



HAL
open science

PROJECTION ESTIMATION OF A QUADRATIC FUNCTIONAL FROM INDIRECT OBSERVATIONS

Ousmane Sacko

► **To cite this version:**

Ousmane Sacko. PROJECTION ESTIMATION OF A QUADRATIC FUNCTIONAL FROM INDIRECT OBSERVATIONS. 2024. hal-04559411

HAL Id: hal-04559411

<https://hal.science/hal-04559411>

Preprint submitted on 25 Apr 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

PROJECTION ESTIMATION OF A QUADRATIC FUNCTIONAL FROM INDIRECT OBSERVATIONS

OUSMANE SACKO¹

ABSTRACT. We consider the convolution model: $Y = X + \varepsilon$, where X and ε are independent. We aim to estimate $\int_{\mathbb{R}} f^2(x)dx$, where f is the unknown density of the signal X from n observations of Y . We introduce a novel projection estimator based on expanding f in the Hermite basis. Convergence rates for f within the Sobolev-Hermite ball are provided for various error types. We also present a novel adaptive procedure inspired by Goldenshluger and Lepski (2011) to select the appropriate space, and we demonstrate an oracle inequality for the adaptive estimator. Numerical experiments are conducted to illustrate the effectiveness of our methodology.

Key words: Deconvolution; Hermite basis; Model selection; Projection estimator; Quadratic functional.

1. INTRODUCTION

Bibliographical context. Consider the convolution model given by:

$$(1) \quad Y_k = X_k + \varepsilon_k, \quad k = 1, \dots, n,$$

and the following assumptions

- (H1) The variables $(X_k)_{k \geq 1}$ are independent and identically distributed (i.i.d.) with an *unknown density* f , with respect to the Lebesgue measure,
- (H2) The variables $(\varepsilon_k)_{k \geq 1}$ are i.i.d. with a *known* density f_ε , with respect to the Lebesgue measure,
- (H3) The variables $(X_k)_{k \geq 1}$ and $(\varepsilon_k)_{k \geq 1}$ are independent.

Our aim is to estimate the parameter

$$\theta(f) = \int_{\mathbb{R}} f^2(x)dx$$

from n copies Y_1, \dots, Y_n with common density $f_Y = f \star f_\varepsilon$, where $g \star h(x) = \int_{\mathbb{R}} g(x-y)h(y)dy$ denotes the convolution product between g and h .

This problem is motivated by applications to adaptive goodness-of-fit tests.

When the variables of interest are available, that is $\varepsilon_k = 0$ almost surely (*a.s.*) in Model (1), nonparametric estimation of a quadratic functional has been studied extensively. For instance, [Bickel and Ritov \(1988\)](#) investigate the estimation of $\int_{\mathbb{R}} (f^{(d)})^2(x)dx$, where $f^{(d)}$ denotes the d -th derivative of f by using a kernel method. They establish that the parametric rate can be reached if the regularity of $f^{(d)}$ is large enough. However, their procedure is non adaptive. [Efromovich and Low \(1996\)](#) propose a projection estimator of $\int_{\mathbb{R}} (f^{(d)})^2(x)dx$ on the Fourier basis. The authors provide convergence rates for $f^{(d)}$ belonging to the Lipschitz class of order β with a logarithmic loss for small β values. But β is unknown. [Laurent \(2005\)](#) develops an adaptive estimator of $\theta(f)$ by considering the Haar basis. The author obtains the same convergence rate as in [Efromovich and Low \(1996\)](#) when the density f belongs to a Besov class. Recently, [Goldenshluger and Lepski \(2022a,b\)](#) establish minimax rates for the \mathbb{L}_p -norm of a multivariate density function with $p \geq 1$ on the Nikolskii space. Furthermore, the problem of estimating a general functional

¹ MODAL'X - Modélisation aléatoire de Paris X, UMR 9023, Université Paris Nanterre.

can be found in [Birgé and Massart \(1995\)](#), [Laurent \(1996\)](#), [Kerkyacharian and Picard \(1996\)](#), [Tribouley \(2000\)](#).

