_n^\circ}}}{\Lambda_{(m_n^\circ)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \\
\leqslant & 4m_n^\circ\exp\left[-K\left(\frac{\psi_n 2 m_n^\circ \overline{\Lambda_{m_n^\circ}}}{\Lambda_{(m_n^\circ)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right].
\end{aligned}$$

Now for EQ. (I.2). As for all $m > G_n^+$, we have $\psi_n m \overline{\Lambda}_m > n\Phi_n^\circ \mathfrak{b}_0^2 \cdot \frac{59}{2}$, fixing $\eta > 59 \cdot 2 \cdot K > \frac{59 \cdot 2 \cdot K}{\Lambda_{(G_n^+)} \cdot \|\lambda\| \cdot \|\theta^\circ\|}$ and using EQ. (I.6)

$$\begin{aligned}
\mathbb{E}^n_{\theta^\circ} & \left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left([\![G_n^+ + 1, n]\!]\right)\right] \\
& \leqslant \sum_{G^+ < m \leqslant n} \exp\left[\frac{n\eta}{2}\left(\frac{57}{8}\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\circ - 2\psi_n \frac{m\overline{\Lambda}_m}{n}\right)\right] + \\
& \quad 3\sum_{G^+ < m \leqslant n} \exp\left[-K\left(\frac{\psi_n 2m\overline{\Lambda_m}}{\Lambda_{(m)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \\
& \leqslant \sum_{G^+ < m \leqslant n} \exp\left[\eta\left(-\frac{\mathfrak{b}_0^2(\theta^\circ)n\Phi_n^\circ}{2} - \frac{\psi_n m\overline{\Lambda}_m}{2}\right)\right] + \\
& \quad \exp\left[-K\left(\frac{\psi_n 2G_n^+\overline{\Lambda_{G_n^+}}}{\Lambda_{(G_n^+)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \cdot \sum_{G^+ < m \leqslant n} \frac{\exp\left[-K\left(\frac{\psi_n 2m\overline{\Lambda_m}}{\Lambda_{(m)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right]}{\exp\left[-K\left(\frac{\psi_n 2G_n^+\overline{\Lambda_{G_n^+}}}{\Lambda_{(G_n^+)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right]} \\
& \leqslant \exp\left[-\frac{\eta\mathfrak{b}_0^2(\theta^\circ)n\Phi_n^\circ}{2}\right]\sum_{G^+ < m \leqslant n} \exp\left[-\frac{\eta\psi_n m\overline{\Lambda}_m}{2}\right] + \\
& \quad \exp\left[-K\left(\frac{\psi_n 2G_n^+\overline{\Lambda_{G_n^+}}}{\Lambda_{(G_n^+)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \cdot \sum_{1 < m \leqslant n} \exp\left[-K\left(\frac{\psi_n 2m\overline{\Lambda_m}}{\Lambda_{(m)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right] \\
& \leqslant C_{\lambda,\theta^\circ} \exp\left[-K\left(\frac{\psi_n 2G_n^+\overline{\Lambda_{G_n^+}}}{\Lambda_{(G_n^+)}\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right].
\end{aligned}$$

Which completes the proof.

□

Proof of Proposition I.1.2

Begin with eq. (I.3).

We use this first decomposition in order to use eq. (I.2) and **??**

$$\begin{aligned}
\sum_{0<|j|\leqslant n} & \Lambda_j\left(\lambda_j\overline{\theta}_j - \lambda_j\theta_j^\circ\right)^2 \mathbb{P}^{n,(\eta)}_{M|Y^n}\left([\![|j|, n]\!]\right) \\
& \leqslant \sum_{0<|j|\leqslant G_n^+} \Lambda_j\left(\lambda_j\overline{\theta}_j - \lambda_j\theta_j^\circ\right)^2 + \sum_{G_n^+<|j|\leqslant n} \Lambda_j\left(\lambda_j\overline{\theta}_j - \lambda_j\theta_j^\circ\right)^2 \mathbb{P}^{n,(\eta)}_{M|Y^n}\left([\![G_n^+ + 1, n]\!]\right) \\
& \leqslant \left\|\Pi_{G_n^+}\left(\theta^\circ - \overline{\theta}\right)\right\|_{l^2}^2 + \left\|\Pi_{G_n^+,n}\left(\theta^\circ - \overline{\theta}\right)\right\|_{l^2}^2 \mathbb{P}^{n,(\eta)}_{M|Y^n}\left([\![G_n^+ + 1, n]\!]\right)
\end{aligned}$$

Considering the result from **??**, and, hence, taking $\delta^\star_{G_n^+,n} \geqslant \sum_{G_n^+\leqslant|j|\leqslant n} \Lambda_j$, we obtain

$$\sum_{0<|j|\leqslant n} \Lambda_j \left(\lambda_j\overline{\theta}_j - \lambda_j\theta^\circ_j\right)^2 \mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket |j|, n\rrbracket\right)$$
$$\leqslant \left\|\Pi_{G_n^+}\left(\theta^\circ - \overline{\theta}\right)\right\|^2_{l^2} + \left(\sup_{t\in\mathbb{B}_{G_n^+,n}} \left|\left\langle t|\overline{\theta} - \theta^\circ\right\rangle_{l^2}\right|^2 - 6\frac{\psi_n\delta^\star_{G_n^+,n}}{n}\right)_+ +$$
$$6\frac{\psi_n\delta^\star_{G_n^+,n}}{n}\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket G_n^+ + 1, n\rrbracket\right)$$

We can hence use EQ. (I.2) and **??** to obtain

$$\sum_{0<|j|\leqslant n} \Lambda_j\, \mathbb{E}^n_{\theta^\circ}\left[\left(\lambda_j\overline{\theta}_j - \lambda_j\theta^\circ_j\right)^2 \mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket |j|, n\rrbracket\right)\right]$$
$$\leqslant \frac{2G_n^+\overline{\Lambda}_{G_n^+}}{n} +$$
$$C\left\{\frac{\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}\Delta^\star_{G_n^+,n}}{n}\exp\left[-\frac{\psi_n\delta^\star_{G_n^+,n}}{6\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}\Delta^\star_{G_n^+,n}}\right] + \frac{\delta^\star_{G_n^+,n}}{n^2}\exp\left[-K\sqrt{n\psi_n}\right]\right\} +$$
$$6\frac{\psi_n\delta^\star_{G_n^+,n}}{n}C_{\lambda,\theta^\circ}\exp\left[-K\left(2\frac{\psi_n m_n^\circ\overline{\Lambda}_{m_n^\circ}}{\Lambda_{(m_n^\circ)}\|\theta^\circ\|_{l^2}\|\lambda\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right].$$

Using the definition of G_n^+ and taking $\Delta^\star_{G_n^+,n} = \Lambda_{(n)}$ and $\delta^\star_{G_n^+,n} = 2n\overline{\Lambda}_n$ and the constraints $C_{\lambda,\theta^\circ} \geqslant C\left[\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2} \vee 1\right]$ we obtain

$$\sum_{0<|j|\leqslant n} \Lambda_j\, \mathbb{E}^n_{\theta^\circ}\left[\left(\lambda_j\overline{\theta}_j - \lambda_j\theta^\circ_j\right)^2 \mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket |j|, n\rrbracket\right)\right]$$
$$\leqslant 28\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\dagger +$$
$$\frac{1}{n}\left\{C\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}\Lambda_{(n)}\exp\left[-\frac{\psi_n 2n\overline{\Lambda}_n}{6\|\lambda\|_{l^2}\|\theta^\circ\|_{l^2}\Lambda_{(n)}}\right] + 2\overline{\Lambda}_n\exp\left[-K\sqrt{n\psi_n}\right]\right\} +$$
$$6\psi_n 2\overline{\Lambda}_n C_{\lambda,\theta^\circ}\exp\left[-K\left(2\frac{\psi_n m_n^\circ\overline{\Lambda}_{m_n^\circ}}{\Lambda_{(m_n^\circ)}\|\theta^\circ\|_{l^2}\|\lambda\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right]$$
$$\leqslant 28\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\dagger +$$
$$\frac{1}{n}C_{\lambda,\theta^\circ}12\psi_n n\Lambda_{(n)}C_{\lambda,\theta^\circ}\exp\left[-K\left(2\frac{\psi_n m_n^\circ\overline{\Lambda}_{m_n^\circ}}{\Lambda_{(m_n^\circ)}\|\theta^\circ\|_{l^2}\|\lambda\|_{l^2}} \wedge \sqrt{n\psi_n}\right)\right];$$

which proves the statement.

Consider now EQ. (I.4).

We split the sum in a similar manner:

$$\begin{aligned}\sum_{0<|j|\leqslant n}\left(\theta_j^\circ\right)^2 \mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left(\llbracket 0, j-1\rrbracket\right)\right]+\sum_{|j|>n}\left(\theta_j^\circ\right)^2 \\ \leqslant \sum_{j\in\llbracket 1,G_n^-\rrbracket}|\theta_j^\circ|^2\,\mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left(\llbracket 1,|j|-1\rrbracket\right)\right]+\sum_{|j|\in\llbracket G_n^-+1,n\rrbracket}|\theta_j^\circ|^2+\sum_{|j|>n}|\theta_j^\circ|^2 \\ \leqslant \|\theta^\circ\|_{l^2}^2\,\mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left(\llbracket 0,G_n^-+1\rrbracket\right)\right]+\mathfrak{b}_{G_n^-}^2(\theta^\circ)\end{aligned}$$

So we now use EQ. (I.1),

$$\begin{aligned}\sum_{0<|j|\leqslant n}\left(\theta_j^\circ\right)^2 \mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left(\llbracket 0, j-1\rrbracket\right)\right]+\sum_{|j|>n}\left(\theta_j^\circ\right)^2 \\ \leqslant \mathfrak{b}_{G_n^-}^2(\theta^\circ)+\frac{1}{n}\|\theta^\circ\|_{l^2}^2\,4m_n^\circ\exp\left[-K\left(\frac{2\psi_n m_n^\circ\overline{\Lambda}_{m_n^\circ}}{\Lambda_{m_n^\circ}\|\theta^\circ\|_{l^2}\|\lambda\|_{l^2}}\wedge\sqrt{n\psi_n}\right)\right]\end{aligned}$$

The proof is completed using the definition of G_n^-.
□

Appendix J

Proof for Theorem 3.5.1

J.1 Intermediate results

Definition 49 Define the following quantities :

$$G_n^- := \min\left\{m \in [\![1, m_n^\dagger]\!] : \quad \mathfrak{b}_m^2 \leqslant 141\mathfrak{b}_0^2\Phi_n^\dagger\right\},$$
$$G_n^+ := \max\left\{m \in [\![m_n^\dagger, n]\!] : \psi_n m \Lambda_m \leqslant \frac{25}{4} n \mathfrak{b}_0^2 \Phi_n^\dagger\right\}.$$

Proposition J.1.1. *Using the notations of* definition 49, *we have*

$$\mathbb{E}_{\theta^\circ}^n\left[\left\|\widehat{\theta} - \theta^\circ\right\|_{l^2}^2\right] \leqslant \sum_{0<|j|\leqslant n} \mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left([\![|j|, n]\!]\right)\left|\left(\overline{\theta}_j - \theta_j^\circ\right)\right|^2\right] + \tag{J.1}$$
$$\mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left(\left[\!\left[1, G_n^- - 1\right]\!\right]\right)\right]\|\theta^\circ\|_{l^2}^2 + 141\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\dagger. \tag{J.2}$$

Proposition J.1.2. Control of eq. (J.10)
Using the notations from definition 49 *and* **??** *we have, under* **??**

$$\sum_{0<|j|\leqslant n} \mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left([\![|j|, n]\!]\right)\left|\overline{\theta}_j - \theta_j^\circ\right|^2\right] \leqslant \frac{25\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\dagger}{2\psi_n} +$$
$$\mathbb{E}_{\theta^\circ}^n\left[\left(\sup_{t \in \mathbb{B}_{G_n^+, n}} \left|\left\langle t|\overline{\theta} - \theta^\circ\right\rangle_{l^2}\right|^2 - 2\Lambda_{(n)}\psi_n\right)_+\right] +$$
$$2\Lambda_{(n)}\psi_n \mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left([\![G_n^+ + 1, n]\!]\right)\right]. \tag{J.3}$$

Proposition J.1.3. *For any n and η in $\mathbb{N}$, and constant $C_{\lambda,\theta^\circ} > \sum_{j=1}^{\infty} \exp\left[-\eta \frac{\psi_n m \Lambda_{(m)}}{2}\right]$ we have*

$$\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket 1, G^-_n - 1\rrbracket\right)\right] \leqslant G^-_n \exp\left[\frac{-\eta n \mathfrak{b}^2_0(\theta^\circ)\Phi^\dagger_n}{2}\right] + \sum_{1\leqslant |j| < G^-_n} \mathbb{P}^n_{\theta^\circ}\left(\sup_{x\in\mathbb{B}_{m,m^\dagger_n}} \left|\left\langle x|\Pi_{m,m^\dagger_n}(\overline{\theta}-\theta^\circ)\right\rangle\right|^2_{l^2} < \frac{2\Lambda_{(m^\dagger_n)} m^\dagger_n \psi_n}{n}\right) \tag{J.4}$$

$$\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}(\llbracket G^+_n + 1, n\rrbracket)\right] \leqslant C_{\theta^\circ\lambda} \exp\left[-\frac{\eta \mathfrak{b}^2_0(\theta^\circ) n \Phi^\dagger_n}{2}\right] + \sum_{G^+_n < |j| \leqslant n} \mathbb{P}^n_{\theta^\circ}\left(\sup_{x\in\mathbb{B}_{m^\dagger_n,m}} \left|\left\langle x|\Pi_{m^\dagger_n,m}(\overline{\theta}-\theta^\circ)\right\rangle\right|^2_{l^2} < \frac{2\Lambda_{(m)} m \psi_n}{n}\right) \tag{J.5}$$

Proposition J.1.4. *For any integer m and l such that $m \leqslant l$, define for any t in $\mathbb{B}_{m,l}$ the functional $\overline{\nu}_t = \langle t|\overline{\theta}-\theta^\circ\rangle_{l^2}$. Under ??, we define*

$$\begin{aligned}
\overline{\nu}^{e,\perp}_t &= \frac{1}{r}\sum_{q=1}^{r}\left(\upsilon_t(E^\perp_q) - \mathbb{E}^n_{\theta^\circ}\left[\upsilon_t(E^\perp_q)\right]\right);\\
\upsilon_t(E^\perp_q) &= \frac{1}{s}\sum_{p=1}^{s}\nu_t(E^\perp_{q,p});\\
\nu_t(E^\perp_{q,p}) &= \sum_{m\leqslant|j|\leqslant l}\left(\frac{t_j}{\overline{\lambda}_j} e_j(E^\perp_{q,p})\right).
\end{aligned}$$

Then, for any sequence $(C_n)_{n\in\mathbb{N}}$, we have the following inequalities:

$$\mathbb{E}^n_{\theta^\circ}\left[\left(\sup_{t\in\mathbb{B}_{m,l}}\left|\langle t|\overline{\theta}-\theta^\circ\rangle_{l^2}\right|^2 - C_n\right)_+\right] \leqslant 2\cdot\mathbb{E}^n_{\theta^\circ}\left[\left(\sup_{t\in\mathbb{B}_{m,l}}|\overline{\nu}^{e,\perp}_t|^2 - C_n\right)_+\right] + 2\cdot\mathbb{E}^n_{\theta^\circ}\left[\sup_{t\in\mathbb{B}_{m,l}}|\overline{\nu}^{e,\perp}_t - \overline{\nu}^e_t|^2\right] \tag{J.6}$$

$$\mathbb{P}^n_{\theta^\circ}\left(\sup_{t\in\mathbb{B}_{m,l}}\left|\langle t|\overline{\theta}-\theta^\circ\rangle_{l^2}\right| \geqslant C_n\right) \leqslant \mathbb{P}^n_{\theta^\circ}\left(\sup_{t\in\mathbb{B}_{m,l}}\left|\overline{\nu}^{e,\perp}_t\right| \geqslant C_n\right) + 3\,\mathbb{P}^n_{\theta^\circ}\left(\overline{\nu}^e_t \neq \overline{\nu}^{e,\perp}_t\right) \tag{J.7}$$

Proposition J.1.5. *For any integers m and l with $m < l$; consider a triplet h^2, H^2 and*

v verifying

$$\begin{aligned} h^2 &\geqslant \sum_{m\leqslant|j|\leqslant l} \Lambda_j; \\ H^2 &\geqslant \frac{\Lambda_{(l)}(l-m+1)(\psi_n+1)}{n}; \\ v &\geqslant \Lambda_{(l)}\|\theta^\circ\|_{l^2}\cdot\|\lambda\|_{l^2}; \end{aligned}$$

then, under **??***, for any* $C>0$*, we have:*

$$\mathbb{E}^n_{\theta^\circ}\left[\left(\sup_{t\in\mathbb{B}_{m,l}}|\overline{\nu}_t^{e,\perp}|^2-6H^2\right)_+\right] \leqslant C\left[\frac{v}{r}\exp\left(\frac{-rH^2}{6v}\right)+\frac{h^2}{r^2}\exp\left(\frac{-rH}{100h}\right)\right]; \tag{J.8}$$

$$\mathbb{P}^n_{\theta^\circ}\left(\sup_{t\in\mathbb{B}_{m,l}}|\overline{\nu}_t^{e,\perp}|\geqslant 6H^2\right) \leqslant 3\left(\exp\left[-\frac{rH^2}{400v}\right]+\exp\left[\frac{-rH}{100h}\right]\right). \tag{J.9}$$

Proposition J.1.6.

$$\mathbb{E}^n_{\theta^\circ}\left[\sup_{t\in\mathbb{B}_{m,l}}\left|\overline{\nu}_t^{e,\perp}-\overline{\nu}_t^{e}\right|^2\right] \leqslant 4\beta_s\sum_{m\leqslant|j|\leqslant l}\Lambda_j \tag{J.10}$$

$$\mathbb{P}^n_{\theta^\circ}\left(\overline{\nu}_t^{e}\neq\overline{\nu}_t^{e,\perp}\right) \leqslant r\beta_s \tag{J.11}$$

J.2 Detailed proofs

Proof for Proposition J.1.1

Using triangular inequality, linearity of the expectation, and the structure of our estimator, straightforward calculus yields:

$$\begin{aligned} \mathbb{E}^n_{\theta^\circ}\left[\left\|\widehat{\theta}-\theta^\circ\right\|^2_{l^2}\right] &= \mathbb{E}^n_{\theta^\circ}\left[\sum_{0<|j|\leqslant n}\left|\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket|j|,n\rrbracket\right)\left(\overline{\theta}_j-\theta^\circ_j\right)-\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket 1,|j|-1\rrbracket\right)\theta^\circ_j\right|^2\right] \\ &\quad+\sum_{|j|>n}|\theta^\circ|^2 \\ &\leqslant \sum_{0<|j|\leqslant n}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket|j|,n\rrbracket\right)\left|\left(\overline{\theta}_j-\theta^\circ_j\right)\right|^2\right]+ \\ &\quad\sum_{0<|j|\leqslant n}\left|\theta^\circ_j\right|^2\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket 1,|j|-1\rrbracket\right)\right]+\sum_{|j|>n}|\theta^\circ_j|^2. \end{aligned}$$

The second term can be decomposed as follows, using the definition of G_n^-:

$$\begin{aligned}
&\sum_{0<|j|\leqslant n} \left|\theta^\circ_j\right|^2 \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket 1,|j|-1\rrbracket\right)\right] + \sum_{|j|>n} \left|\theta^\circ_j\right|^2 \\
&\qquad\leqslant \sum_{0<|j|\leqslant G_n^- -1} \left|\theta^\circ_j\right|^2 \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket 1,|j|-1\rrbracket\right)\right] + \sum_{G_n^-\leqslant|j|\leqslant n} \left|\theta^\circ_j\right|^2 + \sum_{|j|>n} \left|\theta^\circ_j\right|^2 \\
&\qquad\leqslant \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\left\llbracket 1, G_n^- -1\right\rrbracket\right)\right] \|\theta^\circ\|^2_{l^2} + \mathfrak{b}^2_{G_n^-}(\theta^\circ) \\
&\qquad\leqslant \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\left\llbracket 1, G_n^- -1\right\rrbracket\right)\right] \|\theta^\circ\|^2_{l^2} + 141\mathfrak{b}_0^2\Phi_n^\dagger
\end{aligned}$$

.

which proves the statement.

□

Proof of Proposition J.1.2

First decompose the sum between the "good" and "bad" values of the threshold parameter:

$$\sum_{0<|j|\leqslant n} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket |j|,n\rrbracket\right)\left|\overline{\theta}_j-\theta^\circ_j\right|^2\right] \leqslant \sum_{0<|j|\leqslant G_n^+} \mathbb{E}^n_{\theta^\circ}\left[\left|\overline{\theta}_j-\theta^\circ_j\right|^2\right] + \tag{J.12}$$

$$\sum_{G_n^+<|j|\leqslant n} \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket G_n^+ +1,n\rrbracket\right)\left|\overline{\theta}_j-\theta^\circ_j\right|^2\right]. \tag{J.13}$$

We control eq. (J.12) using the definition of G_n^+ and **??** which can be applied thanks to **??**:

$$\begin{aligned}
\sum_{0<|j|\leqslant G_n^+} \mathbb{E}^n_{\theta^\circ}\left[\left|\overline{\theta}_j-\theta^\circ_j\right|^2\right] &= \sum_{0<|j|\leqslant G_n^+} \mathbb{V}^n_{\theta^\circ}\left[\overline{\theta}_j-\theta^\circ_j\right] \\
&= \sum_{0<|j|\leqslant G_n^+} \frac{\Lambda_j}{n^2}\mathbb{V}^n_{\theta^\circ}\left[\sum_{p=1}^{n} e_j(Y_p^n)\right] \\
&\leqslant \frac{\Lambda_{(G_n^+)}2G_n^+}{n} \\
&\leqslant \frac{25\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\dagger}{2\psi_n}
\end{aligned}$$

Consider eq. (J.13)

$$\sum_{G_n^+<|j|\leqslant n} \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left(\llbracket G_n^+ + 1, n\rrbracket\right) \left|\overline{\theta}_j - \theta_j^\circ\right|^2\right] = \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left(\llbracket G_n^+ + 1, n\rrbracket\right) \sum_{G_n^+<|j|\leqslant n} \left|\overline{\theta}_j - \theta_j^\circ\right|^2\right]$$
$$= \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left(\llbracket G_n^+ + 1, n\rrbracket\right) \left\|\Pi_{G_n^+,n}\left(\overline{\theta} - \theta^\circ\right)\right\|_{l^2}^2\right]$$
$$= \mathbb{E}_{\theta^\circ}^n \left[\left(\sup_{t\in\mathbb{B}_{G_n^+,n}} \left|\left\langle t|\overline{\theta} - \theta^\circ\right\rangle_{l^2}\right|^2 - 2\Lambda_{(n)}\psi_n\right)_+\right] + \tag{J.14}$$
$$2\Lambda_{(n)}\psi_n\, \mathbb{E}_{\theta^\circ}^n \left[\mathbb{P}_{M|Y^n}^{n,(\eta)}\left(\llbracket G_n^+ + 1, n\rrbracket\right)\right] \tag{J.15}$$

Proof of Proposition J.1.3

Before considering the two inequalities separately, let us do some observations.
Throughout the proof, m and l will be two positive integers.
For now, assume $m < l$ and let H be a positive real number.
We use ??, ?? in order to obtain, with the event $\mathcal{A}_{m,l} := \left\{\sup_{x\in\mathbb{B}_{m,l}} \left|\left\langle x|\Pi_{m,l}(\overline{\theta} - \theta^\circ)\right\rangle\right|_{l^2}^2 < 6H^2\right\}$

$$\mathbb{P}_{M|Y^n}^{n,(\eta)}(m) \leqslant \exp\left[\eta\left((\mathrm{pen}(l) - \mathrm{pen}(m)) + \frac{n}{2}\left(\frac{21}{4}\sup_{x\in\mathbb{B}_{m,l}} \left|\left\langle x|\Pi_{m,l}(\overline{\theta} - \theta^\circ)\right\rangle\right|_{l^2}^2 - \frac{3}{4}\left\|\Pi_{m,l}\theta^\circ\right\|_{l^2}^2\right)\right)\right]\mathbb{1}_{\mathcal{A}_{m,l}} + \mathbb{1}_{\mathcal{A}_{m,l}^c}$$

which gives us

$$\begin{aligned}
\mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}(m)\right] &\leqslant \mathbb{E}_{\theta^\circ}^n\left[\exp\left[\eta\left((\mathrm{pen}(l) - \mathrm{pen}(m)) + \frac{n}{2}\left(\frac{21}{4}\sup_{x\in\mathbb{B}_{m,l}} \left|\left\langle x|\Pi_{m,l}(\overline{\theta} - \theta^\circ)\right\rangle\right|_{l^2}^2 - \frac{3}{4}\left\|\Pi_{m,l}\theta^\circ\right\|_{l^2}^2\right)\right)\right]\mathbb{1}_{\mathcal{A}_{m,l}}\right] + \mathbb{P}_{\theta^\circ}^n\left(\mathcal{A}_{m,l}^c\right)\\
&\leqslant \exp\left[\eta\left((\mathrm{pen}(l) - \mathrm{pen}(m)) + \frac{n}{2}\left(\frac{63}{2}H^2 - \frac{3}{4}\left\|\Pi_{m,l}\theta^\circ\right\|_{l^2}^2\right)\right)\right] + \mathbb{P}_{\theta^\circ}^n\left(\sup_{x\in\mathbb{B}_{m,l}} \left|\left\langle x|\Pi_{m,l}(\overline{\theta} - \theta^\circ)\right\rangle\right|_{l^2}^2 > 6H^2\right)
\end{aligned}$$

In particular with $l = m_n^\dagger$, set $H^2 = 2\frac{\Lambda_{(m_n^\dagger)} m_n^\dagger \psi_n}{n}$. Using the definitions of pen, and $\Phi_n^\dagger$, and finishing with $\kappa = \frac{41}{4}$, we obtain

$$\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}(m)\right] \leqslant \exp\left[\frac{\eta n}{2}\left(\frac{419}{4}\mathfrak{b}^2_0(\theta^\circ)\Phi^\dagger_n - \frac{3}{4}\mathfrak{b}^2_m(\theta^\circ)\right)\right] + \mathbb{P}^n_{\theta^\circ}\left(\sup_{x\in\mathbb{B}_{m,m^\dagger_n}}\left|\left\langle x|\Pi_{m,m^\dagger_n}(\overline{\theta}-\theta^\circ)\right\rangle\right|^2_{l^2} > 6H^2\right). \tag{J.16}$$

Now, we assume $m > l$.

We use ??, and ?? in order to obtain, with the event $\mathcal{A}_{m,l} := \left\{\sup_{x\in\mathbb{B}_{l,m}}\left|\left\langle x|\Pi_{l,m}(\overline{\theta}-\theta^\circ)\right\rangle\right|^2_{l^2} < 6H^2\right\}$ for some real number H

$$\begin{aligned}\mathbb{P}^{n,(\eta)}_{M|Y^n}(m) \leqslant& \exp\left[\eta\left((\mathrm{pen}(l)-\mathrm{pen}(m))+\right.\right.\\ &\left.\left.\frac{n}{2}\left(\frac{13}{4}\sup_{x\in\mathbb{B}_{l,m}}\left|\left\langle x|(\overline{\theta}-\theta^\circ)\right\rangle_{l^2}\right|^2 + \frac{5}{4}\|\Pi_{l,m}\theta^\circ\|^2_{l^2}\right)\right)\right]\mathbb{1}_{\mathcal{A}_{m,l}} +\\ &\mathbb{1}^c_{\mathcal{A}_{m,l}}\\ \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}(m)\right] \leqslant& \exp\left[\eta\left((\mathrm{pen}(l)-\mathrm{pen}(m))+\right.\right.\\ &\left.\left.\frac{n}{2}\left(\frac{13}{2}3H^2 + \frac{5}{4}\|\Pi_{l,m}\theta^\circ\|^2_{l^2}\right)\right)\right] +\\ &\mathbb{P}^n_{\theta^\circ}\left(\sup_{x\in\mathbb{B}_{l,m}}\left|\left\langle x|\Pi_{l,m}(\overline{\theta}-\theta^\circ)\right\rangle\right|^2_{l^2} > 6H^2\right).\end{aligned}$$

In particular with $l = m^\dagger_n$, set $H^2 = 2\frac{\Lambda_{(m)}m\psi_n}{n}$. Using the definitions of pen and $\Phi^\dagger_n$ and finishing with $\kappa = 41/4$ we obtain

$$\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}(m)\right] \leqslant \exp\left[\frac{\eta n}{2}\left(\left(\frac{169}{4}\right)\mathfrak{b}^2_0(\theta^\circ)\Phi^\dagger_n - 2\frac{m\Lambda_{(m)}\psi_n}{n}\right)\right] + \mathbb{P}^n_{\theta^\circ}\left(\sup_{x\in\mathbb{B}_{m^\dagger_n,m}}\left|\left\langle x|\Pi_{m^\dagger_n,m}(\overline{\theta}-\theta^\circ)\right\rangle\right|^2_{l^2} > 6H^2\right). \tag{J.17}$$

We will now conclude using the definitions of $G^{\dagger-}_n$ and $G^{\dagger+}_n$.

Consider ??. As for all $m < G^{\dagger-}_n$, we have $\mathfrak{b}^2_m(\theta^\circ) > 141\Phi^\dagger_n\mathfrak{b}^2_0(\theta^\circ)$ and using EQ. (J.15)

$$\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}([\![1, G^-_n - 1]\!])\right] \leqslant G^-_n \exp\left[-\frac{\eta n}{2}\left(\mathfrak{b}^2_0(\theta^\circ)\Phi^\dagger_n\right)\right] + \sum_{m=1}^{G^-_n-1}\mathbb{P}^n_{\theta^\circ}\left(\sup_{x\in\mathbb{B}_{m,m^\dagger_n}}\left|\left\langle x|\Pi_{m,m^\dagger_n}(\overline{\theta}-\theta^\circ)\right\rangle\right|^2_{l^2} > 6H^2\right).$$

Consider ??. As for all $m > G_n^{\dagger+}$, we have $m\Lambda_{(m)}\psi_n > \frac{173}{4}\mathfrak{b}_0^2(\theta^\circ)n\Phi_n^\dagger$, with a constant $C_{\theta^\circ,\lambda} > \sum_{m=1}^{\infty} \exp\left[-\eta\frac{m\Lambda_{(m)}\psi_n}{2}\right]$ and using ??

$$\begin{aligned}\mathbb{E}_{\theta^\circ}^n\left[\mathbb{P}_{M|Y^n}^{n,(\eta)}([\![G_n^+ + 1, n]\!])\right] \leqslant \exp\left[-\eta\frac{\mathfrak{b}_0^2(\theta^\circ)n\Phi_n^\dagger}{2}\right]\sum_{m=G_n^+ +1}^{n}\exp\left[-\eta\frac{m\Lambda_{(m)}\psi_n}{2}\right] +\\ \sum_{m=G^+ +1}^{n}\mathbb{P}_{\theta^\circ}^n\left(\sup_{x\in\mathbb{B}_{m_n^\dagger,m}}\left|\left\langle x|\Pi_{m_n^\dagger,m}(\overline{\theta}-\theta^\circ)\right\rangle\right|_{l^2}^2 > 6H^2\right)\\ \leqslant C_{\theta^\circ,\lambda}\exp\left[-\eta\frac{\mathfrak{b}_0^2(\theta^\circ)n\Phi_n^\dagger}{2}\right] +\\ \sum_{m=G^+ +1}^{n}\mathbb{P}_{\theta^\circ}^n\left(\sup_{x\in\mathbb{B}_{m_n^\dagger,m}}\left|\left\langle x|\Pi_{m_n^\dagger,m}(\overline{\theta}-\theta^\circ)\right\rangle\right|_{l^2}^2 > 6H^2\right).\end{aligned}$$

Which proves the statement.

□

Proof of Proposition J.1.3

In this part, let m and l be two positive integers with $m < l$.

We have, for any t in $\mathbb{B}_{m,l}$

$$\begin{aligned}\left\langle t|\overline{\theta}\right\rangle_{l^2} =&\frac{1}{n}\sum_{p=1}^{n}\sum_{m\leqslant|j|\leqslant l}\left(\frac{t_j}{\overline{\lambda}_j}\cdot e_j\left(-Y_p^n\right)\right)\\ =&\frac{1}{n}\sum_{p=1}^{n}\mathcal{F}^{-1}\left(\frac{t}{\overline{\lambda}}\right)\left(-Y_p^n\right).\end{aligned}$$

So we define for any t in $\mathbb{B}_{m,l}$ the functional $\nu_t := \sum_{m\leqslant|j|\leqslant l}\left(\frac{t_j}{\overline{\lambda}_j}\right)e_j = \mathcal{F}^{-1}\left(\frac{t}{\overline{\lambda}}\right)$ and we obtain

$$\overline{\nu}_t := \left\langle t|\overline{\theta}-\theta^\circ\right\rangle_{l^2} = \frac{1}{n}\sum_{p=1}^{n}\left(\nu_t\left(Y_p\right) - \mathbb{E}_{\theta^\circ}^n\left[\nu_t\left(Y_p^n\right)\right]\right)$$

Then, for any t in $\mathbb{B}_{m,l}$ and x in $\mathcal{L}^2$ we define $v_t(x) = \frac{1}{s}\sum_{p=1}^{s}\nu_t(x_p)$, so we can write

$\frac{1}{n}\sum_{p=1}^{n}\nu_t(Y_p^n)=\frac{1}{2}\left\{\frac{1}{r}\sum_{q=1}^{r}\upsilon_t(E_q)+\frac{1}{r}\sum_{q=1}^{r}\upsilon_t(O_q)\right\}$, which gives

$$\begin{aligned}\langle t|\overline{\theta}-\theta^\circ\rangle =&\frac{1}{n}\sum_{p=1}^{n}\left(\nu_t(Y_p^n)-\mathbb{E}^n_{\theta^\circ}\left[\nu_t(Y_p^n)\right]\right)\\ =&\frac{1}{2}\left(\underbrace{\frac{1}{r}\sum_{q=1}^{r}\left(\upsilon_t(E_q)-\mathbb{E}^n_{\theta^\circ}\left[\upsilon_t(E_q)\right]\right)}_{=:\overline{\nu}^e_t}+\underbrace{\frac{1}{r}\sum_{q=1}^{r}\left(\upsilon_t(O_q)-\mathbb{E}^n_{\theta^\circ}\left[\upsilon_t(O_q)\right]\right)}_{=:\overline{\nu}^o_t}\right)\end{aligned}$$

Similarly, we define for any t in $\mathbb{B}_{m,l}$ the quantities $\overline{\nu}^{e,\perp}_t:=\frac{1}{r}\sum_{q=1}^{r}\left(\upsilon_t(E_q^\perp)-\mathbb{E}^n_{\theta^\circ}\left[\upsilon_t(E_q^\perp)\right]\right)$ and $\overline{\nu}^{o,\perp}_t:=\frac{1}{r}\sum_{q=1}^{r}\left(\upsilon_t(O_q^\perp)-\mathbb{E}^n_{\theta^\circ}\left[\upsilon_t(O_q^\perp)\right]\right)$ which combined give $\overline{\nu}^{\perp}_t:=\frac{1}{2}\left(\overline{\nu}^{e,\perp}_t+\overline{\nu}^{o,\perp}_t\right)$.

Consider first EQ. (J.6).

$$\begin{aligned}\mathbb{E}^n_{\theta^\circ}\quad&\left[\left(\sup_{t\in\mathbb{B}_{m,l}}\left|\langle t|\overline{\theta}-\theta^\circ\rangle_{l^2}\right|^2-C_n\right)_+\right]=\mathbb{E}^n_{\theta^\circ}\left[\left(\sup_{t\in\mathbb{B}_{m,l}}|\overline{\nu}_t|^2-C_n\right)_+\right]\\ &\leqslant\mathbb{E}^n_{\theta^\circ}\left[\left(\sup_{t\in\mathbb{B}_{m,l}}|\overline{\nu}^{e,\perp}_t|^2-C_n\right)_+\right]+\mathbb{E}^n_{\theta^\circ}\left[\sup_{t\in\mathbb{B}_{m,l}}|\overline{\nu}^{e,\perp}_t-\overline{\nu}^e_t|^2\right]+\\ &\quad\mathbb{E}^n_{\theta^\circ}\left[\left(\sup_{t\in\mathbb{B}_{m,l}}|\overline{\nu}^{o,\perp}_t|^2-C_n\right)_+\right]+\mathbb{E}^n_{\theta^\circ}\left[\sup_{t\in\mathbb{B}_{m,l}}|\overline{\nu}^{o,\perp}_t-\overline{\nu}^o_t|^2\right]\\ &\leqslant 2\cdot\mathbb{E}^n_{\theta^\circ}\left[\left(\sup_{t\in\mathbb{B}_{m,l}}|\overline{\nu}^{e,\perp}_t|^2-C_n\right)_+\right]+\\ &\quad 2\cdot\mathbb{E}^n_{\theta^\circ}\left[\sup_{t\in\mathbb{B}_{m,l}}|\overline{\nu}^{e,\perp}_t-\overline{\nu}^e_t|^2\right]\end{aligned}$$

Which proves the statement.