In privacy constraints framework, [Butucea et al. \(2023\)](#) investigate minimax rates and propose adaptive estimators (up to a logarithmic factor) of $\theta(f) = \int_0^1 f^2(x)dx$ based on the Haar wavelet basis. Obviously, the support of f is assumed known and is equal to $[0, 1]$ in the last case.

In convolution models (1), to our knowledge, [Butucea \(2007\)](#) is the first to consider the problem of estimating of $\theta(f) = \int_{\mathbb{R}} f^2(x)dx$ from Model (1). The author proposes a kernel estimator and proves minimax results under certain assumptions on the density of noise f_ε . [Loubes and Marteau \(2014\)](#) apply the estimation of $\theta(f) = \int_{\mathbb{R}} f^2(x)dx$ in context of goodness-of-fit tests. [Chesneau \(2011\)](#) considers the estimation of quadratic functional of f in the cases where the convolution of the signal f with some known function is contaminated by Gaussian noise. Recently, [Schlottenhofer and Johannes \(2020\)](#) provide minimax results for quadratic functional estimation problems in a circular version of Model (1).

It is worth noting the projection method and adaptation issues from Model (1) have not yet been addressed in the literature. In the present article, we delve into how the properties of the *Hermite basis* contribute to resolving the estimation problem of $\theta(f) = \int_{\mathbb{R}} f^2(x)dx$ from Model (1). The Hermite basis is an orthonormal basis on \mathbb{R} , which then leads to the construction of an unconstrained estimator on f 's support. The properties of this basis are described in [Belomestny et al. \(2019\)](#). For instance, the Hermite basis has low complexity, requiring only a few coefficients for accurate estimation. We also address the issue of adaptation, which is a also novel aspect of our work.

Contributions. Our approach involves introducing a new projection estimator constructed from the expansion of f in the Hermite basis. We provide convergence rates for f belonging to the Sobolev-Hermite ball and for various error types. The rates of convergence coincide with the one obtained by [Butucea \(2007\)](#). Moreover, if $\varepsilon_k = 0$ *a.s.*, we recover the classical rate of convergence as described in [Laurent \(2005\)](#). An innovative adaptive procedure inspired by [Goldenshluger and Lepski \(2011\)](#)'s methods to select the relevant dimension of the projection space is presented. We demonstrate that the adaptive estimator satisfies an oracle inequality and achieves the optimal rate, with at least a logarithmic factor. Finally, we illustrate our procedure through numerical experiments.

Organization of paper. In Section 2, we recall the definition of the Hermite basis, regularity spaces, classical assumption on the noise and we describe our methodology. In Section 3, we discuss the rates of convergence over Sobolev Hermite ball. Section 4 is devoted to the adaptive estimation. The results of a simulation study are detailed in Section 5. Finally, all proofs are presented in Section 7, while concentration tools are given in the Appendix.

2. HERMITE BASIS AND REGULARITY SPACES, METHODOLOGY

2.1. Hermite basis and regularity spaces.

2.1.1. *Notation.* For a, b two real numbers, denote $a \vee b = \max(a, b)$, $a \wedge b = \min(a, b)$ and $a_+ = \max(0, a)$. Let z be the complex number, we denote by \bar{z} the conjugate of z . For g and h two functions in $\mathbb{L}^2(\mathbb{R})$, we denote $\langle g, h \rangle = \int_{-\infty}^{+\infty} g(x)\overline{h(x)}dx$ the scalar product on $\mathbb{L}^2(\mathbb{R})$ and $\|g\| = (\int_{-\infty}^{+\infty} |g(x)|^2 dx)^{1/2}$ the norm on $\mathbb{L}^2(\mathbb{R})$. The Fourier transform of g is defined by $g^*(u) = \int e^{iux}g(x)dx$. We recall the Plancherel-Parseval equality $\langle g, h \rangle = (2\pi)^{-1}\langle g^*, h^* \rangle$. Let u_n and v_n be two real sequences, denote $u_n \lesssim v_n$ if there exists a positive constant c such that for all $n \in \mathbb{N}$, $u_n \leq cv_n$ and $u_n \asymp v_n$ if $u_n \lesssim v_n$ and $v_n \lesssim u_n$.

We begin by recalling the definition of the Hermite basis and regularity space.

2.1.2. *The Hermite basis.* Define the Hermite basis $(\varphi_j)_{j \geq 0}$ from Hermite polynomials $(H_j)_{j \geq 0}$ by:

$$(2) \quad \varphi_j(x) = c_j H_j(x) e^{-x^2/2}, \quad H_j(x) = (-1)^j e^{x^2} \frac{d^j}{dx^j} (e^{-x^2}), \quad c_j = (2^j j! \sqrt{\pi})^{-1/2}, \quad x \in \mathbb{R}, j \geq 0.$$

The family $(H_j)_{j \geq 0}$ is orthogonal with respect to the weight function e^{-x^2} : $\int_{\mathbb{R}} H_j(x) H_k(x) e^{-x^2} dx = 2^j j! \sqrt{\pi} \delta_{j,k}$, where $\delta_{k,j}$ is the Kronecker symbol (see [Abramowitz and Stegun \(1964\)](#), Chap. 22.2.14). It follows that the sequence $(\varphi_j)_{j \geq 0}$ is an orthonormal basis on \mathbb{R} such that

$$(3) \quad \|\varphi_j\|_{\infty} = \sup_{x \in \mathbb{R}} |\varphi_j(x)| \leq \phi_0, \quad \text{with } \phi_0 = \pi^{-1/4},$$

(see [Abramowitz and Stegun \(1964\)](#), chap. 22.14.17 and [Indritz \(1961\)](#)). From [Askey and Wainger \(1965\)](#), it yields

$$(4) \quad |\varphi_j(x)| < C'_{\infty} e^{-\xi x^2}, \quad |x| \geq \sqrt{2j+1}, \quad C'_{\infty} > 0,$$

where ξ is a positive constant independent on x and $0 < \xi < \frac{1}{2}$. The Fourier transform of $(\varphi_j)_{j \geq 0}$ satisfies:

$$(5) \quad \varphi_j^* = \sqrt{2\pi} (i)^j \varphi_j.$$

2.1.3. *Regularity spaces.* We consider the regularity spaces associated to the Hermite basis called Sobolev-Hermite spaces. In the sequel, it is used to evaluate the bias of our estimator.

Definition 2.1 (Sobolev-Hermite ball). *For $s > 0$, define the Sobolev-Hermite ball of regularity s and radius D by:*

$$(6) \quad W_H^s(D) = \{\theta \in \mathbb{L}^2(\mathbb{R}), \sum_{k \geq 0} k^s a_k^2(\theta) \leq D\}, \quad a_k(\theta) = \langle \theta, \varphi_k \rangle = \int_{\mathbb{R}} \theta(x) \varphi_k(x) dx.$$

When $s \geq 1$ is an integer, [Bongioanni and Torrea \(2006\)](#) showed that $\theta \in W_H^s(D)$ if and only if θ is s -times differentiable and the functions $\theta, \dots, \theta^{(\ell)}, x \mapsto x^{s-\ell} \theta^{(\ell)}$ belong to $\mathbb{L}^2(\mathbb{R})$ for $\ell = 0, \dots, s-1$.