Consider now EQ. (J.7).

$$
\begin{aligned}
\mathbb{P}^n_{\theta^\circ}\left(\sup_{t\in\mathbb{B}_{m,l}}\left|\left\langle t|\overline{\theta}-\theta^\circ\right\rangle_{l^2}\right|\geqslant C'_n\right) &= \mathbb{P}^n_{\theta^\circ}\left(\sup_{t\in\mathbb{B}_{m,l}}|\overline{\nu}_t|\geqslant C'_n\right)\\
&= \mathbb{P}^n_{\theta^\circ}\left(\sup_{t\in\mathbb{B}_{m,l}}\left|\frac{1}{2}\left(\overline{\nu}^e_t-\overline{\nu}^{e,\perp}_t+\overline{\nu}^{e,\perp}_t+\overline{\nu}^o_t-\overline{\nu}^{o,\perp}_t+\overline{\nu}^{o,\perp}_t\right)\right|\geqslant C'_n\right)\\
&= \mathbb{P}^n_{\theta^\circ}\left(\left\{\sup_{t\in\mathbb{B}_{m,l}}\left|\frac{1}{2}\left(\overline{\nu}^e_t-\overline{\nu}^{e,\perp}_t+\overline{\nu}^{e,\perp}_t+\overline{\nu}^o_t-\overline{\nu}^{o,\perp}_t+\overline{\nu}^{o,\perp}_t\right)\right|\geqslant C'_n\right\}\cap\right.\\
&\qquad\left.\left\{\overline{\nu}^e_t=\overline{\nu}^{e,\perp}_t,\overline{\nu}^o_t=\overline{\nu}^{o,\perp}_t\right\}\right)+\\
&\qquad\mathbb{P}^n_{\theta^\circ}\left(\left\{\sup_{t\in\mathbb{B}_{m,l}}\left|\frac{1}{2}\left(\overline{\nu}^e_t-\overline{\nu}^{e,\perp}_t+\overline{\nu}^{e,\perp}_t+\overline{\nu}^o_t-\overline{\nu}^{o,\perp}_t+\overline{\nu}^{o,\perp}_t\right)\right|\geqslant C'_n\right\}\right.\\
&\qquad\left.\cap\left\{\overline{\nu}^e_t\neq\overline{\nu}^{e,\perp}_t,\overline{\nu}^o_t\neq\overline{\nu}^{o,\perp}_t\right\}\right)+\\
&\qquad\mathbb{P}^n_{\theta^\circ}\left(\left\{\sup_{t\in\mathbb{B}_{m,l}}\left|\frac{1}{2}\left(\overline{\nu}^e_t-\overline{\nu}^{e,\perp}_t+\overline{\nu}^{e,\perp}_t+\overline{\nu}^o_t-\overline{\nu}^{o,\perp}_t+\overline{\nu}^{o,\perp}_t\right)\right|\geqslant C'_n\right\}\right.\\
&\qquad\left.\cap\left\{\overline{\nu}^e_t\neq\overline{\nu}^{e,\perp}_t,\overline{\nu}^o_t=\overline{\nu}^{o,\perp}_t\right\}\right)+\\
&\qquad\mathbb{P}^n_{\theta^\circ}\left(\left\{\sup_{t\in\mathbb{B}_{m,l}}\left|\frac{1}{2}\left(\overline{\nu}^e_t-\overline{\nu}^{e,\perp}_t+\overline{\nu}^{e,\perp}_t+\overline{\nu}^o_t-\overline{\nu}^{o,\perp}_t+\overline{\nu}^{o,\perp}_t\right)\right|\geqslant C'_n\right\}\right.\\
&\qquad\left.\cap\left\{\overline{\nu}^e_t=\overline{\nu}^{e,\perp}_t,\overline{\nu}^o_t\neq\overline{\nu}^{o,\perp}_t\right\}\right)+\\
&\leqslant \mathbb{P}^n_{\theta^\circ}\left(\sup_{t\in\mathbb{B}_{m,l}}\left|\frac{1}{2}\left(\overline{\nu}^{e,\perp}_t+\overline{\nu}^{o,\perp}_t\right)\right|\geqslant C'_n\right)+\mathbb{P}^n_{\theta^\circ}\left(\overline{\nu}^e_t\neq\overline{\nu}^{e,\perp}_t,\overline{\nu}^o_t\neq\overline{\nu}^{o,\perp}_t\right)+\\
&\qquad\mathbb{P}^n_{\theta^\circ}\left(\overline{\nu}^e_t=\overline{\nu}^{e,\perp}_t,\overline{\nu}^o_t\neq\overline{\nu}^{o,\perp}_t\right)+\\
&\qquad\mathbb{P}^n_{\theta^\circ}\left(\overline{\nu}^e_t\neq\overline{\nu}^{e,\perp}_t,\overline{\nu}^o_t=\overline{\nu}^{o,\perp}_t\right)\\
&\leqslant \mathbb{P}^n_{\theta^\circ}\left(\sup_{t\in\mathbb{B}_{m,l}}\left|\max\left\{\overline{\nu}^{e,\perp}_t,\overline{\nu}^{o,\perp}_t\right\}\right|\geqslant C'_n\right)+\\
&\qquad\min\left\{\mathbb{P}^n_{\theta^\circ}\left(\overline{\nu}^e_t\neq\overline{\nu}^{e,\perp}_t\right),\mathbb{P}\left(\overline{\nu}^o_t\neq\overline{\nu}^{o,\perp}_t\right)\right\}+\\
&\qquad\min\left\{\mathbb{P}^n_{\theta^\circ}\left(\overline{\nu}^e_t=\overline{\nu}^{e,\perp}_t\right),\mathbb{P}\left(\overline{\nu}^o_t\neq\overline{\nu}^{o,\perp}_t\right)\right\}+\\
&\qquad\min\left\{\mathbb{P}^n_{\theta^\circ}\left(\overline{\nu}^e_t\neq\overline{\nu}^{e,\perp}_t\right),\mathbb{P}\left(\overline{\nu}^o_t=\overline{\nu}^{o,\perp}_t\right)\right\}\\
&\leqslant \mathbb{P}^n_{\theta^\circ}\left(\sup_{t\in\mathbb{B}_{m,l}}\left|\max\left\{\overline{\nu}^{e,\perp}_t,\overline{\nu}^{o,\perp}_t\right\}\right|\geqslant C'_n|\overline{\nu}^{e,\perp}_t\geqslant\overline{\nu}^{o,\perp}_t\right)\cdot\mathbb{P}\left(\overline{\nu}^{e,\perp}_t\geqslant\overline{\nu}^{o,\perp}_t\right)+\\
&\qquad\mathbb{P}^n_{\theta^\circ}\left(\sup_{t\in\mathbb{B}_{m,l}}\left|\max\left\{\overline{\nu}^{e,\perp}_t,\overline{\nu}^{o,\perp}_t\right\}\right|\geqslant C'_n|\overline{\nu}^{e,\perp}_t\leqslant\overline{\nu}^{o,\perp}_t\right)\cdot\mathbb{P}\left(\overline{\nu}^{e,\perp}_t\leqslant\overline{\nu}^{o,\perp}_t\right)+\\
&\qquad 3\,\mathbb{P}^n_{\theta^\circ}\left(\overline{\nu}^e_t\neq\overline{\nu}^{e,\perp}_t\right)
\end{aligned}
$$

$$
\begin{aligned}
&\leqslant \mathbb{P}^n_{\theta^\circ}\left(\sup_{t\in\mathbb{B}_{m,l}}\left|\overline{\nu}_t^{e,\perp}\right| \geqslant C_n'\right)\cdot \mathbb{P}\left(\overline{\nu}_t^{e,\perp} \geqslant \overline{\nu}_t^{o,\perp}\right) + \\
&\quad \mathbb{P}^n_{\theta^\circ}\left(\sup_{t\in\mathbb{B}_{m,l}}\left|\overline{\nu}_t^{o,\perp}\right| \geqslant C_n'\right)\cdot \mathbb{P}\left(\overline{\nu}_t^{e,\perp} \leqslant \overline{\nu}_t^{o,\perp}\right) + \\
&\quad 3\,\mathbb{P}^n_{\theta^\circ}\left(\overline{\nu}_t^{e} \neq \overline{\nu}_t^{e,\perp}\right) \\
&\leqslant \mathbb{P}^n_{\theta^\circ}\left(\sup_{t\in\mathbb{B}_{m,l}}\left|\overline{\nu}_t^{e,\perp}\right| \geqslant C_n'\right) + 3\,\mathbb{P}^n_{\theta^\circ}\left(\overline{\nu}_t^{e} \neq \overline{\nu}_t^{e,\perp}\right)
\end{aligned}
$$

Which completes the proof.

□

Proof of Proposition J.1.4

We will use Talagrand's inequality with, for any t in $\mathbb{B}_{m,l}$

$$
\begin{aligned}
\overline{\nu}_t^{e,\perp} &= \frac{1}{r}\sum_{q=1}^{r}\left(v_t(E_q^\perp) - \mathbb{E}^n_{\theta^\circ}\left[v_t(E_q^\perp)\right]\right) \\
v_t(E_q^\perp) &= \frac{1}{s}\sum_{p=1}^{s}\nu_t(E_{q,p}^\perp) \\
\nu_t(E_{q,p}^\perp) &= \sum_{m\leqslant|j|\leqslant l}\left(\frac{t_j}{\lambda_j}e_j(E_{q,p}^\perp)\right)
\end{aligned}
$$

$$
\begin{aligned}
\sup_{t\in\mathbb{B}_{m,l}}\sup_{x\in[0,1]^s}|v_t(x)|^2 &= \sup_{t\in\mathbb{B}_{m,l}}\sup_{x\in[0,1]^s}\left|\frac{1}{s}\sum_{p=1}^{s}\nu_t(x_p)\right|^2 \\
&= \sup_{t\in\mathbb{B}_{m,l}}\sup_{x\in[0,1]^s}\left|\frac{1}{s}\sum_{p=1}^{s}\sum_{m\leqslant|j|\leqslant l}\left(\frac{t_j}{\lambda_j}e_j(x_p)\right)\right|^2 \\
&\leqslant \sup_{t\in\mathbb{B}_{m,l}}\sup_{x\in[0,1]^s}\frac{1}{s^2}\sum_{m\leqslant|j|\leqslant l}|t_j|^2\,\Lambda_j\underbrace{\left|\sum_{p=1}^{s}e_j(x_p)\right|^2}_{\leqslant s^2} \\
&\leqslant \sup_{t\in\mathbb{B}_{m,l}}\sum_{m\leqslant|j|\leqslant l}|t_j|^2\,\Lambda_j \\
&\leqslant \sum_{m\leqslant|j|\leqslant l}\Lambda_j.
\end{aligned}
$$

Hence we define $h^2 := \delta^\star_{m,l} \geqslant \sum_{m\leqslant|j|\leqslant l}\Lambda_j$.

To define H^2, we define the following objects: $\overline{e}_j(E_q^\perp) := \frac{1}{s}\sum_{p=1}^{s} e_j(E_{q,p}^\perp)$ and $\overline{e}(E_q^\perp) = \left(\overline{e}_j(E_q^\perp)\right)_{j\in\mathbb{Z}}$.

We use **??** in the last line to conclude:

$$
\begin{aligned}
\mathbb{E}_{\theta^\circ}\left[\sup_{t\in\mathbb{B}_{m,l}}\left|\overline{\nu}_t^{e,\perp}\right|^2\right] = \quad & \mathbb{E}_{\theta^\circ}\left[\sup_{t\in\mathbb{B}_{m,l}}\left|\frac{1}{r}\sum_{q=1}^{r}\left(v_t(E_q^\perp)-\mathbb{E}_{\theta^\circ}^n\left[v_t(E_q^\perp)\right]\right)\right|^2\right]\\
= \quad \mathbb{E}_{\theta^\circ} \quad & \left[\sup_{t\in\mathbb{B}_{m,l}}\left|\frac{1}{r}\sum_{q=1}^{r}\left(\frac{1}{s}\sum_{p=1}^{s}\nu_t(E_{q,p}^\perp)-\mathbb{E}_{\theta^\circ}^n\left[\frac{1}{s}\sum_{p=1}^{s}\nu_t(E_{q,p}^\perp)\right]\right)\right|^2\right]\\
= \quad \mathbb{E}_{\theta^\circ} \quad & \left[\sup_{t\in\mathbb{B}_{m,l}}\left|\frac{1}{r}\sum_{q=1}^{r}\left(\frac{1}{s}\sum_{p=1}^{s}\sum_{m\leqslant|j|\leqslant l}\left(\frac{t_j}{\lambda_j}e_j(E_{q,p}^\perp)\right)\right.\right.\right.\\
& \left.\left.\left.-\,\mathbb{E}_{\theta^\circ}^n\left[\frac{1}{s}\sum_{p=1}^{s}\sum_{m\leqslant|j|\leqslant l}\left(\frac{t_j}{\lambda_j}e_j(E_{q,p}^\perp)\right)\right]\right)\right|^2\right]\\
= \quad \mathbb{E}_{\theta^\circ} \quad & \left[\sup_{t\in\mathbb{B}_{m,l}}\frac{1}{r^2}\left|\sum_{q=1}^{r}\left(\sum_{m\leqslant|j|\leqslant l}\frac{t_j}{\lambda_j}\sum_{p=1}^{s}\frac{1}{s}e_j(E_{q,p}^\perp)\right.\right.\right.\\
& \left.\left.\left.-\,\mathbb{E}_{\theta^\circ}^n\left[\sum_{m\leqslant|j|\leqslant l}\frac{t_j}{\lambda_j}\sum_{p=1}^{s}\frac{1}{s}e_j(E_{q,p}^\perp)\right]\right)\right|^2\right]\\
= \quad \mathbb{E}_{\theta^\circ} \quad & \left[\sup_{t\in\mathbb{B}_{m,l}}\frac{1}{r^2}\left|\sum_{q=1}^{r}\left\langle t\middle|\frac{\overline{e}(-E_q^\perp)}{\lambda}-\overline{\theta^\circ}\right\rangle\right|^2\right]\\
\leqslant \quad \mathbb{E}_{\theta^\circ} \quad & \left[\sup_{t\in\mathbb{B}_{m,l}}\frac{1}{r^2}\|t\|_{l^2}^2\sum_{q=1}^{r}\left\|\frac{\overline{e}(-E_q^\perp)}{\lambda}-\overline{\theta^\circ}\right\|_{l^2}^2\right]\\
\leqslant \quad \mathbb{E}_{\theta^\circ} \quad & \left[\frac{1}{r^2}\sum_{q=1}^{r}\left\|\Pi_{m,l}\left(\frac{\overline{e}(-E_q^\perp)}{\lambda}-\overline{\theta^\circ}\right)\right\|_{l^2}^2\right]\\
\leqslant \quad \frac{1}{r}\mathbb{E}_{\theta^\circ} \quad & \left[\left\|\Pi_{m,l}\left(\frac{\overline{e}(-E_1^\perp)}{\lambda}-\overline{\theta^\circ}\right)\right\|_{l^2}^2\right]\\
\leqslant \quad \frac{1}{r}\mathbb{E}_{\theta^\circ} \quad & \left[\sum_{m\leqslant|j|\leqslant l}\left|\frac{\overline{e}_j(E_1^\perp)}{\lambda_j}-\theta_j^\circ\right|^2\right]\\
\leqslant \quad \frac{1}{r}\sum_{m\leqslant|j|\leqslant l}\mathbb{E}_{\theta^\circ} \quad & \left[\left|\frac{\overline{e}_j(E_1^\perp)}{\lambda_j}-\theta_j^\circ\right|^2\right]\\
\leqslant \quad \frac{1}{r}\sum_{m\leqslant|j|\leqslant l}\left(\frac{\Lambda_j}{s^2}\mathbb{V}_{\theta^\circ}\right. \quad & \left.\left[\sum_{p=1}^{s}e_j(E_{1,p}^\perp)\right]+\left|\frac{1}{\lambda_j}\mathbb{E}_{\theta^\circ}\left[e_j(E_{1,1}^\perp)\right]-\theta_j^\circ\right|^2\right)\\
\leqslant \quad \frac{\Lambda_{(l)}}{rs} \quad & \left(2(l-m+1)\left\{1+2\left[\gamma_{\theta^\circ\lambda}\frac{S}{\sqrt{2(l-m+1)}}+\right.\right.\right.\\
& \left.\left.\left.2\sum_{p=S+1}^{s-1}\beta(E_{1,0}^\perp,E_{1,K}^\perp)\right]\right\}\right)
\end{aligned}
$$

In particular, we set $S := \left\lfloor \frac{\psi_n \sqrt{2(l-m+1)}}{\gamma_{\theta^\circ\lambda}} \right\rfloor \leqslant s-1$ so it tends to infinity and $\sum_{p=S+1}^{s-1} \beta(E_{1,0}^{\perp}, E_{1,K}^{\perp}) \leqslant \frac{1}{8}$ for $n \geqslant n^\circ$ and

$$\mathbb{E}_{\theta^\circ}\left[\sup_{t\in\mathbb{B}_{m,l}} \left|\overline{\nu}_t^{e,\perp}\right|^2\right] \leqslant \frac{6\Lambda_{(l)}(l-m+1)(\psi_n+1)}{n}$$

So we set $H^2 \geqslant \frac{6\Lambda_{(l)}(l-m+1)(\psi_n+1)}{n}$.
Finally we control v in the same way as in **??** and hence $v \geqslant \Lambda_{(l)}\|\theta^\circ\|_{l^2} \cdot \|\lambda\|_{l^2}$.
Using Talagrand's inequality gives us the result.
□

Proof of Proposition J.1.5
Both inequalities are verified using $\mathbb{P}_{\theta^\circ}^n\left(E_q \neq E_q^{\perp}\right) \leqslant \beta_s$ and, as it was proven in appendix J.2, $\|v_t\|_\infty^2 \leqslant \sum_{m\leqslant|j|\leqslant l} \Lambda_j \leqslant h^2$.

Consider first eq. (J.10).

$$\begin{aligned}
\mathbb{E}_{\theta^\circ}^n\left[\sup_{t\in\mathbb{B}_{m,l}} \left|\overline{\nu}_t^{e,\perp} - \overline{\nu}_t^{e}\right|^2\right] =& \frac{1}{r^2}\,\mathbb{E}_{\theta^\circ}^n\left[\sup_{t\in\mathbb{B}_{m,l}} \left|\sum_{q=1}^{r} v_t(E_q^{\perp}) - v_t(E_q)\right|^2\right] \\
\leqslant& 4\sup_{t\in\mathbb{B}_{m,l}} \|v_t\|_\infty^2 \beta_s \\
\leqslant& 4\beta_s \sum_{m\leqslant|j|\leqslant l} \Lambda_j
\end{aligned}$$

Consider now eq. (J.11)

$$\begin{aligned}
\mathbb{P}_{\theta^\circ}^n\left(\overline{\nu}_t^{e} \neq \overline{\nu}_t^{e,\perp}\right) \leqslant& \mathbb{P}_{\theta^\circ}^n\left(\bigcup_{q=1}^{r} E_q^{\perp} \neq E_q\right) \\
\leqslant& r\beta_s
\end{aligned}$$

Which completes the proof.
□

Proof of ??
We use eq. (J.15) and **??** in eq. (J.11) to obtain for any m and l such that $m < l$

$$\mathbb{P}_{\theta^\circ}^n\left(\sup_{t\in\mathbb{B}_{m,l}} \left|\langle t|\overline{\theta} - \theta^\circ\rangle_{l^2}\right|^2 \geqslant H^2\right) \leqslant 3\left(\exp\left[-\frac{rH^2}{400v}\right] + \exp\left[\frac{-rH}{100h}\right]\right) + 3r\beta_s.$$

We inject this result in eq. (J.12) with $H^2 = 2\frac{\Lambda_{(m_n^\dagger)} m_n^\dagger \psi_n}{n}$, $h^2 = \Lambda_{(m_n^\dagger)}$ and $v = \Lambda_{(m_n^\dagger)\|\theta^\circ\|_{l^2}\cdot\|\lambda\|_{l^2}}$

to obtain

$$\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket 1, G_n^- - 1\rrbracket\right)\right] \leqslant G_n^- \exp\left[\frac{-\eta n \mathfrak{b}_0^2(\theta^\circ)\Phi_n^\dagger}{2}\right] + 3\sum_{1\leqslant|j|<G_n^-}\left(\exp\left[-\frac{m_n^\dagger\psi_n}{400s\|\theta^\circ\|_{l^2}\cdot\|\lambda\|_{l^2}}\right] + \exp\left[\frac{-\sqrt{rm_n^\dagger\psi_n}}{100\sqrt{s}}\right] + r\beta_s\right) \tag{J.18}$$

and in similarly in EQ. (J.13) with $H^2 = 2\frac{\Lambda_{(m)}m\psi_n}{n}$, $h^2 = \Lambda_{(m)}$ and $v = \Lambda_{(m)\|\theta^\circ\|_{l^2}\cdot\|\lambda\|_{l^2}}$ and fixing $C_{\theta^\circ,\lambda} \geqslant \sum_{1<|j|\leqslant\infty}\left(\exp\left[-\frac{m\psi_n}{400s\|\theta^\circ\|_{l^2}\cdot\|\lambda\|_{l^2}}\right] + \exp\left[\frac{-\sqrt{rm\psi_n}}{100\sqrt{s}}\right]\right)$ to obtain

$$\begin{aligned}\mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}(\llbracket G_n^+ + 1, n\rrbracket)\right] \leqslant & C_{\theta^\circ\lambda}\exp\left[-\frac{\eta\mathfrak{b}_0^2(\theta^\circ)n\Phi_n^\dagger}{2}\right] + \\ & 3\left(\exp\left[-\frac{G_n^+\psi_n}{400s\|\theta^\circ\|_{l^2}\cdot\|\lambda\|_{l^2}}\right] + \exp\left[\frac{-\sqrt{rG_n^+\psi_n}}{100\sqrt{s}}\right] + G_n^+ r\beta_s\right)\cdot \\ & \sum_{1<|j|\leqslant\infty}\left(\exp\left[-\frac{m\psi_n}{400s\|\theta^\circ\|_{l^2}\cdot\|\lambda\|_{l^2}}\right] + \exp\left[\frac{-\sqrt{rm\psi_n}}{100\sqrt{s}}\right]\right) \\ \leqslant & C_{\theta^\circ\lambda}\exp\left[-\frac{\eta\mathfrak{b}_0^2(\theta^\circ)n\Phi_n^\dagger}{2}\right] + \\ & 3C_{\theta^\circ,\lambda}\left(\exp\left[-\frac{G_n^+\psi_n}{400s\|\theta^\circ\|_{l^2}\cdot\|\lambda\|_{l^2}}\right] + \exp\left[\frac{-\sqrt{rG_n^+\psi_n}}{100\sqrt{s}}\right] + G_n^+ r\beta_s\right)\end{aligned} \tag{J.19}$$

We use first EQ. (J.18) in EQ. (J.7) to obtain

$$\begin{aligned}\mathbb{E}^n_{\theta^\circ} & \left[\mathbb{P}^{n,(\eta)}_{M|Y^n}\left(\llbracket 1, G_n^- - 1\rrbracket\right)\right]\|\theta^\circ\|_{l^2}^2 + 141\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\dagger \\ & \leqslant G_n^- \exp\left[\frac{-\eta n\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\dagger}{2}\right] + \\ & 3\sum_{1\leqslant|j|<G_n^-}\left(\exp\left[-\frac{m_n^\dagger\psi_n}{400s\|\theta^\circ\|_{l^2}\cdot\|\lambda\|_{l^2}}\right] + \exp\left[\frac{-\sqrt{rm_n^\dagger\psi_n}}{100\sqrt{s}}\right] + r\beta_s\right).\end{aligned} \tag{J.20}$$

And we use ?? and EQ. (J.11) followed by ?? and EQ. (J.15) in EQ. (J.12) with $H^2 =$

$2\frac{\Lambda_{(n)} n\psi_n}{r}$, $h^2 = n\Lambda_n$ and $v = \Lambda_{(n)}\|\theta^\circ\|_{l^2} \cdot \|\lambda\|_{l^2}$ to obtain

$$\begin{aligned}\sum_{0<|j|\leqslant n} \quad & \mathbb{E}^n_{\theta^\circ}\left[\mathbb{P}^{n,(\eta)}_{M|Y^n}(\llbracket |j|, n\rrbracket)\,|\overline{\theta}_j - \theta^\circ_j|^2\right] \\ & \leqslant \frac{25\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\dagger}{2\psi_n} + \\ & 2 \cdot C\left[\frac{\Lambda_{(n)}\|\theta^\circ\|_{l^2} \cdot \|\lambda\|_{l^2}}{r}\exp\left(\frac{-n\Lambda_{(n)}\psi_n}{3\|\theta^\circ\|_{l^2} \cdot \|\lambda\|_{l^2}}\right) + \frac{n\Lambda_{(n)}}{r^2}\exp\left(\frac{-\sqrt{2r\psi_n}}{100}\right)\right] + \\ & 2 \cdot 4\beta_s \sum_{m\leqslant|j|\leqslant l} \Lambda_j + \\ & 2\Lambda_{(n)}\psi_n C_{\theta^\circ\lambda}\exp\left[-\frac{\eta\mathfrak{b}_0^2(\theta^\circ)n\Phi_n^\dagger}{2}\right] + \\ & 3C_{\theta^\circ,\lambda}\left(\exp\left[-\frac{G_n^+\psi_n}{400s\|\theta^\circ\|_{l^2} \cdot \|\lambda\|_{l^2}}\right] + \exp\left[\frac{-\sqrt{rG_n^+\psi_n}}{100\sqrt{s}}\right] + G_n^+ r\beta_s\right). \qquad \text{(J.21)}\end{aligned}$$

Finally, we combine ?? and ?? in Proposition J.1.1 to obtain the result:

$$\begin{aligned}\mathbb{E}^n_{\theta^\circ}\left[\left\|\widehat{\theta} - \theta^\circ\right\|^2_{l^2}\right] \leqslant & \frac{25\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\dagger}{2\psi_n} + \\ & 2 \cdot C\left[\frac{\Lambda_{(n)}\|\theta^\circ\|_{l^2} \cdot \|\lambda\|_{l^2}}{r}\exp\left(\frac{-n\Lambda_{(n)}\psi_n}{3\|\theta^\circ\|_{l^2} \cdot \|\lambda\|_{l^2}}\right) + \frac{n\Lambda_{(n)}}{r^2}\exp\left(\frac{-\sqrt{2r\psi_n}}{100}\right)\right] + \\ & 2 \cdot 4\beta_s \sum_{m\leqslant|j|\leqslant l} \Lambda_j + \\ & 2\Lambda_{(n)}\psi_n C_{\theta^\circ\lambda}\exp\left[-\frac{\eta\mathfrak{b}_0^2(\theta^\circ)n\Phi_n^\dagger}{2}\right] + \\ & 3C_{\theta^\circ,\lambda}\left(\exp\left[-\frac{G_n^+\psi_n}{400s\|\theta^\circ\|_{l^2} \cdot \|\lambda\|_{l^2}}\right] + \exp\left[\frac{-\sqrt{rG_n^+\psi_n}}{100\sqrt{s}}\right] + G_n^+ r\beta_s\right) + \\ & \left(G_n^- \exp\left[\frac{-\eta n\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\dagger}{2}\right] + \right. \\ & \left. 3\sum_{1\leqslant|j|<G_n^-}\left(\exp\left[-\frac{m_n^\dagger\psi_n}{400s\|\theta^\circ\|_{l^2} \cdot \|\lambda\|_{l^2}}\right] + \exp\left[\frac{-\sqrt{rm_n^\dagger\psi_n}}{100\sqrt{s}}\right] + r\beta_s\right)\right)\|\theta^\circ\|^2_{l^2} + \\ & 141\mathfrak{b}_0^2(\theta^\circ)\Phi_n^\dagger.\end{aligned}$$

which completes the proof.

□

Lemma J.2.1. *If $f = e_0$ then there is a finite numerical constant $\mathcal{C}$ such that for all $n \in \mathbb{N}$ we have $\mathbb{E}^n_Y\|\widetilde{f} - f\|^2_{L^2} \leqslant \mathcal{C}\Delta^\Lambda_{n_o} n^{-1}$ with $n_o := \lceil 15(\frac{300}{\sqrt{\kappa}})^4\rceil$.*

Proof of Lemma J.2.1.

Let $n_o := \lceil 15(\frac{300}{\sqrt{\kappa}})^4 \rceil$. We destinguish for $n \in \mathbb{N}$ the following two cases cases, (a) $n \in [\![1, n_o[\![$ and (b) $n \geqslant n_o$.

Consider (a). We select $m_+ = n \leqslant n_o$ and thus keeping in mind that $f = e_0$, and hence $\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 = 0$ from Lemma G.1.1 follows for all $n \in [\![1, n_o[\![$

$$\mathbb{E}_Y^n \|\widetilde{f} - f\|_{L^2}^2 \leqslant 2\mathbb{E}_Y^n \|\widetilde{f}n - \theta^{\circ,n}\|_{L^2}^2 \leqslant 4n\overline{\Lambda}_n n^{-1} \leqslant 4n_o \overline{\Lambda}_{n_o} n^{-1} \leqslant 4\Delta_{n_o}^{\Lambda} n^{-1}. \tag{J.22}$$

Consider (b), i.e., $n \geqslant n_o$. We select $m_+^{\diamond} := n_o \in [\![1, n]\!]$. Note that $\|\phi\|_{\ell^1} = 1$ and hence, $m_+ \geqslant m_+^{\diamond} \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$. Therefore, for all $n \geqslant n_o \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ due to Proposition G.1.1 follows

$$\begin{aligned}
\mathbb{E}_Y^n \|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\big\{[1 \vee \|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2] \mathcal{R}_n^{\diamond}(m_+^{\diamond}, f, \Lambda) + n^{-1}\big\} \\
+ 2\|\Pi_{\mathbb{U}_0^{\perp}} f\|_{L^2}^2 \mathfrak{b}_{m_-}^2(f) + 6\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 \exp\big(\tfrac{-\kappa \psi_{m_-^{\diamond}}^{\Lambda} m_-^{\diamond}}{400\|\phi\|_{\ell^1}}\big) \\
+ 2\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 \,[m_- - 1] \exp\big(-\kappa\kappa n \mathcal{R}_n^{\diamond}(m_-^{\diamond}, f, \Lambda) - \tfrac{\kappa \|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2}{4} n \mathfrak{b}_{[m_- - 1]}^2(f)\big).
\end{aligned}$$

Since $\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 = 0$, and thus $\mathcal{R}_n^{\diamond}(m_+^{\diamond}, f, \Lambda) = \Delta_{m_+^{\diamond}}^{\Lambda}/n = \Delta_{n_o}^{\Lambda}/n$, there is a numerical constant $\mathcal{C}$ such that $\mathbb{E}_Y^n \|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\Delta_{n_o}^{\Lambda} n^{-1}$ for all $n \geqslant n_o$. Combining the upper bounds for the two cases (a) and (b) we obtain the assertion which completes the proof.

Lemma J.2.2. *Assume there is $K \in \mathbb{N}$ with $1 \geqslant \mathfrak{b}_{[K-1]}(f) > 0$ and $\mathfrak{b}_K(f) = 0$. Set $K_h := K \vee 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$, $c_f := \frac{2\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 + 484\kappa}{\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 \mathfrak{b}_{[K-1]}^2(f)}$ and $n_{f,\Lambda} = \lceil c_f \Delta_{K_h}^{\Lambda} \vee 15(\frac{300}{\sqrt{\kappa}})^4 \rceil$. If $n \leqslant n_{f,\Lambda}$ then let $m_n^{\bullet} := K_h(\log n)$, and otherwise if $n > n_{f,\Lambda}$ then let $m_n^{\bullet} := \max\{m \in [\![K, n]\!] : c_f\, \Delta_m^{\Lambda} < n\}$ where the defining set contains K_h and thus it is not empty. There is a finite numerical constant $\mathcal{C}$ such that for all $n \in \mathbb{N}$ holds*

$$\mathbb{E}_Y^n \|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\{\Delta_{n_{f,\Lambda}}^{\Lambda} + \|\Pi_{\mathbb{U}_0^{\perp}} f\|_{L^2}^2 n_{f,\Lambda} + \|\phi\|_{\ell^1}^2\} n^{-1} + 6\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 \{\exp\big(\tfrac{-\kappa \psi_{m_n^{\bullet}}^{\Lambda} m_n^{\bullet}}{400\|\phi\|_{\ell^1}}\big) - \tfrac{1}{n}\}. \tag{J.23}$$

If there is $\widetilde{n}_{f,\Lambda} \in \mathbb{N}$ such that for all $n \geqslant \widetilde{n}_{f,\Lambda}$ in addition $\psi_{m_n^{\bullet}}^{\Lambda} m_n^{\bullet} \geqslant K_h(\log n)$ holds true then

$$\mathbb{E}_Y^n \|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\{\Delta_{[n_{f,\Lambda} \vee \widetilde{n}_{f,\Lambda}]}^{\Lambda} + \|\Pi_{\mathbb{U}_0^{\perp}} f\|_{L^2}^2 [n_{f,\Lambda} \vee \widetilde{n}_{f,\Lambda}] + \|\phi\|_{\ell^1}^2\} n^{-1}. \tag{J.24}$$

Proof of Lemma J.2.2.
Given $K \in \mathbb{N}$ with $1 \geqslant \mathfrak{b}_{[K-1]}(f) > 0$ and $\mathfrak{b}_m(f) = 0$ for all $m \geqslant K$ let $K_h := K \vee 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$, $c_f := \frac{2\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 + 484\kappa}{\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 \mathfrak{b}_{[K-1]}^2(f)}$ and $n_{f,\Lambda} = \lceil c_f \Delta_{K_h}^{\Lambda} \vee 15(\frac{300}{\sqrt{\kappa}})^4 \rceil$ we distinguish for $n \in \mathbb{N}$ the following two cases, (a) $n \in [\![1, n_{f,\Lambda}[\![$ and (b) $n > n_{f,\Lambda}$.

Firstly, consider (a), let $n \in [\![1, n_{f,\Lambda}[\![$, then setting $m_- = 1$ and $m_+ = n$ from Lemma

G.1.1 follows

$$\begin{aligned}\mathbb{E}_Y^n\|\widetilde{f}-f\|_{L^2}^2 \leqslant 2\mathbb{E}_Y^n\|\widetilde{f}n-\theta^{\circ,n}\|_{L^2}^2+2\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2\mathfrak{b}_1^2(f)\\ \leqslant 4n\overline{\Lambda}_n n^{-1}+2\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2 \leqslant 4n_{f,\Lambda}\overline{\Lambda}_{n_{f,\Lambda}}n^{-1}+2\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2 n_{f,\Lambda}n^{-1}\\ \leqslant (4\Delta_{n_{f,\Lambda}}^{\Lambda}+2\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2 n_{f,\Lambda})n^{-1}.\end{aligned} \tag{J.25}$$

Secondly, consider (b), i.e., $n > n_{f,\Lambda}$. Setting $m_+^\diamond := K_h \leqslant \Delta_{K_h}^\Lambda \leqslant n_{f,\Lambda}$, i.e., $m_+^\diamond \in [\![1,n]\!]$ from $m_+^\diamond = K_h \geqslant K$ follows $\mathfrak{b}_{m_+^\diamond}(f)=0$ and hence $\mathcal{R}_n^\diamond(m_+^\diamond,f,\Lambda)=\Delta_{K_h}^\Lambda n^{-1}$. Keeping in mind that $m_+^\diamond \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ and $n \geqslant n_o \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ from Proposition G.1.1 follows

$$\begin{aligned}\mathbb{E}_Y^n\|\widetilde{f}-f\|_{L^2}^2 \leqslant \mathcal{C}\{[1\vee\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2]\Delta_{K_h}^\Lambda n^{-1}+\|\phi\|_{\ell^1}^2 n^{-1}\}\\ +2\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2\mathfrak{b}_{m_-}^2(f)+6\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\exp\big(\tfrac{-\kappa\psi_{m_-^\diamond}^\Lambda m_-^\diamond}{400\|\phi\|_{\ell^1}}\big)\\ +2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2[m_- -1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond(m_-^\diamond,f,\Lambda)-\tfrac{\kappa\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2}{4}n\mathfrak{b}_{[m_- -1]}^2(f)\big).\end{aligned} \tag{J.26}$$

Since $n > n_{f,\Lambda} \geqslant c_f\Delta_{K_h}^\Lambda$ with $c_f = \frac{2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2+484\kappa}{\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\mathfrak{b}_{[K-1]}^2(f)}$ the defining set of $m_n^\bullet := \max\{m \in [\![K,n]\!] : n > c_{f,\Lambda}\Delta_m^\Lambda\}$ evenutally containing K_h is not empty. Consequently, $m_n^\bullet \geqslant K$ and $\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\mathfrak{b}_{[K-1]}^2(f) > [2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2+484\kappa]\Delta_{m_n^\bullet}^\Lambda/n = [2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2+484\kappa]\mathcal{R}_n^\diamond(m_n^\bullet,f,\Lambda)$. Therefore, setting $m_-^\diamond := m_n^\bullet$ the definition (G.10) implies $m_- = K$ and hence $\mathfrak{b}_{m_-}^2(f) = \mathfrak{b}_K^2(f)=0$, $\mathfrak{b}_{[m_- -1]}^2(f)=\mathfrak{b}_{[K-1]}^2(f)>0$. From (J.26) follows for all $n > n_{f,\Lambda}$ thus

$$\begin{aligned}\mathbb{E}_Y^n\|\widetilde{f}-f\|_{L^2}^2 \leqslant \mathcal{C}\{[1\vee\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2]\Delta_{K_h}^\Lambda n^{-1}+\|\phi\|_{\ell^1}^2 n^{-1}\}\\ +6\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\exp\big(\tfrac{-\kappa\psi_{m_n^\bullet}^\Lambda m_n^\bullet}{400\|\phi\|_{\ell^1}}\big)\\ +2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 m_n^\bullet\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond(m_-^\diamond,f,\Lambda)-\tfrac{\kappa\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2}{4}n\mathfrak{b}_{[K-1]}^2(f)\big)\\ \leqslant \mathcal{C}\{[1\vee\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2]\Delta_{K_h}^\Lambda n^{-1}+\|\phi\|_{\ell^1}^2 n^{-1}\}\\ +6\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\exp\big(\tfrac{-\kappa\psi_{m_n^\bullet}^\Lambda m_n^\bullet}{400\|\phi\|_{\ell^1}}\big)\\ +2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2[K-1]\underbrace{\exp\big(-\tfrac{\kappa\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2}{4}n\mathfrak{b}_{[m_- -1]}^2(f)\big)}_{\leqslant\frac{4}{\kappa\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\mathfrak{b}_{[K-1]}^2(f)}n^{-1}\exp(-1)}\end{aligned} \tag{J.27}$$

Note that $\Delta_{K_h}^\Lambda \leqslant n_{f,\Lambda}$ and $\frac{8[K-1]}{e\kappa\mathfrak{b}_{[K-1]}^2(f)} \leqslant \frac{1}{\kappa}\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 n_{f,\Lambda}$. Thereby, we obtain

$$\begin{aligned}\mathbb{E}_Y^n\|\widetilde{f}-f\|_{L^2}^2 \leqslant \mathcal{C}_2\{\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 n_{f,\Lambda}+\|\phi\|_{\ell^1}^2\}n^{-1}\\ +6\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\{\exp\big(\tfrac{-\kappa\psi_{m_n^\bullet}^\Lambda m_n^\bullet}{400\|\phi\|_{\ell^1}}\big)-\tfrac{1}{n}\}\end{aligned} \tag{J.28}$$

for some finite numerical constant $\mathcal{C}_2$.
Combining the upper bounds (J.25) and (J.28) for the two cases (a) and (b) we obtain the assertion (J.23), that is, there is a finite numerical constant $\mathcal{C}$ such that for all $n \in \mathbb{N}$ holds

$$\mathbb{E}_Y^n \|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\{\Delta_{n_{f,\Lambda}}^{\Lambda} + \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 n_{f,\Lambda} + \|\phi\|_{\ell^1}^2\} n^{-1}$$
$$+ 6\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 \{\exp\big(\tfrac{-\kappa \psi_{m_n^\bullet}^{\Lambda} m_n^\bullet}{400\|\phi\|_{\ell^1}}\big) - \tfrac{1}{n}\} \quad \text{(J.29)}$$

Assume finally, that there is in addition $\widetilde{n}_{f,\Lambda} \in \mathbb{N}$ such that $\psi_{m_n^\bullet}^{\Lambda} m_n^\bullet \geqslant K_h(\log n)$ for all $n \geqslant \widetilde{n}_{f,\Lambda}$. We shall use without further reference that then $\exp\big(\tfrac{-\kappa \psi_{m_n^\bullet}^{\Lambda} m_n^\bullet}{400\|\phi\|_{\ell^1}}\big) \leqslant n^{-1}$ for all $n \geqslant \widetilde{n}_{f,\Lambda}$ since $K_h \geqslant \frac{400\|\phi\|_{\ell^1}}{\kappa}$. Following line by line the proof of (J.29) using $\widetilde{n}_{f,\Lambda} \vee n_{f,\Lambda}$ rather than $n_{f,\Lambda}$ we obtain the assertion, that is, $\mathbb{E}_Y^n \|\widetilde{f} - f\|_{L^2}^2 \leqslant \mathcal{C}\{\Delta_{[n_{f,\Lambda} \vee \widetilde{n}_{f,\Lambda}]}^{\Lambda} + \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 [n_{f,\Lambda} \vee \widetilde{n}_{f,\Lambda}] + \|\phi\|_{\ell^1}^2\} n^{-1}$, which completes the proof.

Appendix K

Proof for Theorem 3.6.1

K.1 Intermediate results

K.2 Detailed proofs

Proof of ??.
We distinguish the cases (a) and (b) of Assumption 19.

Firstly, consider (a). Due to Lemma L.2.5 we have $\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant \mathcal{C}\{\Delta^{\Lambda}_{K_h} n^{-1} + K_h \Lambda^2_{(K_h)} q^{-1}\}$ since $n_{f,\Lambda} > \lceil 15(\frac{300}{\sqrt{\kappa}})^4 \rceil$.