Recall also the definition of the classical Sobolev ball

Definition 2.2 (Usual Sobolev-ball). *A Sobolev class with smoothness $s > 0$ and radius $D > 0$ is defined by:*

$$W^s(D) = \{\theta \in \mathbb{L}^2(\mathbb{R}), \int_{\mathbb{R}} (1+u^2)^s |\theta^*(u)|^2 du < D\}.$$

For any $s > 0$, we know in [Bongioanni and Torrea \(2006\)](#) that $W_H^s(D) \subset W^s(D')$. Moreover, if $s \geq 1$ is an integer the space $W^s(D)$ is given by:

$$W^s(D) = \{\theta \in \mathbb{L}^2(\mathbb{R}), \theta \text{ is } s\text{-times differentiable and } \sum_{\ell=1}^s \|\theta^{(\ell)}\|^2 < \infty\}.$$

Therefore, if θ has a compact support, $\theta \in W_H^s(D)$ is equivalent to $\theta \in W^s(D')$, where D and D' are related constants.

2.1.4. *Assumption on the noise.* The following assumptions on the noise are considered in convolution context.

(H4) The Fourier transform of the noise density is such that $f_{\varepsilon}^* \neq 0$.

Assume also that: there exist $c_1 \geq c'_1 > 0$, and $\gamma \geq 0, \mu \geq 0, \delta \geq 0$ (with $\gamma > 0$ if $\delta = 0$) such that

$$(7) \quad c'_1 (1+t^2)^{\gamma} e^{\mu|t|^{\delta}} \leq |f_{\varepsilon}^*(t)|^{-2} \leq c_1 (1+t^2)^{\gamma} e^{\mu|t|^{\delta}}, \quad \forall t \in \mathbb{R}.$$

In the sequel, we consider two types of errors.

- If $\delta = 0$, we say that ε and f_ε are ordinary smooth.
- When $\delta > 0$, they are called super smooth noises or functions.

Let us note that Condition (7) implies Assumption (H4) and is fulfilled by some classical densities. Examples of ordinary smooth noises are: Gamma, Laplace and Gamma symmetric distributions. For super smooth noises, we can cite Gaussian, Mixed Gaussian and Cauchy distributions.

Let us now present our methodology.

2.2. Methodology. Let $(Y_k)_{k \in \{1, \dots, n\}}$ be n i.i.d. observations drawn from Model (1). Our purpose is to estimate $\theta(f) = \int_{\mathbb{R}} f^2(x) dx$ from indirect observations Y_1, \dots, Y_n . For any integer $m \geq 1$, define $S_m = \text{Span}\{\varphi_0, \dots, \varphi_{m-1}\}$, the linear spaces generated by $\varphi_0, \dots, \varphi_{m-1}$. Assuming that $f \in \mathbb{L}^2(\mathbb{R})$, we decompose f in an orthonormal basis (here in the Hermite basis) and we define its orthogonal projection on S_m by:

$$(8) \quad f_m = \sum_{j=0}^{m-1} a_j(f) \varphi_j, \quad a_j(f) = \langle f, \varphi_j \rangle = \int_{\mathbb{R}} f(x) \varphi_j(x) dx.$$

Set

$$\theta_m(f) = \|f_m\|^2 = \sum_{j=0}^{m-1} a_j^2(f).$$

We will estimate $\theta_m(f)$. Under (H4) and using Plancherel-Parseval's theorem, we have

$$(9) \quad a_j(f) = \frac{1}{2\pi} \langle \frac{f_Y^*}{f_\varepsilon^*}, \varphi_j^* \rangle.$$

Since $f_Y^*(u) = \int_{\mathbb{R}} e^{iux} f_Y(x) dx = \mathbb{E}[e^{iuY_1}]$, we estimate $f_Y^*(u)$ by:

$$(10) \quad \hat{f}_Y^*(u) = \frac{1}{n} \sum_{k=1}^n e^{iuY_k}.$$

Consider the random variables $(U_t^*(Y_k))_{k \in \{1, \dots, n\}}$ defined as:

$$(11) \quad U_t^*(Y_k) = \frac{1}{2\pi} \langle \frac{e^{iY_k}}{f_\varepsilon^*}, t^* \rangle, \quad U_t(x) = \frac{1}{2\pi} \frac{t^*(-x)}{f_\varepsilon^*(x)}.$$

Hence, injecting (10) into (9) and assuming that the ratio $\varphi_j/f_\varepsilon^*$ is integrable over \mathbb{R} for $j \in \{0, \dots, m-1\}$, we define the following estimator:

$$(12) \quad \hat{\theta}_m(f) = \sum_{j=0}^{m-1} \widehat{a_j^2(f)}, \quad \text{where} \quad \widehat{a_j^2(f)} = \frac{1}{n(n-1)} \sum_{k \neq \ell=1}^n U_{\varphi_j}^*(Y_k) U_{\varphi_j}^*(Y_\ell).$$

The Hermite basis ensures that the ratio $\varphi_j/f_\varepsilon^*$ is integrable in many cases, in particular, for any ordinary smooth noise or super smooth noise with $\delta \leq 2$. Consequently, $\hat{\theta}_m(f)$ is well defined for several function classes satisfying Condition (7). Furthermore, by the Plancherel-Parseval theorem, it is evident that $\mathbb{E}[U_{\varphi_j}^*(Y_k)] = a_j(f)$. Hence, $\hat{\theta}_m(f)$ is an unbiased estimator of $\theta_m(f) = \sum_{j=0}^{m-1} a_j^2(f)$. Moreover, for any real function t , it holds that $\overline{U_t^*(x)} = U_t^*(x)$. Thus, our estimator $\hat{\theta}_m(f)$ is a real random variable, which is crucial since we are estimating a real number.

3. RATES OF CONVERGENCE

3.1. Risk bound for fixed m . Let $m \geq 1$ an integer and $\rho > 0$ an absolute constant. Define

$$(13) \quad \Delta(m) := \int_{|u| \leq \sqrt{\rho m}} \frac{du}{|f_\varepsilon^*(u)|^2}, \quad \Lambda(m) := \sum_{j=0}^{m-1} \int_{|u| > \sqrt{\rho m}} \frac{|\varphi_j(u)|^2}{|f_\varepsilon^*(u)|^2} du,$$

and

$$(14) \quad \Delta^{(1)}(m) := \sup_{|u| \leq \sqrt{\rho m}} |f_\varepsilon^*(u)|^{-2}.$$

We can establish the following bound on the quadratic risk.

Proposition 3.1. *Let $(Y_k)_{k \in \{1, \dots, n\}}$ be n i.i.d. observations drawn from Model (1) and assumptions (H1) to (H4) hold. Assume that f is bounded ($\|f\|_\infty < \infty$) and f_ε is square integrable. Consider the estimator $\hat{\theta}_m(f)$ defined in (12).*

(i) *Then, for any $\rho > 0$ and $C_0 = 4\|f\|_\infty^{\frac{1}{2}}\|f_\varepsilon\| < \infty$, it holds*

$$\begin{aligned} \mathbb{E}[(\hat{\theta}_m(f) - \theta(f))^2] &\leq \|f - f_m\|^4 + \frac{C_0}{n(n-1)} \Delta(m) \left(\Delta^{(1)}(m) + \Lambda(m) \right) + \frac{C_0^2}{n(n-1)} \Lambda^2(m) \\ &\quad + \frac{4\|f\|_\infty^2}{n(n-1)} + \frac{4}{n} \mathbb{E}[|U_{f_m}^*(Y_1)|^2]. \end{aligned}$$

(ii) *Moreover, if $\rho \geq 2$ and f_ε satisfies (7) with $0 \leq \delta < 2$ or ($\delta = 2$, with $\mu < \xi$), where ξ is defined in (4), then, we get*

$$\mathbb{E}[(\hat{\theta}_m(f) - \theta(f))^2] \leq \|f - f_m\|^4 + \frac{C_1}{n(n-1)} \Delta(m) \Delta^{(1)}(m) + \frac{4}{n} \mathbb{E}[|U_{f_m}^*(Y_1)|^2] + \frac{C_2}{n},$$

where C_1 and C_2 are constants depending on C_0 , $\|f_Y\|_\infty$, ρ and constants appearing in (7)-(4).