Secondly, consider (b). Since under Assumption 19 for all $n \geqslant n_{f,\Lambda}$ and $q \geqslant q_{f,\Lambda}$ $\psi^{\Lambda}_{m_n^\bullet} m_n^\bullet \geqslant K_h(\log n)$ and $\psi^{\Lambda}_{m_q^\bullet} m_q^\bullet \geqslant K_h(\log q)$ hold true from Lemma L.2.6 (L.24) follows for all $n, q \in \mathbb{N}$

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant \mathcal{C}(1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2})(\Delta^{\Lambda}_{n_{f,\Lambda}} n^{-1} + K_h \Lambda_{(K_h)} q^{-1} + q_{f,\Lambda} q^{-1})$$

Keeping in mind that $\Delta^{\Lambda}_{n_{f,\Lambda}} \geqslant n_{f,\Lambda} \geqslant \Delta^{\Lambda}_{K_h}$ a combination of the upper bounds for both cases (a) and (b) implies the assertion (3.2), which completes the proof.

Lemma K.2.1. *If $f = e_0$ then there is a finite numerical constant $\mathcal{C}$ such that for all $n, q \in \mathbb{N}$ we have $\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant \mathcal{C}\{\Delta^{\Lambda}_{n_o} n^{-1} + n_o \Lambda^2_{(n_o)} q^{-1}\}$ with $n_o := \lceil 15\frac{300^4}{\kappa^2} \vee 3\frac{800^2}{\kappa^2} \rceil$.*

Proof of Lemma L.2.5.
Let $n_o := \lceil 15\frac{300^4}{\kappa^2} \vee 3\frac{800^2}{\kappa^2} \rceil$. We destinguish for $n \in \mathbb{N}$ the following two cases cases, (a) $n \in \llbracket 1, n_o \llbracket$ and (b) $n \geqslant n_o$. Consider (a). We select $m_+ = n \leqslant n_o$ and thus keeping in mind that $f = e_0$, and hence $\theta^{\circ}_j = 0$ and $\phi_j = 0$ for all $|j| \in \mathbb{N}$, from (L.4) (cf. proof of Lemma L.2.1) together with Lemma 1.5.1 (i) follows for all $q \in \mathbb{N}$

$$\begin{aligned}\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} &= 2 \sum_{j \in \llbracket 1,n \rrbracket} \mathbb{E}^{n,q}_{Y,\varepsilon}|\widehat{[g]}^{+}_j (\widehat{[h]}_j - \phi_j)\mathbb{P}_w(\llbracket j,n \rrbracket)|^2 \\ &\leqslant 2 \sum_{j \in \llbracket 1,n \rrbracket} \Lambda_j \mathbb{E}^q_{\varepsilon}|\lambda_j \widehat{[g]}^{+}_j|^2 \mathbb{E}^n_Y|\widehat{[h]}_j - \phi_j|^2 \leqslant 8 \sum_{j \in \llbracket 1,n \rrbracket} \Lambda_j n^{-1} = 8n\overline{\Lambda}_n n^{-1} \\ &\leqslant 8 n_o \overline{\Lambda}_{n_o} n^{-1} \leqslant 8\Delta^{\Lambda}_{n_o} n^{-1}. \quad \text{(K.1)}\end{aligned}$$

Consider (b), i.e., $n \geqslant n_o$. We select $m_+^\diamond := n_o \in [\![1,n]\!]$. Since $\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 = 0$, and thus $\|\phi\|_{\ell^1}^2 = 1$, $\mathcal{R}_n^\diamond(m_+^\diamond, f, \Lambda) = \Delta_{m_+^\diamond}^\Lambda/n = \Delta_{n_o}^\Lambda/n$ and $\mathcal{R}_q^\star(f,\Lambda) = 0$, from Proposition L.2.1 keeping in mind that by construction $m_+^\diamond \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2 = 3(\frac{800}{\kappa})^2$ and $n \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ follows

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ} - f\|_{L^2}^2 \leqslant \mathcal{C}\{\Delta_{n_o}^\Lambda n^{-1} + n^{-1}\} + 3q\mathbb{P}_\varepsilon^q(\mho_{n_o}^c) \tag{K.2}$$

Setting $q_o := \lceil 9\Lambda_{(n_o)}/4\rceil \in \mathbb{N}$ by employing ?? ?? we have $\mathbb{P}_\varepsilon^q(\mho_{n_o}^c) \leqslant 555 n_o q_o^2 q^{-2}$ for all $q \in \mathbb{N}$, where $q_o^2 \leqslant \frac{81}{4}\Lambda_{(n_o)}^2$. Combining the upper bounds (L.21) and (L.22) for the two cases (a) and (b) there is a finite numerical constant $\mathcal{C}$ such that $\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ} - f\|_{L^2}^2 \leqslant \mathcal{C}\{\Delta_{n_o}^\Lambda n^{-1} + n_o\Lambda_{(n_o)}^2 q^{-1}\}$ for all $n, q \in \mathbb{N}$, which shows the assertion and completes the proof.

Lemma K.2.2. *Consider the aggregated OSE* $\widehat{\theta^\circ} = \sum_{m=1}^n w_m \widehat{\theta^\circ}^m$ *with weights* $w_m \in [0,1]$, $m \in [\![1,n]\!]$, *satisfying* $\sum_{m=1}^n w_m = 1$ *and given* $m \in \mathbb{N}$ *let* $\check{f}^m := \sum_{j=-m}^m \widehat{[g]}_j^+ \phi_j e_j$. *For any* $m_- \in [\![1,n]\!]$ *and* $m_+ \in [\![1,n]\!]$ *holds*

$$\begin{aligned}
\|\widehat{\theta^\circ} - f\|_{L^2}^2 \leqslant 3\|\widehat{\theta^\circ}^{m_+} - \check{f}^{m_+}\|_{L^2}^2 + 3\sum_{l\in]\!]m_+,n]\!]}\left(\|\widehat{\theta^\circ}^l - \check{f}^l\|_{L^2}^2 - \mathcal{C}_1 12\kappa\Lambda_l^{\widehat{\Phi}}/n\right)_+\\
+3(\mathcal{C}_1\kappa/n)\sum_{l\in]\!]m_+,n]\!]}\Lambda_l^{\widehat{\Phi}}\mathbb{1}_{\{\|\widehat{\theta^\circ}^l - \check{f}^l\|_{L^2}^2 \geqslant 12\kappa\Lambda_l^{\widehat{\Phi}}/n\}} + 3(\kappa/n)\sum_{l\in]\!]m_+,n]\!]}\Lambda_l^{\widehat{\Phi}} w_l \mathbb{1}_{\{\|\widehat{\theta^\circ}^l - \check{f}^l\|_{L^2}^2 < 12\kappa\Lambda_l^{\widehat{\Phi}}/n\}}\\
+ 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2\{\mathbb{P}_w([\![1,m_-[\![) + \mathfrak{b}_{m_-}^2(f)\}\\
+ 6\sum_{j\in[\![1,n]\!]}|\widehat{[g]}_j^+|^2|\lambda_j - \widehat{[g]}_j|^2|\theta_j^\circ|^2 + 2\sum_{j\in[\![1,n]\!]}\mathbb{1}\mathcal{X}_j^c|\theta_j^\circ|^2
\end{aligned} \tag{K.3}$$

Lemma K.2.3. *Assume there is* $K \in \mathbb{N}$ *with* $1 \geqslant \mathfrak{b}_{[K-1]}(f) > 0$ *and* $\mathfrak{b}_K(f) = 0$. *Set* $K_h := K \vee 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$, $c_f := \frac{2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 + 7576\kappa}{\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\mathfrak{b}_{[K-1]}^2(f)} \geqslant 1$, $n_{f,\Lambda} = \lceil c_f\Delta_{K_h}^\Lambda \vee 15(\frac{300}{\sqrt{\kappa}})^4\rceil$ *and* $q_{f,\Lambda} := \lceil 289\log(K_h+2)\psi_{K_h}^\Lambda\Lambda_{(K_h)}\rceil$. *For all* $n, q \in \mathbb{N}$ *we set* $m_n^\bullet := K_h(\log n)$ *for* $n \leqslant n_{f,\Lambda}$, $m_n^\bullet := \max\{m \in [\![K_h, n]\!] : c_f\Delta_m^\Lambda < n\}$ *for* $n > n_{f,\Lambda}$, $m_q^\bullet := K_h(\log q)$ *for* $q \leqslant q_{K,\Lambda}$, *and* $m_q^\bullet := \max\{m \in [\![K_h, q]\!] : 289\log(m+2)\psi_m^\Lambda\Lambda_{(m)} \leqslant q\}$ *for* $q > q_{K,\Lambda}$, *where the defining sets contain* K_h *and thus they are not empty. There is a finite numerical constant* $\mathcal{C}$ *such that for all* $n, q \in \mathbb{N}$ *holds*

$$\begin{aligned}
\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ} - f\|_{L^2}^2 \leqslant \mathcal{C}(1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2)(\Delta_{n_{f,\Lambda}}^\Lambda n^{-1} + K_h\Lambda_{(K_h)}^2 q^{-1} + q_{f,\Lambda}q^{-1})\\
+ 9\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2\{\exp\left(\frac{-3\kappa\psi_{m_n^\bullet\wedge m_q^\bullet}^\Lambda(m_n^\bullet \wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\right) - \tfrac{1}{n} \vee \tfrac{1}{q}\}
\end{aligned} \tag{K.4}$$

If there are $\widetilde{n}_{f,\Lambda}, \widetilde{q}_{f,\Lambda} \in \mathbb{N}$ *such that additionally for all* $n \geqslant \widetilde{n}_{f,\Lambda}$, $\psi_{m_n^\bullet}^\Lambda m_n^\bullet \geqslant K_h(\log n)$, *and for all* $q \geqslant \widetilde{q}_{f,\Lambda}$, $\psi_{m_q^\bullet}^\Lambda m_q^\bullet \geqslant K_h(\log q)$ *hold true, then for all* $n, q \in \mathbb{N}$ *we have*

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ} - f\|_{L^2}^2 \leqslant \mathcal{C}(1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2)(\Delta_{[\widetilde{n}_{f,\Lambda}\vee n_{f,\Lambda}]}^\Lambda n^{-1} + K_h\Lambda_{(K_h)}^2 q^{-1} + [\widetilde{q}_{f,\Lambda} \vee q_{f,\Lambda}]q^{-1}) \tag{K.5}$$

Proof of Lemma L.2.6.

Given $K \in \mathbb{N}$ with $1 \geqslant \mathfrak{b}_{[K-1]}(f) > 0$ and $\mathfrak{b}_m(f) = 0$ for all $m \geqslant K$ we note that $\mathcal{R}_q^\star(f,\Lambda) = \sum_{j=1}^{K-1} |\theta_j^\circ|^2[1 \wedge q^{-1}\Lambda_j] \leqslant q^{-1}\Lambda_{(K)}\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2$. Let $K_h := K \vee 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$, $c_f := \frac{2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 + 7576\kappa}{\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 \mathfrak{b}_{[K-1]}^2(f)}$, $n_{f,\Lambda} = \lceil c_f \Delta_{K_h}^\Lambda \vee 15(\frac{300}{\sqrt{\kappa}})^4\rceil$ and $q_{f,\Lambda} := \lceil 289\log(K_h+2)\psi_{K_h}^\Lambda \Lambda_{(K_h)}\rceil$.

We distinguish for $n \in \mathbb{N}$ the following two cases, (a) $n \in [\![1, n_{f,\Lambda}[\![$ and (b) $n \geqslant n_{f,\Lambda}$.

Firstly, consider (a), let $n \in [\![1, n_{f,\Lambda}[\![$, then setting $m_- = 1$ and $m_+ = n$ from Lemma L.2.1 follows

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ} - f\|_{L^2}^2 \leqslant 3\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ}^n - \breve{f}^n\|_{L^2}^2 + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_1^2(f) \\ + 6\sum_{j\in[\![1,n]\!]} \mathbb{E}_\varepsilon^q|\widehat{[g]}_j^+|^2|\lambda_j - \widehat{[g]}_j|^2|\theta_j^\circ|^2 + 2\sum_{j\in[\![1,n]\!]} \mathbb{P}_\varepsilon^q(\mathcal{X}_j^c)|\theta_j^\circ|^2$$

Exploiting Lemma 1.5.1 we obtain from (i) $\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ}^n - \breve{f}^n\|_{L^2}^2 = 2\sum_{j\in[\![1,n]\!]}\Lambda_j\mathbb{E}_\varepsilon^q|\lambda_j\widehat{[g]}_j^+|^2\mathbb{E}_Y^n|\widehat{[h]}_j - \phi_j|^2 \leqslant 8n\overline{\Lambda}_n n^{-1} \leqslant 8\Delta_{n_{f,\Lambda}}^\Lambda n^{-1}$, from (iii) $\sum_{j\in[\![1,n]\!]}|\theta_j^\circ|^2\mathbb{E}_\varepsilon^q|\widehat{[g]}_j^+|^2|\lambda_j - \widehat{[g]}_j|^2 \leqslant 4\mathcal{C}_4\mathcal{R}_q^\star(f,\Lambda)$ and from (ii) $\sum_{j\in[\![1,n]\!]}\mathbb{P}_\varepsilon^q(\mathcal{X}_j^c)|\theta_j^\circ|^2 \leqslant 4\mathcal{R}_q^\star(f,\Lambda)$ where $\mathcal{R}_q^\star(f,\Lambda) := \sum_{j\in[\![1,K[\![}\theta_j^{\circ 2}[1\wedge\Lambda_j/q] \leqslant \|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\Lambda_{(K)}q^{-1}$. The last bounds together with $\mathfrak{b}_1^2(f) \leqslant 1$, $n < n_{f,\Lambda} \leqslant \Delta_{n_{f,\Lambda}}^\Lambda$ and $\Lambda_{(K)} \leqslant \Lambda_{(K_h)}$ imply for all $q \in \mathbb{N}$

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ} - f\|_{L^2}^2 \leqslant 24\Delta_{n_{f,\Lambda}}^\Lambda n^{-1} + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 + (24\mathcal{C}_4 + 8)\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\Lambda_{(K)}q^{-1} \\ \leqslant \mathcal{C}(1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2)\{\Delta_{n_{f,\Lambda}}^\Lambda n^{-1} + \Lambda_{(K_h)}q^{-1}\} \quad (K.6)$$

where $\mathcal{C}$ is a finite numerical constant.

Secondly, consider (b), i.e., $n \geqslant n_{f,\Lambda}$. Setting $m_+^\diamond := K_h \leqslant \Delta_{K_h}^\Lambda \leqslant n_{f,\Lambda}$, i.e., $m_+^\diamond \in [\![1,n]\!]$ from $m_+^\diamond = K_h \geqslant K$ follows $\mathfrak{b}_{m_+^\diamond}(f) = 0$ and hence $\mathcal{R}_n^\diamond(m_+^\diamond, f, \Lambda) = \Delta_{K_h}^\Lambda n^{-1}$. Consequently, from Proposition L.2.1 (keep in mind that $\mathcal{R}_q^\star(f,\Lambda) \leqslant \|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\Lambda_{(K)}q^{-1}$ and by construction $m_+^\diamond \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ and $n \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$) there is a numerical constant $\mathcal{C}$ such that

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ} - f\|_{L^2}^2 \leqslant \mathcal{C}\big\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2]\Delta_{K_h}^\Lambda n^{-1} + \|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2\Lambda_{(K)}q^{-1} + \|\phi\|_{\ell^1}^2 n^{-1}\big\} \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2\mathfrak{b}_{m_-}^2(f) + 9\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \exp\big(\frac{-3\kappa\psi_{m_-^\diamond}^\Lambda m_-^\diamond}{4000\|\phi\|_{\ell^1}}\big) \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond[m_-^\diamond, f, \Lambda] - \frac{\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{16}n\mathfrak{b}_{[m_- - 1]}^2(f)\big) \\ + 3q\mathbb{P}_\varepsilon^q(\mho_{K_h}^c) + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2\mathbb{P}_\varepsilon^q(\mho_{m_-^\diamond}^c)$$

Setting $q_h := \lceil 9\Lambda_{(K_h)}/4\rceil \in \mathbb{N}$ by employing **??** **??** we have $\mathbb{P}_\varepsilon^q(\mho_{K_h}^c) \leqslant 555K_h q_h^2 q^{-2}$ for all

$q \in \mathbb{N}$. Since $q_h^2 \leqslant \frac{81}{4}\Lambda_{(K_h)}^2$ and $\|\phi\|_{\ell^1}^2 \leqslant \kappa^2 K_h \leqslant \kappa^2 \Delta_{K_h}^\Lambda$ we obtain

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ} - f\|_{L^2}^2 \leqslant \mathcal{C}_1[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2](\Delta_{K_h}^\Lambda n^{-1} + K_h \Lambda_{(K_h)}^2 q^{-1})\\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_{m_-}^2(f) + 9\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \exp\big(\frac{-3\kappa\psi_{m_-^\diamond}^\Lambda m_-^\diamond}{4000\|\phi\|_{\ell^1}}\big)\\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 [m_- - 1]\exp\big(-\kappa\kappa n \mathcal{R}_n^\diamond[m_-^\diamond, f, \Lambda] - \frac{\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{16} n \mathfrak{b}_{[m_- -1]}^2(f)\big)\\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathbb{P}_\varepsilon^q(\mho_{m_-^\diamond}^c) \quad (K.7)\end{aligned}$$

where $\mathcal{C}_1$ is a numerical constant. In order to control the terms involving $m_-^\diamond$ and m_- we destinguish for $q \in \mathbb{N}$ the following two cases cases, (b-i) $q \in [\![1, q_{f,\Lambda}[\![$ and (b-ii) $q \geqslant q_{f,\Lambda}$ where $q_{f,\Lambda} = \lceil 289 \log(K_h + 2)\psi_{K_h}^\Lambda \Lambda_{(K_h)}\rceil$.

Consider (b-i). We set $m_-^\diamond = 1$ and hence $m_- = 1$. Thereby, from (L.26) together with $\mathfrak{b}_1^2(f) \leqslant 1$ and $q < q_{f,\Lambda}$ follows

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ} - f\|_{L^2}^2 &\leqslant \mathcal{C}_2[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2](\Delta_{K_h}^\Lambda n^{-1} + K_h \Lambda_{(K_h)}^2 q^{-1}) + 15\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2\\ &\leqslant \mathcal{C}_3[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2](\Delta_{K_h}^\Lambda n^{-1} + K_h \Lambda_{(K_h)}^2 q^{-1} + q_{f,\Lambda} q^{-1}) \quad (K.8)\end{aligned}$$

where $\mathcal{C}_3$ is a numerical constant.

Consider (b-ii). Note that for all $q \geqslant q_{K,\Lambda}$ the defining set of $m_q^\bullet := \max\{m \in [\![K_h, q]\!] : 289\log(m+2)\psi_m^\Lambda \Lambda_{(m)} \leqslant q\}$ is not empty, where obviously for each $m_-^\diamond \in [\![K_h, m_q^\bullet]\!]$ holds $q \geqslant 289\log(m_-^\diamond + 2)\psi_{m_-^\diamond}^\Lambda \Lambda_{(m_-^\diamond)}$, and thus from ?? ?? follows $\mathbb{P}_\varepsilon^q(\mho_{m_-^\diamond}^c) \leqslant 53q^{-1}$. Recall that $c_f := \frac{2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 + 7576\kappa}{\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 \mathfrak{b}_{[K-1]}^2(f)}$ and $n > n_{f,\Lambda} \geqslant c_f \Delta_{K_h}^\Lambda$. Thereby, the defining set of $m_n^\bullet := \max\{m \in [\![K_h, n]\!] : n > c_f \Delta_m^\Lambda\}$ contains K_h and it is not empty. Consequently, for all $m_-^\diamond \in [\![K_h, m_n^\bullet]\!]$ there hold $m_-^\diamond \geqslant K_h \geqslant K$, hence $\mathcal{R}_n^\diamond[m_-^\diamond, f, \Lambda] = \Delta_{m_-^\diamond}^\Lambda/n$ and thus $\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 \mathfrak{b}_{[K-1]}^2(f) > [2\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 + 7576\kappa]\mathcal{R}_n^\diamond[m_-^\diamond, f, \Lambda]$, which in turn employing the definition (??) implies $m_- = K$ and hence $\mathfrak{b}_{m_-}^2(f) = \mathfrak{b}_K^2(f) = 0$, $\mathfrak{b}_{[m_- -1]}^2(f) = \mathfrak{b}_{[K-1]}^2(f) > 0$. Selecting $m_-^\diamond := m_n^\bullet \wedge m_q^\bullet$ we have $\mathbb{P}_\varepsilon^q(\mho_{m_-^\diamond}^c) \leqslant 53q^{-1}$ and $m_- = K$,

$\mathfrak{b}^2_{m_-}(f) = 0$, $\mathfrak{b}^2_{[m_- -1]}(f) = \mathfrak{b}^2_{[K-1]}(f) > 0$, such that from (L.26) follows

$$\begin{aligned}
\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^\circ - f\|^2_{L^2} &\leqslant \mathcal{C}_2[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^\Lambda_{K_h} n^{-1} + K_h \Lambda^2_{(K_h)} q^{-1}) \\
&\quad + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2} \exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet \wedge m_q^\bullet}(m_n^\bullet \wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) \\
&\quad + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}[K-1] \exp\big(-\kappa\kappa n \mathcal{R}^\diamond_n[m_n^\bullet \wedge m_q^\bullet, f, \Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}}{16} n \mathfrak{b}^2_{[K-1]}(f)\big) \\
&\quad + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2} 53 q^{-1} \\
&\leqslant \mathcal{C}_3[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^\Lambda_{K_h} n^{-1} + K_h \Lambda^2_{(K_h)} q^{-1}) \\
&\quad + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2} \exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet \wedge m_q^\bullet}(m_n^\bullet \wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}[K-1] \underbrace{\exp\big(-\tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}}{16} n \mathfrak{b}^2_{[K-1]}(f)\big)}_{\leqslant \frac{16}{e\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{[K-1]}(f)} n^{-1}} \\
&\leqslant \mathcal{C}_3[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^\Lambda_{K_h} n^{-1} + K_h \Lambda^2_{(K_h)} q^{-1}) \\
&\quad + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2} \exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet \wedge m_q^\bullet}(m_n^\bullet \wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) + 3\tfrac{16[K-1]}{e\kappa\mathfrak{b}^2_{[K-1]}(f)} n^{-1} \\
&\leqslant \mathcal{C}_4[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\tfrac{\Delta^\Lambda_{K_h}}{\mathfrak{b}^2_{[K-1]}(f)} n^{-1} + K_h \Lambda^2_{(K_h)} q^{-1}) \\
&\quad + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2} \exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet \wedge m_q^\bullet}(m_n^\bullet \wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) \qquad \text{(K.9)}
\end{aligned}$$

where $\mathcal{C}_4$ is a numerical constant.

Combining the upper bounds (L.25), (L.27) and(L.28) for the three cases (a), (b-i) and (b-ii) and keeping in mind the definition of $m_n^\bullet$ and $m_q^\bullet$ there is a finite numerical constant $\mathcal{C}$ such that for all $n, q \in \mathbb{N}$ holds

$$\begin{aligned}
\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^\circ - f\|^2_{L^2} &\leqslant \mathcal{C}(1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2})(\Delta^\Lambda_{n_{f,\Lambda}} n^{-1} + \Lambda_{(K_h)} q^{-1}) \\
&\quad + \mathcal{C}[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^\Lambda_{K_h} n^{-1} + K_h \Lambda^2_{(K_h)} q^{-1} + q_{f,\Lambda} q^{-1}) \\
&\quad + \mathcal{C}[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\tfrac{\Delta^\Lambda_{K_h}}{\mathfrak{b}^2_{[K-1]}(f)} n^{-1} + K_h \Lambda^2_{(K_h)} q^{-1}) \\
&\quad + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2} \exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet \wedge m_q^\bullet}(m_n^\bullet \wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) \\
&\leqslant \mathcal{C}(1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2})(\Delta^\Lambda_{n_{f,\Lambda}} n^{-1} + K_h \Lambda^2_{(K_h)} q^{-1} + q_{f,\Lambda} q^{-1}) \\
&\quad + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2} \{\exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet \wedge m_q^\bullet}(m_n^\bullet \wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) - (n^{-1} \vee q^{-1})\} \qquad \text{(K.10)}
\end{aligned}$$

where we used that $\frac{\Delta^\Lambda_{K_h}}{\mathfrak{b}^2_{[K-1]}(f)} \leqslant n_{f,\Lambda} \leqslant \Delta^\Lambda_{n_{f,\Lambda}}$.

Assume finally, that there are in addition $\widetilde{n}_{f,\Lambda}, \widetilde{q}_{f,\Lambda} \in \mathbb{N}$ such that for all $n \geqslant \widetilde{n}_{f,\Lambda}$, $\psi^\Lambda_{m_n^\bullet} m_n^\bullet \geqslant K_h(\log n)$, and for all $q \geqslant \widetilde{q}_{f,\Lambda}$, $\psi^\Lambda_{m_q^\bullet} m_q^\bullet \geqslant K_h(\log q)$. We shall use without further reference that then $\exp\big(\frac{-3\kappa\psi^\Lambda_{m_n^\bullet \wedge m_q^\bullet}(m_n^\bullet \wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) \leqslant n^{-1} + q^{-1}$ for all $n \geqslant \widetilde{n}_{f,\Lambda}$ and

$q \geqslant \widetilde{q}_{f,\Lambda}$ since $K_h \geqslant \frac{4000\|\phi\|_{\ell^1}}{3\kappa}$. Following line by line the proof of (L.29) using $\widetilde{n}_{f,\Lambda} \vee n_{f,\Lambda}$ and $\widetilde{q}_{f,\Lambda} \vee q_{f,\Lambda}$ rather than $n_{f,\Lambda}$ and $q_{f,\Lambda}$, respectively, we obtain the assertion, that is, s $\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^\circ} - f\|^2_{L^2} \leqslant \mathcal{C}(1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2})(\Delta^{\Lambda}_{[\widetilde{n}_{f,\Lambda} \vee n_{f,\Lambda}]} n^{-1} + K_h \Lambda^2_{(K_h)} q^{-1} + [\widetilde{q}_{f,\Lambda} \vee q_{f,\Lambda}] q^{-1})$, which completes the proof.

Proof of Lemma L.2.1.

We start the proof with the observation that $\overline{[\widehat{\theta^\circ}]}_j - \overline{[f]}_j = \widehat{\theta^\circ}_{-j} - \theta^\circ_{-j}$ for all $j \in \mathbb{Z}$, $\widehat{\theta^\circ}_0 - \theta^\circ_0 = 0$ and $\widehat{\theta^\circ}_j - \theta^\circ_j = -\theta^\circ_j$ for all $j > n$, while for all $j \in [\![1,n]\!]$ with $\mathcal{X}_j := \{|\widehat{[g]}_j|^2 \geqslant 1/q\}$ and $\mathcal{X}^c_j := \{|\widehat{[g]}_j|^2 < 1/q\}$ holds

$$\begin{aligned}\widehat{\theta^\circ}_j - \theta^\circ_j &= (\widehat{[g]}^{+}_j \widehat{[h]}_j - \theta^\circ_j)\mathbb{P}_w([\![j,n]\!]) - \theta^\circ_j \mathbb{P}_w([\![1,j[\![)\\ &= \widehat{[g]}^{+}_j(\widehat{[h]}_j - \phi_j)\mathbb{P}_w([\![j,n]\!]) + \widehat{[g]}^{+}_j(\lambda_j - \widehat{[g]}_j)\theta^\circ_j\mathbb{P}_w([\![j,n]\!]) - \mathbb{1}_{\mathcal{X}_j}\theta^\circ_j\mathbb{P}_w([\![1,j[\![) - \mathbb{1}_{\mathcal{X}^c_j}\theta^\circ_j\end{aligned}$$

Consequently, we have

$$\begin{aligned}\|\widehat{\theta^\circ} - f\|^2_{L^2} &= 2\sum_{j\in[\![1,n]\!]} |\widehat{[g]}^{+}_j(\widehat{[h]}_j - \phi_j)\mathbb{P}_w([\![j,n]\!]) + \widehat{[g]}^{+}_j(\lambda_j - \widehat{[g]}_j)\theta^\circ_j\mathbb{P}_w([\![j,n]\!]) - \theta^\circ_j\mathbb{P}_w([\![1,j[\![)|^2\mathbb{1}_{\mathcal{X}_j}\\ &\qquad + 2\sum_{j\in[\![1,n]\!]}\mathbb{1}_{\mathcal{X}^c_j}|\theta^\circ_j|^2 + 2\sum_{j>n}|\theta^\circ_j|^2\\ &\leqslant 6\sum_{j\in[\![1,n]\!]}|\widehat{[g]}^{+}_j|^2|\widehat{[h]}_j - \phi_j|^2\mathbb{P}_w([\![j,n]\!]) + 6\sum_{j\in[\![1,n]\!]}\mathbb{1}_{\mathcal{X}_j}|\theta^\circ_j|^2\mathbb{P}_w([\![1,j[\![) + 2\sum_{j>n}|\theta^\circ_j|^2\\ &\qquad + 6\sum_{j\in[\![1,n]\!]}|\widehat{[g]}^{+}_j|^2|\lambda_j - \widehat{[g]}_j|^2|\theta^\circ_j|^2 + 2\sum_{j\in[\![1,n]\!]}\mathbb{1}_{\mathcal{X}^c_j}|\theta^\circ_j|^2. \end{aligned} \tag{K.11}$$

Consider the first r.hs. term in (L.4). We split the sum into two parts which we bound separately. Precisely, given $\breve{f}^m = \sum_{j=-m}^{m}\widehat{[g]}^{+}_j\phi_j e_j$ where $\|\widehat{\theta^\circ}^m - \breve{f}^m\|^2_{L^2} = 2\sum_{j\in[\![1,m]\!]}|\widehat{\theta^\circ}^m_j - \breve{f}^m_j|^2 = 2\sum_{j\in[\![1,m]\!]}|\widehat{[g]}^{+}_j|^2|\widehat{[h]}_j - \phi_j|^2$ it follows

$$\begin{aligned}&2\sum_{j\in[\![1,n]\!]}|\widehat{[g]}^{+}_j|^2(\widehat{[h]}_j - \phi_j)^2\mathbb{P}_w([\![j,n]\!])\\ &\quad\leqslant 2\sum_{j\in[\![1,m_+]\!]}|\widehat{[g]}^{+}_j|^2|(\widehat{[h]}_j - \phi_j)^2 + 2\sum_{j\in]\!]m_+,n]\!]}|\widehat{[g]}^{+}_j|^2(\widehat{[h]}_j - \phi_j)^2\sum_{l\in[\![j,n]\!]}w_l\\ &\quad= 2\sum_{j\in[\![1,m_+]\!]}|\widehat{[g]}^{+}_j|^2|(\widehat{[h]}_j - \phi_j)^2 + 2\sum_{l\in]\!]m_+,n]\!]}w_l\sum_{j\in]\!]m_+,l]\!]}|\widehat{[g]}^{+}_j|^2(\widehat{[h]}_j - \phi_j)^2\\ &\quad\leqslant \|\widehat{\theta^\circ}^{m_+} - \breve{f}^{m_+}\|^2_{L^2} + \sum_{l\in]\!]m_+,n]\!]}w_l\|\widehat{\theta^\circ}^l - \breve{f}^l\|^2_{L^2}\\ &\quad\leqslant \|\widehat{\theta^\circ}^{m_+} - \breve{f}^{m_+}\|^2_{L^2} + \sum_{l\in]\!]m_+,n]\!]}(\|\widehat{\theta^\circ}^l - \breve{f}^l\|^2_{L^2} - \mathcal{C}_1 12\kappa\widehat{\Lambda}^{\Phi}_l/n)_+\\ &+(\mathcal{C}_1\kappa/n)\sum_{l\in]\!]m_+,n]\!]}\Lambda^{\widehat{\Phi}}_l\mathbb{1}_{\{\|\widehat{\theta^\circ}^l - \breve{f}^l\|^2_{L^2} \geqslant 12\kappa\Lambda^{\widehat{\Phi}}_l/n\}} + (\kappa/n)\sum_{l\in]\!]m_+,n]\!]}\Lambda^{\widehat{\Phi}}_l w_l\mathbb{1}_{\{\|\widehat{\theta^\circ}^l - \breve{f}^l\|^2_{L^2} < 12\kappa\Lambda^{\widehat{\Phi}}_l/n\}}\end{aligned} \tag{K.12}$$

Consider the second and third r.hs. term in (L.4). Splitting the first sum into two parts we obtain

$$
\begin{aligned}
2\sum_{j\in[\![1,n]\!]} \mathbb{1}_{\mathcal{X}_j}|\theta^\circ_j|^2\mathbb{P}_w([\![1,j[\![) + 2\sum_{j>n}|\theta^\circ_j|^2 \\
\leqslant 2\sum_{j\in[\![1,m_-]\!]}|\theta^\circ_j|^2\mathbb{1}_{\mathcal{X}_j}\mathbb{P}_w([\![1,j[\![) + 2\sum_{j\in]\!]m_-,n]\!]}|\theta^\circ_j|^2 + 2\sum_{j>n}|\theta^\circ_j|^2 \\
\leqslant \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\{\mathbb{P}_w([\![1,m_-[\![) + \mathfrak{b}^2_{m_-}(f)\} \quad \text{(K.13)}
\end{aligned}
$$

Combining (L.4) and the upper bounds (L.5) and (L.6) we obtain the assertion, which completes the proof.

Lemma K.2.4. *For any $l \in [\![1,n]\!]$*

(i) *with $\mathcal{R}^\diamond_n[l,f,\Lambda] = [\mathfrak{b}^2_l(f)\vee\Delta^\Lambda_l n^{-1}]$ for all $k\in[\![1,l[\![$ we have*
$$\widehat{w}_k\mathbb{1}\Big\{\|\widehat{\theta^\circ}^l - \check{f}^l\|^2_{L^2} < 12\kappa\Lambda^{\widehat{\Phi}}_l/n\Big\}\cap\Big\{1/2\leqslant|\lambda_j\widehat{[g]}^+_j|\leqslant 3/2,\ \forall\, j\in[\![1,l]\!]\Big\}$$
$$\leqslant \exp\Big(\kappa n\Big\{-\frac{\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}}{8}\mathfrak{b}^2_k(f) + [\tfrac{630}{16}*12\kappa + \frac{\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}}{8}]\mathcal{R}^\diamond_n[l,f,\Lambda]\Big\}\Big)$$

(ii) *with $\|\Pi^\perp_{\mathbb{U}_l}\check{f}^n\|^2_{L^2} = 2\sum_{j=l+1}^n|\lambda_j\widehat{[g]}^+_j|^2|\theta^\circ_j|^2$ for all $k\in]\!]l,n]\!]$ we have*
$$\widehat{w}_k\mathbb{1}\{\|\widehat{\theta^\circ}^k - \check{f}^k\|^2_{L^2} < 12\kappa\Lambda^{\widehat{\Phi}}_k/n\} \leqslant \exp\Big(\kappa n\Big\{-12\kappa\Lambda^{\widehat{\Phi}}_k/n + [\tfrac{3}{2}\|\Pi^\perp_{\mathbb{U}_l}\check{f}^n\|^2_{L^2} + \tfrac{9}{2}12\kappa\Lambda^{\widehat{\Phi}}_l/n]\Big\}\Big).$$

Proof of Lemma L.2.2.
Given $m,l\in[\![1,n]\!]$ and an event Ω_{ml} (to be specified below) it clearly follows

$$
\begin{aligned}
\widehat{w}_m\mathbb{1}_{\Omega_{ml}} &= \frac{\exp(-\kappa n\{-\|\widehat{\theta^\circ}^m\|^2_{L^2} + \frac{9}{2}12\kappa\Lambda^{\widehat{\Phi}}_m/n\})}{\sum_{s\in[\![1,n]\!]}\exp(-\kappa n\{-\|\widehat{\theta^\circ}^s\|^2_{L^2} + \frac{9}{2}12\kappa\Lambda^{\widehat{\Phi}}_s/n\})}\mathbb{1}_{\Omega_{ml}} \\
&\leqslant \exp\Big(\kappa n\Big\{\|\widehat{\theta^\circ}^m\|^2_{L^2} - \|\widehat{\theta^\circ}^l\|^2_{L^2} + \tfrac{9}{2}(12\kappa\Lambda^{\widehat{\Phi}}_l/n - 12\kappa\Lambda^{\widehat{\Phi}}_m/n)\Big\}\Big)\mathbb{1}_{\Omega_{ml}} \quad \text{(K.14)}
\end{aligned}
$$

We distinguish the two cases $m<l$ and $m>l$. Consider first that $m<l$. From *(i)* in corollary G.1.1 (with $\check{f}^\bullet = \widehat{\theta^\circ}^n$ and $f = \check{f}^n = \sum_{j\in[\![-n,n]\!]}\widehat{[g]}^+_j\phi_j e_j$) follows that

$$
\begin{aligned}
\widehat{w}_m\mathbb{1}_{\Omega_{ml}} &\leqslant \exp\Big(\kappa n\Big\{\|\widehat{\theta^\circ}^m\|^2_{L^2} - \|\widehat{\theta^\circ}^l\|^2_{L^2} + \tfrac{9}{2}(12\kappa\Lambda^{\widehat{\Phi}}_l/n - 12\kappa\Lambda^{\widehat{\Phi}}_m/n)\Big\}\Big)\mathbb{1}_{\Omega_{ml}} \\
&\leqslant \exp\Big(\kappa n\Big\{\tfrac{11}{2}\|\widehat{\theta^\circ}^l - \check{f}^l\|^2_{L^2} - \tfrac{1}{2}\|\Pi_{\mathbb{U}_{kl}}\check{f}^n\|^2_{L^2} + \tfrac{9}{2}(12\kappa\Lambda^{\widehat{\Phi}}_l/n - 12\kappa\Lambda^{\widehat{\Phi}}_k/n)\Big\}\Big)\mathbb{1}_{\Omega_{kl}} \quad \text{(K.15)}
\end{aligned}
$$

Note that on the event $\mho_l := \Big\{1/2\leqslant|\lambda_j\widehat{[g]}^+_j|\leqslant 3/2,\ \forall\, j\in[\![1,l]\!]\Big\}$ we have

$$\|\Pi_{\mathbb{U}_{kl}}\check{f}^n\|^2_{L^2}\mathbb{1}_{\mho_l} \geqslant \tfrac{1}{4}\|\Pi_{\mathbb{U}_{kl}} f\|^2_{L^2} = \tfrac{1}{4}\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}(\mathfrak{b}^2_k(f) - \mathfrak{b}^2_l(f))$$
$$\widehat{\Phi}_{(l)}\mathbb{1}_{\mho_l} = \max\Big\{\widehat{\Phi}_j = (\widehat{[g]}^+_j)^2, j\in[\![1,l]\!]\Big\}\mathbb{1}_{\mho_l} \leqslant \tfrac{9}{4}\max\{\Lambda_j = \lambda_j^{-2}, j\in[\![1,l]\!]\} = \tfrac{9}{4}\Lambda_{(l)}$$
$$\widehat{\Phi}_{(l)}\mathbb{1}_{\mho_l} \geqslant \tfrac{1}{4}\Lambda_{(l)}$$

Thus on $\mho_l$ holds $\frac{1}{4}l\Lambda_{(l)}\vee(l+2) \leqslant l\widehat{\Phi}_{(l)}\vee(l+2) \leqslant \frac{9}{4}l\Lambda_{(l)}\vee(l+2)$. Since $\psi^\Lambda_l = \frac{\log(l\Lambda_{(l)}\vee(l+2))}{\log(l+2)} \geqslant$