Let us make some comments on the bounds obtained in Proposition 3.1. The first term in the right-hand side ($\|f - f_m\|^4$) is the classical bias term: it is decreasing with m . The others are variance terms. It has been established (refer to Proposition 3.1 in Sacko (2020)) that $\Lambda(m)$ becomes negligible under Condition (7). More precisely, under the assumptions outlined in Proposition 3.1 (ii), we have:

$$(15) \quad \Lambda(m) = \sum_{j=0}^{m-1} \int_{|u| > \sqrt{\rho m}} \frac{|\varphi_j(u)|^2}{|f_\varepsilon^*(u)|^2} du \lesssim e^{-\frac{\xi \rho m}{2}}.$$

Equation (15) is a key property of the Hermite basis and what makes it so relevant in the context of inverse problems. Consequently, the main variance term is $\Delta(m)\Delta^{(1)}(m)/n(n-1) + \mathbb{E}[|U_{f_m}^*(Y_1)|^2]/n$. The first is clearly increasing with m contrary of $\mathbb{E}[|U_{f_m}^*(Y_1)|^2]/n$, which is a mixed of bias and variance (see Lemma 3.1). It can be negligible for ordinary smooth noises with an adequate choice of m . The examination of $\mathbb{E}[|U_{f_m}^*(Y_1)|^2]$ entails technical intricacies. Let us proceed to bound it in two distinct ways. Consider the following assumption:

(H5) The density f_Y is bounded.

Observe that if $(f$ and f_ε belong to $\mathbb{L}^2(\mathbb{R})$) or $(f$ or f_ε is bounded), (H5) is automatically satisfied.

Lemma 3.1. *Assume that (H5) holds, $\|f\|_\infty < \infty$ and set $C'_0 = \|f\|_\infty \|f_Y\|_\infty < \infty$.*

(i) For any $m \geq 1$, it holds

$$\mathbb{E}[|U_{f_m}^*(Y_1)|^2] \leq 2\|f_Y\|_\infty \left(\Delta^{(1)}(m)\|f - f_m\|^2 + \frac{1}{2\pi} \int_{|u| \leq \sqrt{\rho m}} \frac{|f^*(u)|^2}{|f_\varepsilon^*(u)|^2} du \right) + C'_0 \Lambda(m).$$

(ii) Furthermore, it yields

$$\mathbb{E}[|U_{f_m}^*(Y_1)|^2] \leq C'_0 \left(\Delta^{(1)}(m) + \Lambda(m) \right).$$

Lemma 3.1 (i) is the key to achieve the parametric rate: $1/n$. It also indicates that the term $\mathbb{E}[|U_{f_m}^*(Y_1)|^2]/n$ will influence the rate of convergence if f_ε is ordinary smooth. Moreover, using

$$2 \frac{\|f_Y\|_\infty \Delta^{(1)}(m)}{n} \|f - f_m\|^2 \leq \|f - f_m\|^4 + \frac{\|f_Y\|_\infty^2 (\Delta^{(1)}(m))^2}{n^2},$$

it holds by Lemma 3.1 (i):

$$(16) \quad \frac{\mathbb{E}[|U_{f_m}^*(Y_1)|^2]}{n} \leq \|f - f_m\|^4 + \frac{\|f_Y\|_\infty^2}{n(n-1)} \Delta(m) \Delta^{(1)}(m) + \frac{\|f_Y\|_\infty}{\pi n} \int_{|u| \leq \sqrt{\rho m}} \frac{|f^*(u)|^2}{|f_\varepsilon^*(u)|^2} du + \frac{C'_0}{n} \Lambda(m).$$