1 for all $l\in\mathbb{N}$ hold $\frac{\log(\frac{1}{4}l\Lambda_{(l)}\vee(l+2))}{\log(l+2)}\geqslant\psi_l^\Lambda\frac{\log(3/4)}{\log 3}\geqslant\frac{3}{10}\psi_l^\Lambda$ and $\frac{\log(\frac{9}{4}l\Lambda_{(l)}\vee(l+2))}{\log(l+2)}\leqslant\psi_l^\Lambda\frac{\log(27/4)}{\log 3}\leqslant\frac{7}{4}\psi_l^\Lambda$ which together with $\Delta_l^\Lambda=l\psi_l^\Lambda\Lambda_{(l)}$ imply

$$\frac{3}{10}\psi_l^\Lambda\leqslant\delta_l^{\widehat{\Phi}}\mathbb{1}_{\mho_l}\leqslant\frac{7}{4}\psi_l^\Lambda$$
$$\frac{3}{40}\Delta_l^\Lambda=l\,\frac{3}{10}\psi_l^\Lambda\,\frac{1}{4}\Lambda_{(l)}\leqslant l\,\delta_l^{\widehat{\Phi}}\,\widehat{\Phi}_{(l)}\mathbb{1}_{\mho_l}=\Lambda_l^{\widehat{\Phi}}\mathbb{1}_{\mho_l}\leqslant l\,\frac{7}{4}\psi_l^\Lambda\,\frac{9}{4}\Lambda_{(l)}=\frac{63}{16}\Delta_l^\Lambda\tag{K.16}$$

If we define $\Omega_{kl}:=\{\|\widehat{\theta^\circ}^l-\check{f}^l\|_{L^2}^2<12\kappa\Lambda_l^{\widehat{\Phi}}/n\}\cap\mho_l$ then the last bounds imply

$$\begin{aligned}&\widehat{w}_k\mathbb{1}_{\{\|\widehat{\theta^\circ}^l-\check{f}^l\|_{L^2}^2<12\kappa\Lambda_l^{\widehat{\Phi}}/n\}}\mathbb{1}_{\mho_l}\\&\leqslant\exp\big(\kappa n\{\tfrac{11}{2}12\kappa\Lambda_l^{\widehat{\Phi}}/n-\tfrac{1}{8}\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2(\mathfrak{b}_k^2(f)-\mathfrak{b}_l^2(f))+\tfrac{9}{2}(12\kappa\Lambda_l^{\widehat{\Phi}}/n-12\kappa\Lambda_k^{\widehat{\Phi}}/n)\}\big)\\&=\exp\big(\kappa n\{10*12\kappa\Lambda_l^{\widehat{\Phi}}/n-\tfrac{1}{8}\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2(\mathfrak{b}_k^2(f)-\mathfrak{b}_l^2(f))-\tfrac{9}{2}*12\kappa\Lambda_k^{\widehat{\Phi}}/n\}\big)\\&\leqslant\exp\big(\kappa n\{\tfrac{630}{16}*12\kappa\Delta_l^\Lambda/n-\tfrac{1}{8}\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2(\mathfrak{b}_k^2(f)-\mathfrak{b}_l^2(f))\}\big)\end{aligned}$$

and hence, by exploiting that $12\kappa\Lambda_k^{\widehat{\Phi}}/n\geqslant 0$ and $\mathcal{R}_n^\diamond[l,f,\Lambda]=[\mathfrak{b}_l^2(f)\vee\Delta_l^\Lambda n^{-1}]$ follows the assertion ***(i)***, that is

$$\widehat{w}_k\mathbb{1}_{\{\|\widehat{\theta^\circ}^l-\check{f}^l\|_{L^2}^2<12\kappa\Lambda_l^{\widehat{\Phi}}/n\}}\mathbb{1}_{\mho_l}\leqslant\exp\big(\kappa n\{-\frac{\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2}{8}\mathfrak{b}_k^2(f)+[\tfrac{630}{16}*12\kappa+\frac{\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2}{8}]\mathcal{R}_n^\diamond[l,f,\Lambda])\}\big).$$

Consider secondly that $k>l$. From ***(ii)*** in corollary G.1.1 (with $\check{f}^\bullet=\widehat{\theta^\circ}^n$ and $f=\check{f}^n=\sum_{j\in[\![-n,n]\!]}\widehat{[g]}_j^+\phi_je_j$) and (L.7) follows

$$\begin{aligned}&\widehat{w}_k\mathbb{1}_{\Omega_{lk}}\leqslant\exp\big(\kappa n\{\|\widehat{\theta^\circ}^m\|_{L^2}^2-\|\widehat{\theta^\circ}^l\|_{L^2}^2+\tfrac{9}{2}(12\kappa\Lambda_l^{\widehat{\Phi}}/n-12\kappa\Lambda_m^{\widehat{\Phi}}/n)\}\big)\mathbb{1}_{\Omega_{ml}}\\&\quad\leqslant\exp\big(\kappa n\{\tfrac{7}{2}\|\widehat{\theta^\circ}^k-\check{f}^k\|_{L^2}^2+\tfrac{3}{2}\|\Pi_{\mathbb{U}_{lk}}\check{f}^n\|_{L^2}^2+\tfrac{9}{2}(12\kappa\Lambda_l^{\widehat{\Phi}}/n-12\kappa\Lambda_k^{\widehat{\Phi}}/n)\}\big)\mathbb{1}_{\Omega_{lk}}\end{aligned}\tag{K.17}$$

Keep in mind that $\|\Pi_{\mathbb{U}_{lk}}\check{f}^n\|_{L^2}^2\mathbb{1}_{\mho_l}=2\sum_{j=l+1}^k(\lambda_j\widehat{[g]}_j^+)^2|\theta_j^\circ|^2\leqslant2\sum_{j=l+1}^n(\lambda_j\widehat{[g]}_j^+)^2|\theta_j^\circ|^2=\|\Pi_{\mathbb{U}_l}^\perp\check{f}^n\|_{L^2}^2$. If we set $\Omega_{lk}:=\{\|\widehat{\theta^\circ}^k-\check{f}^k\|_{L^2}^2<12\kappa\Lambda_k^{\widehat{\Phi}}/n\}$ then we clearly have ***(ii)***, that is

$$\begin{aligned}\widehat{w}_k\mathbb{1}_{\{\|\widehat{\theta^\circ}^k-\check{f}^k\|_{L^2}^2<12\kappa\Lambda_k^{\widehat{\Phi}}/n\}}&\leqslant\exp\big(\kappa n\{\tfrac{7}{2}*12\kappa\Lambda_k^{\widehat{\Phi}}/n+\tfrac{3}{2}\|\Pi_{\mathbb{U}_l}^\perp\check{f}^n\|_{L^2}^2+\tfrac{9}{2}(12\kappa\Lambda_l^{\widehat{\Phi}}/n-12\kappa\Lambda_k^{\widehat{\Phi}}/n)\}\big)\\&=\exp\big(\kappa n\{-12\kappa\Lambda_k^{\widehat{\Phi}}/n+[\tfrac{3}{2}\|\Pi_{\mathbb{U}_l}^\perp\check{f}^n\|_{L^2}^2+\tfrac{9}{2}12\kappa\Lambda_l^{\widehat{\Phi}}/n]\}\big)\end{aligned}$$

which completes the proof.

Lemma K.2.5. *Given $m_+^\diamond,m_-^\diamond\in[\![1,n]\!]$ let m_+ and m_- as in (??). Let $\Lambda_m^{\widehat{\Phi}}=\delta_m^{\widehat{\Phi}}m\widehat{\Phi}_{(m)}$, $\sqrt{\delta_m^{\widehat{\Phi}}}=\frac{\log(m\widehat{\Phi}_{(m)}\vee(m+2))}{\log(m+2)}\geqslant1$, $\kappa\geqslant1$ and $\kappa\geqslant1$. If $\mathcal{R}_n^\diamond[m,f,\Lambda]=[\mathfrak{b}_m^2(f)\vee\Lambda_m^{|\Phi}/n]$ for any $m\in[\![1,n]\!]$ and $\mho_{m_-^\diamond}:=\{1/2\leqslant|\lambda_j\widehat{[g]}_j^+|\leqslant3/2,\ \forall\,j\in[\![1,m_-^\diamond]\!]\}$, then*

$$\begin{aligned}\textbf{(i)}\ \ \mathbb{P}_{\widehat{w}}([\![1,m_-[\![)\leqslant[m_--1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond[m_-^\diamond,f,\Lambda]-\frac{\kappa\|\Pi_{\mathbb{U}_0^\perp}\theta\|_{L^2}^2}{16}n\mathfrak{b}_{[m_--1]}^2(f)\big)\\+\mathbb{1}_{\{\|\widehat{\theta^\circ}^{m_-^\diamond}-\check{f}^{m_-^\diamond}\|_{L^2}^2\geqslant12\kappa\Lambda_{m_-^\diamond}^{\widehat{\Phi}}/n\}\cap\mho_{m_-^\diamond}}+\mathbb{1}_{\mho_{m_-^\diamond}^c};\end{aligned}$$

(ii) $\sum_{m=m_++1}^{n} \Lambda_m^{\widehat{\Phi}} \widehat{w}_m \mathbb{1}_{\{\|\widehat{\theta^\circ}^k - \check{f}^k\|_{L^2}^2 < 12\kappa\Lambda_m^{\widehat{\Phi}}/n\}} \leqslant \frac{1}{36\kappa^2\kappa^2} + \frac{1}{3\kappa\kappa}$.

PROOF OF LEMMA L.2.3.

Consider *(i)*. From Lemma G.1.6 *(i)* with $l = m_-^\diamond$ follows for all $m < m_- \leqslant m_-^\diamond$, and hence $\mathfrak{b}_m \geqslant \mathfrak{b}_{m_--1}$ that

$$\begin{aligned}
&\widehat{w}_k \mathbb{1}_{\left\{\|\widehat{\theta^\circ}^l - \check{f}^l\|_{L^2}^2 < 12\kappa\Lambda_l^{\widehat{\Phi}}/n\right\} \cap \mho_l} \\
&\qquad\leqslant \exp\left(\kappa n\left\{ - \frac{\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{8}\mathfrak{b}_k^2(f) + [\tfrac{630}{16} * 12\kappa + \frac{\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{8}]\mathcal{R}_n^\diamond[l, f, \Lambda]\right\}\right) \\
&\qquad= \exp\Big(\kappa n\Big\{ \underbrace{-\frac{\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{16}\mathfrak{b}_m^2(f) + [\tfrac{630*3+4}{4}\kappa + \frac{2\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{16}]\mathcal{R}_n^\diamond[m_-^\diamond, f, \Lambda]}_{\leqslant 0}\Big\}\Big) \\
&\qquad\qquad \times \exp\left(\kappa n\left\{ - \kappa\mathcal{R}_n^\diamond[m_-^\diamond, f, \Lambda] - \frac{\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{16}\mathfrak{b}_m^2(f)\right\}\right) \\
&\qquad\leqslant \exp\left(- \kappa n\kappa\mathcal{R}_n^\diamond[m_-^\diamond, f, \Lambda] - \kappa n\frac{\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{16}\mathfrak{b}_{m_-^\diamond-1}^2(f)\right)
\end{aligned}$$

which in turn implies *(i)*, that is,

$$\begin{aligned}
\mathbb{P}_{\widehat{w}}(\llbracket 1, m_- \llbracket) &\leqslant \mathbb{P}_{\widehat{w}}(\llbracket 1, m_- \llbracket)\mathbb{1}_{\left\{\|\widehat{\theta^\circ}^{m_-^\diamond} - \check{f}^{m_-^\diamond}\|_{L^2}^2 < 12\kappa\Lambda_{m_-^\diamond}^{\widehat{\Phi}}/n\right\} \cap \mho_{m_-^\diamond}} + \mathbb{1}_{\{\|\widehat{\theta^\circ}^{m_-^\diamond} - \check{\theta}_{m_-^\diamond}\|_{\mathbb{H}}^2 < 12\kappa\Lambda_{m_-^\diamond}^{\widehat{\Phi}}/n\}^c \cup \mho_{m_-^\diamond}^c} \\
&\leqslant [m_- - 1]\exp\left(- \kappa\kappa n\mathcal{R}_n^\diamond[m_-^\diamond, f, \Lambda] - \kappa n\frac{\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{16}\mathfrak{b}_{[m_--1]}^2(f)\right) \\
&\qquad + \mathbb{1}_{\{\|\widehat{\theta^\circ}^{m_-^\diamond} - \check{f}^{m_-^\diamond}\|_{L^2}^2 \geqslant 12\kappa\Lambda_{m_-^\diamond}^{\widehat{\Phi}}/n\} \cap \mho_{m_-^\diamond}} + \mathbb{1}_{\mho_{m_-^\diamond}^c}
\end{aligned}$$

Consider *(ii)*. From Lemma L.2.2 *(ii)* with $l = m_+^\diamond$ follows for all $m > m_+ \geqslant m_+^\diamond$

$$\begin{aligned}
\widehat{w}_m \mathbb{1}_{\{\|\widehat{\theta^\circ}^k - \check{f}^k\|_{L^2}^2 < 12\kappa\Lambda_m^{\widehat{\Phi}}/n\}} &\leqslant \exp\left(\kappa n\left\{ - 12\kappa\Lambda_m^{\widehat{\Phi}}/n + [\tfrac{3}{2}\|\Pi_{\mathbb{U}_l}^\perp \check{f}^n\|_{L^2}^2 + \tfrac{9}{2} * 12\kappa\Lambda_l^{\widehat{\Phi}}/n]\right\}\right) \\
&= \exp\Big(\kappa n\Big\{ - \tfrac{1}{2}12\kappa\Lambda_m^{\widehat{\Phi}}/n \underbrace{- \tfrac{1}{2} * 12\kappa\Lambda_m^{\widehat{\Phi}}/n + [\tfrac{3}{2}\|\Pi_{\mathbb{U}_l}^\perp \check{f}^n\|_{L^2}^2 + \tfrac{9}{2} * 12\kappa\Lambda_l^{\widehat{\Phi}}/n]}_{\leqslant 0}\Big\}\Big) \\
&\leqslant \exp\left(- \kappa 6\kappa\Lambda_m^{\widehat{\Phi}}\right).
\end{aligned}$$

Note that $|\widehat{[g]}_j|^2 \leqslant 1$ for all $j \in \mathbb{Z}$, hence if $|\widehat{[g]}_j|^2 \geqslant 1/q$ then $\widehat{\Phi}_j = |\widehat{[g]}_j^+|^2 \geqslant 1$. Thereby, $\widehat{\Phi}_j = |\widehat{[g]}_j^+|^2 < 1$ implies $|\widehat{[g]}_j|^2 < 1/q$ and hence $\widehat{\Phi}_j = |\widehat{[g]}_j^+|^2 = 0$. Thereby $1 > \widehat{\Phi}_{(m)} = \max\{|\widehat{[g]}_j^+|^2, j \in \llbracket 1, m \rrbracket\}$ implies $\widehat{\Phi}_{(m)} = 0$, that is,

$$\{\widehat{\Phi}_{(m)} < 1\} = \{\widehat{\Phi}_{(m)} = 0\}. \tag{K.18}$$

Consequently, with $\Lambda_m^{\widehat{\Phi}} = \delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)}$ follows

$$\begin{aligned}\sum_{m=m_++1}^{n} \Lambda_m^{\widehat{\Phi}} \widehat{w}_m \mathbb{1}_{\{\|\widehat{\theta^\circ}^k - \check{f}^k\|_{L^2}^2 \leqslant 12\kappa\Lambda_m^{\widehat{\Phi}}/n\}} &\leqslant \sum_{k=m_++1}^{n} \Lambda_m^{\widehat{\Phi}} \exp\big(-\kappa 6\kappa \Lambda_m^{\widehat{\Phi}}\big) \\ &= \sum_{k=m_++1}^{n} \delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)} \exp\big(-6\kappa\kappa\delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)}\big)\mathbb{1}_{\{\widehat{\Phi}_{(m)}\geqslant 1\}} \\ &\quad + \sum_{k=m_++1}^{n} \delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)} \exp\big(-6\kappa\kappa\delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)}\big)\mathbb{1}_{\{\widehat{\Phi}_{(m)}< 1\}} \\ &= \sum_{k=m_++1}^{n} \delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)} \exp\big(-6\kappa\kappa\delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)}\big)\mathbb{1}_{\{\widehat{\Phi}_{(m)}\geqslant 1\}}. \quad \text{(K.19)}\end{aligned}$$

Exploiting that $\sqrt{\delta_m^{\widehat{\Phi}}} = \frac{\log(m\widehat{\Phi}_{(m)}\vee(m+2))}{\log(m+2)} \geqslant 1$ and $6\kappa \geqslant 2\log(3e) \approx 4.2$ let $\kappa \geqslant 1$, then for all $k \in \mathbb{N}$ we have $6\kappa\kappa k - \log(k+2) \geqslant 1$ and thus exploiting $a\exp(-ab) \leqslant \exp(-b)$ for all $a, b \geqslant 1$ we have

$$\begin{aligned}\delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)} \exp\big(-6\kappa\kappa\delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)}\big)\mathbb{1}_{\{\widehat{\Phi}_{(m)}\geqslant 1\}} &\leqslant \delta_m^{\widehat{\Phi}} \exp\big(-6\kappa\kappa\delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)} + \delta_m^{\widehat{\Phi}}\log(m+2)\big)\mathbb{1}_{\{\widehat{\Phi}_{(m)}\geqslant 1\}} \\ &\leqslant \delta_m^{\widehat{\Phi}} \exp\big(-\delta_m^{\widehat{\Phi}}(6\kappa\kappa m - \log(m+2))\big) \leqslant \exp\big(-(6\kappa\kappa m - \log(m+2))\big) \\ &= (m+2)\exp\big(-6\kappa\kappa m\big)\end{aligned}$$

Since $\sum_{m\in\mathbb{N}} \mu m \exp(-\mu m) \leqslant (\mu)^{-1}$ und $\sum_{m\in\mathbb{N}} \mu \exp(-\mu m) \leqslant 1$ follows

$$\begin{aligned}\sum_{k=m_++1}^{n} \delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)} \exp\big(-6\kappa\kappa\delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)}\big)\mathbb{1}_{\{\widehat{\Phi}_{(m)}\geqslant 1\}} &\leqslant \sum_{k=m_++1}^{\infty} (m+2)\exp\big(-6\kappa\kappa m\big) \\ &\leqslant \tfrac{1}{36\kappa^2\kappa^2} + \tfrac{1}{3\kappa\kappa}. \quad \text{(K.20)}\end{aligned}$$

Thereby, combining (L.12) and (L.13) we obtain ***(ii)***, that is

$$\sum_{m=m_++1}^{n} \Lambda_m^{\widehat{\Phi}} \widehat{w}_m \mathbb{1}_{\{\|\widehat{\theta^\circ}^k - \check{f}^k\|_{L^2}^2 \leqslant 12\kappa\Lambda_m^{\widehat{\Phi}}/n\}} \leqslant \tfrac{1}{36\kappa^2\kappa^2} + \tfrac{1}{3\kappa\kappa},$$

which completes the proof.

Lemma K.2.6. *Consider $\widehat{\theta^\circ}^m - \check{f}^m = \sum_{j\in[\![-m,m]\!]} \widehat{[g]}_j^+ (\widehat{[h]}_j - \phi_j)e_j$. Conditionally on $\varepsilon_1, \dots, \varepsilon_q$ the r.v.'s $Y_1, \dots, Y_n$ are iid. and we denote by $\mathbb{P}^n_{Y|\varepsilon}$ and $\mathbb{E}^n_{Y|\varepsilon}$ their conditional distribution and expectation, respectively. Let $\widehat{\Phi}_j = |\widehat{[g]}_j^+|^2$, $\overline{\widehat{\Phi}}_m = \frac{1}{m}\sum_{j\in[\![1,m]\!]}\widehat{\Phi}_j$, $\widehat{\Phi}_{(m)} = \max\{\widehat{\Phi}_j, j \in [\![1,m]\!]\}$, $\kappa \geqslant 1$, $\Lambda_m^{\widehat{\Phi}} = \delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)}$ and $\sqrt{\delta_m^{\widehat{\Phi}}} = \frac{\log(m\widehat{\Phi}_{(m)}\vee(m+2))}{\log(m+2)} \geqslant 1$. Then there is a numerical constant $\mathcal{C}$ such that for all $n \in \mathbb{N}$ and $m \in [\![1, n[\![$ holds*

(i) *if $m_+ \geqslant 3(\frac{12\|\phi\|_{\ell^1}}{\kappa})^2$ and $n \geqslant 15(\frac{200}{\sqrt{\kappa}})^4$ (alternatively $\sqrt{\kappa}/200 \geqslant \sqrt{3}$) then*
$\sum_{m=1+m_+}^{n} \mathbb{E}^n_{Y|\varepsilon}\big(\|\widehat{\theta^\circ}^m - \check{f}^m\|_{L^2}^2 - 12\kappa\Lambda_m^{\widehat{\Phi}}/n\big)_+ \leqslant \mathcal{C}\big[\frac{12\|\phi\|_{\ell^1}^2}{\kappa} + 4\big]n^{-1}$

(ii) *if* $m_+ \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ *and* $n \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ *(alternatively* $\sqrt{\kappa}/300 \geqslant \sqrt{3}$*) then* $\sum_{m=1+m_+}^{n} \Delta_m^{\Lambda} \mathbb{P}_{Y|\varepsilon}^{n}\big(\|\widehat{\theta}^{\circ}{}^{m} - \check{f}^{m}\|_{L^2}^2 \geqslant 12\kappa\Lambda_m^{\widehat{\Phi}}/n\big) \leqslant 3\big[(\frac{400\|\phi\|_{\ell^1}}{\kappa})^2 + \frac{800\|\phi\|_{\ell^1}}{\kappa} + 1\big]$.

Proof of Lemma L.2.4.

Since $\delta_m^{\widehat{\Phi}} \geqslant 1$ for $m \geqslant 3(\frac{12\|\phi\|_{\ell^1}}{\kappa})^2$ holds $\frac{\kappa\sqrt{\psi_m^{\Lambda}}m}{12\|\phi\|_{\ell^1}} - \log(m+2) \geqslant 0$ and hence

$$\widehat{\Phi}_{(m)} \exp\big(\tfrac{-\kappa\delta_m^{\widehat{\Phi}} m}{6\|\phi\|_{\ell^1}}\big) \leqslant \exp\big(\tfrac{-\kappa\delta_m^{\widehat{\Phi}} m}{12\|\phi\|_{\ell^1}}\big) \exp\Big(-\sqrt{\delta_m^{\widehat{\Phi}}}\big[\tfrac{\kappa\sqrt{\delta_m^{\widehat{\Phi}}}m}{12\|\phi\|_{\ell^1}} - \log(m+2)\big]\Big)$$
$$\leqslant \exp\big(\tfrac{-\kappa\delta_m^{\widehat{\Phi}} m}{12\|\phi\|_{\ell^1}}\big) \leqslant \exp\big(-\tfrac{\kappa}{12\|\phi\|_{\ell^1}} m\big)$$

consequently, if $m_+ \geqslant 3(\frac{12\|\phi\|_{\ell^1}}{\kappa})^2$ then

$$\sum_{m=1+m_+}^{n} \widehat{\Phi}_{(m)} \exp\big(\tfrac{-\kappa\delta_m^{\widehat{\Phi}} m}{6\|\phi\|_{\ell^1}}\big) \leqslant \sum_{m=1+m_+}^{n} \exp\big(-\tfrac{\kappa}{12\|\phi\|_{\ell^1}} m\big) \leqslant \frac{12\|\phi\|_{\ell^1}}{\kappa}$$

For all $n \geqslant 15(\frac{200}{\sqrt{\kappa}})^4$ holds $\sqrt{n} \geqslant \frac{200}{\sqrt{\kappa}} \log(n+2)$ (or using $\sqrt{3x} \geqslant \log(x+2)$ for all $x \geqslant 1$ and $\sqrt{\kappa}/200 \geqslant \sqrt{3}$) thereby

$$\frac{m^2\widehat{\Phi}_{(m)}^2}{n} \exp\big(\tfrac{-\sqrt{n\kappa\delta_m^{\widehat{\Phi}}}}{100}\big) \leqslant \frac{1}{n}\exp\Big(-2\sqrt{\delta_m^{\widehat{\Phi}}}\big[\tfrac{\sqrt{n\kappa}}{200} - \log(m+2)\big]\Big) \leqslant \frac{1}{n}$$

consequently,

$$\sum_{m=1+m_+}^{n} \frac{m^2\widehat{\Phi}_{(m)}^2}{n} \exp\big(\tfrac{-\sqrt{n\kappa\delta_m^{\widehat{\Phi}}}}{100}\big) \leqslant \sum_{m=1+m_+}^{n} \frac{1}{n} \leqslant 1$$

Combining the last two bounds and **??** we obtain *(i)*, that is

$$\sum_{m=1+m_+}^{n} \mathbb{E}_{Y|\varepsilon}^{n}\big(\|\widehat{\theta}^{\circ}{}^{m} - \check{f}^{m}\|_{L^2}^2 - 12\kappa\Lambda_m^{\widehat{\Phi}}/n\big)_+$$
$$\leqslant \mathcal{C}\Big[\frac{\|\phi\|_{\ell^1}}{n}\sum_{m=1+m_+}^{n} \widehat{\Phi}_{(m)} \exp\big(\tfrac{-\kappa\delta_m^{\widehat{\Phi}} m}{6\|\phi\|_{\ell^1}}\big) + \frac{4}{n}\sum_{m=1+m_+}^{n} \frac{m^2\widehat{\Phi}_{(m)}^2}{n}\exp\big(\tfrac{-\sqrt{n\kappa\delta_m^{\widehat{\Phi}}}}{100}\big)\Big]$$
$$\leqslant \mathcal{C}n^{-1}\big[\tfrac{12\|\phi\|_{\ell^1}^2}{\kappa} + 4\big]$$

If $m \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ then $m \geqslant (\frac{800\|\phi\|_{\ell^1}}{\kappa})\log(m+2)$ and hence $m - \frac{400\|\phi\|_{\ell^1}}{\kappa}\log(m+2) \geqslant \frac{400\|\phi\|_{\ell^1}}{\kappa}\log(m+2)$ or equivalently, $\frac{\kappa}{400\|\phi\|_{\ell^1}}m - \log(m+2) \geqslant \log(m+2) \geqslant 1$ and thus

$$m\delta_m^{\widehat{\Phi}}\widehat{\Phi}_{(m)} \exp\big(\tfrac{-\kappa\delta_m^{\widehat{\Phi}} m}{400\|\phi\|_{\ell^1}}\big) \leqslant \delta_m^{\widehat{\Phi}} \exp\Big(-\delta_m^{\widehat{\Phi}}\big[\tfrac{\kappa}{400\|\phi\|_{\ell^1}}m - \log(m+2)\big]\Big)$$
$$\leqslant (m+2)\exp\big(-\tfrac{\kappa}{400\|\phi\|_{\ell^1}}m\big)$$

consequently, if $m_+ \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ exploiting $\sum_{m\in\mathbb{N}}(m+2)\exp(-\mu m) \leqslant \mu^{-2} + 2\mu^{-1}$ follows

$$\sum_{m=1+m_+}^{n} m\delta_m^{\widehat{\Phi}}\widehat{\Phi}_{(m)}\exp\big(\tfrac{-\kappa\delta_m^{\widehat{\Phi}}m}{400\|\phi\|_{\ell^1}}\big) \leqslant \sum_{m=1+m_+}^{n}(k+2)\exp\big(-\tfrac{\kappa}{400\|\phi\|_{\ell^1}}m\big)$$
$$\leqslant (\frac{400\|\phi\|_{\ell^1}}{\kappa})^2 + \frac{800\|\phi\|_{\ell^1}}{\kappa}$$

Keep in mind that $\{\widehat{\Phi}_{(m)} < 1\} = \{\widehat{\Phi}_{(m)} = 0\}$ (cf. (L.11) in the proof of Lemma L.2.3). Since $\log(m\widehat{\Phi}_{(m)})\mathbb{1}\{\widehat{\Phi}_{(m)}\geqslant 1\} \leqslant (1/e)m\widehat{\Phi}_{(m)}\mathbb{1}\{\widehat{\Phi}_{(m)}\geqslant 1\}$ follows $\delta_m^{\widehat{\Phi}}\mathbb{1}\{\widehat{\Phi}_{(m)}\geqslant 1\} \leqslant m\widehat{\Phi}_{(m)}\mathbb{1}\{\widehat{\Phi}_{(m)}\geqslant 1\}$, and using for all $n \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ holds $\sqrt{n} \geqslant \frac{300}{\sqrt{\kappa}}\log(n+2)$ (alternatively: $\sqrt{3x} \geqslant \log(x+2)$ for all $x \geqslant 1$ and $\sqrt{\kappa}/300 \geqslant \sqrt{3}$)

$$m\delta_m^{\widehat{\Phi}}\widehat{\Phi}_{(m)}\exp\big(\tfrac{-\sqrt{n\kappa\delta_m^{\widehat{\Phi}}}}{100}\big)\mathbb{1}\{\widehat{\Phi}_{(m)}\geqslant 1\} \leqslant m^2\widehat{\Phi}_{(m)}^2\exp\big(\tfrac{-\sqrt{n\kappa\delta_m^{\widehat{\Phi}}}}{100}\big)\mathbb{1}\{\widehat{\Phi}_{(m)}\geqslant 1\}$$
$$\leqslant \tfrac{1}{n}\exp\big(-\sqrt{\psi_m^{\Lambda}}[\tfrac{\sqrt{n\kappa}}{100} - 2\log(m+2)] + \log(n+2)\big)$$
$$\leqslant \tfrac{1}{n}\exp\big(-3\sqrt{\delta_m^{\widehat{\Phi}}}[\tfrac{\sqrt{n\kappa}}{300} - \log(n+2)]\big) \leqslant \tfrac{1}{n}$$

consequently,

$$\sum_{m=1+m_+}^{n} m\delta_m^{\widehat{\Phi}}\widehat{\Phi}_{(m)}\exp\big(\tfrac{-\sqrt{n\kappa\delta_m^{\widehat{\Phi}}}}{100}\big) = \sum_{m=1+m_+}^{n} m\delta_m^{\widehat{\Phi}}\widehat{\Phi}_{(m)}\exp\big(\tfrac{-\sqrt{n\kappa\delta_m^{\widehat{\Phi}}}}{100}\big)\mathbb{1}\{\widehat{\Phi}_{(m)}\geqslant 1\}$$
$$\leqslant \sum_{m=1+m_+}^{n}\tfrac{1}{n} \leqslant 1$$

Combining the last two bounds and **??** **??** we obtain ***(ii)***, that is

$$\sum_{m=1+m_+}^{n}\delta_m^{\widehat{\Phi}}m\widehat{\Phi}_{(m)}\mathbb{P}^n_{Y|\varepsilon}\big(\|\widehat{\theta}^{\circ m} - \breve{f}^m\|^2_{L^2} \geqslant 12\kappa\Lambda_m^{\widehat{\Phi}}/n\big)$$
$$\leqslant 3\sum_{m=1+m_+}^{n}\delta_m^{\widehat{\Phi}}m\widehat{\Phi}_{(m)}\Big[\exp\big(\tfrac{-\kappa\delta_m^{\widehat{\Phi}}m}{400\|\phi\|_{\ell^1}}\big) + \exp\big(\tfrac{-\sqrt{n\kappa\delta_m^{\widehat{\Phi}}}}{100}\big)\Big]$$
$$\leqslant 3\Big[(\frac{400\|\phi\|_{\ell^1}}{\kappa})^2 + \frac{800\|\phi\|_{\ell^1}}{\kappa} + 1\Big]$$

which implies the result and completes the proof.

Proposition K.2.1. *Let $\kappa \geqslant 1$, $\kappa \geqslant 1$, $\Delta_m^{\Lambda} = \psi_m^{\Lambda}m\Lambda_{(m)}$ with $\sqrt{\psi_m^{\Lambda}} = \frac{\log(m\Lambda_{(m)}\vee(m+2))}{\log(m+2)} \geqslant 1$, $\mathcal{R}_n^{\diamond}[l,f,\Lambda] = [\mathfrak{b}_l^2(f)\vee\Delta_l^{\Lambda}n^{-1}]$ and $\mathcal{R}_q^{\star}(f,\Lambda) := \sum_{j\in\mathbb{N}}|\theta_j^{\circ}|^2[1\wedge q^{-1}\Lambda_j]$. Given $m_+^{\diamond}, m_-^{\diamond} \in [\![1,n]\!]$ let m_- as in (**??**). If $m_+^{\diamond} \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ and $n \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ then there is an universal* *numerical constant $\mathcal{C}$ such that for all $q \in \mathbb{N}$*

$$\begin{aligned}\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant \mathcal{C}\big\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}]\mathcal{R}^{\diamond}_n[m^{\diamond}_+, f, \Lambda] + \mathcal{R}^{\star}_q(f,\Lambda) + \|\phi\|^2_{\ell^1} n^{-1}\big\}\\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\exp\big(\frac{-3\kappa\psi^{\Lambda}_{m^{\diamond}_-} m^{\diamond}_-}{4000\|\phi\|_{\ell^1}}\big)\\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^{\diamond}_n[m^{\diamond}_-, f, \Lambda] - \frac{\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}}{16} n\mathfrak{b}^2_{[m_- - 1]}(f)\big)\\ + 3q\mathbb{P}^q_{\varepsilon}(\mho^c_{m^{\diamond}_+}) + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathbb{P}^q_{\varepsilon}(\mho^c_{m^{\diamond}_-})\end{aligned}$$

Proof of Proposition L.2.1.

Given $m^{\diamond}_+, m^{\diamond}_- \in [\![1,n]\!]$ let m_+ and m_- as in (**??**). Keep in mind that m_- is not random and m_+ does depend on the error sample only. From Lemma L.2.1 together with Lemma L.2.3 follows

$$\begin{aligned}\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant 3\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta^{\circ}}^{m_+} - \breve{f}^{m_+}\|^2_{L^2} + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + [\tfrac{1}{12\kappa\kappa^2} + \tfrac{1}{\kappa}]n^{-1}\\ + 3\sum_{l \in]\!]m_+,n]\!]}\mathbb{E}^n_{Y|\varepsilon}\big(\|\widehat{\theta^{\circ}}^{l} - \breve{f}^{l}\|^2_{L^2} - \mathcal{C}_1 12\kappa\Lambda^{\widehat{\Phi}}_l/n\big)_+\\ + 3(\mathcal{C}_1\kappa/n)\sum_{l \in]\!]m_+,n]\!]}\Lambda^{\widehat{\Phi}}_l\mathbb{P}^n_{Y|\varepsilon}\Big(\|\widehat{\theta^{\circ}}^{l} - \breve{f}^{l}\|^2_{L^2} \geqslant 12\kappa\Lambda^{\widehat{\Phi}}_l/n\Big)\\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\{[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^{\diamond}_n[m^{\diamond}_-, f, \Lambda] - \frac{\kappa\|\Pi_{\mathbb{U}_0^\perp}\theta\|^2_{L^2}}{16} n\mathfrak{b}^2_{[m_- - 1]}(f)\big)\\ + \mathbb{P}^n_{Y|\varepsilon}\Big(\|\widehat{\theta^{\circ}}^{m^{\diamond}_-} - \breve{f}^{m^{\diamond}_-}\|^2_{L^2} \geqslant 12\kappa\Lambda^{\widehat{\Phi}}_{m^{\diamond}_-}/n\Big)\mathbb{1}_{\mho_{m^{\diamond}_-}} + \mathbb{1}_{\mho^c_{m^{\diamond}_-}}\}\\ + 6\sum_{j \in [\![1,n]\!]}|\widehat{[g]}^+_j|^2|\lambda_j - \widehat{[g]}_j|^2|\theta^{\circ}_j|^2 + 2\sum_{j \in [\![1,n]\!]}\mathbb{1}_{\mathcal{X}^c_j}|\theta^{\circ}_j|^2 \qquad \text{(K.21)}\end{aligned}$$

Since $m_+ \geqslant m^{\diamond}_+ \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ and $n \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ due to Lemma L.2.4 ***(i)*** and ***(ii)*** there is a finite numerical constant $\mathcal{C}$ such that

$$\begin{aligned}\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant 3\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta^{\circ}}^{m_+} - \breve{f}^{m_+}\|^2_{L^2} + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + [\tfrac{1}{12\kappa\kappa^2} + \tfrac{1}{\kappa}]n^{-1}\\ + \mathcal{C}\big[\frac{26\|\phi\|^2_{\ell^1}}{\kappa} + 12\big]n^{-1} + \mathcal{C}_1\big[\frac{9*400^2\|\phi\|^2_{\ell^1}}{\kappa} + 9*800\|\phi\|_{\ell^1} + 9\kappa\big]n^{-1}\\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^{\diamond}_n[m^{\diamond}_-, f, \Lambda] - \frac{\kappa\|\Pi_{\mathbb{U}_0^\perp}\theta\|^2_{L^2}}{16} n\mathfrak{b}^2_{[m_- - 1]}(f)\big)\\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathbb{P}^n_{Y|\varepsilon}\Big(\|\widehat{\theta^{\circ}}^{m^{\diamond}_-} - \breve{f}^{m^{\diamond}_-}\|^2_{L^2} \geqslant 12\kappa\Lambda^{\widehat{\Phi}}_{m^{\diamond}_-}/n\Big)\mathbb{1}_{\mho_{m^{\diamond}_-}} + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathbb{1}_{\mho^c_{m^{\diamond}_-}}\\ + 6\sum_{j \in [\![1,n]\!]}|\widehat{[g]}^+_j|^2|\lambda_j - \widehat{[g]}_j|^2|\theta^{\circ}_j|^2 + 2\sum_{j \in [\![1,n]\!]}\mathbb{1}_{\mathcal{X}^c_j}|\theta^{\circ}_j|^2. \qquad \text{(K.22)}\end{aligned}$$

and together with ?? ?? we obtain

$$\begin{aligned}\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta^\circ} - f\|^2_{L^2} \leqslant 3\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta^\circ}^{m_+} - \breve{f}^{m_+}\|^2_{L^2} + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + [\tfrac{1}{12\kappa\kappa^2} + \tfrac{1}{\kappa}]n^{-1}\\ + \mathcal{C}\big[\tfrac{26\|\phi\|^2_{\ell^1}}{\kappa} + 12\big]n^{-1} + \mathcal{C}_1\big[\tfrac{9*400^2\|\phi\|^2_{\ell^1}}{\kappa} + 9*800\|\phi\|_{\ell^1} + 9\kappa\big]n^{-1}\\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^\diamond_n[m^\diamond_-, f, \Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp}\theta\|^2_{L^2}}{16}n\mathfrak{b}^2_{[m_- -1]}(f)\big)\\ + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\Big[\exp\big(\tfrac{-\kappa\delta^{\widehat{\Phi}}_{m^\diamond_-}m^\diamond_-}{400\|\phi\|_{\ell^1}}\big) + \exp\big(\tfrac{-\sqrt{n\kappa\delta^{\widehat{\Phi}}_{m^\diamond_-}}}{100}\big)\Big]\mathbb{1}\mho_{m^\diamond_-}\\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathbb{1}\mho^c_{m^\diamond_-} + 6\sum_{j\in[\![1,n]\!]}|\widehat{[g]}^+_j|^2|\lambda_j - \widehat{[g]}_j|^2|\theta^\circ_j|^2 + 2\sum_{j\in[\![1,n]\!]}\mathbb{1}\mathcal{X}^c_j|\theta^\circ_j|^2.\end{aligned}$$

Moreover, for $n > n_{f,\Lambda} \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ holds $\sqrt{n} \geqslant \frac{300}{\sqrt{\kappa}}\log(n+2) \geqslant \frac{100}{\sqrt{\kappa}}\log(n+2)$ which in turn together with $\delta^{\widehat{\Phi}}_{m_-} \geqslant 1$ implies $n\exp\big(-\sqrt{n}\frac{\sqrt{\kappa\delta^{\widehat{\Phi}}_{m^\diamond_-}}}{100}\big) \leqslant \exp\big(-\sqrt{n}\frac{\sqrt{\kappa}}{100} + \log(n+2)\big) \leqslant 1$, and thus

$$\begin{aligned}\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta^\circ} - f\|^2_{L^2} \leqslant 3\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta^\circ}^{m_+} - \breve{f}^{m_+}\|^2_{L^2} + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + [\tfrac{1}{12\kappa\kappa^2} + \tfrac{1}{\kappa}]n^{-1}\\ + \mathcal{C}\big[\tfrac{26\|\phi\|^2_{\ell^1}}{\kappa} + 12\big]n^{-1} + \mathcal{C}_1\big[\tfrac{9*400^2\|\phi\|^2_{\ell^1}}{\kappa} + 9*800\|\phi\|_{\ell^1} + 9\kappa\big]n^{-1}\\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^\diamond_n[m^\diamond_-, f, \Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp}\theta\|^2_{L^2}}{16}n\mathfrak{b}^2_{[m_- -1]}(f)\big)\\ + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\Big[\exp\big(\tfrac{-\kappa\delta^{\widehat{\Phi}}_{m^\diamond_-}m^\diamond_-}{400\|\phi\|_{\ell^1}}\big) + n^{-1}\Big]\mathbb{1}\mho_{m^\diamond_-}\\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathbb{1}\mho^c_{m^\diamond_-} + 6\sum_{j\in[\![1,n]\!]}|\widehat{[g]}^+_j|^2|\lambda_j - \widehat{[g]}_j|^2|\theta^\circ_j|^2 + 2\sum_{j\in[\![1,n]\!]}\mathbb{1}\mathcal{X}^c_j|\theta^\circ_j|^2. \quad \text{(K.23)}\end{aligned}$$

As shown in (L.9) (cf. proof of Lemma L.2.2) holds $\frac{3}{10}\psi^\Lambda_{m^\diamond_-} \leqslant \delta^{\widehat{\Phi}}_{m^\diamond_-}\mathbb{1}\mho_{m^\diamond_-}$ and thus

$$\begin{aligned}\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta^\circ} - f\|^2_{L^2} \leqslant 3\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta^\circ}^{m_+} - \breve{f}^{m_+}\|^2_{L^2} + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + [\tfrac{1}{12\kappa\kappa^2} + \tfrac{1}{\kappa}]n^{-1}\\ + \mathcal{C}\big[\tfrac{26\|\phi\|^2_{\ell^1}}{\kappa} + 12\big]n^{-1} + \mathcal{C}_1\big[\tfrac{9*400^2\|\phi\|^2_{\ell^1}}{\kappa} + 9*800\|\phi\|_{\ell^1} + 9\kappa\big]n^{-1} + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}n^{-1}\\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^\diamond_n[m^\diamond_-, f, \Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp}\theta\|^2_{L^2}}{16}n\mathfrak{b}^2_{[m_- -1]}(f)\big)\\ + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\exp\big(\tfrac{-\kappa\frac{3}{10}\psi^\Lambda_{m^\diamond_-}m^\diamond_-}{400\|\phi\|_{\ell^1}}\big)\\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathbb{1}\mho^c_{m^\diamond_-} + 6\sum_{j\in[\![1,n]\!]}|\widehat{[g]}^+_j|^2|\lambda_j - \widehat{[g]}_j|^2|\theta^\circ_j|^2 + 2\sum_{j\in[\![1,n]\!]}\mathbb{1}\mathcal{X}^c_j|\theta^\circ_j|^2. \quad \text{(K.24)}\end{aligned}$$

Consider $\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta^\circ}^{m_+} - \breve{f}^{m_+}\|^2_{L^2} = 2\sum_{j=1}^{m_+}(\widehat{[g]}^+_j)^2/n = 2\sum_{j=1}^{m_+}\widehat{\Phi}_j/n = 2m_+\overline{\widehat{\Phi}_{m_+}}/n$ where $\Lambda^{\widehat{\Phi}}_{m_+} \geqslant \overline{\widehat{\Phi}_{m_+}}$ and $q \geqslant m_+\overline{\widehat{\Phi}_{m_+}}/n$ exploiting $(\widehat{[g]}^+_j)^2 \leqslant q$ and $m_+ \leqslant n$. Considering the event $\mho_{m^\diamond_+}$ and its complement $\mho^c_{m^\diamond_+}$ it follows $\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta^\circ}^{m_+} - \breve{f}^{m_+}\|^2_{L^2} \leqslant q\mathbb{1}\mho^c_{m^\diamond_+} + \Lambda^{\widehat{\Phi}}_{m_+}/n\mathbb{1}\mho_{m^\diamond_+}$.

Taking into account the definition (??) of m_+ we obtain $\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta^\circ}^{m_+} - \breve{f}^{m_+}\|^2_{L^2} \leqslant q\mathbb{1}\mho^c_{m^\diamond_+} + [\frac{1}{4\kappa}\|\Pi^\perp_{\mathbb{U}_{m^\diamond_+}}\breve{f}^n\|^2_{L^2} + 9\Lambda^{\widehat{\Phi}}_{m^\diamond_+}/n]\mathbb{1}\mho_{m^\diamond_+}$. Due to (L.9) (cf. proof of Lemma L.2.2) we have $\Lambda^{\widehat{\Phi}}_{m^\diamond_+}\mathbb{1}\mho_{m^\diamond_+} \leqslant \frac{63}{16}\Delta^\Lambda_{m^\diamond_+}$, which implies $\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta^\circ}^{m_+} - \breve{f}^{m_+}\|^2_{L^2} \leqslant q\mathbb{1}\mho^c_{m^\diamond_+} + \frac{1}{4\kappa}\|\Pi^\perp_{\mathbb{U}_{m^\diamond_+}}\breve{f}^n\|^2_{L^2} + \frac{9*63}{16}\Delta^\Lambda_{m^\diamond_+}/n$. Thereby, from (L.17) together with Lemma L.2.4 follows now

$$\begin{aligned}\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta^\circ} - f\|^2_{L^2} \leqslant 3q\mathbb{1}\mho^c_{m^\diamond_+} + \tfrac{3}{4\kappa}\|\Pi^\perp_{\mathbb{U}_{m^\diamond_+}}\breve{f}^n\|^2_{L^2} + \tfrac{3*9*63}{16}\Delta^\Lambda_{m^\diamond_+}/n + 3\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + [\tfrac{1}{12\kappa\kappa^2} + \tfrac{1}{\kappa}]n^{-1}\\ + \mathcal{C}\big[\tfrac{26\|\phi\|^2_{\ell^1}}{\kappa} + 12\big]n^{-1} + \mathcal{C}_1\big[\tfrac{9*400^2\|\phi\|^2_{\ell^1}}{\kappa} + 9*800\|\phi\|_{\ell^1} + 9\kappa\big]n^{-1} + 9\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}n^{-1}\\ + 3\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^\diamond_n[m^\diamond_-, f, \Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}}{16}n\mathfrak{b}^2_{[m_- - 1]}(f)\big)\\ + 9\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\exp\big(\tfrac{-\kappa 3\psi^\Lambda_{m^\diamond_-}m^\diamond_-}{4000\|\phi\|_{\ell^1}}\big)\\ + 3\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\mathbb{1}\mho^c_{m^\diamond_-} + 6\sum_{j\in[\![1,n]\!]}|\widehat{[g]}^+_j|^2|\lambda_j - \widehat{[g]}_j|^2|\theta^\circ_j|^2 + 2\sum_{j\in[\![1,n]\!]}\mathbb{1}\mathcal{X}^c_j|\theta^\circ_j|^2. \quad \text{(K.25)}\end{aligned}$$

Exploiting Lemma 1.5.1 we obtain from (i) $\mathbb{E}^q_\varepsilon\|\Pi^\perp_{\mathbb{U}_{m^\diamond_+}}\breve{f}^n\|^2_{L^2} \leqslant 4\sum_{|j|\in]\!]m_+,n]\!]}|\theta^\circ_j|^2 \leqslant 4\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\mathfrak{b}^2_{m^\diamond_+}(f)$, from (iii) $\sum_{j\in[\![1,n]\!]}|\theta^\circ_j|^2\mathbb{E}^q_\varepsilon|\widehat{[g]}^+_j|^2|\lambda_j - \widehat{[g]}_j|^2 \leqslant 4\mathcal{C}_4\mathcal{R}^\star_q(f,\Lambda)$ and from (ii) $\sum_{j\in[\![1,n]\!]}\mathbb{P}^q_\varepsilon(\mathcal{X}^c_j)|\theta^\circ_j|^2 \leqslant 4\mathcal{R}^\star_q(f,\Lambda)$ where $\mathcal{R}^\star_q(f,\Lambda) := \sum_{j\in\mathbb{N}}\theta^{\circ 2}_j[1\wedge\Lambda_j/q]$. The last bounds imply

$$\begin{aligned}\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^\circ} - f\|^2_{L^2} \leqslant \tfrac{3*9*63}{16}\Delta^\Lambda_{m^\diamond_+}n^{-1} + \tfrac{3}{\kappa}\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\mathfrak{b}^2_{m^\diamond_+}(f) + 3\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f)\\ + [\tfrac{1}{12\kappa\kappa^2} + \tfrac{1}{\kappa}]n^{-1} + \mathcal{C}\big[\tfrac{26\|\phi\|^2_{\ell^1}}{\kappa} + 12\big]n^{-1} + \mathcal{C}_1\big[\tfrac{9*400^2\|\phi\|^2_{\ell^1}}{\kappa} + 9*800\|\phi\|_{\ell^1} + 9\kappa\big]n^{-1}\\ + 9\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}n^{-1}\\ + 3\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^\diamond_n[m^\diamond_-, f, \Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}}{16}n\mathfrak{b}^2_{[m_- - 1]}(f)\big)\\ + 9\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\exp\big(\tfrac{-3\kappa\psi^\Lambda_{m^\diamond_-}m^\diamond_-}{4000\|\phi\|_{\ell^1}}\big)\\ + 3q\mathbb{P}^q_\varepsilon(\mho^c_{m^\diamond_+}) + 3\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\mathbb{P}^q_\varepsilon(\mho^c_{m^\diamond_-}) + (24\mathcal{C}_4 + 8)\mathcal{R}^\star_q(f,\Lambda). \quad \text{(K.26)}\end{aligned}$$

Recalling that $\mathcal{R}^\diamond_n[l, f, \Lambda] = [\mathfrak{b}^2_l(f) \vee \Delta^\Lambda_l n^{-1}] \geqslant n^{-1}$ there is a finite numerical constant $\mathcal{C}$ such that

$$\begin{aligned}\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^\circ} - f\|^2_{L^2} \leqslant \mathcal{C}\big\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}]\mathcal{R}^\diamond_n[m^\diamond_+, f, \Lambda] + \mathcal{R}^\star_q(f,\Lambda) + \|\phi\|^2_{\ell^1}n^{-1}\big\}\\ + 3\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + 9\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\exp\big(\tfrac{-3\kappa\psi^\Lambda_{m^\diamond_-}m^\diamond_-}{4000\|\phi\|_{\ell^1}}\big)\\ + 3\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^\diamond_n[m^\diamond_-, f, \Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}}{16}n\mathfrak{b}^2_{[m_- - 1]}(f)\big)\\ + 3q\mathbb{P}^q_\varepsilon(\mho^c_{m^\diamond_+}) + 3\|\Pi_{\mathbb{U}_0^\perp}f\|^2_{L^2}\mathbb{P}^q_\varepsilon(\mho^c_{m^\diamond_-}), \quad \text{(K.27)}\end{aligned}$$

which shows the assertion and completes the proof.

LEMMA K.2.7. *Let f have an infinite series expansion as definied in (np), i.e., $\mathfrak{b}_m(f) > 0$ for all $m \in \mathbb{N}$. Set $m_h := \lceil 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2\rceil$, $\widetilde{m}_h := \min\{m \in \mathbb{N} : \mathfrak{b}_{m_h}(f) > \mathfrak{b}_m(f)\}$, $n_{f,\Lambda} := \lceil \frac{\Delta^\Lambda_{\widetilde{m}_h}}{\mathfrak{b}^2_{\widetilde{m}_h}(f)} \vee 15\frac{300^4}{\kappa^2}\rceil$ and $q_{f,\Lambda} := \lceil 289\log(m_h+2)\psi^\Lambda_{m_h}\Lambda_{(m_h)}\rceil$.*

For all $n \in \mathbb{N}$ let $m^\bullet_n \in [\![m^\diamond_n, n]\!]$ and for all $q \in \mathbb{N}$ we set $m^\bullet_q := \max\{m \in [\![K_h, q]\!] : 289\log(m+2)\psi^\Lambda_m\Lambda_{(m)} \leqslant q\}$ for $q > q_{f,\Lambda}$, where the defining set containing K_h is not empty, and $m^\bullet_q := m^\diamond_n$ for $q \leqslant q_{f,\Lambda}$. There is a finite numerical constant $\mathcal{C}$ such that for all $n, q \in \mathbb{N}$ holds

$$\begin{aligned}\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^\circ - f\|^2_{L^2} \leqslant +\mathcal{C}\big\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}]\,\mathcal{R}^\diamond_n[m^\bullet_n \wedge m^\bullet_q, f, \Lambda] + \mathcal{R}^\star_q(f,\Lambda)\\ + [1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^\Lambda_{n_{f,\Lambda}} n^{-1} + q_{f,\Lambda} q^{-1}) + \|\phi\|^2_{\ell^1} n^{-1} + q^2_{f,\Lambda} q^{-1}\big\}\\ + 12\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\{\exp\big(\frac{-3\kappa\psi^\Lambda_{m^\bullet_n\wedge m^\bullet_q}[m^\bullet_n \wedge m^\bullet_q]}{4000\|\phi\|_{\ell^1}}\big) - [\mathcal{R}^\diamond_n(f,\Lambda) \vee \mathcal{R}^\star_q(f,\Lambda)]\}\end{aligned} \quad (K.28)$$

If there are $\widetilde{n}_{f,\Lambda}, \widetilde{q}_{f,\Lambda} \in \mathbb{N}$ such that additionaly (i) $\psi^\Lambda_{m^\diamond_n} m^\diamond_n \geqslant m_h(-\log\mathcal{R}^\diamond_n(f,\Lambda))$ for all $n \geqslant \widetilde{n}_{f,\Lambda}$ and (ii) $\psi^\Lambda_{m^\bullet_q} m^\bullet_q \geqslant m_h(-\log\mathcal{R}^\star_q(f,\Lambda))$ for all $q \geqslant \widetilde{q}_{f,\Lambda}$, then for all $n, q \in \mathbb{N}$ holds

$$\begin{aligned}\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^\circ - f\|^2_{L^2} \leqslant \mathcal{C}\big\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\mathcal{R}^\diamond_n[m^\diamond_n \wedge m^\bullet_q, f, \Lambda] + \mathcal{R}^\star_q(f,\Lambda))\\ + [1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^\Lambda_{[\widetilde{n}_{f,\Lambda}\vee n_{f,\Lambda}]} n^{-1} + [\widetilde{q}_{f,\Lambda} \vee q_{f,\Lambda}] q^{-1}) + \|\phi\|^2_{\ell^1} n^{-1} + [\widetilde{q}_{f,\Lambda} \vee q_{f,\Lambda}]^2 q^{-1}\big\}\end{aligned} \quad (K.29)$$

REMINDER. *Keep in mind that $m^\diamond_n := \arg\min\{\mathcal{R}^\diamond_n[m,f,\Lambda], m \in [\![1,n]\!]\}$ with $\mathcal{R}^\diamond_n[m,f,\Lambda] = [\mathfrak{b}^2_m(f) \vee \Delta^\Lambda_m n^{-1}]$ and $\mathcal{R}^\diamond_n(f,\Lambda) := \mathcal{R}^\diamond_n[m^\diamond_n, f, \Lambda]$. Considering $\widetilde{m}_h = \min\{m \in \mathbb{N} : \mathfrak{b}_{m_h}(h) > \mathfrak{b}_m(h)\}$ as defined in Lemma L.2.7 we note that the defining set is not empty since $\mathfrak{b}_m(f) > 0$ for all $m \in \mathbb{N}$ and $\lim_{m\to\infty}\mathfrak{b}_m(f) = 0$, where $\widetilde{m}_h > m_h$ due to the monotonicity of $\mathfrak{b}_m(f)$. Moreover, for all $n \geqslant n_{f,\Lambda} := \lceil \frac{\Delta^\Lambda_{\widetilde{m}_h}}{\mathfrak{b}^2_{\widetilde{m}_h}(f)}\rceil$ holds $\mathcal{R}^\circ_n[m_h, f, \Lambda] \geqslant \mathfrak{b}^2_{m_h}(f) > \mathfrak{b}^2_{\widetilde{m}_h}(f) = \mathcal{R}^\circ_n[\widetilde{m}_h, f, \Lambda]$ and hence, for all $n \geqslant n_{f,\Lambda}$ we have $m^\bullet_n \geqslant m^\diamond_n > m_h$. Furthermore, for all $q \geqslant q_{f,\Lambda} := \lceil 289\log(m_h+2)\psi^\Lambda_{m_h}\Lambda_{(m_h)}\rceil$ is the defining set of $m^\bullet_q = \max\{m \in [\![m_h, q]\!] : 289\log(m+2)\psi^\Lambda_m\Lambda_{(m)} \leqslant q\}$ not empty. Consequently, for all $n \geqslant n_{f,\Lambda}$, $q \geqslant q_{f,\Lambda}$ follows $m^\bullet_n \wedge m^\bullet_q \geqslant m_h$. We use these preliminary findings in the proof of Lemma L.2.7 without further reference.*

PROOF OF LEMMA L.2.7.
Let $m_h := \lceil 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2\rceil$, $\widetilde{m}_h := \min\{m \in \mathbb{N} : \mathfrak{b}_{m_h}(f) > \mathfrak{b}_m(f)\}$, $n_{f,\Lambda} := \lceil \frac{\widetilde{m}_\lambda \delta^{|\Phi}_{\widetilde{m}_\lambda} |\Phi_{(\widetilde{m}_\lambda)}}{\mathfrak{b}^2_{\widetilde{m}_\lambda}(\theta)} \vee 15\frac{300^4}{\kappa^2}\rceil$ and $q_{f,\Lambda} := \lceil 289\log(m_h+2)\psi^\Lambda_{m_h}\Lambda_{(m_h)}\rceil$.

We distinguish for $n \in \mathbb{N}$ the following two cases, (a) $n \in [\![1, n_{f,\Lambda}[\![$ and (b) $n \geqslant n_{f,\Lambda}$.

Firstly, consider (a), let $n \in [\![1, n_{f,\Lambda}[\![$, then setting $m_- = 1$ and $m_+ = n$ from Lemma

L.2.1 follows

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant 3\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ n} - \breve{f}^n\|^2_{L^2} + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}_1^2(f)$$
$$+ 6\sum_{j\in[\![1,n]\!]} \mathbb{E}^q_\varepsilon|\widehat{[g]}_j^+|^2|\lambda_j - \widehat{[g]}_j|^2|\theta^\circ_j|^2 + 2\sum_{j\in[\![1,n]\!]} \mathbb{P}^q_\varepsilon(\mathcal{X}^c_j)|\theta^\circ_j|^2$$

Exploiting Lemma 1.5.1 we obtain from (i) $\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ n} - \breve{f}^n\|^2_{L^2} = 2\sum_{j\in[\![1,n]\!]}\Lambda_j\mathbb{E}^q_\varepsilon|\lambda_j\widehat{[g]}_j^+|^2\mathbb{E}^n_Y|\widehat{[h]}_j - \phi_j|^2 \leqslant 8n\overline{\Lambda}_n n^{-1} \leqslant 8\Delta^\Lambda_{n_{f,\Lambda}}n^{-1}$, from (iii) $\sum_{j\in[\![1,n]\!]}|\theta^\circ_j|^2\mathbb{E}^q_\varepsilon|\widehat{[g]}_j^+|^2|\lambda_j - \widehat{[g]}_j|^2 \leqslant 4\mathcal{C}_4\mathcal{R}^\star_q(f,\Lambda)$ and from (ii) $\sum_{j\in[\![1,n]\!]}\mathbb{P}^q_\varepsilon(\mathcal{X}^c_j)|\theta^\circ_j|^2 \leqslant 4\mathcal{R}^\star_q(f,\Lambda)$ where $\mathcal{R}^\star_q(f,\Lambda) := \sum_{j\in[\![1,K[\![}{\theta^\circ_j}^2[1\wedge\Lambda_j/q]$. The last bounds together with $\mathfrak{b}_1^2(f) \leqslant 1$ and $n < n_{f,\Lambda} \leqslant \Delta^\Lambda_{n_{f,\Lambda}}$ imply for all $q \in \mathbb{N}$

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant 24\Delta^\Lambda_{n_{f,\Lambda}}n^{-1} + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2} + (24\mathcal{C}_4 + 8)\mathcal{R}^\star_q(f,\Lambda)$$
$$\leqslant \mathcal{C}\{(1\vee\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2})\Delta^\Lambda_{n_{f,\Lambda}}n^{-1} + \mathcal{R}^\star_q(f,\Lambda)\} \quad \text{(K.30)}$$

where $\mathcal{C}$ is a finite numerical constant.

Secondly, consider (b), i.e., $n \geqslant n_{f,\Lambda} \geqslant 15\frac{300^4}{\kappa^2}$ where by construction $m^\bullet_n \geqslant m^\diamond_n > m_h \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ (see appendix L.2). We destinguish further for $q \in \mathbb{N}$ the following two cases cases, (b-i) $q \in [\![1, q_{f,\Lambda}[\![$ and (b-ii) $q \geqslant q_{f,\Lambda}$ where $q_{f,\Lambda} = \lceil 289\log(m_h+2)\psi^\Lambda_{m_h}\Lambda_{(m_h)}\rceil$.

Consider (b-i). Since $n \geqslant n_{f,\Lambda} \geqslant 15\frac{300^4}{\kappa^2}$ and $m^\bullet_n > m_h \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ we can employ Proposition L.2.1 with $m^\diamond_+ := m^\bullet_n$, and thus

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant \mathcal{C}\{[1\vee\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}]\mathcal{R}^\diamond_n[m^\bullet_n, f,\Lambda] + \mathcal{R}^\star_q(f,\Lambda) + \|\phi\|^2_{\ell^1}n^{-1}\}$$
$$+ 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\exp\big(\frac{-3\kappa\psi^\Lambda_{m^\diamond_-}m^\diamond_-}{4000\|\phi\|_{\ell^1}}\big)$$
$$+ 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^\diamond_n[m^\diamond_-, f,\Lambda] - \frac{\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}}{16}n\mathfrak{b}^2_{[m_- - 1]}(f)\big)$$
$$+ 3q\mathbb{P}^q_\varepsilon(\mho^c_{m^\bullet_n}) + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathbb{P}^q_\varepsilon(\mho^c_{m^\diamond_-})$$

Setting $m^\diamond_- = 1$, and hence $m_- = 1$, by employing $\mathfrak{b}^2_1(f) \leqslant 1$ and $q \leqslant q_{f,\Lambda}$ it follows

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant \mathcal{C}\{[1\vee\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}]\mathcal{R}^\diamond_n[m^\bullet_n, f,\Lambda] + \mathcal{R}^\star_q(f,\Lambda) + \|\phi\|^2_{\ell^1}n^{-1}\}$$
$$+ 15\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2} + 3q$$
$$\leqslant \mathcal{C}_1\{[1\vee\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\mathcal{R}^\diamond_n[m^\bullet_n, f,\Lambda] + q_{f,\Lambda}q^{-1}) + \mathcal{R}^\star_q(f,\Lambda) + q^2_{f,\Lambda}q^{-1} + \|\phi\|^2_{\ell^1}n^{-1}\} \quad \text{(K.31)}$$

where $\mathcal{C}_1$ is a finite numerical constant.

Consider (b-ii). We note that for all $q \geqslant q_{f,\Lambda} := \lceil 289\log(m_h+2)\psi^\Lambda_{m_h}\Lambda_{(m_h)}\rceil$ is the defining set of $m^\bullet_q = \max\{m \in [\![m_h, q]\!] : 289\log(m+2)\psi^\Lambda_m\Lambda_{(m)} \leqslant q\}$ not empty. By construction, for all $n \geqslant n_{f,\Lambda}$, $q \geqslant q_{f,\Lambda}$ follows $m^\bullet_n \wedge m^\bullet_q > m_h \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$. Since also $n \geqslant n_{f,\Lambda} \geqslant 15\frac{300^4}{\kappa^2}$

we can employ Proposition L.2.1 with $m_+^\diamond := m_n^\bullet \wedge m_q^\bullet$, and thus

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2]\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda] + \mathcal{R}_q^\star(f,\Lambda) + \|\phi\|_{\ell^1}^2 n^{-1}\} \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_{m_-}^2(f) + 9\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \exp\big(\tfrac{-3\kappa\psi_{m_-^\diamond}^\Lambda m_-^\diamond}{4000\|\phi\|_{\ell^1}}\big) \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 [m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond[m_-^\diamond, f, \Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{16} n\mathfrak{b}_{[m_- - 1]}^2(f)\big) \\ + 3q\mathbb{P}_\varepsilon^q(\mho_{m_n^\bullet \wedge m_q^\bullet}^c) + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathbb{P}_\varepsilon^q(\mho_{m_-^\diamond}^c)\end{aligned}$$

Noting that $m_+^\diamond = m_n^\bullet \wedge m_q^\bullet \in [\![m_h, m_q^\bullet]\!]$ and $289\log(m_+^\diamond + 2)\psi_{m_+^\diamond}^\Lambda \Lambda_{(m_+^\diamond)} \leqslant q$ by construction of $m_q^\bullet$ from ?? ?? follows $q\mathbb{P}_\varepsilon^q(\mho_{m_n^\bullet \wedge m_q^\bullet}^c) \leqslant 11226q^{-1}$ and thus

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}_1\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2]\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda] + \mathcal{R}_q^\star(f,\Lambda) + \|\phi\|_{\ell^1}^2 n^{-1} + q^{-1}\} \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_{m_-}^2(f) + 9\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \exp\big(\tfrac{-3\kappa\psi_{m_-^\diamond}^\Lambda m_-^\diamond}{4000\|\phi\|_{\ell^1}}\big) \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 [m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond[m_-^\diamond, f, \Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{16} n\mathfrak{b}_{[m_- - 1]}^2(f)\big) \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathbb{P}_\varepsilon^q(\mho_{m_-^\diamond}^c)\end{aligned}$$

where $\mathcal{C}_1$ is a finite numerical constant. Setting $m_-^\diamond := m_n^\bullet \wedge m_q^\bullet$ and m_- as in definition (??), that is $m_- \leqslant m_-^\diamond = m_n^\bullet \wedge m_q^\bullet$ and $\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_{m_-}^2(f) \leqslant [2\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 + 7576\kappa]\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda]$, it follows

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}_2\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2]\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda] + \mathcal{R}_q^\star(f,\Lambda) + \|\phi\|_{\ell^1}^2 n^{-1} + q^{-1}\} \\ + 9\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \exp\big(\tfrac{-3\kappa\psi_{m_n^\bullet \wedge m_q^\bullet}^\Lambda [m_n^\diamond \wedge m_q^\bullet]}{4000\|\phi\|_{\ell^1}}\big) \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 [m_n^\bullet \wedge m_q^\bullet]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda]\big) \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathbb{P}_\varepsilon^q(\mho_{m_n^\bullet \wedge m_q^\bullet}^c)\end{aligned}$$

where $\mathcal{C}_2$ is a finite numerical constant. Noting that $m_-^\diamond = m_n^\bullet \wedge m_q^\bullet \in [\![m_h, m_q^\bullet]\!]$ and $289\log(m_-^\diamond + 2)\psi_{m_-^\diamond}^\Lambda \Lambda_{(m_-^\diamond)} \leqslant q$ by construction of $m_q^\bullet$ from ?? ?? follows $\mathbb{P}_\varepsilon^q(\mho_{m_n^\bullet \wedge m_q^\bullet}^c) \leqslant 53q^{-1}$ and together with $\mathcal{R}_q^\star(f,\Lambda) \geqslant \frac{1}{2}\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 q^{-1}$ thus

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}_3\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2]\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda] + \mathcal{R}_q^\star(f,\Lambda) + \|\phi\|_{\ell^1}^2 n^{-1} + q^{-1}\} \\ + 9\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \exp\big(\tfrac{-3\kappa\psi_{m_n^\bullet \wedge m_q^\bullet}^\Lambda [m_n^\bullet \wedge m_q^\bullet]}{4000\|\phi\|_{\ell^1}}\big) \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 [m_n^\bullet \wedge m_q^\bullet]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda]\big)\end{aligned}$$

where $\mathcal{C}_3$ is a finite numerical constant. Taking further into account that for all $m \in \mathbb{N}$ hold $n\mathcal{R}_n^\diamond[m, f, \Lambda] \geqslant \Delta_m^\Lambda \geqslant \psi_m^\Lambda m$ and (keeping in mind $\psi_m^\Lambda \geqslant 1$, $\kappa \geqslant 1$, $\kappa \geqslant 1$ and $\|\phi\|_{\ell^1} \geqslant 1$)

hence

$$m\exp(-\kappa\kappa n\mathcal{R}_n^\diamond[m,f,\Lambda]) \leqslant \tfrac{2}{\kappa\kappa}\tfrac{\kappa\kappa}{2}\psi_m^\Lambda m\exp\big(-\tfrac{\kappa\kappa}{2}\psi_m^\Lambda m\big)\exp\big(-\tfrac{\kappa\kappa}{2}\psi_m^\Lambda m\big) \\ \leqslant \tfrac{2}{e\kappa\kappa}\exp\big(-\tfrac{\kappa\kappa}{2}\psi_m^\Lambda m\big) \leqslant \tfrac{1}{\kappa\kappa}\exp\big(\tfrac{-3\kappa\psi_m^\Lambda m}{4000\|\phi\|_{\ell^1}}\big).$$

The last bound implies that

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}_3\big\{[1\vee\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2]\mathcal{R}_n^\diamond[m_n^\bullet\wedge m_q^\bullet,f,\Lambda] + \mathcal{R}_q^\star(f,\Lambda) + \|\phi\|_{\ell^1}^2 n^{-1} + q^{-1}\big\} \\ + 12\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2\exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet\wedge m_q^\bullet}[m_n^\bullet\wedge m_q^\bullet]}{4000\|\phi\|_{\ell^1}}\big) \quad \text{(K.32)}$$

Combining the upper bounds (L.32), (L.33) and(L.34) for the three cases (a), (b-i) and (b-ii) and keeping in mind that $\mathcal{R}_n^\diamond[m_n^\bullet\wedge m_q^\bullet,f,\Lambda] \geqslant \mathcal{R}_n^\diamond(f,\Lambda) = \min\{\mathcal{R}_n^\diamond[m,f,\Lambda], m\in\mathbb{N}\}$ there is a finite numerical constant $\mathcal{C}$ such that for all $n,q\in\mathbb{N}$ holds

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}\big\{[1\vee\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2]\,\mathcal{R}_n^\diamond[m_n^\bullet\wedge m_q^\bullet,f,\Lambda] + \mathcal{R}_q^\star(f,\Lambda) \\ + [1\vee\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2](\Delta^\Lambda_{n_{f,\Lambda}}n^{-1} + q_{f,\Lambda}q^{-1}) + \|\phi\|_{\ell^1}^2 n^{-1} + q_{f,\Lambda}^2 q^{-1}\big\} \\ + 12\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2\{\exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet\wedge m_q^\bullet}[m_n^\bullet\wedge m_q^\bullet]}{4000\|\phi\|_{\ell^1}}\big) - [\mathcal{R}_n^\diamond(f,\Lambda)\vee\mathcal{R}_q^\star(f,\Lambda)]\} \quad \text{(K.33)}$$

which shows the assertion (L.30).

Assume finally, that there are in addition $\widetilde{n}_{f,\Lambda},\widetilde{q}_{f,\Lambda}\in\mathbb{N}$ such that for all $n\geqslant\widetilde{n}_{f,\Lambda}$, $\psi^\Lambda_{m_n^\diamond}m_n^\diamond \geqslant m_h(-\log\mathcal{R}_n^\diamond(f,\Lambda))$, and for all $q\geqslant\widetilde{q}_{f,\Lambda}$, $\psi^\Lambda_{m_q^\bullet}m_q^\bullet \geqslant m_h(-\log\mathcal{R}_q^\star(f,\Lambda))$. We shall use without further reference that then $\exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\diamond\wedge m_q^\bullet}(m_n^\diamond\wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) \leqslant \mathcal{R}_n^\diamond(f,\Lambda)\vee\mathcal{R}_q^\star(f,\Lambda) \leqslant \mathcal{R}_n^\diamond[m_n^\diamond\wedge m_q^\bullet,f,\Lambda]\vee\mathcal{R}_q^\star(f,\Lambda)$ for all $n\geqslant\widetilde{n}_{f,\Lambda}$ and $q\geqslant\widetilde{q}_{f,\Lambda}$ since $m_h\geqslant\frac{4000\|\phi\|_{\ell^1}}{3\kappa}$. Following line by line the proof of (L.29) using $\widetilde{n}_{f,\Lambda}\vee n_{f,\Lambda}$ and $\widetilde{q}_{f,\Lambda}\vee q_{f,\Lambda}$ rather than $n_{f,\Lambda}$ and $q_{f,\Lambda}$, respectively, we obtain the assertion (L.31), that is,

$$\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}\big\{[1\vee\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2](\mathcal{R}_n^\diamond[m_n^\diamond\wedge m_q^\bullet,f,\Lambda] + \mathcal{R}_q^\star(f,\Lambda)) \\ + [1\vee\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2](\Delta^\Lambda_{[\widetilde{n}_{f,\Lambda}\vee n_{f,\Lambda}]}n^{-1} + [\widetilde{q}_{f,\Lambda}\vee q_{f,\Lambda}]q^{-1}) + \|\phi\|_{\ell^1}^2 n^{-1} + [\widetilde{q}_{f,\Lambda}\vee q_{f,\Lambda}]^2 q^{-1}\big\}, \quad \text{(K.34)}$$

which completes the proof.

Appendix L

Proofs of Theorem 3.6.2

L.1 Intermediate results

L.2 Detailed proofs

Proof of ??.
Due to Lemma L.2.7 under Assumption 20 there is a finite numerical constant $\mathcal{C}$ such that for all $n, q \in \mathbb{N}$ holds

$$\begin{aligned}\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant \mathcal{C}\big\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}]\,\mathcal{R}^{\diamond}_n[m^{\bullet}_n \wedge m^{\bullet}_q, f, \Lambda] + \mathcal{R}^{\star}_q(f,\Lambda)\\ + [1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^{\Lambda}_{n_{f,\Lambda}} n^{-1} + q_{f,\Lambda} q^{-1}) + \|\phi\|^2_{\ell^1} n^{-1} + q^2_{f,\Lambda} q^{-1}\big\}\\ + 12\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\{\exp\big(\tfrac{-3\kappa\psi^{\Lambda}_{m^{\bullet}_n\wedge m^{\bullet}_q}[m^{\bullet}_n\wedge m^{\bullet}_q]}{4000\|\phi\|_{\ell^1}}\big) - [\mathcal{R}^{\diamond}_n(f,\Lambda) \vee \mathcal{R}^{\star}_q(f,\Lambda)]\}.\end{aligned} \quad \text{(L.1)}$$

Since under Assumption 20 for all $n, q \in \mathbb{N}$ hold $\psi^{\Lambda}_{m^{\bullet}_n} m^{\bullet}_n \geqslant m_h|\log \mathcal{R}^{\diamond}_n(f,\Lambda)|$ and $\psi^{\Lambda}_{m^{\bullet}_q} m^{\bullet}_q \geqslant m_h[|\log \mathcal{R}^{\diamond}_n(f,\Lambda)| \wedge |\log \mathcal{R}^{\star}_q(f,\Lambda)|]$ true it follows

$$\begin{aligned}\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant \mathcal{C}\big\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}]\,\mathcal{R}^{\diamond}_n[m^{\bullet}_n \wedge m^{\bullet}_q, f, \Lambda] + \mathcal{R}^{\star}_q(f,\Lambda)\\ + [1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^{\Lambda}_{n_{f,\Lambda}} n^{-1} + q_{f,\Lambda} q^{-1}) + \|\phi\|^2_{\ell^1} n^{-1} + q^2_{f,\Lambda} q^{-1}\big\}.\end{aligned} \quad \text{(L.2)}$$

which shows the assertion (3.3) and completes the proof.

Lemma L.2.1. *Consider the aggregated OSE $\widehat{\theta}^{\circ} = \sum_{m=1}^n w_m \widehat{\theta}^{\circ^m}$ with weights $w_m \in [0,1]$, $m \in [\![1,n]\!]$, satisfying $\sum_{m=1}^n w_m = 1$ and given $m \in \mathbb{N}$ let $\breve{f}^m := \sum_{j=-m}^{m} \widehat{[g]}^+_j \phi_j e_j$. For* *any* $m_- \in [\![1,n]\!]$ *and* $m_+ \in [\![1,n]\!]$ *holds*

$$\|\widehat{\theta^\circ} - f\|_{L^2}^2 \leqslant 3\|\widehat{\theta^\circ}^{m_+} - \breve{f}^{m_+}\|_{L^2}^2 + 3\sum_{l\in]\!]m_+,n]\!]}\left(\|\widehat{\theta^\circ}^l - \breve{f}^l\|_{L^2}^2 - \mathcal{C}_1 12\kappa\Lambda_l^{\widehat{\Phi}}/n\right)_+$$
$$+3(\mathcal{C}_1\kappa/n)\sum_{l\in]\!]m_+,n]\!]}\Lambda_l^{\widehat{\Phi}}\mathbb{1}_{\{\|\widehat{\theta^\circ}^l - \breve{f}^l\|_{L^2}^2 \geqslant 12\kappa\Lambda_l^{\widehat{\Phi}}/n\}} + 3(\kappa/n)\sum_{l\in]\!]m_+,n]\!]}\Lambda_l^{\widehat{\Phi}} w_l \mathbb{1}_{\{\|\widehat{\theta^\circ}^l - \breve{f}^l\|_{L^2}^2 < 12\kappa\Lambda_l^{\widehat{\Phi}}/n\}}$$
$$+ 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2\{\mathbb{P}_w([\![1,m_-[\![) + \mathfrak{b}_{m_-}^2(f)\}$$
$$+ 6\sum_{j\in[\![1,n]\!]}|\widehat{[g]}_j^+|^2|\lambda_j - \widehat{[g]}_j|^2|\theta_j^\circ|^2 + 2\sum_{j\in[\![1,n]\!]}\mathbb{1}_{\mathcal{X}_j^c}|\theta_j^\circ|^2 \quad \text{(L.3)}$$

PROOF OF LEMMA L.2.1.
We start the proof with the observation that $\overline{[\widehat{\theta^\circ}]_j} - \overline{[f]_j} = \widehat{\theta^\circ}_{-j} - \theta^\circ_{-j}$ for all $j \in \mathbb{Z}$, $\widehat{\theta^\circ}_0 - \theta_0^\circ = 0$ and $\widehat{\theta^\circ}_j - \theta_j^\circ = -\theta_j^\circ$ for all $j > n$, while for all $j \in [\![1,n]\!]$ with $\mathcal{X}_j := \{|\widehat{[g]}_j|^2 \geqslant 1/q\}$ and $\mathcal{X}_j^c := \{|\widehat{[g]}_j|^2 < 1/q\}$ holds

$$\begin{aligned}\widehat{\theta^\circ}_j - \theta_j^\circ &= (\widehat{[g]}_j^+\widehat{[h]}_j - \theta_j^\circ)\mathbb{P}_w([\![j,n]\!]) - \theta_j^\circ\mathbb{P}_w([\![1,j[\![)\\ &= \widehat{[g]}_j^+(\widehat{[h]}_j - \phi_j)\mathbb{P}_w([\![j,n]\!]) + \widehat{[g]}_j^+(\lambda_j - \widehat{[g]}_j)\theta_j^\circ\mathbb{P}_w([\![j,n]\!]) - \mathbb{1}_{\mathcal{X}_j}\theta_j^\circ\mathbb{P}_w([\![1,j[\![) - \mathbb{1}_{\mathcal{X}_j^c}\theta_j^\circ\end{aligned}$$

Consequently, we have

$$\begin{aligned}\|\widehat{\theta^\circ} - f\|_{L^2}^2 &= 2\sum_{j\in[\![1,n]\!]}|\widehat{[g]}_j^+(\widehat{[h]}_j - \phi_j)\mathbb{P}_w([\![j,n]\!]) + \widehat{[g]}_j^+(\lambda_j - \widehat{[g]}_j)\theta_j^\circ\mathbb{P}_w([\![j,n]\!]) - \theta_j^\circ\mathbb{P}_w([\![1,j[\![)|^2\mathbb{1}_{\mathcal{X}_j}\\ &\qquad + 2\sum_{j\in[\![1,n]\!]}\mathbb{1}_{\mathcal{X}_j^c}|\theta_j^\circ|^2 + 2\sum_{j>n}|\theta_j^\circ|^2\\ &\leqslant 6\sum_{j\in[\![1,n]\!]}|\widehat{[g]}_j^+|^2|\widehat{[h]}_j - \phi_j|^2\mathbb{P}_w([\![j,n]\!]) + 6\sum_{j\in[\![1,n]\!]}\mathbb{1}_{\mathcal{X}_j}|\theta_j^\circ|^2\mathbb{P}_w([\![1,j[\![) + 2\sum_{j>n}|\theta_j^\circ|^2\\ &\qquad + 6\sum_{j\in[\![1,n]\!]}|\widehat{[g]}_j^+|^2|\lambda_j - \widehat{[g]}_j|^2|\theta_j^\circ|^2 + 2\sum_{j\in[\![1,n]\!]}\mathbb{1}_{\mathcal{X}_j^c}|\theta_j^\circ|^2. \quad \text{(L.4)}\end{aligned}$$

Consider the first r.hs. term in (L.4). We split the sum into two parts which we bound separately. Precisely, given $\breve{f}^m = \sum_{j=-m}^m \widehat{[g]}_j^+\phi_j e_j$ where $\|\widehat{\theta^\circ}^m - \breve{f}^m\|_{L^2}^2 = 2\sum_{j\in[\![1,m]\!]}|\widehat{\theta^\circ}_j^m -$

$\check{f}^m_j|^2 = 2\sum_{j\in[\![1,m]\!]} |\widehat{[g]}_j^+|^2|\widehat{[h]}_j - \phi_j|^2$ it follows

$$
\begin{aligned}
2\sum_{j\in[\![1,n]\!]} &|\widehat{[g]}_j^+|^2(\widehat{[h]}_j - \phi_j)^2\mathbb{P}_{w}([\![j,n]\!])\\
&\leqslant 2\sum_{j\in[\![1,m_+]\!]} |\widehat{[g]}_j^+|^2|(\widehat{[h]}_j - \phi_j)^2 + 2\sum_{j\in]\!]m_+,n]\!]} |\widehat{[g]}_j^+|^2(\widehat{[h]}_j - \phi_j)^2\sum_{l\in[\![j,n]\!]} w_l\\
&= 2\sum_{j\in[\![1,m_+]\!]} |\widehat{[g]}_j^+|^2|(\widehat{[h]}_j - \phi_j)^2 + 2\sum_{l\in]\!]m_+,n]\!]} w_l \sum_{j\in]\!]m_+,l]\!]} |\widehat{[g]}_j^+|^2(\widehat{[h]}_j - \phi_j)^2\\
&\leqslant \|\widehat{\theta^\circ}^{m_+} - \check{f}^{m_+}\|_{L^2}^2 + \sum_{l\in]\!]m_+,n]\!]} w_l\|\widehat{\theta^\circ}^l - \check{f}^l\|_{L^2}^2\\
&\leqslant \|\widehat{\theta^\circ}^{m_+} - \check{f}^{m_+}\|_{L^2}^2 + \sum_{l\in]\!]m_+,n]\!]} (\|\widehat{\theta^\circ}^l - \check{f}^l\|_{L^2}^2 - \mathcal{C}_1 12\kappa\Lambda_l^{\widehat{\Phi}}/n)_+\\
&+(\mathcal{C}_1\kappa/n)\sum_{l\in]\!]m_+,n]\!]} \Lambda_l^{\widehat{\Phi}}\mathbb{1}_{\{\|\widehat{\theta^\circ}^l - \check{f}^l\|_{L^2}^2 \geqslant 12\kappa\Lambda_l^{\widehat{\Phi}}/n\}} + (\kappa/n)\sum_{l\in]\!]m_+,n]\!]} \Lambda_l^{\widehat{\Phi}} w_l \mathbb{1}_{\{\|\widehat{\theta^\circ}^l - \check{f}^l\|_{L^2}^2 < 12\kappa\Lambda_l^{\widehat{\Phi}}/n\}}
\end{aligned}
\tag{L.5}
$$

Consider the second and third r.hs. term in (L.4). Splitting the first sum into two parts we obtain

$$
\begin{aligned}
2\sum_{j\in[\![1,n]\!]} &\mathbb{1}\mathcal{X}_j|\theta^\circ_j|^2\mathbb{P}_{w}([\![1,j[\![) + 2\sum_{j>n}|\theta^\circ_j|^2\\
&\leqslant 2\sum_{j\in[\![1,m_-]\!]} |\theta^\circ_j|^2\mathbb{1}\mathcal{X}_j\mathbb{P}_{w}([\![1,j[\![) + 2\sum_{j\in]\!]m_-,n]\!]} |\theta^\circ_j|^2 + 2\sum_{j>n}|\theta^\circ_j|^2\\
&\leqslant \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2\{\mathbb{P}_{w}([\![1,m_-[\![) + \mathfrak{b}^2_{m_-}(f)\}
\end{aligned}
\tag{L.6}
$$

Combining (L.4) and the upper bounds (L.5) and (L.6) we obtain the assertion, which completes the proof.

Lemma L.2.2. *For any $l \in [\![1,n]\!]$*

(i) *with $\mathcal{R}^\diamond_n[l,f,\Lambda] = [\mathfrak{b}^2_l(f) \vee \Delta^\Lambda_l\, n^{-1}]$ for all $k \in [\![1,l[\![$ we have*
$$\widehat{w}_k\mathbb{1}\Big\{\|\widehat{\theta^\circ}^l - \check{f}^l\|_{L^2}^2 < 12\kappa\Lambda_l^{\widehat{\Phi}}/n\Big\}\cap\Big\{1/2 \leqslant |\lambda_j\widehat{[g]}_j^+| \leqslant 3/2,\ \forall\, j\in[\![1,l]\!]\Big\}$$
$$\leqslant \exp\Big(\kappa n\Big\{-\frac{\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{8}\mathfrak{b}^2_k(f) + [\tfrac{630}{16} * 12\kappa + \frac{\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{8}]\mathcal{R}^\diamond_n[l,f,\Lambda]\Big\}\Big)$$

(ii) *with $\|\Pi_{\mathbb{U}_l}^\perp \check{f}^n\|_{L^2}^2 = 2\sum_{j=l+1}^n |\lambda_j\widehat{[g]}_j^+|^2|\theta^\circ_j|^2$ for all $k \in]\!]l,n]\!]$ we have*
$$\widehat{w}_k\mathbb{1}_{\{\|\widehat{\theta^\circ}^k - \check{f}^k\|_{L^2}^2 < 12\kappa\Lambda_k^{\widehat{\Phi}}/n\}} \leqslant \exp\Big(\kappa n\Big\{-12\kappa\Lambda_k^{\widehat{\Phi}}/n + [\tfrac{3}{2}\|\Pi_{\mathbb{U}_l}^\perp \check{f}^n\|_{L^2}^2 + \tfrac{9}{2}12\kappa\Lambda_l^{\widehat{\Phi}}/n]\Big\}\Big).$$

Proof of Lemma L.2.2.
Given $m, l \in [\![1,n]\!]$ and an event Ω_{ml} (to be specified below) it clearly follows

$$
\begin{aligned}
\widehat{w}_m\mathbb{1}_{\Omega_{ml}} &= \frac{\exp(-\kappa n\{-\|\widehat{\theta^\circ}^m\|_{L^2}^2 + \frac{9}{2}12\kappa\Lambda_m^{\widehat{\Phi}}/n\})}{\sum_{s\in[\![1,n]\!]}\exp(-\kappa n\{-\|\widehat{\theta^\circ}^s\|_{L^2}^2 + \frac{9}{2}12\kappa\Lambda_s^{\widehat{\Phi}}/n\})}\mathbb{1}_{\Omega_{ml}}\\
&\leqslant \exp\big(\kappa n\{\|\widehat{\theta^\circ}^m\|_{L^2}^2 - \|\widehat{\theta^\circ}^l\|_{L^2}^2 + \tfrac{9}{2}(12\kappa\Lambda_l^{\widehat{\Phi}}/n - 12\kappa\Lambda_m^{\widehat{\Phi}}/n)\}\big)\mathbb{1}_{\Omega_{ml}}
\end{aligned}
\tag{L.7}
$$

We distinguish the two cases $m < l$ and $m > l$. Consider first that $m < l$. From ***(i)*** in corollary G.1.1 (with $\breve{f}^\bullet = \widehat{\theta^\circ}^n$ and $f = \breve{f}^n = \sum_{j\in[\![-n,n]\!]} \widehat{[g]}_j^+ \phi_j e_j$) follows that

$$\begin{aligned}\widehat{w}_m \mathbb{1}_{\Omega_{ml}} &\leqslant \exp\big(\kappa n\{\|\widehat{\theta^\circ}^m\|_{L^2}^2 - \|\widehat{\theta^\circ}^l\|_{L^2}^2 + \tfrac{9}{2}(12\kappa\Lambda_l^{\widehat{\Phi}}/n - 12\kappa\Lambda_m^{\widehat{\Phi}}/n)\}\big)\mathbb{1}_{\Omega_{ml}} \\ &\leqslant \exp\big(\kappa n\{\tfrac{11}{2}\|\widehat{\theta^\circ}^l - \breve{f}^l\|_{L^2}^2 - \tfrac{1}{2}\|\Pi_{\mathbb{U}_{kl}}\breve{f}^n\|_{L^2}^2 + \tfrac{9}{2}(12\kappa\Lambda_l^{\widehat{\Phi}}/n - 12\kappa\Lambda_k^{\widehat{\Phi}}/n)\}\big)\mathbb{1}_{\Omega_{kl}} \quad (L.8)\end{aligned}$$

Note that on the event $\mho_l := \Big\{1/2 \leqslant |\lambda_j\widehat{[g]}_j^+| \leqslant 3/2,\ \forall\, j \in [\![1,l]\!]\Big\}$ we have

$$\begin{aligned}&\|\Pi_{\mathbb{U}_{kl}}\breve{f}^n\|_{L^2}^2\mathbb{1}_{\mho_l} \geqslant \tfrac{1}{4}\|\Pi_{\mathbb{U}_{kl}} f\|_{L^2}^2 = \tfrac{1}{4}\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2(\mathfrak{b}_k^2(f) - \mathfrak{b}_l^2(f)) \\ &\widehat{\Phi}_{(l)}\mathbb{1}_{\mho_l} = \max\Big\{\widehat{\Phi}_j = (\widehat{[g]}_j^+)^2, j \in [\![1,l]\!]\Big\}\mathbb{1}_{\mho_l} \leqslant \tfrac{9}{4}\max\{\Lambda_j = \lambda_j^{-2}, j\in[\![1,l]\!]\} = \tfrac{9}{4}\Lambda_{(l)} \\ &\widehat{\Phi}_{(l)}\mathbb{1}_{\mho_l} \geqslant \tfrac{1}{4}\Lambda_{(l)}\end{aligned}$$

Thus on $\mho_l$ holds $\frac{1}{4}l\Lambda_{(l)}\vee(l+2) \leqslant l\widehat{\Phi}_{(l)}\vee(l+2) \leqslant \frac{9}{4}l\Lambda_{(l)}\vee(l+2)$. Since $\psi_l^\Lambda = \frac{\log(l\Lambda_{(l)}\vee(l+2))}{\log(l+2)} \geqslant 1$ for all $l\in\mathbb{N}$ hold $\frac{\log(\frac{1}{4}l\Lambda_{(l)}\vee(l+2))}{\log(l+2)} \geqslant \psi_l^\Lambda\frac{\log(3/4)}{\log 3} \geqslant \frac{3}{10}\psi_l^\Lambda$ and $\frac{\log(\frac{9}{4}l\Lambda_{(l)}\vee(l+2))}{\log(l+2)} \leqslant \psi_l^\Lambda\frac{\log(27/4)}{\log 3} \leqslant \frac{7}{4}\psi_l^\Lambda$ which together with $\Delta_l^\Lambda = l\psi_l^\Lambda\Lambda_{(l)}$ imply

$$\begin{aligned}&\tfrac{3}{10}\psi_l^\Lambda \leqslant \delta_l^{\widehat{\Phi}}\mathbb{1}_{\mho_l} \leqslant \tfrac{7}{4}\psi_l^\Lambda \\ &\tfrac{3}{40}\Delta_l^\Lambda = l\,\tfrac{3}{10}\psi_l^\Lambda\,\tfrac{1}{4}\Lambda_{(l)} \leqslant l\,\delta_l^{\widehat{\Phi}}\,\widehat{\Phi}_{(l)}\mathbb{1}_{\mho_l} = \Lambda_l^{\widehat{\Phi}}\mathbb{1}_{\mho_l} \leqslant l\,\tfrac{7}{4}\psi_l^\Lambda\,\tfrac{9}{4}\Lambda_{(l)} = \tfrac{63}{16}\Delta_l^\Lambda \quad (L.9)\end{aligned}$$

If we define $\Omega_{kl} := \{\|\widehat{\theta^\circ}^l - \breve{f}^l\|_{L^2}^2 < 12\kappa\Lambda_l^{\widehat{\Phi}}/n\}\cap\mho_l$ then the last bounds imply

$$\begin{aligned}&\widehat{w}_k\mathbb{1}\{\|\widehat{\theta^\circ}^l - \breve{f}^l\|_{L^2}^2 < 12\kappa\Lambda_l^{\widehat{\Phi}}/n\}\mathbb{1}_{\mho_l} \\ &\quad\leqslant \exp\big(\kappa n\{\tfrac{11}{2}12\kappa\Lambda_l^{\widehat{\Phi}}/n - \tfrac{1}{8}\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2(\mathfrak{b}_k^2(f) - \mathfrak{b}_l^2(f)) + \tfrac{9}{2}(12\kappa\Lambda_l^{\widehat{\Phi}}/n - 12\kappa\Lambda_k^{\widehat{\Phi}}/n)\}\big) \\ &\quad= \exp\big(\kappa n\{10 * 12\kappa\Lambda_l^{\widehat{\Phi}}/n - \tfrac{1}{8}\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2(\mathfrak{b}_k^2(f) - \mathfrak{b}_l^2(f)) - \tfrac{9}{2} * 12\kappa\Lambda_k^{\widehat{\Phi}}/n\}\big) \\ &\quad\leqslant \exp\big(\kappa n\{\tfrac{630}{16} * 12\kappa\Delta_l^\Lambda/n - \tfrac{1}{8}\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2(\mathfrak{b}_k^2(f) - \mathfrak{b}_l^2(f))\}\big)\end{aligned}$$

and hence, by exploiting that $12\kappa\Lambda_k^{\widehat{\Phi}}/n \geqslant 0$ and $\mathcal{R}_n^\diamond[l,f,\Lambda] = [\mathfrak{b}_l^2(f)\vee\Delta_l^\Lambda n^{-1}]$ follows the assertion ***(i)***, that is

$$\widehat{w}_k\mathbb{1}\{\|\widehat{\theta^\circ}^l - \breve{f}^l\|_{L^2}^2 < 12\kappa\Lambda_l^{\widehat{\Phi}}/n\}\mathbb{1}_{\mho_l} \leqslant \exp\big(\kappa n\{-\tfrac{\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{8}\mathfrak{b}_k^2(f) + [\tfrac{630}{16}*12\kappa + \tfrac{\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{8}]\mathcal{R}_n^\diamond[l,f,\Lambda])\}\big).$$

Consider secondly that $k > l$. From ***(ii)*** in corollary G.1.1 (with $\breve{f}^\bullet = \widehat{\theta^\circ}^n$ and $f = \breve{f}^n = \sum_{j\in[\![-n,n]\!]}\widehat{[g]}_j^+\phi_j e_j$) and (L.7) follows

$$\begin{aligned}\widehat{w}_k\mathbb{1}_{\Omega_{lk}} &\leqslant \exp\big(\kappa n\{\|\widehat{\theta^\circ}^m\|_{L^2}^2 - \|\widehat{\theta^\circ}^l\|_{L^2}^2 + \tfrac{9}{2}(12\kappa\Lambda_l^{\widehat{\Phi}}/n - 12\kappa\Lambda_m^{\widehat{\Phi}}/n)\}\big)\mathbb{1}_{\Omega_{ml}} \\ &\leqslant \exp\big(\kappa n\{\tfrac{7}{2}\|\widehat{\theta^\circ}^k - \breve{f}^k\|_{L^2}^2 + \tfrac{3}{2}\|\Pi_{\mathbb{U}_{lk}}\breve{f}^n\|_{L^2}^2 + \tfrac{9}{2}(12\kappa\Lambda_l^{\widehat{\Phi}}/n - 12\kappa\Lambda_k^{\widehat{\Phi}}/n)\}\big)\mathbb{1}_{\Omega_{lk}} \quad (L.10)\end{aligned}$$

Keep in mind that $\|\Pi_{\mathbb{U}_{lk}}\breve{f}^n\|_{L^2}^2\mathbb{1}\mho_l = 2\sum_{j=l+1}^{k}(\lambda_j\widehat{[g]}_j^+)^2|\theta_j^\circ|^2 \leqslant 2\sum_{j=l+1}^{n}(\lambda_j\widehat{[g]}_j^+)^2|\theta_j^\circ|^2 = \|\Pi_{\mathbb{U}_l}^\perp\breve{f}^n\|_{L^2}^2$. If we set $\Omega_{lk} := \{\|\widehat{\theta^\circ}^k - \breve{f}^k\|_{L^2}^2 < 12\kappa\Lambda_k^{\widehat{\Phi}}/n\}$ then we clearly have *(ii)*, that is

$$\widehat{w}_k\mathbb{1}_{\{\|\widehat{\theta^\circ}^k - \breve{f}^k\|_{L^2}^2 < 12\kappa\Lambda_k^{\widehat{\Phi}}/n\}} \leqslant \exp\big(\kappa n\{\tfrac{7}{2}*12\kappa\Lambda_k^{\widehat{\Phi}}/n+\tfrac{3}{2}\|\Pi_{\mathbb{U}_l}^\perp\breve{f}^n\|_{L^2}^2+\tfrac{9}{2}(12\kappa\Lambda_l^{\widehat{\Phi}}/n-12\kappa\Lambda_k^{\widehat{\Phi}}/n)\}\big)$$
$$= \exp\big(\kappa n\{-12\kappa\Lambda_k^{\widehat{\Phi}}/n + [\tfrac{3}{2}\|\Pi_{\mathbb{U}_l}^\perp\breve{f}^n\|_{L^2}^2 + \tfrac{9}{2}12\kappa\Lambda_l^{\widehat{\Phi}}/n]\}\big)$$

which completes the proof.

Lemma L.2.3. *Given $m_+^\diamond, m_-^\diamond \in [\![1,n]\!]$ let m_+ and m_- as in (??). Let $\Lambda_m^{\widehat{\Phi}} = \delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)}$, $\sqrt{\delta_m^{\widehat{\Phi}}} = \frac{\log(m\widehat{\Phi}_{(m)}\vee(m+2))}{\log(m+2)} \geqslant 1$, $\kappa \geqslant 1$ and $\kappa \geqslant 1$. If $\mathcal{R}_n^\diamond[m,f,\Lambda] = [\mathfrak{b}_m^2(f)\vee\Lambda_m^{|\Phi}/n]$ for any $m \in [\![1,n]\!]$ and $\mho_{m_-^\diamond} := \{1/2 \leqslant |\lambda_j\widehat{[g]}_j^+| \leqslant 3/2,\ \forall\, j \in [\![1,m_-^\diamond]\!]\}$, then*

(i) $$\mathbb{P}_{\widehat{w}}([\![1,m_-[\![) \leqslant [m_- -1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond[m_-^\diamond,f,\Lambda] - \frac{\kappa\|\Pi_{\mathbb{U}_0^\perp}\theta\|_{L^2}^2}{16}n\mathfrak{b}_{[m_--1]}^2(f)\big) + \mathbb{1}_{\{\|\widehat{\theta^\circ}^{m_-^\diamond} - \breve{f}^{m_-^\diamond}\|_{L^2}^2 \geqslant 12\kappa\Lambda_{m_-^\diamond}^{\widehat{\Phi}}/n\}\cap\mho_{m_-^\diamond}} + \mathbb{1}\mho_{m_-^\diamond}^c;$$

(ii) $\sum_{m=m_++1}^n \Lambda_m^{\widehat{\Phi}}\widehat{w}_m\mathbb{1}_{\{\|\widehat{\theta^\circ}^k - \breve{f}^k\|_{L^2}^2 < 12\kappa\Lambda_m^{\widehat{\Phi}}/n\}} \leqslant \frac{1}{36\kappa^2\kappa^2} + \frac{1}{3\kappa\kappa}$.

Proof of Lemma L.2.3.
Consider *(i)*. From Lemma G.1.6 *(i)* with $l = m_-^\diamond$ follows for all $m < m_- \leqslant m_-^\diamond$, and hence $\mathfrak{b}_m \geqslant \mathfrak{b}_{m_--1}$ that

$$\widehat{w}_k\mathbb{1}\Big\{\|\widehat{\theta^\circ}^l - \breve{f}^l\|_{L^2}^2 < 12\kappa\Lambda_l^{\widehat{\Phi}}/n\Big\}\cap\mho_l$$
$$\leqslant \exp\big(\kappa n\{-\frac{\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2}{8}\mathfrak{b}_k^2(f) + [\tfrac{630}{16}*12\kappa + \frac{\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2}{8}]\mathcal{R}_n^\diamond[l,f,\Lambda]\}\big)$$
$$= \exp\big(\kappa n\{\underbrace{-\frac{\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2}{16}\mathfrak{b}_m^2(f) + [\tfrac{630*3+4}{4}\kappa + \frac{2\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2}{16}]\mathcal{R}_n^\diamond[m_-^\diamond,f,\Lambda]}_{\leqslant 0}\}\big)$$
$$\times\exp\big(\kappa n\{-\kappa\mathcal{R}_n^\diamond[m_-^\diamond,f,\Lambda] - \frac{\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2}{16}\mathfrak{b}_m^2(f)\}\big)$$
$$\leqslant \exp\big(-\kappa n\kappa\mathcal{R}_n^\diamond[m_-^\diamond,f,\Lambda] - \kappa n\frac{\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2}{16}\mathfrak{b}_{m_-^\diamond-1}^2(f)\big)$$

which in turn implies *(i)*, that is,

$$\mathbb{P}_{\widehat{w}}([\![1,m_-[\![) \leqslant \mathbb{P}_{\widehat{w}}([\![1,m_-[\![)\mathbb{1}\Big\{\|\widehat{\theta^\circ}^{m_-^\diamond} - \breve{f}^{m_-^\diamond}\|_{L^2}^2 < 12\kappa\Lambda_{m_-^\diamond}^{\widehat{\Phi}}/n\Big\}\cap\mho_{m_-^\diamond} + \mathbb{1}_{\{\|\widehat{\theta^\circ}^{m_-^\diamond} - \breve{\theta}_{m_-^\diamond}\|_{\mathbb{H}}^2 < 12\kappa\Lambda_{m_-^\diamond}^{\widehat{\Phi}}/n\}^c\cup\mho_{m_-^\diamond}^c}$$
$$\leqslant [m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond[m_-^\diamond,f,\Lambda] - \kappa n\frac{\|\Pi_{\mathbb{U}_0^\perp}f\|_{L^2}^2}{16}\mathfrak{b}_{[m_--1]}^2(f)\big)$$
$$+\mathbb{1}_{\{\|\widehat{\theta^\circ}^{m_-^\diamond} - \breve{f}^{m_-^\diamond}\|_{L^2}^2 \geqslant 12\kappa\Lambda_{m_-^\diamond}^{\widehat{\Phi}}/n\}\cap\mho_{m_-^\diamond}} + \mathbb{1}\mho_{m_-^\diamond}^c$$

Consider *(ii)*. From Lemma L.2.2 *(ii)* with $l = m_+^\diamond$ follows for all $m > m_+ \geqslant m_+^\diamond$

$$\begin{aligned}\widehat{w}_m \mathbb{1}\{\|\widehat{\theta^\circ}^k - \check{f}^k\|_{L^2}^2 < 12\kappa\Lambda_m^{\widehat{\Phi}}/n\} &\leqslant \exp\big(\kappa n\{-12\kappa\Lambda_m^{\widehat{\Phi}}/n + [\tfrac{3}{2}\|\Pi_{\mathbb{U}_l}^\perp \check{f}^n\|_{L^2}^2 + \tfrac{9}{2} * 12\kappa\Lambda_l^{\widehat{\Phi}}/n]\}\big) \\ &= \exp\big(\kappa n\{-\tfrac{1}{2}12\kappa\Lambda_m^{\widehat{\Phi}}/n \underbrace{-\tfrac{1}{2} * 12\kappa\Lambda_m^{\widehat{\Phi}}/n + [\tfrac{3}{2}\|\Pi_{\mathbb{U}_l}^\perp \check{f}^n\|_{L^2}^2 + \tfrac{9}{2} * 12\kappa\Lambda_l^{\widehat{\Phi}}/n]}_{\leqslant 0}\}\big) \\ &\leqslant \exp\big(-\kappa 6\kappa\Lambda_m^{\widehat{\Phi}}\big).\end{aligned}$$

Note that $|\widehat{[g]}_j|^2 \leqslant 1$ for all $j \in \mathbb{Z}$, hence if $|\widehat{[g]}_j|^2 \geqslant 1/q$ then $\widehat{\Phi}_j = |\widehat{[g]}_j^+|^2 \geqslant 1$. Thereby, $\widehat{\Phi}_j = |\widehat{[g]}_j^+|^2 < 1$ implies $|\widehat{[g]}_j|^2 < 1/q$ and hence $\widehat{\Phi}_j = |\widehat{[g]}_j^+|^2 = 0$. Thereby $1 > \widehat{\Phi}_{(m)} = \max\{|\widehat{[g]}_j^+|^2, j \in [\![1, m]\!]\}$ implies $\widehat{\Phi}_{(m)} = 0$, that is,

$$\{\widehat{\Phi}_{(m)} < 1\} = \{\widehat{\Phi}_{(m)} = 0\}. \tag{L.11}$$

Consequently, with $\Lambda_m^{\widehat{\Phi}} = \delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)}$ follows

$$\begin{aligned}\sum_{m=m_++1}^n \Lambda_m^{\widehat{\Phi}}\widehat{w}_m \mathbb{1}_{\{\|\widehat{\theta^\circ}^k - \check{f}^k\|_{L^2}^2 \leqslant 12\kappa\Lambda_m^{\widehat{\Phi}}/n\}} &\leqslant \sum_{k=m_++1}^n \Lambda_m^{\widehat{\Phi}} \exp\big(-\kappa 6\kappa\Lambda_m^{\widehat{\Phi}}\big) \\ = \sum_{k=m_++1}^n \delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)} \exp\big(-6\kappa\kappa\delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)}\big)\mathbb{1}_{\{\widehat{\Phi}_{(m)} \geqslant 1\}} \\ + \sum_{k=m_++1}^n \delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)} \exp\big(-6\kappa\kappa\delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)}\big)\mathbb{1}_{\{\widehat{\Phi}_{(m)} < 1\}} \\ = \sum_{k=m_++1}^n \delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)} \exp\big(-6\kappa\kappa\delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)}\big)\mathbb{1}_{\{\widehat{\Phi}_{(m)} \geqslant 1\}}.\end{aligned} \tag{L.12}$$

Exploiting that $\sqrt{\delta_m^{\widehat{\Phi}}} = \frac{\log(m\widehat{\Phi}_{(m)} \vee (m+2))}{\log(m+2)} \geqslant 1$ and $6\kappa \geqslant 2\log(3e) \approx 4.2$ let $\kappa \geqslant 1$, then for all $k \in \mathbb{N}$ we have $6\kappa\kappa k - \log(k+2) \geqslant 1$ and thus exploiting $a\exp(-ab) \leqslant \exp(-b)$ for all $a, b \geqslant 1$ we have

$$\begin{aligned}\delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)} \exp\big(-6\kappa\kappa\delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)}\big)\mathbb{1}_{\{\widehat{\Phi}_{(m)} \geqslant 1\}} &\leqslant \delta_m^{\widehat{\Phi}} \exp\big(-6\kappa\kappa\delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)} + \delta_m^{\widehat{\Phi}}\log(m+2)\big)\mathbb{1}_{\{\widehat{\Phi}_{(m)} \geqslant 1\}} \\ \leqslant \delta_m^{\widehat{\Phi}} \exp\big(-\delta_m^{\widehat{\Phi}}(6\kappa\kappa m - \log(m+2))\big) &\leqslant \exp\big(-(6\kappa\kappa m - \log(m+2))\big) \\ &= (m+2)\exp\big(-6\kappa\kappa m\big)\end{aligned}$$

Since $\sum_{m\in\mathbb{N}} \mu m \exp(-\mu m) \leqslant (\mu)^{-1}$ und $\sum_{m\in\mathbb{N}} \mu\exp(-\mu m) \leqslant 1$ follows

$$\begin{aligned}\sum_{k=m_++1}^n \delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)} \exp\big(-6\kappa\kappa\delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)}\big)\mathbb{1}_{\{\widehat{\Phi}_{(m)} \geqslant 1\}} &\leqslant \sum_{k=m_++1}^\infty (m+2)\exp\big(-6\kappa\kappa m\big) \\ &\leqslant \tfrac{1}{36\kappa^2\kappa^2} + \tfrac{1}{3\kappa\kappa}.\end{aligned} \tag{L.13}$$

Thereby, combining (L.12) and (L.13) we obtain ***(ii)***, that is

$$\sum_{m=m_{+}+1}^{n} \Lambda_m^{\widehat{\Phi}} \widehat{w}_m \mathbb{1}_{\{\|\widehat{\theta^\circ}^k - \check{f}^k\|_{L^2}^2 \leqslant 12\kappa\Lambda_m^{\widehat{\Phi}}/n\}} \leqslant \tfrac{1}{36\kappa^2\kappa^2} + \tfrac{1}{3\kappa\kappa},$$

which completes the proof.

Lemma L.2.4. *Consider $\widehat{\theta^\circ}^m - \check{f}^m = \sum_{j\in[\![-m,m]\!]} \widehat{[g]}_j^+ (\widehat{[h]}_j - \phi_j)e_j$. Conditionally on $\varepsilon_1,\dots,\varepsilon_q$ the r.v.'s $Y_1,\dots,Y_n$ are iid. and we denote by $\mathbb{P}^n_{Y|\varepsilon}$ and $\mathbb{E}^n_{Y|\varepsilon}$ their conditional distribution and expectation, respectively. Let $\widehat{\Phi}_j = |\widehat{[g]}_j^+|^2$, $\overline{\widehat{\Phi}}_m = \frac{1}{m}\sum_{j\in[\![1,m]\!]}\widehat{\Phi}_j$, $\widehat{\Phi}_{(m)} = \max\{\widehat{\Phi}_j, j\in[\![1,m]\!]\}$, $\kappa \geqslant 1$, $\Lambda_m^{\widehat{\Phi}} = \delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)}$ and $\sqrt{\delta_m^{\widehat{\Phi}}} = \frac{\log(m\widehat{\Phi}_{(m)}\vee(m+2))}{\log(m+2)} \geqslant 1$. Then there is a numerical constant $\mathcal{C}$ such that for all $n\in\mathbb{N}$ and $m\in[\![1,n[\![$ holds*

(i) *if $m_+ \geqslant 3(\frac{12\|\phi\|_{\ell^1}}{\kappa})^2$ and $n \geqslant 15(\frac{200}{\sqrt{\kappa}})^4$ (alternatively $\sqrt{\kappa}/200 \geqslant \sqrt{3}$) then* $\sum_{m=1+m_+}^{n} \mathbb{E}^n_{Y|\varepsilon}\big(\|\widehat{\theta^\circ}^m - \check{f}^m\|_{L^2}^2 - 12\kappa\Lambda_m^{\widehat{\Phi}}/n\big)_+ \leqslant \mathcal{C}\big[\frac{12\|\phi\|_{\ell^1}^2}{\kappa} + 4\big]n^{-1}$

(ii) *if $m_+ \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ and $n \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ (alternatively $\sqrt{\kappa}/300 \geqslant \sqrt{3}$) then* $\sum_{m=1+m_+}^{n} \Delta_m^{\Lambda}\mathbb{P}^n_{Y|\varepsilon}\big(\|\widehat{\theta^\circ}^m - \check{f}^m\|_{L^2}^2 \geqslant 12\kappa\Lambda_m^{\widehat{\Phi}}/n\big) \leqslant 3\big[(\frac{400\|\phi\|_{\ell^1}}{\kappa})^2 + \frac{800\|\phi\|_{\ell^1}}{\kappa} + 1\big].$

Proof of Lemma L.2.4.
Since $\delta_m^{\widehat{\Phi}} \geqslant 1$ for $m \geqslant 3(\frac{12\|\phi\|_{\ell^1}}{\kappa})^2$ holds $\frac{\kappa\sqrt{\psi_m^{\Lambda}m}}{12\|\phi\|_{\ell^1}} - \log(m+2) \geqslant 0$ and hence

$$\begin{aligned}\widehat{\Phi}_{(m)}\exp\big(\tfrac{-\kappa\delta_m^{\widehat{\Phi}}m}{6\|\phi\|_{\ell^1}}\big) &\leqslant \exp\big(\tfrac{-\kappa\delta_m^{\widehat{\Phi}}m}{12\|\phi\|_{\ell^1}}\big)\exp\big(-\sqrt{\delta_m^{\widehat{\Phi}}}[\tfrac{\kappa\sqrt{\delta_m^{\widehat{\Phi}}m}}{12\|\phi\|_{\ell^1}} - \log(m+2)]\big)\\ &\leqslant \exp\big(\tfrac{-\kappa\delta_m^{\widehat{\Phi}}m}{12\|\phi\|_{\ell^1}}\big) \leqslant \exp\big(-\tfrac{\kappa}{12\|\phi\|_{\ell^1}}m\big)\end{aligned}$$

consequently, if $m_+ \geqslant 3(\frac{12\|\phi\|_{\ell^1}}{\kappa})^2$ then

$$\sum_{m=1+m_+}^{n} \widehat{\Phi}_{(m)}\exp\big(\tfrac{-\kappa\delta_m^{\widehat{\Phi}}m}{6\|\phi\|_{\ell^1}}\big) \leqslant \sum_{m=1+m_+}^{n}\exp\big(-\tfrac{\kappa}{12\|\phi\|_{\ell^1}}m\big) \leqslant \frac{12\|\phi\|_{\ell^1}}{\kappa}$$

For all $n \geqslant 15(\frac{200}{\sqrt{\kappa}})^4$ holds $\sqrt{n} \geqslant \frac{200}{\sqrt{\kappa}}\log(n+2)$ (or using $\sqrt{3x} \geqslant \log(x+2)$ for all $x \geqslant 1$ and $\sqrt{\kappa}/200 \geqslant \sqrt{3}$) thereby

$$\frac{m^2\widehat{\Phi}_{(m)}^2}{n}\exp\big(\tfrac{-\sqrt{n\kappa\delta_m^{\widehat{\Phi}}}}{100}\big) \leqslant \tfrac{1}{n}\exp\big(-2\sqrt{\delta_m^{\widehat{\Phi}}}[\tfrac{\sqrt{n\kappa}}{200} - \log(m+2)]\big) \leqslant \tfrac{1}{n}$$

consequently,

$$\sum_{m=1+m_+}^{n} \frac{m^2\widehat{\Phi}_{(m)}^2}{n}\exp\big(\tfrac{-\sqrt{n\kappa\delta_m^{\widehat{\Phi}}}}{100}\big) \leqslant \sum_{m=1+m_+}^{n}\tfrac{1}{n} \leqslant 1$$

Combining the last two bounds and **??** we obtain *(i)*, that is

$$\sum_{m=1+m_+}^{n} \mathbb{E}^n_{Y|\varepsilon}\big(\|\widehat{\theta^\circ}^m - \breve{f}^m\|^2_{L^2} - 12\kappa\Lambda^{\widehat{\Phi}}_m/n\big)_+$$
$$\leqslant \mathcal{C}\Big[\frac{\|\phi\|_{\ell^1}}{n}\sum_{m=1+m_+}^{n}\widehat{\Phi}_{(m)}\exp\big(\tfrac{-\kappa\delta^{\widehat{\Phi}}_m m}{6\|\phi\|_{\ell^1}}\big) + \frac{4}{n}\sum_{m=1+m_+}^{n}\frac{m^2\widehat{\Phi}^2_{(m)}}{n}\exp\big(\tfrac{-\sqrt{n\kappa\delta^{\widehat{\Phi}}_m}}{100}\big)\Big]$$
$$\leqslant \mathcal{C}n^{-1}\big[\tfrac{12\|\phi\|^2_{\ell^1}}{\kappa} + 4\big]$$

If $m \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ then $m \geqslant (\frac{800\|\phi\|_{\ell^1}}{\kappa})\log(m+2)$ and hence $m - \frac{400\|\phi\|_{\ell^1}}{\kappa}\log(m+2) \geqslant \frac{400\|\phi\|_{\ell^1}}{\kappa}\log(m+2)$ or equivalently, $\frac{\kappa}{400\|\phi\|_{\ell^1}}m - \log(m+2) \geqslant \log(m+2) \geqslant 1$ and thus

$$m\delta^{\widehat{\Phi}}_m\widehat{\Phi}_{(m)}\exp\big(\tfrac{-\kappa\delta^{\widehat{\Phi}}_m m}{400\|\phi\|_{\ell^1}}\big) \leqslant \delta^{\widehat{\Phi}}_m\exp\big(-\delta^{\widehat{\Phi}}_m\big[\tfrac{\kappa}{400\|\phi\|_{\ell^1}}m - \log(m+2)\big]\big)$$
$$\leqslant (m+2)\exp\big(-\tfrac{\kappa}{400\|\phi\|_{\ell^1}}m\big)$$

consequently, if $m_+ \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ exploiting $\sum_{m\in\mathbb{N}}(m+2)\exp(-\mu m) \leqslant \mu^{-2} + 2\mu^{-1}$ follows

$$\sum_{m=1+m_+}^{n} m\delta^{\widehat{\Phi}}_m\widehat{\Phi}_{(m)}\exp\big(\tfrac{-\kappa\delta^{\widehat{\Phi}}_m m}{400\|\phi\|_{\ell^1}}\big) \leqslant \sum_{m=1+m_+}^{n}(k+2)\exp\big(-\tfrac{\kappa}{400\|\phi\|_{\ell^1}}m\big)$$
$$\leqslant (\tfrac{400\|\phi\|_{\ell^1}}{\kappa})^2 + \tfrac{800\|\phi\|_{\ell^1}}{\kappa}$$

Keep in mind that $\{\widehat{\Phi}_{(m)} < 1\} = \{\widehat{\Phi}_{(m)} = 0\}$ (cf. (L.11) in the proof of Lemma L.2.3). Since $\log(m\widehat{\Phi}_{(m)})\mathbb{1}_{\{\widehat{\Phi}_{(m)}\geqslant 1\}} \leqslant (1/e)m\widehat{\Phi}_{(m)}\mathbb{1}_{\{\widehat{\Phi}_{(m)}\geqslant 1\}}$ follows $\delta^{\widehat{\Phi}}_m\mathbb{1}_{\{\widehat{\Phi}_{(m)}\geqslant 1\}} \leqslant m\widehat{\Phi}_{(m)}\mathbb{1}_{\{\widehat{\Phi}_{(m)}\geqslant 1\}}$, and using for all $n \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ holds $\sqrt{n} \geqslant \frac{300}{\sqrt{\kappa}}\log(n+2)$ (alternatively: $\sqrt{3x} \geqslant \log(x+2)$ for all $x \geqslant 1$ and $\sqrt{\kappa}/300 \geqslant \sqrt{3}$)

$$m\delta^{\widehat{\Phi}}_m\widehat{\Phi}_{(m)}\exp\big(\tfrac{-\sqrt{n\kappa\delta^{\widehat{\Phi}}_m}}{100}\big)\mathbb{1}_{\{\widehat{\Phi}_{(m)}\geqslant 1\}} \leqslant m^2\widehat{\Phi}^2_{(m)}\exp\big(\tfrac{-\sqrt{n\kappa\delta^{\widehat{\Phi}}_m}}{100}\big)\mathbb{1}_{\{\widehat{\Phi}_{(m)}\geqslant 1\}}$$
$$\leqslant \tfrac{1}{n}\exp\Big(-\sqrt{\psi^{\Lambda}_m}\big[\tfrac{\sqrt{n\kappa}}{100} - 2\log(m+2)\big] + \log(n+2)\Big)$$
$$\leqslant \tfrac{1}{n}\exp\Big(-3\sqrt{\delta^{\widehat{\Phi}}_m}\big[\tfrac{\sqrt{n\kappa}}{300} - \log(n+2)\big]\Big) \leqslant \tfrac{1}{n}$$

consequently,

$$\sum_{m=1+m_+}^{n} m\delta^{\widehat{\Phi}}_m\widehat{\Phi}_{(m)}\exp\big(\tfrac{-\sqrt{n\kappa\delta^{\widehat{\Phi}}_m}}{100}\big) = \sum_{m=1+m_+}^{n} m\delta^{\widehat{\Phi}}_m\widehat{\Phi}_{(m)}\exp\big(\tfrac{-\sqrt{n\kappa\delta^{\widehat{\Phi}}_m}}{100}\big)\mathbb{1}_{\{\widehat{\Phi}_{(m)}\geqslant 1\}}$$
$$\leqslant \sum_{m=1+m_+}^{n}\tfrac{1}{n} \leqslant 1$$

Combining the last two bounds and **??** **??** we obtain *(ii)*, that is

$$\sum_{m=1+m_+}^{n} \delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)} \mathbb{P}^n_{Y|\varepsilon}\big(\|\widehat{\theta^{\circ}}^m - \check{f}^m\|^2_{L^2} \geqslant 12\kappa\Lambda_m^{\widehat{\Phi}}/n\big)$$
$$\leqslant 3 \sum_{m=1+m_+}^{n} \delta_m^{\widehat{\Phi}} m\widehat{\Phi}_{(m)} \Big[\exp\big(\tfrac{-\kappa\delta_m^{\widehat{\Phi}} m}{400\|\phi\|_{\ell^1}}\big) + \exp\big(\tfrac{-\sqrt{n\kappa\delta_m^{\widehat{\Phi}}}}{100}\big)\Big]$$
$$\leqslant 3\Big[\big(\frac{400\|\phi\|_{\ell^1}}{\kappa}\big)^2 + \frac{800\|\phi\|_{\ell^1}}{\kappa} + 1\Big]$$

which implies the result and completes the proof.

Proposition L.2.1. *Let $\kappa \geqslant 1$, $\kappa \geqslant 1$, $\Delta_m^{\Lambda} = \psi_m^{\Lambda} m\Lambda_{(m)}$ with $\sqrt{\psi_m^{\Lambda}} = \frac{\log(m\Lambda_{(m)} \vee (m+2))}{\log(m+2)} \geqslant 1$, $\mathcal{R}_n^{\diamond}[l,f,\Lambda] = [\mathfrak{b}_l^2(f) \vee \Delta_l^{\Lambda}\, n^{-1}]$ and $\mathcal{R}_q^{\star}(f,\Lambda) := \sum_{j\in\mathbb{N}} |\theta_j^{\circ}|^2[1\wedge q^{-1}\Lambda_j]$. Given $m_+^{\diamond}, m_-^{\diamond} \in [\![1,n]\!]$ let m_- as in (**??**). If $m_+^{\diamond} \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ and $n \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ then there is an universal numerical constant $\mathcal{C}$ such that for all $q \in \mathbb{N}$*

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^{\circ}} - f\|^2_{L^2} \leqslant \mathcal{C}\big\{[1 \vee \|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}]\mathcal{R}_n^{\diamond}[m_+^{\diamond}, f, \Lambda] + \mathcal{R}_q^{\star}(f,\Lambda) + \|\phi\|^2_{\ell^1} n^{-1}\big\}$$
$$+ 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + 9\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2} \exp\big(\frac{-3\kappa\psi^{\Lambda}_{m_-^{\diamond}} m_-^{\diamond}}{4000\|\phi\|_{\ell^1}}\big)$$
$$+ 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}_n^{\diamond}[m_-^{\diamond}, f, \Lambda] - \frac{\kappa\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}}{16} n\mathfrak{b}^2_{[m_- - 1]}(f)\big)$$
$$+ 3q\mathbb{P}^q_{\varepsilon}(\mho^c_{m_+^{\diamond}}) + 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\mathbb{P}^q_{\varepsilon}(\mho^c_{m_-^{\diamond}})$$

Proof of Proposition L.2.1.

Given $m_+^{\diamond}, m_-^{\diamond} \in [\![1,n]\!]$ let m_+ and m_- as in (**??**). Keep in mind that m_- is not random and m_+ does depend on the error sample only. From Lemma L.2.1 together with Lemma L.2.3 follows

$$\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta^{\circ}} - f\|^2_{L^2} \leqslant 3\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta^{\circ}}^{m_+} - \check{f}^{m_+}\|^2_{L^2} + 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + [\tfrac{1}{12\kappa\kappa^2} + \tfrac{1}{\kappa}]n^{-1}$$
$$+ 3\sum_{l\in\rrbracket m_+,n\rrbracket} \mathbb{E}^n_{Y|\varepsilon}\big(\|\widehat{\theta^{\circ}}^l - \check{f}^l\|^2_{L^2} - \mathcal{C}_1 12\kappa\Lambda_l^{\widehat{\Phi}}/n\big)_+$$
$$+ 3(\mathcal{C}_1\kappa/n)\sum_{l\in\rrbracket m_+,n\rrbracket} \Lambda_l^{\widehat{\Phi}}\mathbb{P}^n_{Y|\varepsilon}\Big(\|\widehat{\theta^{\circ}}^l - \check{f}^l\|^2_{L^2} \geqslant 12\kappa\Lambda_l^{\widehat{\Phi}}/n\Big)$$
$$+ 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\{[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}_n^{\diamond}[m_-^{\diamond}, f, \Lambda] - \frac{\kappa\|\Pi_{\mathbb{U}_0^{\perp}} \theta\|^2_{L^2}}{16} n\mathfrak{b}^2_{[m_- - 1]}(f)\big)$$
$$+ \mathbb{P}^n_{Y|\varepsilon}\Big(\|\widehat{\theta^{\circ}}^{m_-^{\diamond}} - \check{f}^{m_-^{\diamond}}\|^2_{L^2} \geqslant 12\kappa\Lambda^{\widehat{\Phi}}_{m_-^{\diamond}}/n\Big)\mathbb{1}_{\mho_{m_-^{\diamond}}} + \mathbb{1}_{\mho^c_{m_-^{\diamond}}}\}$$
$$+ 6\sum_{j\in[\![1,n]\!]} |\widehat{[g]}_j^{+}|^2|\lambda_j - \widehat{[g]}_j|^2|\theta_j^{\circ}|^2 + 2\sum_{j\in[\![1,n]\!]} \mathbb{1}_{\mathcal{X}_j^c}|\theta_j^{\circ}|^2 \qquad \text{(L.14)}$$

Since $m_+ \geqslant m_+^{\diamond} \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ and $n \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ due to Lemma L.2.4 *(i)* and *(ii)* there is a finite numerical constant $\mathcal{C}$ such that

$$\begin{aligned}\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant 3\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta}^{\circ}{}^{m_+} - \check{f}^{m_+}\|^2_{L^2} + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + [\tfrac{1}{12\kappa\kappa^2} + \tfrac{1}{\kappa}]n^{-1} \\ + \mathcal{C}\big[\tfrac{26\|\phi\|^2_{\ell^1}}{\kappa} + 12\big]n^{-1} + \mathcal{C}_1\big[\tfrac{9*400^2\|\phi\|^2_{\ell^1}}{\kappa} + 9*800\|\phi\|_{\ell^1} + 9\kappa\big]n^{-1} \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^{\diamond}_n[m^{\diamond}_-, f, \Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp}\theta\|^2_{L^2}}{16}n\mathfrak{b}^2_{[m_- - 1]}(f)\big) \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathbb{P}^n_{Y|\varepsilon}\Big(\|\widehat{\theta}^{\circ}{}^{m^{\diamond}_-} - \check{f}^{m^{\diamond}_-}\|^2_{L^2} \geqslant 12\kappa\Lambda^{\widehat{\Phi}}_{m^{\diamond}_-}/n\Big)\mathbb{1}\mho_{m^{\diamond}_-} + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathbb{1}\mho^c_{m^{\diamond}_-} \\ + 6\sum_{j\in[\![1,n]\!]}|\widehat{[g]}^+_j|^2|\lambda_j - \widehat{[g]}_j|^2|\theta^{\circ}_j|^2 + 2\sum_{j\in[\![1,n]\!]}\mathbb{1}\mathcal{X}^c_j|\theta^{\circ}_j|^2. \quad \text{(L.15)}\end{aligned}$$

and together with **??** **??** we obtain

$$\begin{aligned}\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant 3\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta}^{\circ}{}^{m_+} - \check{f}^{m_+}\|^2_{L^2} + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + [\tfrac{1}{12\kappa\kappa^2} + \tfrac{1}{\kappa}]n^{-1} \\ + \mathcal{C}\big[\tfrac{26\|\phi\|^2_{\ell^1}}{\kappa} + 12\big]n^{-1} + \mathcal{C}_1\big[\tfrac{9*400^2\|\phi\|^2_{\ell^1}}{\kappa} + 9*800\|\phi\|_{\ell^1} + 9\kappa\big]n^{-1} \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^{\diamond}_n[m^{\diamond}_-, f, \Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp}\theta\|^2_{L^2}}{16}n\mathfrak{b}^2_{[m_- - 1]}(f)\big) \\ + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\left[\exp\Big(\frac{-\kappa\delta^{\widehat{\Phi}}_{m^{\diamond}_-}m^{\diamond}_-}{400\|\phi\|_{\ell^1}}\Big) + \exp\Big(\frac{-\sqrt{n\kappa\delta^{\widehat{\Phi}}_{m^{\diamond}_-}}}{100}\Big)\right]\mathbb{1}\mho_{m^{\diamond}_-} \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathbb{1}\mho^c_{m^{\diamond}_-} + 6\sum_{j\in[\![1,n]\!]}|\widehat{[g]}^+_j|^2|\lambda_j - \widehat{[g]}_j|^2|\theta^{\circ}_j|^2 + 2\sum_{j\in[\![1,n]\!]}\mathbb{1}\mathcal{X}^c_j|\theta^{\circ}_j|^2.\end{aligned}$$

Moreover, for $n > n_{f,\Lambda} \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$ holds $\sqrt{n} \geqslant \frac{300}{\sqrt{\kappa}}\log(n+2) \geqslant \frac{100}{\sqrt{\kappa}}\log(n+2)$ which in turn together with $\delta^{\widehat{\Phi}}_{m_-} \geqslant 1$ implies $n\exp\big(-\sqrt{n}\frac{\sqrt{\kappa\delta^{\widehat{\Phi}}_{m^{\diamond}_-}}}{100}\big) \leqslant \exp\big(-\sqrt{n}\frac{\sqrt{\kappa}}{100} + \log(n+2)\big) \leqslant 1$, and thus

$$\begin{aligned}\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant 3\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta}^{\circ}{}^{m_+} - \check{f}^{m_+}\|^2_{L^2} + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + [\tfrac{1}{12\kappa\kappa^2} + \tfrac{1}{\kappa}]n^{-1} \\ + \mathcal{C}\big[\tfrac{26\|\phi\|^2_{\ell^1}}{\kappa} + 12\big]n^{-1} + \mathcal{C}_1\big[\tfrac{9*400^2\|\phi\|^2_{\ell^1}}{\kappa} + 9*800\|\phi\|_{\ell^1} + 9\kappa\big]n^{-1} \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^{\diamond}_n[m^{\diamond}_-, f, \Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp}\theta\|^2_{L^2}}{16}n\mathfrak{b}^2_{[m_- - 1]}(f)\big) \\ + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\left[\exp\Big(\frac{-\kappa\delta^{\widehat{\Phi}}_{m^{\diamond}_-}m^{\diamond}_-}{400\|\phi\|_{\ell^1}}\Big) + n^{-1}\right]\mathbb{1}\mho_{m^{\diamond}_-} \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathbb{1}\mho^c_{m^{\diamond}_-} + 6\sum_{j\in[\![1,n]\!]}|\widehat{[g]}^+_j|^2|\lambda_j - \widehat{[g]}_j|^2|\theta^{\circ}_j|^2 + 2\sum_{j\in[\![1,n]\!]}\mathbb{1}\mathcal{X}^c_j|\theta^{\circ}_j|^2. \quad \text{(L.16)}\end{aligned}$$

As shown in (L.9) (cf. proof of Lemma L.2.2) holds $\frac{3}{10}\psi^{\Lambda}_{m^{\diamond}_{-}} \leqslant \delta^{\widehat{\Phi}}_{m^{\diamond}_{-}} \mathbb{1}\mho_{m^{\diamond}_{-}}$ and thus

$$\begin{aligned}\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} &\leqslant 3\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta}^{\circ}{}^{m_+} - \check{f}^{m_+}\|^2_{L^2} + 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + [\tfrac{1}{12\kappa\kappa^2} + \tfrac{1}{\kappa}]n^{-1} \\ &+ \mathcal{C}\big[\tfrac{26\|\phi\|^2_{\ell^1}}{\kappa} + 12\big]n^{-1} + \mathcal{C}_1\big[\tfrac{9*400^2\|\phi\|^2_{\ell^1}}{\kappa} + 9*800\|\phi\|_{\ell^1} + 9\kappa\big]n^{-1} + 9\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}n^{-1} \\ &+ 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^{\diamond}_n[m^{\diamond}_{-}, f, \Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^{\perp}}\theta\|^2_{L^2}}{16} n\mathfrak{b}^2_{[m_- -1]}(f)\big) \\ &+ 9\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\exp\big(\tfrac{-\kappa\frac{3}{10}\psi^{\Lambda}_{m^{\diamond}_{-}} m^{\diamond}_{-}}{400\|\phi\|_{\ell^1}}\big) \\ &+ 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\mathbb{1}\mho^c_{m^{\diamond}_{-}} + 6\sum_{j\in[\![1,n]\!]}|\widehat{[g]}^{+}_j|^2|\lambda_j - \widehat{[g]}_j|^2|\theta^{\circ}_j|^2 + 2\sum_{j\in[\![1,n]\!]}\mathbb{1}\mathcal{X}^c_j|\theta^{\circ}_j|^2. \quad (L.17)\end{aligned}$$

Consider $\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta}^{\circ}{}^{m_+} - \check{f}^{m_+}\|^2_{L^2} = 2\sum_{j=1}^{m_+}(\widehat{[g]}^{+}_j)^2/n = 2\sum_{j=1}^{m_+}\widehat{\Phi}_j/n = 2m_+\overline{\widehat{\Phi}_{m_+}}/n$ where $\Lambda^{\widehat{\Phi}}_{m_+} \geqslant \overline{\widehat{\Phi}_{m_+}}$ and $q \geqslant m_+\overline{\widehat{\Phi}_{m_+}}/n$ exploiting $(\widehat{[g]}^{+}_j)^2 \leqslant q$ and $m_+ \leqslant n$. Considering the event $\mho_{m^{\diamond}_{+}}$ and its complement $\mho^c_{m^{\diamond}_{+}}$ it follows $\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta}^{\circ}{}^{m_+} - \check{f}^{m_+}\|^2_{L^2} \leqslant q\mathbb{1}\mho^c_{m^{\diamond}_{+}} + \Lambda^{\widehat{\Phi}}_{m_+}/n\mathbb{1}\mho_{m^{\diamond}_{+}}$. Taking into account the definition (??) of m_+ we obtain $\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta}^{\circ}{}^{m_+} - \check{f}^{m_+}\|^2_{L^2} \leqslant q\mathbb{1}\mho^c_{m^{\diamond}_{+}} + [\frac{1}{4\kappa}\|\Pi^{\perp}_{\mathbb{U}_{m^{\diamond}_{+}}}\check{f}^n\|^2_{L^2} + 9\Lambda^{\widehat{\Phi}}_{m^{\diamond}_{+}}/n]\mathbb{1}\mho_{m^{\diamond}_{+}}$. Due to (L.9) (cf. proof of Lemma L.2.2) we have $\Lambda^{\widehat{\Phi}}_{m^{\diamond}_{+}}\mathbb{1}\mho_{m^{\diamond}_{+}} \leqslant \frac{63}{16}\Delta^{\Lambda}_{m^{\diamond}_{+}}$, which implies $\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta}^{\circ}{}^{m_+} - \check{f}^{m_+}\|^2_{L^2} \leqslant q\mathbb{1}\mho^c_{m^{\diamond}_{+}} + \frac{1}{4\kappa}\|\Pi^{\perp}_{\mathbb{U}_{m^{\diamond}_{+}}}\check{f}^n\|^2_{L^2} + \frac{9*63}{16}\Delta^{\Lambda}_{m^{\diamond}_{+}}/n$. Thereby, from (L.17) together with Lemma L.2.4 follows now

$$\begin{aligned}\mathbb{E}^n_{Y|\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} &\leqslant 3q\mathbb{1}\mho^c_{m^{\diamond}_{+}} + \tfrac{3}{4\kappa}\|\Pi^{\perp}_{\mathbb{U}_{m^{\diamond}_{+}}}\check{f}^n\|^2_{L^2} + \tfrac{3*9*63}{16}\Delta^{\Lambda}_{m^{\diamond}_{+}}/n + 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + [\tfrac{1}{12\kappa\kappa^2} + \tfrac{1}{\kappa}]n^{-1} \\ &+ \mathcal{C}\big[\tfrac{26\|\phi\|^2_{\ell^1}}{\kappa} + 12\big]n^{-1} + \mathcal{C}_1\big[\tfrac{9*400^2\|\phi\|^2_{\ell^1}}{\kappa} + 9*800\|\phi\|_{\ell^1} + 9\kappa\big]n^{-1} + 9\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}n^{-1} \\ &+ 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^{\diamond}_n[m^{\diamond}_{-}, f, \Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}}{16} n\mathfrak{b}^2_{[m_- -1]}(f)\big) \\ &+ 9\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\exp\big(\tfrac{-\kappa 3\psi^{\Lambda}_{m^{\diamond}_{-}} m^{\diamond}_{-}}{4000\|\phi\|_{\ell^1}}\big) \\ &+ 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\mathbb{1}\mho^c_{m^{\diamond}_{-}} + 6\sum_{j\in[\![1,n]\!]}|\widehat{[g]}^{+}_j|^2|\lambda_j - \widehat{[g]}_j|^2|\theta^{\circ}_j|^2 + 2\sum_{j\in[\![1,n]\!]}\mathbb{1}\mathcal{X}^c_j|\theta^{\circ}_j|^2. \quad (L.18)\end{aligned}$$

Exploiting Lemma 1.5.1 we obtain from (i) $\mathbb{E}^q_{\varepsilon}\|\Pi^{\perp}_{\mathbb{U}_{m^{\diamond}_{+}}}\check{f}^n\|^2_{L^2} \leqslant 4\sum_{|j|\in]\!]m_+,n]\!]}|\theta^{\circ}_j|^2 \leqslant 4\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\mathfrak{b}^2_{m^{\diamond}_{+}}(f)$, from (iii) $\sum_{j\in[\![1,n]\!]}|\theta^{\circ}_j|^2\mathbb{E}^q_{\varepsilon}|\widehat{[g]}^{+}_j|^2|\lambda_j - \widehat{[g]}_j|^2 \leqslant 4\mathcal{C}_4\mathcal{R}^{\star}_q(f,\Lambda)$ and from (ii) $\sum_{j\in[\![1,n]\!]}\mathbb{P}^q_{\varepsilon}(\mathcal{X}^c_j)|\theta^{\circ}_j|^2 \leqslant$

$4\mathcal{R}_q^\star(f,\Lambda)$ where $\mathcal{R}_q^\star(f,\Lambda) := \sum_{j\in\mathbb{N}} \theta_j^{\circ 2}[1\wedge\Lambda_j/q]$. The last bounds imply

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ} - f\|_{L^2}^2 &\leqslant \tfrac{3*9*63}{16}\Delta_{m_+^\diamond}^\Lambda n^{-1} + \tfrac{3}{\kappa}\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_{m_+^\diamond}^2(f) + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_{m_-}^2(f)\\ &+ \big[\tfrac{1}{12\kappa\kappa^2} + \tfrac{1}{\kappa}\big]n^{-1} + \mathcal{C}\big[\tfrac{26\|\phi\|_{\ell^1}^2}{\kappa} + 12\big]n^{-1} + \mathcal{C}_1\big[\tfrac{9*400^2\|\phi\|_{\ell^1}^2}{\kappa} + 9*800\|\phi\|_{\ell^1} + 9\kappa\big]n^{-1}\\ &\qquad + 9\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 n^{-1}\\ &+ 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond[m_-^\diamond, f,\Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{16} n\mathfrak{b}_{[m_- - 1]}^2(f)\big)\\ &\qquad + 9\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \exp\big(\tfrac{-3\kappa\psi_{m_-^\diamond}^\Lambda m_-^\diamond}{4000\|\phi\|_{\ell^1}}\big)\\ &+ 3q\mathbb{P}_\varepsilon^q(\mho_{m_+^\diamond}^c) + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathbb{P}_\varepsilon^q(\mho_{m_-^\diamond}^c) + (24\mathcal{C}_4 + 8)\mathcal{R}_q^\star(f,\Lambda). \end{aligned} \tag{L.19}$$

Recalling that $\mathcal{R}_n^\diamond[l,f,\Lambda] = [\mathfrak{b}_l^2(f) \vee \Delta_l^\Lambda n^{-1}] \geqslant n^{-1}$ there is a finite numerical constant $\mathcal{C}$ such that

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ} - f\|_{L^2}^2 &\leqslant \mathcal{C}\big\{[1\vee\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2]\mathcal{R}_n^\diamond[m_+^\diamond, f,\Lambda] + \mathcal{R}_q^\star(f,\Lambda) + \|\phi\|_{\ell^1}^2 n^{-1}\big\}\\ &\qquad + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_{m_-}^2(f) + 9\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \exp\big(\tfrac{-3\kappa\psi_{m_-^\diamond}^\Lambda m_-^\diamond}{4000\|\phi\|_{\ell^1}}\big)\\ &+ 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond[m_-^\diamond, f,\Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{16} n\mathfrak{b}_{[m_- - 1]}^2(f)\big)\\ &\qquad + 3q\mathbb{P}_\varepsilon^q(\mho_{m_+^\diamond}^c) + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathbb{P}_\varepsilon^q(\mho_{m_-^\diamond}^c), \end{aligned} \tag{L.20}$$

which shows the assertion and completes the proof.

Lemma L.2.5. *If $f = e_0$ then there is a finite numerical constant $\mathcal{C}$ such that for all $n, q \in \mathbb{N}$ we have $\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ} - f\|_{L^2}^2 \leqslant \mathcal{C}\{\Delta_{n_o}^\Lambda n^{-1} + n_o\Lambda_{(n_o)}^2 q^{-1}\}$ with $n_o := \lceil 15\frac{300^4}{\kappa^2} \vee 3\frac{800^2}{\kappa^2}\rceil$.*

Proof of Lemma L.2.5.
Let $n_o := \lceil 15\frac{300^4}{\kappa^2} \vee 3\frac{800^2}{\kappa^2}\rceil$. We destinguish for $n\in\mathbb{N}$ the following two cases cases, (a) $n \in [\![1, n_o[\![$ and (b) $n \geqslant n_o$. Consider (a). We select $m_+ = n \leqslant n_o$ and thus keeping in mind that $f = e_0$, and hence $\theta_j^\circ = 0$ and $\phi_j = 0$ for all $|j|\in\mathbb{N}$, from (L.4) (cf. proof of Lemma L.2.1) together with Lemma 1.5.1 (i) follows for all $q\in\mathbb{N}$

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta^\circ} - f\|_{L^2}^2 &= 2\sum_{j\in[\![1,n]\!]} \mathbb{E}_{Y,\varepsilon}^{n,q}|\widehat{[g]}_j^+(\widehat{[h]}_j - \phi_j)\mathbb{P}_w([\![j,n]\!])|^2\\ &\leqslant 2\sum_{j\in[\![1,n]\!]} \Lambda_j \mathbb{E}_\varepsilon^q|\lambda_j\widehat{[g]}_j^+|^2 \mathbb{E}_Y^n|\widehat{[h]}_j - \phi_j|^2 \leqslant 8\sum_{j\in[\![1,n]\!]}\Lambda_j n^{-1} = 8n\overline{\Lambda}_n n^{-1}\\ &\leqslant 8n_o\overline{\Lambda}_{n_o} n^{-1} \leqslant 8\Delta_{n_o}^\Lambda n^{-1}. \end{aligned} \tag{L.21}$$

Consider (b), i.e., $n \geqslant n_o$. We select $m_+^\diamond := n_o \in [\![1,n]\!]$. Since $\|\Pi_{\mathbb{U}_0}^\perp f\|_{L^2}^2 = 0$, and thus $\|\phi\|_{\ell^1}^2 = 1$, $\mathcal{R}_n^\diamond(m_+^\diamond, f,\Lambda) = \Delta_{m_+^\diamond}^\Lambda/n = \Delta_{n_o}^\Lambda/n$ and $\mathcal{R}_q^\star(f,\Lambda) = 0$, from Proposition L.2.1 keeping in mind that by construction $m_+^\diamond \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2 = 3(\frac{800}{\kappa})^2$ and $n \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$

follows

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant \mathcal{C}\{\Delta^{\Lambda}_{n_o} n^{-1} + n^{-1}\} + 3q\mathbb{P}^q_{\varepsilon}(\mho^c_{n_o}) \tag{L.22}$$

Setting $q_o := \lceil 9\Lambda_{(n_o)}/4\rceil \in \mathbb{N}$ by employing ?? ?? we have $\mathbb{P}^q_{\varepsilon}(\mho^c_{n_o}) \leqslant 555 n_o q_o^2 q^{-2}$ for all $q \in \mathbb{N}$, where $q_o^2 \leqslant \frac{81}{4}\Lambda^2_{(n_o)}$. Combining the upper bounds (L.21) and (L.22) for the two cases (a) and (b) there is a finite numerical constant $\mathcal{C}$ such that $\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant \mathcal{C}\{\Delta^{\Lambda}_{n_o} n^{-1} + n_o\Lambda^2_{(n_o)} q^{-1}\}$ for all $n, q \in \mathbb{N}$, which shows the assertion and completes the proof.

Lemma L.2.6. *Assume there is $K \in \mathbb{N}$ with $1 \geqslant \mathfrak{b}_{[K-1]}(f) > 0$ and $\mathfrak{b}_K(f) = 0$. Set $K_h := K \vee 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$, $c_f := \frac{2\|\Pi^{\perp}_{\mathbb{U}_0} f\|^2_{L^2} + 7576\kappa}{\|\Pi^{\perp}_{\mathbb{U}_0} f\|^2_{L^2}\mathfrak{b}^2_{[K-1]}(f)} \geqslant 1$, $n_{f,\Lambda} = \lceil c_f \Delta^{\Lambda}_{K_h} \vee 15(\frac{300}{\sqrt{\kappa}})^4\rceil$ and $q_{f,\Lambda} := \lceil 289\log(K_h+2)\psi^{\Lambda}_{K_h}\Lambda_{(K_h)}\rceil$. For all $n, q \in \mathbb{N}$ we set $m^{\bullet}_n := K_h(\log n)$ for $n \leqslant n_{f,\Lambda}$, $m^{\bullet}_n := \max\{m \in [\![K_h, n]\!] : c_f\Delta^{\Lambda}_m < n\}$ for $n > n_{f,\Lambda}$, $m^{\bullet}_q := K_h(\log q)$ for $q \leqslant q_{K,\Lambda}$, and $m^{\bullet}_q := \max\{m \in [\![K_h, q]\!] : 289\log(m+2)\psi^{\Lambda}_m\Lambda_{(m)} \leqslant q\}$ for $q > q_{K,\Lambda}$, where the defining sets contain K_h and thus they are not empty. There is a finite numerical constant $\mathcal{C}$ such that for all $n, q \in \mathbb{N}$ holds*

$$\begin{aligned}\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} &\leqslant \mathcal{C}(1 \vee \|\Pi_{\mathbb{U}^{\perp}_0} f\|^2_{L^2})(\Delta^{\Lambda}_{n_{f,\Lambda}} n^{-1} + K_h\Lambda^2_{(K_h)} q^{-1} + q_{f,\Lambda} q^{-1}) \\ &\quad + 9\|\Pi_{\mathbb{U}^{\perp}_0} f\|^2_{L^2}\{\exp\big(\tfrac{-3\kappa\psi^{\Lambda}_{m^{\bullet}_n \wedge m^{\bullet}_q}(m^{\bullet}_n \wedge m^{\bullet}_q)}{4000\|\phi\|_{\ell^1}}\big) - \tfrac{1}{n} \vee \tfrac{1}{q}\}\end{aligned} \tag{L.23}$$

If there are $\widetilde{n}_{f,\Lambda}, \widetilde{q}_{f,\Lambda} \in \mathbb{N}$ such that additionally for all $n \geqslant \widetilde{n}_{f,\Lambda}$, $\psi^{\Lambda}_{m^{\bullet}_n} m^{\bullet}_n \geqslant K_h(\log n)$, and for all $q \geqslant \widetilde{q}_{f,\Lambda}$, $\psi^{\Lambda}_{m^{\bullet}_q} m^{\bullet}_q \geqslant K_h(\log q)$ hold true, then for all $n, q \in \mathbb{N}$ we have

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant \mathcal{C}(1 \vee \|\Pi_{\mathbb{U}^{\perp}_0} f\|^2_{L^2})(\Delta^{\Lambda}_{[\widetilde{n}_{f,\Lambda} \vee n_{f,\Lambda}]} n^{-1} + K_h\Lambda^2_{(K_h)} q^{-1} + [\widetilde{q}_{f,\Lambda} \vee q_{f,\Lambda}] q^{-1}) \tag{L.24}$$

Proof of Lemma L.2.6.
Given $K \in \mathbb{N}$ with $1 \geqslant \mathfrak{b}_{[K-1]}(f) > 0$ and $\mathfrak{b}_m(f) = 0$ for all $m \geqslant K$ we note that $\mathcal{R}^{\star}_q(f, \Lambda) = \sum_{j=1}^{K-1} |\theta^{\circ}_j|^2[1 \wedge q^{-1}\Lambda_j] \leqslant q^{-1}\Lambda_{(K)}\|\Pi^{\perp}_{\mathbb{U}_0} f\|^2_{L^2}$. Let $K_h := K \vee 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$, $c_f := \frac{2\|\Pi^{\perp}_{\mathbb{U}_0} f\|^2_{L^2} + 7576\kappa}{\|\Pi^{\perp}_{\mathbb{U}_0} f\|^2_{L^2}\mathfrak{b}^2_{[K-1]}(f)}$, $n_{f,\Lambda} = \lceil c_f\Delta^{\Lambda}_{K_h} \vee 15(\frac{300}{\sqrt{\kappa}})^4\rceil$ and $q_{f,\Lambda} := \lceil 289\log(K_h+2)\psi^{\Lambda}_{K_h}\Lambda_{(K_h)}\rceil$. We distinguish for $n \in \mathbb{N}$ the following two cases, (a) $n \in [\![1, n_{f,\Lambda}[\![$ and (b) $n \geqslant n_{f,\Lambda}$.

Firstly, consider (a), let $n \in [\![1, n_{f,\Lambda}[\![$, then setting $m_- = 1$ and $m_+ = n$ from Lemma L.2.1 follows

$$\begin{aligned}\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} &\leqslant 3\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ n} - \breve{f}^n\|^2_{L^2} + 3\|\Pi_{\mathbb{U}^{\perp}_0} f\|^2_{L^2}\mathfrak{b}^2_1(f) \\ &\quad + 6\sum_{j\in[\![1,n]\!]}\mathbb{E}^q_{\varepsilon}|\widehat{[g]}^{+}_j|^2|\lambda_j - \widehat{[g]}_j|^2|\theta^{\circ}_j|^2 + 2\sum_{j\in[\![1,n]\!]}\mathbb{P}^q_{\varepsilon}(\mathcal{X}^c_j)|\theta^{\circ}_j|^2\end{aligned}$$

Exploiting Lemma 1.5.1 we obtain from (i) $\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ n} - \breve{f}^n\|^2_{L^2} = 2\sum_{j\in[\![1,n]\!]}\Lambda_j\mathbb{E}^q_{\varepsilon}|\lambda_j\widehat{[g]}^{+}_j|^2\mathbb{E}^n_Y|\widehat{[h]}_j - \phi_j|^2 \leqslant 8n\overline{\Lambda}_n n^{-1} \leqslant 8\Delta^{\Lambda}_{n_{f,\Lambda}} n^{-1}$, from (iii) $\sum_{j\in[\![1,n]\!]}|\theta^{\circ}_j|^2\mathbb{E}^q_{\varepsilon}|\widehat{[g]}^{+}_j|^2|\lambda_j - \widehat{[g]}_j|^2 \leqslant 4\mathcal{C}_4\mathcal{R}^{\star}_q(f, \Lambda)$ and from (ii) $\sum_{j\in[\![1,n]\!]}\mathbb{P}^q_{\varepsilon}(\mathcal{X}^c_j)|\theta^{\circ}_j|^2 \leqslant 4\mathcal{R}^{\star}_q(f, \Lambda)$ where $\mathcal{R}^{\star}_q(f, \Lambda) := \sum_{j\in[\![1,K[\![}\theta^{\circ 2}_j[1 \wedge \Lambda_j/q] \leqslant$

$\|\Pi_{\mathbb{U}_0}^{\perp} f\|_{L^2}^2 \Lambda_{(K)} q^{-1}$. The last bounds together with $\mathfrak{b}_1^2(f) \leqslant 1$, $n < n_{f,\Lambda} \leqslant \Delta^{\Lambda}_{n_{f,\Lambda}}$ and $\Lambda_{(K)} \leqslant \Lambda_{(K_h)}$ imply for all $q \in \mathbb{N}$

$$\begin{aligned}\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^{\circ}} - f\|^2_{L^2} &\leqslant 24\Delta^{\Lambda}_{n_{f,\Lambda}} n^{-1} + 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2} + (24\mathcal{C}_4 + 8)\|\Pi_{\mathbb{U}_0}^{\perp} f\|^2_{L^2}\Lambda_{(K)} q^{-1} \\ &\leqslant \mathcal{C}(1 \vee \|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2})\{\Delta^{\Lambda}_{n_{f,\Lambda}} n^{-1} + \Lambda_{(K_h)} q^{-1}\} \quad (L.25)\end{aligned}$$

where $\mathcal{C}$ is a finite numerical constant.
Secondly, consider (b), i.e., $n \geqslant n_{f,\Lambda}$. Setting $m_+^{\diamond} := K_h \leqslant \Delta^{\Lambda}_{K_h} \leqslant n_{f,\Lambda}$, i.e., $m_+^{\diamond} \in [\![1, n]\!]$ from $m_+^{\diamond} = K_h \geqslant K$ follows $\mathfrak{b}_{m_+^{\diamond}}(f) = 0$ and hence $\mathcal{R}^{\diamond}_n(m_+^{\diamond}, f, \Lambda) = \Delta^{\Lambda}_{K_h} n^{-1}$. Consequenly, from Proposition L.2.1 (keep in mind that $\mathcal{R}^{\star}_q(f, \Lambda) \leqslant \|\Pi_{\mathbb{U}_0}^{\perp} f\|^2_{L^2}\Lambda_{(K)} q^{-1}$ and by construction $m_+^{\diamond} \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ and $n \geqslant 15(\frac{300}{\sqrt{\kappa}})^4$) there is a numerical constant $\mathcal{C}$ such that

$$\begin{aligned}\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^{\circ}} - f\|^2_{L^2} \leqslant \mathcal{C}\big\{[1 \vee \|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}]\Delta^{\Lambda}_{K_h} n^{-1} + \|\Pi_{\mathbb{U}_0}^{\perp} f\|^2_{L^2}\Lambda_{(K)} q^{-1} + \|\phi\|^2_{\ell^1} n^{-1}\big\} \\ + 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + 9\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\exp\big(\frac{-3\kappa\psi^{\Lambda}_{m_-^{\diamond}} m_-^{\diamond}}{4000\|\phi\|_{\ell^1}}\big) \\ + 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^{\diamond}_n[m_-^{\diamond}, f, \Lambda] - \frac{\kappa\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}}{16} n\mathfrak{b}^2_{[m_- - 1]}(f)\big) \\ + 3q\mathbb{P}^q_{\varepsilon}(\mho^c_{K_h}) + 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\mathbb{P}^q_{\varepsilon}(\mho^c_{m_-^{\diamond}})\end{aligned}$$

Setting $q_h := \lceil 9\Lambda_{(K_h)}/4\rceil \in \mathbb{N}$ by employing **??** **??** we have $\mathbb{P}^q_{\varepsilon}(\mho^c_{K_h}) \leqslant 555K_h q_h^2 q^{-2}$ for all $q \in \mathbb{N}$. Since $q_h^2 \leqslant \frac{81}{4}\Lambda^2_{(K_h)}$ and $\|\phi\|^2_{\ell^1} \leqslant \kappa^2 K_h \leqslant \kappa^2\Delta^{\Lambda}_{K_h}$ we obtain

$$\begin{aligned}\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^{\circ}} - f\|^2_{L^2} \leqslant \mathcal{C}_1[1 \vee \|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}](\Delta^{\Lambda}_{K_h} n^{-1} + K_h\Lambda^2_{(K_h)} q^{-1}) \\ + 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\mathfrak{b}^2_{m_-}(f) + 9\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\exp\big(\frac{-3\kappa\psi^{\Lambda}_{m_-^{\diamond}} m_-^{\diamond}}{4000\|\phi\|_{\ell^1}}\big) \\ + 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}^{\diamond}_n[m_-^{\diamond}, f, \Lambda] - \frac{\kappa\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}}{16} n\mathfrak{b}^2_{[m_- - 1]}(f)\big) \\ + 3\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}\mathbb{P}^q_{\varepsilon}(\mho^c_{m_-^{\diamond}}) \quad (L.26)\end{aligned}$$

where $\mathcal{C}_1$ is a numerical constant. In order to control the terms involving $m_-^{\diamond}$ and m_- we destinguish for $q \in \mathbb{N}$ the following two cases cases, (b-i) $q \in [\![1, q_{f,\Lambda}[\![$ and (b-ii) $q \geqslant q_{f,\Lambda}$ where $q_{f,\Lambda} = \lceil 289\log(K_h + 2)\psi^{\Lambda}_{K_h}\Lambda_{(K_h)}\rceil$.
Consider (b-i). We set $m_-^{\diamond} = 1$ and hence $m_- = 1$. Thereby, from (L.26) together with $\mathfrak{b}_1^2(f) \leqslant 1$ and $q < q_{f,\Lambda}$ follows

$$\begin{aligned}\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^{\circ}} - f\|^2_{L^2} &\leqslant \mathcal{C}_2[1 \vee \|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}](\Delta^{\Lambda}_{K_h} n^{-1} + K_h\Lambda^2_{(K_h)} q^{-1}) + 15\|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2} \\ &\leqslant \mathcal{C}_3[1 \vee \|\Pi_{\mathbb{U}_0^{\perp}} f\|^2_{L^2}](\Delta^{\Lambda}_{K_h} n^{-1} + K_h\Lambda^2_{(K_h)} q^{-1} + q_{f,\Lambda} q^{-1}) \quad (L.27)\end{aligned}$$

where $\mathcal{C}_3$ is a numerical constant.
Consider (b-ii). Note that for all $q \geqslant q_{K,\Lambda}$ the defining set of $m_q^{\bullet} := \max\{m \in [\![K_h, q]\!] : 289\log(m + 2)\psi^{\Lambda}_m\Lambda_{(m)} \leqslant q\}$ is not empty, where obviously for each $m_-^{\diamond} \in [\![K_h, m_q^{\bullet}]\!]$

holds $q \geqslant 289\log(m_-^\diamond + 2)\psi^\Lambda_{m_-^\diamond}\Lambda_{(m_-^\diamond)}$, and thus from ?? ?? follows $\mathbb{P}^q_\varepsilon(\mho^c_{m_-^\diamond}) \leqslant 53q^{-1}$. Recall that $c_f := \frac{2\|\Pi_{\mathbb{U}_0}^\perp f\|^2_{L^2}+7576\kappa}{\|\Pi_{\mathbb{U}_0}^\perp f\|^2_{L^2}\mathfrak{b}^2_{[K-1]}(f)}$ and $n > n_{f,\Lambda} \geqslant c_f\Delta^\Lambda_{K_h}$. Thereby, the defining set of $m_n^\bullet := \max\{m \in [\![K_h, n]\!] : n > c_f\Delta^\Lambda_m\}$ contains K_h and it is not empty. Consequently, for all $m_-^\diamond \in [\![K_h, m_n^\bullet]\!]$ there hold $m_-^\diamond \geqslant K_h \geqslant K$, hence $\mathcal{R}^\diamond_n[m_-^\diamond, f, \Lambda] = \Delta^\Lambda_{m_-^\diamond}/n$ and thus $\|\Pi_{\mathbb{U}_0}^\perp f\|^2_{L^2}\mathfrak{b}^2_{[K-1]}(f) > [2\|\Pi_{\mathbb{U}_0}^\perp f\|^2_{L^2} + 7576\kappa]\mathcal{R}^\diamond_n[m_-^\diamond, f, \Lambda]$, which in turn employing the definition (??) implies $m_- = K$ and hence $\mathfrak{b}^2_{m_-}(f) = \mathfrak{b}^2_K(f) = 0$, $\mathfrak{b}^2_{[m_--1]}(f) = \mathfrak{b}^2_{[K-1]}(f) > 0$. Selecting $m_-^\diamond := m_n^\bullet \wedge m_q^\bullet$ we have $\mathbb{P}^q_\varepsilon(\mho^c_{m_-^\diamond}) \leqslant 53q^{-1}$ and $m_- = K$, $\mathfrak{b}^2_{m_-}(f) = 0$, $\mathfrak{b}^2_{[m_--1]}(f) = \mathfrak{b}^2_{[K-1]}(f) > 0$, such that from (L.26) follows

$$
\begin{aligned}
\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^\circ} - f\|^2_{L^2} &\leqslant \mathcal{C}_2[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^\Lambda_{K_h}n^{-1} + K_h\Lambda^2_{(K_h)}q^{-1}) \\
&\quad + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet\wedge m_q^\bullet}(m_n^\bullet\wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) \\
&\quad + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}[K-1]\exp\big(-\kappa\kappa n\mathcal{R}^\diamond_n[m_n^\bullet \wedge m_q^\bullet, f, \Lambda] - \tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}}{16}n\mathfrak{b}^2_{[K-1]}(f)\big) \\
&\quad + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}53q^{-1} \\
&\leqslant \mathcal{C}_3[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^\Lambda_{K_h}n^{-1} + K_h\Lambda^2_{(K_h)}q^{-1}) \\
&\quad + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet\wedge m_q^\bullet}(m_n^\bullet\wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}[K-1]\underbrace{\exp\big(-\tfrac{\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}}{16}n\mathfrak{b}^2_{[K-1]}(f)\big)}_{\leqslant \frac{16}{e\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_{[K-1]}(f)}n^{-1}} \\
&\leqslant \mathcal{C}_3[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^\Lambda_{K_h}n^{-1} + K_h\Lambda^2_{(K_h)}q^{-1}) \\
&\quad + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet\wedge m_q^\bullet}(m_n^\bullet\wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) + 3\tfrac{16[K-1]}{e\kappa\mathfrak{b}^2_{[K-1]}(f)}n^{-1} \\
&\leqslant \mathcal{C}_4[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\tfrac{\Delta^\Lambda_{K_h}}{\mathfrak{b}^2_{[K-1]}(f)}n^{-1} + K_h\Lambda^2_{(K_h)}q^{-1}) \\
&\quad + 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet\wedge m_q^\bullet}(m_n^\bullet\wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) \qquad \text{(L.28)}
\end{aligned}
$$

where $\mathcal{C}_4$ is a numerical constant.

Combining the upper bounds (L.25), (L.27) and(L.28) for the three cases (a), (b-i) and (b-ii) and keeping in mind the definition of $m_n^\bullet$ and $m_q^\bullet$ there is a finite numerical constant

$\mathcal{C}$ such that for all $n, q \in \mathbb{N}$ holds

$$
\begin{aligned}
\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant \mathcal{C}(1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2})(\Delta^\Lambda_{n_{f,\Lambda}} n^{-1} + \Lambda_{(K_h)} q^{-1}) \\
+ \mathcal{C}[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^\Lambda_{K_h} n^{-1} + K_h \Lambda^2_{(K_h)} q^{-1} + q_{f,\Lambda} q^{-1}) \\
+ \mathcal{C}[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\frac{\Delta^\Lambda_{K_h}}{\mathfrak{b}^2_{[K-1]}(f)} n^{-1} + K_h \Lambda^2_{(K_h)} q^{-1}) \\
+ 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2} \exp\big(\frac{-3\kappa\psi^\Lambda_{m_n^\bullet \wedge m_q^\bullet}(m_n^\bullet \wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) \\
\leqslant \mathcal{C}(1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2})(\Delta^\Lambda_{n_{f,\Lambda}} n^{-1} + K_h \Lambda^2_{(K_h)} q^{-1} + q_{f,\Lambda} q^{-1}) \\
+ 9\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2} \{\exp\big(\frac{-3\kappa\psi^\Lambda_{m_n^\bullet \wedge m_q^\bullet}(m_n^\bullet \wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) - (n^{-1} \vee q^{-1})\}
\end{aligned}
\tag{L.29}
$$

where we used that $\frac{\Delta^\Lambda_{K_h}}{\mathfrak{b}^2_{[K-1]}(f)} \leqslant n_{f,\Lambda} \leqslant \Delta^\Lambda_{n_{f,\Lambda}}$.

Assume finally, that there are in addition $\widetilde{n}_{f,\Lambda}, \widetilde{q}_{f,\Lambda} \in \mathbb{N}$ such that for all $n \geqslant \widetilde{n}_{f,\Lambda}$, $\psi^\Lambda_{m_n^\bullet} m_n^\bullet \geqslant K_h(\log n)$, and for all $q \geqslant \widetilde{q}_{f,\Lambda}$, $\psi^\Lambda_{m_q^\bullet} m_q^\bullet \geqslant K_h(\log q)$. We shall use without further reference that then $\exp\big(\frac{-3\kappa\psi^\Lambda_{m_n^\bullet \wedge m_q^\bullet}(m_n^\bullet \wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) \leqslant n^{-1} + q^{-1}$ for all $n \geqslant \widetilde{n}_{f,\Lambda}$ and $q \geqslant \widetilde{q}_{f,\Lambda}$ since $K_h \geqslant \frac{4000\|\phi\|_{\ell^1}}{3\kappa}$. Following line by line the proof of (L.29) using $\widetilde{n}_{f,\Lambda} \vee n_{f,\Lambda}$ and $\widetilde{q}_{f,\Lambda} \vee q_{f,\Lambda}$ rather than $n_{f,\Lambda}$ and $q_{f,\Lambda}$, respectively, we obtain the assertion, that is, s $\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant \mathcal{C}(1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2})(\Delta^\Lambda_{[\widetilde{n}_{f,\Lambda} \vee n_{f,\Lambda}]} n^{-1} + K_h \Lambda^2_{(K_h)} q^{-1} + [\widetilde{q}_{f,\Lambda} \vee q_{f,\Lambda}] q^{-1})$, which completes the proof.

Lemma L.2.7. *Let f have an infinite series expansion as definied in (np), i.e., $\mathfrak{b}_m(f) > 0$ for all $m \in \mathbb{N}$. Set $m_h := \lceil 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2 \rceil$, $\widetilde{m}_h := \min\{m \in \mathbb{N} : \mathfrak{b}_{m_h}(f) > \mathfrak{b}_m(f)\}$, $n_{f,\Lambda} := \lceil \frac{\Delta^\Lambda_{\widetilde{m}_h}}{\mathfrak{b}^2_{\widetilde{m}_h}(f)} \vee 15\frac{300^4}{\kappa^2} \rceil$ and $q_{f,\Lambda} := \lceil 289 \log(m_h + 2)\psi^\Lambda_{m_h} \Lambda_{(m_h)} \rceil$.*

For all $n \in \mathbb{N}$ let $m_n^\bullet \in [\![m_n^\diamond, n]\!]$ and for all $q \in \mathbb{N}$ we set $m_q^\bullet := \max\{m \in [\![K_h, q]\!] : 289 \log(m+2)\psi^\Lambda_m \Lambda_{(m)} \leqslant q\}$ for $q > q_{f,\Lambda}$, where the defining set containing K_h is not empty, and $m_q^\bullet := m_n^\diamond$ for $q \leqslant q_{f,\Lambda}$. There is a finite numerical constant $\mathcal{C}$ such that for all $n, q \in \mathbb{N}$ holds

$$
\begin{aligned}
\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ} - f\|^2_{L^2} \leqslant +\mathcal{C}\big\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}]\, \mathcal{R}^\diamond_n[m_n^\bullet \wedge m_q^\bullet, f, \Lambda] + \mathcal{R}^\star_q(f, \Lambda) \\
+ [1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^\Lambda_{n_{f,\Lambda}} n^{-1} + q_{f,\Lambda} q^{-1}) + \|\phi\|^2_{\ell^1} n^{-1} + q^2_{f,\Lambda} q^{-1}\big\} \\
+ 12\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2} \{\exp\big(\frac{-3\kappa\psi^\Lambda_{m_n^\bullet \wedge m_q^\bullet}[m_n^\bullet \wedge m_q^\bullet]}{4000\|\phi\|_{\ell^1}}\big) - [\mathcal{R}^\diamond_n(f, \Lambda) \vee \mathcal{R}^\star_q(f, \Lambda)]\}
\end{aligned}
\tag{L.30}
$$

If there are $\widetilde{n}_{f,\Lambda}, \widetilde{q}_{f,\Lambda} \in \mathbb{N}$ such that additionaly (i) $\psi^\Lambda_{m_n^\diamond} m_n^\diamond \geqslant m_h(-\log \mathcal{R}^\diamond_n(f, \Lambda))$ for all

$n \geqslant \widetilde{n}_{f,\Lambda}$ and (ii) *$\psi^{\Lambda}_{m^\bullet_q} m^\bullet_q \geqslant m_h(-\log \mathcal{R}^\star_q(f,\Lambda))$ for all $q \geqslant \widetilde{q}_{f,\Lambda}$, then for all $n, q \in \mathbb{N}$ holds*

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^\circ - f\|^2_{L^2} \leqslant \mathcal{C}\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\mathcal{R}^\diamond_n[m^\diamond_n \wedge m^\bullet_q, f, \Lambda] + \mathcal{R}^\star_q(f,\Lambda)) + [1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^\Lambda_{[\widetilde{n}_{f,\Lambda} \vee n_{f,\Lambda}]} n^{-1} + [\widetilde{q}_{f,\Lambda} \vee q_{f,\Lambda}] q^{-1}) + \|\phi\|^2_{\ell^1} n^{-1} + [\widetilde{q}_{f,\Lambda} \vee q_{f,\Lambda}]^2 q^{-1}\} \tag{L.31}$$

Reminder. *Keep in mind that $m^\diamond_n := \arg\min\{\mathcal{R}^\diamond_n[m,f,\Lambda], m \in [\![1,n]\!]\}$ with $\mathcal{R}^\diamond_n[m,f,\Lambda] = [\mathfrak{b}^2_m(f) \vee \Delta^\Lambda_m n^{-1}]$ and $\mathcal{R}^\diamond_n(f,\Lambda) := \mathcal{R}^\diamond_n[m^\diamond_n, f, \Lambda]$. Considering $\widetilde{m}_h = \min\{m \in \mathbb{N} : \mathfrak{b}_{m_h}(h) > \mathfrak{b}_m(h)\}$ as defined in Lemma L.2.7 we note that the defining set is not empty since $\mathfrak{b}_m(f) > 0$ for all $m \in \mathbb{N}$ and $\lim_{m\to\infty} \mathfrak{b}_m(f) = 0$, where $\widetilde{m}_h > m_h$ due to the monotonicity of $\mathfrak{b}_m(f)$. Moreover, for all $n \geqslant n_{f,\Lambda} := \lceil \frac{\Delta^\Lambda_{\widetilde{m}_h}}{\mathfrak{b}^2_{\widetilde{m}_h}(f)} \rceil$ holds $\mathcal{R}^\circ_n[m_h, f, \Lambda] \geqslant \mathfrak{b}^2_{m_h}(f) > \mathfrak{b}^2_{\widetilde{m}_h}(f) = \mathcal{R}^\circ_n[\widetilde{m}_h, f, \Lambda]$ and hence, for all $n \geqslant n_{f,\Lambda}$ we have $m^\bullet_n \geqslant m^\diamond_n > m_h$. Furthermore, for all $q \geqslant q_{f,\Lambda} := \lceil 289 \log(m_h + 2)\psi^\Lambda_{m_h}\Lambda_{(m_h)}\rceil$ is the defining set of $m^\bullet_q = \max\{m \in [\![m_h, q]\!] : 289\log(m+2)\psi^\Lambda_m \Lambda_{(m)} \leqslant q\}$ not empty. Consequently, for all $n \geqslant n_{f,\Lambda}$, $q \geqslant q_{f,\Lambda}$ follows $m^\bullet_n \wedge m^\bullet_q \geqslant m_h$. We use these preliminary findings in the proof of Lemma L.2.7 without further reference.*

Proof of Lemma L.2.7.
Let $m_h := \lceil 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2\rceil$, $\widetilde{m}_h := \min\{m \in \mathbb{N} : \mathfrak{b}_{m_h}(f) > \mathfrak{b}_m(f)\}$, $n_{f,\Lambda} := \lceil \frac{\widetilde{m}_\lambda \delta^{|\Phi}_{\widetilde{m}_\lambda} |\Phi_{(\widetilde{m}_\lambda)}}{\mathfrak{b}^2_{\widetilde{m}_\lambda}(\theta)} \vee 15\frac{300^4}{\kappa^2}\rceil$ and $q_{f,\Lambda} := \lceil 289 \log(m_h+2)\psi^\Lambda_{m_h}\Lambda_{(m_h)}\rceil$.
We distinguish for $n \in \mathbb{N}$ the following two cases, (a) $n \in [\![1, n_{f,\Lambda}[\![$ and (b) $n \geqslant n_{f,\Lambda}$.

Firstly, consider (a), let $n \in [\![1, n_{f,\Lambda}[\![$, then setting $m_- = 1$ and $m_+ = n$ from Lemma L.2.1 follows

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^\circ - f\|^2_{L^2} \leqslant 3\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ n} - \breve{f}^n\|^2_{L^2} + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}\mathfrak{b}^2_1(f) + 6\sum_{j\in[\![1,n]\!]} \mathbb{E}^q_\varepsilon|\widehat{[g]}^+_j|^2|\lambda_j - \widehat{[g]}_j|^2|\theta^\circ_j|^2 + 2\sum_{j\in[\![1,n]\!]} \mathbb{P}^q_\varepsilon(\mathcal{X}^c_j)|\theta^\circ_j|^2$$

Exploiting Lemma 1.5.1 we obtain from (i) $\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^{\circ n} - \breve{f}^n\|^2_{L^2} = 2\sum_{j\in[\![1,n]\!]}\Lambda_j \mathbb{E}^q_\varepsilon|\lambda_j\widehat{[g]}^+_j|^2\mathbb{E}^n_Y|\widehat{[h]}_j - \phi_j|^2 \leqslant 8n\overline{\Lambda}_n n^{-1} \leqslant 8\Delta^\Lambda_{n_{f,\Lambda}} n^{-1}$, from (iii) $\sum_{j\in[\![1,n]\!]}|\theta^\circ_j|^2\mathbb{E}^q_\varepsilon|\widehat{[g]}^+_j|^2|\lambda_j - \widehat{[g]}_j|^2 \leqslant 4\mathcal{C}_4\mathcal{R}^\star_q(f,\Lambda)$ and from (ii) $\sum_{j\in[\![1,n]\!]}\mathbb{P}^q_\varepsilon(\mathcal{X}^c_j)|\theta^\circ_j|^2 \leqslant 4\mathcal{R}^\star_q(f,\Lambda)$ where $\mathcal{R}^\star_q(f,\Lambda) := \sum_{j\in[\![1,K[\![}\theta^{\circ 2}_j[1 \wedge \Lambda_j/q]$. The last bounds together with $\mathfrak{b}^2_1(f) \leqslant 1$ and $n < n_{f,\Lambda} \leqslant \Delta^\Lambda_{n_{f,\Lambda}}$ imply for all $q \in \mathbb{N}$

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta}^\circ - f\|^2_{L^2} \leqslant 24\Delta^\Lambda_{n_{f,\Lambda}} n^{-1} + 3\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2} + (24\mathcal{C}_4 + 8)\mathcal{R}^\star_q(f,\Lambda) \leqslant \mathcal{C}\{(1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2})\Delta^\Lambda_{n_{f,\Lambda}} n^{-1} + \mathcal{R}^\star_q(f,\Lambda)\} \tag{L.32}$$

where $\mathcal{C}$ is a finite numerical constant.
Secondly, consider (b), i.e., $n \geqslant n_{f,\Lambda} \geqslant 15\frac{300^4}{\kappa^2}$ where by construction $m^\bullet_n \geqslant m^\diamond_n > m_h \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ (see appendix L.2). We destinguish further for $q \in \mathbb{N}$ the following two cases cases, (b-i) $q \in [\![1, q_{f,\Lambda}[\![$ and (b-ii) $q \geqslant q_{f,\Lambda}$ where $q_{f,\Lambda} = \lceil 289\log(m_h+2)\psi^\Lambda_{m_h}\Lambda_{(m_h)}\rceil$.

Consider (b-i). Since $n \geqslant n_{f,\Lambda} \geqslant 15\frac{300^4}{\kappa^2}$ and $m_n^\bullet > m_h \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$ we can employ Proposition L.2.1 with $m_+^\diamond := m_n^\bullet$, and thus

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2]\mathcal{R}_n^\diamond[m_n^\bullet, f, \Lambda] + \mathcal{R}_q^\star(f,\Lambda) + \|\phi\|_{\ell^1}^2 n^{-1}\} \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_{m_-}^2(f) + 9\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \exp\big(\frac{-3\kappa\psi_{m_-^\diamond}^\Lambda m_-^\diamond}{4000\|\phi\|_{\ell^1}}\big) \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond[m_-^\diamond, f, \Lambda] - \frac{\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{16} n\mathfrak{b}_{[m_- - 1]}^2(f)\big) \\ + 3q\mathbb{P}_\varepsilon^q(\mho_{m_n^\bullet}^c) + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathbb{P}_\varepsilon^q(\mho_{m_-^\diamond}^c)\end{aligned}$$

Setting $m_-^\diamond = 1$, and hence $m_- = 1$, by employing $\mathfrak{b}_1^2(f) \leqslant 1$ and $q \leqslant q_{f,\Lambda}$ it follows

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2]\mathcal{R}_n^\diamond[m_n^\bullet, f, \Lambda] + \mathcal{R}_q^\star(f,\Lambda) + \|\phi\|_{\ell^1}^2 n^{-1}\} \\ + 15\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 + 3q \\ \leqslant \mathcal{C}_1\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2](\mathcal{R}_n^\diamond[m_n^\bullet, f, \Lambda] + q_{f,\Lambda} q^{-1}) + \mathcal{R}_q^\star(f,\Lambda) + q_{f,\Lambda}^2 q^{-1} + \|\phi\|_{\ell^1}^2 n^{-1}\} \quad \text{(L.33)}\end{aligned}$$

where $\mathcal{C}_1$ is a finite numerical constant.

Consider (b-ii). We note that for all $q \geqslant q_{f,\Lambda} := \lceil 289\log(m_h+2)\psi_{m_h}^\Lambda \Lambda_{(m_h)}\rceil$ is the defining set of $m_q^\bullet = \max\{m \in [\![m_h, q]\!] : 289\log(m+2)\psi_m^\Lambda \Lambda_{(m)} \leqslant q\}$ not empty. By construction, for all $n \geqslant n_{f,\Lambda}$, $q \geqslant q_{f,\Lambda}$ follows $m_n^\bullet \wedge m_q^\bullet > m_h \geqslant 3(\frac{800\|\phi\|_{\ell^1}}{\kappa})^2$. Since also $n \geqslant n_{f,\Lambda} \geqslant 15\frac{300^4}{\kappa^2}$ we can employ Proposition L.2.1 with $m_+^\diamond := m_n^\bullet \wedge m_q^\bullet$, and thus

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2]\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda] + \mathcal{R}_q^\star(f,\Lambda) + \|\phi\|_{\ell^1}^2 n^{-1}\} \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_{m_-}^2(f) + 9\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \exp\big(\frac{-3\kappa\psi_{m_-^\diamond}^\Lambda m_-^\diamond}{4000\|\phi\|_{\ell^1}}\big) \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond[m_-^\diamond, f, \Lambda] - \frac{\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{16} n\mathfrak{b}_{[m_- - 1]}^2(f)\big) \\ + 3q\mathbb{P}_\varepsilon^q(\mho_{m_n^\bullet \wedge m_q^\bullet}^c) + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathbb{P}_\varepsilon^q(\mho_{m_-^\diamond}^c)\end{aligned}$$

Noting that $m_+^\diamond = m_n^\bullet \wedge m_q^\bullet \in [\![m_h, m_q^\bullet]\!]$ and $289\log(m_+^\diamond+2)\psi_{m_+^\diamond}^\Lambda \Lambda_{(m_+^\diamond)} \leqslant q$ by construction of $m_q^\bullet$ from ?? ?? follows $q\mathbb{P}_\varepsilon^q(\mho_{m_n^\bullet \wedge m_q^\bullet}^c) \leqslant 11226 q^{-1}$ and thus

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}_1\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2]\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda] + \mathcal{R}_q^\star(f,\Lambda) + \|\phi\|_{\ell^1}^2 n^{-1} + q^{-1}\} \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_{m_-}^2(f) + 9\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \exp\big(\frac{-3\kappa\psi_{m_-^\diamond}^\Lambda m_-^\diamond}{4000\|\phi\|_{\ell^1}}\big) \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2[m_- - 1]\exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond[m_-^\diamond, f, \Lambda] - \frac{\kappa\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2}{16} n\mathfrak{b}_{[m_- - 1]}^2(f)\big) \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathbb{P}_\varepsilon^q(\mho_{m_-^\diamond}^c)\end{aligned}$$

where $\mathcal{C}_1$ is a finite numerical constant. Setting $m_-^\diamond := m_n^\bullet \wedge m_q^\bullet$ and m_- as in definition (??), that is $m_- \leqslant m_-^\diamond = m_n^\bullet \wedge m_q^\bullet$ and $\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathfrak{b}_{m_-}^2(f) \leqslant [2\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 + 7576\kappa]\mathcal{R}_n^\diamond[m_n^\bullet \wedge$

$m_q^\bullet, f, \Lambda]$, it follows

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}_2\big\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2]\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda] + \mathcal{R}_q^\star(f,\Lambda) + \|\phi\|_{\ell^1}^2 n^{-1} + q^{-1}\big\} \\ + 9\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet\wedge m_q^\bullet}[m_n^\diamond \wedge m_q^\bullet]}{4000\|\phi\|_{\ell^1}}\big) \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 [m_n^\bullet \wedge m_q^\bullet] \exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda]\big) \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \mathbb{P}_\varepsilon^q(\mho^c_{m_n^\bullet\wedge m_q^\bullet})\end{aligned}$$

where $\mathcal{C}_2$ is a finite numerical constant. Noting that $m_-^\diamond = m_n^\bullet \wedge m_q^\bullet \in [\![m_h, m_q^\bullet]\!]$ and $289\log(m_-^\diamond + 2)\psi^\Lambda_{m_-^\diamond}\Lambda_{(m_-^\diamond)} \leqslant q$ by construction of $m_q^\bullet$ from **??** **??** follows $\mathbb{P}_\varepsilon^q(\mho^c_{m_n^\bullet\wedge m_q^\bullet}) \leqslant 53q^{-1}$ and together with $\mathcal{R}_q^\star(f,\Lambda) \geqslant \frac{1}{2}\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 q^{-1}$ thus

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}_3\big\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2]\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda] + \mathcal{R}_q^\star(f,\Lambda) + \|\phi\|_{\ell^1}^2 n^{-1} + q^{-1}\big\} \\ + 9\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet\wedge m_q^\bullet}[m_n^\bullet \wedge m_q^\bullet]}{4000\|\phi\|_{\ell^1}}\big) \\ + 3\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 [m_n^\bullet \wedge m_q^\bullet] \exp\big(-\kappa\kappa n\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda]\big)\end{aligned}$$

where $\mathcal{C}_3$ is a finite numerical constant. Taking further into account that for all $m \in \mathbb{N}$ hold $n\mathcal{R}_n^\diamond[m, f, \Lambda] \geqslant \Delta_m^\Lambda \geqslant \psi_m^\Lambda m$ and (keeping in mind $\psi_m^\Lambda \geqslant 1$, $\kappa \geqslant 1$, $\kappa \geqslant 1$ and $\|\phi\|_{\ell^1} \geqslant 1$) hence

$$\begin{aligned}m\exp(-\kappa\kappa n\mathcal{R}_n^\diamond[m, f, \Lambda]) \leqslant \tfrac{2}{\kappa\kappa}\tfrac{\kappa\kappa}{2}\psi_m^\Lambda m \exp\big(-\tfrac{\kappa\kappa}{2}\psi_m^\Lambda m\big)\exp\big(-\tfrac{\kappa\kappa}{2}\psi_m^\Lambda m\big) \\ \leqslant \tfrac{2}{e\kappa\kappa}\exp\big(-\tfrac{\kappa\kappa}{2}\psi_m^\Lambda m\big) \leqslant \tfrac{1}{\kappa\kappa}\exp\big(\tfrac{-3\kappa\psi_m^\Lambda m}{4000\|\phi\|_{\ell^1}}\big).\end{aligned}$$

The last bound implies that

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}_3\big\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2]\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda] + \mathcal{R}_q^\star(f,\Lambda) + \|\phi\|_{\ell^1}^2 n^{-1} + q^{-1}\big\} \\ + 12\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet\wedge m_q^\bullet}[m_n^\bullet \wedge m_q^\bullet]}{4000\|\phi\|_{\ell^1}}\big)\end{aligned} \tag{L.34}$$

Combining the upper bounds (L.32), (L.33) and(L.34) for the three cases (a), (b-i) and (b-ii) and keeping in mind that $\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda] \geqslant \mathcal{R}_n^\diamond(f,\Lambda) = \min\{\mathcal{R}_n^\diamond[m, f, \Lambda], m \in \mathbb{N}\}$ there is a finite numerical constant $\mathcal{C}$ such that for all $n, q \in \mathbb{N}$ holds

$$\begin{aligned}\mathbb{E}_{Y,\varepsilon}^{n,q}\|\widehat{\theta}^\circ - f\|_{L^2}^2 \leqslant \mathcal{C}\big\{[1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2]\,\mathcal{R}_n^\diamond[m_n^\bullet \wedge m_q^\bullet, f, \Lambda] + \mathcal{R}_q^\star(f,\Lambda) \\ + [1 \vee \|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2](\Delta^\Lambda_{n_{f,\Lambda}} n^{-1} + q_{f,\Lambda} q^{-1}) + \|\phi\|_{\ell^1}^2 n^{-1} + q_{f,\Lambda}^2 q^{-1}\big\} \\ + 12\|\Pi_{\mathbb{U}_0^\perp} f\|_{L^2}^2 \{\exp\big(\tfrac{-3\kappa\psi^\Lambda_{m_n^\bullet\wedge m_q^\bullet}[m_n^\bullet \wedge m_q^\bullet]}{4000\|\phi\|_{\ell^1}}\big) - [\mathcal{R}_n^\diamond(f,\Lambda) \vee \mathcal{R}_q^\star(f,\Lambda)]\}\end{aligned} \tag{L.35}$$

which shows the assertion (L.30).

Assume finally, that there are in addition $\widetilde{n}_{f,\Lambda}, \widetilde{q}_{f,\Lambda} \in \mathbb{N}$ such that for all $n \geqslant \widetilde{n}_{f,\Lambda}$,

$\psi^{\Lambda}_{m_n^\diamond} m_n^\diamond \geqslant m_h(-\log \mathcal{R}_n^\diamond(f,\Lambda))$, and for all $q \geqslant \widetilde{q}_{f,\Lambda}$, $\psi^{\Lambda}_{m_q^\bullet} m_q^\bullet \geqslant m_h(-\log \mathcal{R}_q^\star(f,\Lambda))$. We shall use without further reference that then $\exp\big(\frac{-3\kappa\psi^{\Lambda}_{m_n^\diamond\wedge m_q^\bullet}(m_n^\diamond\wedge m_q^\bullet)}{4000\|\phi\|_{\ell^1}}\big) \leqslant \mathcal{R}_n^\diamond(f,\Lambda)\vee\mathcal{R}_q^\star(f,\Lambda) \leqslant \mathcal{R}_n^\diamond[m_n^\diamond\wedge m_q^\bullet, f,\Lambda]\vee\mathcal{R}_q^\star(f,\Lambda)$ for all $n \geqslant \widetilde{n}_{f,\Lambda}$ and $q \geqslant \widetilde{q}_{f,\Lambda}$ since $m_h \geqslant \frac{4000\|\phi\|_{\ell^1}}{3\kappa}$. Following line by line the proof of (L.29) using $\widetilde{n}_{f,\Lambda}\vee n_{f,\Lambda}$ and $\widetilde{q}_{f,\Lambda}\vee q_{f,\Lambda}$ rather than $n_{f,\Lambda}$ and $q_{f,\Lambda}$, respectively, we obtain the assertion (L.31), that is,

$$\mathbb{E}^{n,q}_{Y,\varepsilon}\|\widehat{\theta^\circ} - f\|^2_{L^2} \leqslant \mathcal{C}\big\{[1\vee\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\mathcal{R}_n^\diamond[m_n^\diamond\wedge m_q^\bullet, f,\Lambda] + \mathcal{R}_q^\star(f,\Lambda))$$
$$+ [1\vee\|\Pi_{\mathbb{U}_0^\perp} f\|^2_{L^2}](\Delta^{\Lambda}_{[\widetilde{n}_{f,\Lambda}\vee n_{f,\Lambda}]} n^{-1} + [\widetilde{q}_{f,\Lambda}\vee q_{f,\Lambda}]q^{-1}) + \|\phi\|^2_{\ell^1} n^{-1} + [\widetilde{q}_{f,\Lambda}\vee q_{f,\Lambda}]^2 q^{-1}\big\}, \quad \text{(L.36)}$$

which completes the proof.

Appendix M

Simulation skim

Appendix N

R Code

Bibliography

Barron, A., Birgé, L., and Massart, P. (1999). Risk bounds for model selection via penalization. *Probability Theory and Related Fields*, 113(3):301–413.

Bigot, J. and Van Bellgem, S. (2009). Log-density deconvolution by wavelet thresholding. *Scandinavian Journal of Statistics*, 36:749–763.

Billingsley, P. (1986). *Probability and Measure*. John Wiley and Sons, second edition.

Birgé, L. (2001). An alternative point of view on lepski's method. In Monogr., editor, *State of the art in probability and statistics*, volume 36 of *IMS Lecture Notes*, pages 113–133. (Leiden 1999).

BUNKE, O. and JOHANNES, J. (2005). Selfinformative limits of bayes estimates and generalized maximum likelihood. *Statistics*, 39(6):483–502.

Carroll, R. J. and Hall, P. (1988). Optimal rates of convergence for deconvolving a density. *Journal of the American Statistical Association*, 83:1184–1186.

Cavalier, L. and Hengartner, N. (2005). Adaptive estimation for inverse problems with noisy operators. *Inverse Problems*, 21:1345–1361.

Cochran, W. W., Mouritsen, H., and Wikelski, M. (2004). Migrating songbirds recalibrate their magnetic compass daily from twylight cues. *Science*, 304:405.

Comte, F. and Taupin, M.-L. (2003). Adaptive density deconvolution for circular data. Prépublication map5 2003-10, Université Paris Descartes.

Curray, J. R. (1956). The analysis of two-dimensional orientation data. *The Journal of Geology*, 64(2):117–131.

Devroye, L. (1989). Consistent deconvolution in density estimation. *Canadian Journal of Statistics*, 17:235–239.

Efromovich, S. (1997). Density estimation for the case of supersmooth measurement error. *Journal of the American Statistical Association*, 92:526–535.

Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problems. *Annals of Statistics*, 19:1257–1272.

Fan, J. (1992). Deconvolution with supersmooth distributions. *Canadian Journal of Statistics*, 20:734–747.

Fan, J. and Koo, J. Y. (2002). Wavelet deconvolution. *IEEE Transactions on Information Theory*, 48:734–747.

Fisher, N. (1993). *Statistical analysis of circular data.* Cambridge University Press.

Ghosal, S., Ghosh, J. K., Van Der Vaart, A. W., et al. (2000). Convergence rates of posterior distributions. *Annals of Statistics*, 28(2):500–531.

Goldenshluger, A. (1999). On pointwise adaptive nonparametric deconvolution. *Bernoulli*, 5:907–925.

Goldenshluger, A. (2000). Density deconvolution in the circular structural model. *Journal of Multivariate Analysis*, 81:360–375.

Goldenshluger, A. and Lepski, O. (2011). Bandwidth selection in kernel density estimation: Oracle inequalities and adaptive minimax optimality. *The Annals of Statistics*, 39:1608–1632.

Hadamard, J. (1902). Sur les problèmes aux dérivés partielles et leur signification physique. *Princeton University Bulletin*, 13:49–52.

Heinz, E. (1951). Beiträge zur Störungstheorie der Spektralzerlegung. *Mathematische Annalen*, 123:415–438.

Johannes, J. and Schenk, R. (2013). On rate optimal local estimation in functional linear regression. *Electronic Journal of Statistics*, 7:191–216.

JOHANNES, J. and SCHWARZ, M. (2012). Adaptive gaussian inverse regression with partially unknown operator.

Johannes, J. and Schwarz, M. (2013a). Adaptive circular deconvolution by model selection under unknown error distribution. *Bernoulli*, 19(5A):1576–1611.

Johannes, J. and Schwarz, M. (2013b). Adaptive gaussian inverse regression with partially unknown operator. *Communications in Statistics - Theory and Methods*, 42(7):1343–1362.

Johannes, J., Simoni, A., and Schenk, R. (2014). Adaptive bayesian estimation in gaussian sequence space models. *Discussion paper at Université catholique de Louvain.*

JOHANNES, J., SIMONI, A., and SCHENK, R. (2016). Adaptive bayesian estimation in indirect gaussian sequence space model.

Kim, P. T. and Koo, J. Y. (2002). Optimal spherical deconvolution. *Journal of Multivariate Analysis*, 80:21–42.

Koo, J. Y. and Park, B. U. (1996). B-spline deconvolution based on the EM algorithm. *Journal of Statistical Computation and Simulation.*, 54:275–288.

Laurent, B., Loubes, J., and Marteau, C. (2012). Non asymptotic minimax rates of testing in signal detection with heterogeneous variances. *Electronic Journal of Statistics*, 6:91–122.

Liptser, R. S. and Shiryaev, A. N. (2013). *Statistics of random processes: I. General theory*, volume 5. Springer Science & Business Media.

Mardia, K. (1972). *Statistics of directional data.* Probability and Mathematical Statistics. A Series of Monographs and Textbooks. Vol. 13. Academic Press.

MASSART, P. (2003). Concentration inequalities and model selection. In Springer, editor, *Ecole d'Été de Probabilités de Saint-Flour XXXIII*, volume 1869.

Meister, A. (2009). *Deconvolution problems in nonparametric statistics.* Lecture Notes in Statistics 193. Berlin: Springer.

Mendelsohn, J. and Rice, J. (1982). Deconvolution of micro-fluorometric histograms with B-splines. *Journal of the American Statistical Association*, 77:748–753.

Neumann, M. H. (1997). On the effect of estimating the error density in nonparametric deconvolution. *Journal of Nonparametric Statistics*, 7:307–330.

Pensky, M. and Vidakovic, B. (1999). Adaptive wavelet estimator for nonparametric density deconvolution. *Annals of Statistics*, 27:2033–2053.

Petrov, V. V. (1995). *Limit theorems of probability theory. Sequences of independent random variables.* Oxford Studies in Probability. Clarendon Press., Oxford, 4. edition.

Rigollet, P. and Tsybakov, A. B. (2007). Linear and convex aggregation of density estimators. *Mathematical Methods of Statistics*, 16(3):260–280.

Stefanski, L. A. (1990). Rate of convergence of some estimators in a class of deconvolution problems. *Statistics and Probability Letters*, 9:229–235.

Zhang, C.-H. (1990). Fourier methods for estimating mixing densities and distributions. *Annals of Statistics*, 18:806–831.