The first two terms are the same as in Proposition 3.1 (ii) with two additional terms $\frac{\|f_Y\|_\infty}{\pi n} \int_{|u| \leq \sqrt{\rho m}} \frac{|f^*(u)|^2}{|f_\varepsilon^*(u)|^2} du + \frac{C'_0}{n} \Lambda(m)$. Under regularity conditions on f , the term $\frac{1}{n} \int_{|u| \leq \sqrt{\rho m}} \frac{|f^*(u)|^2}{|f_\varepsilon^*(u)|^2} du$ is negligible for an adequate choice of m compared to $\Delta(m) \Delta^{(1)}(m)/n(n-1)$. The last term $\Lambda(m)/n$ has order $1/n$ for any value of $m \in \mathbb{N}$ and is therefore negligible. The second part is applicable in the super smooth case and show that the rate depends on $\Delta^{(1)}(m)/n$.

Remark 1 (Rate of convergence from direct observations). *When X_k are available, $\varepsilon_k = 0$ a.s. and the terms $\Delta^{(1)}(m)$ and $\Lambda(m)$ are bounded uniformly in m . Thus, $\mathbb{E}[|U_{f_m}^*(Y_1)|^2]/n$ is a residual term compared to the term $C_1 \Delta(m) \Delta^{(1)}(m)/n(n-1) \lesssim \sqrt{m}/n(n-1)$. For $f \in W_H^s(L)$ (see Definition 2.1), we recover the rate $n^{-8s/(4s+1)} \vee n^{-1}$ computed respectively over Lipschitz and Besov regularity spaces by [Bickel and Ritov \(1988\)](#), [Laurent \(2005\)](#). Our results contain the case of direct observation.*

3.2. Rate of convergence on Sobolev-Hermite ball. In order to derive rates of convergence, we consider the Sobolev-Hermite regularity spaces, defined in Definition 2.1.

Proposition 3.2 (Order of bias and variance). *Let $(Y_k)_{k \in \{1, \dots, n\}}$ be i.i.d. observations drawn from Model (1) and assumptions (H1) to (H5) hold. Assume that $\|f\|_\infty < \infty$, f belongs to $W_H^s(L)$ and f_ε is square integrable. Consider $\hat{\theta}_m(f)$ be defined in (12).*

(i) If $\rho \geq 2$ and f_ε is ordinary smooth (that is f_ε satisfies (7) with $\delta = 0$), we have

$$\mathbb{E}[(\hat{\theta}_m(f) - \theta(f))^2] \leq 2Lm^{-2s} + \mathfrak{C} \left(\frac{1}{n(n-1)} m^{2\gamma + \frac{1}{2}} + \frac{1}{n} m^{-(s-\gamma)} \mathbf{1}_{s \leq \gamma} + \frac{1}{n} \right),$$

where \mathfrak{C} is a constant depending on f_ε , L , $\|f\|_\infty$ and $\|f_Y\|_\infty$.

(ii) If $\rho \geq 2$ and f_ε is super smooth (that is f_ε satisfies (7) with $0 \leq \delta < 2$ or $(\delta = 2, \text{ with } \mu < \xi)$), where ξ is defined in (4), it holds

$$\mathbb{E}[(\hat{\theta}_m(f) - \theta(f))^2] \leq Lm^{-2s} + \mathfrak{C}_0 \left(\frac{1}{n(n-1)} m^{2\gamma + \frac{1}{2}} e^{2\mu(m\rho)^{\frac{\delta}{2}}} + \frac{1}{n} m^\gamma e^{\mu(m\rho)^{\frac{\delta}{2}}} + \frac{1}{n} \right),$$

where \mathfrak{C}_0 is also a constant depending on f_ε , L , $\|f\|_\infty$ and $\|f_Y\|_\infty$.

Proposition 3.2 is a consequence of Proposition 3.1 and Lemma 3.1. By realizing a bias-variance compromise with respect to m in the risk bound in Proposition 3.2, we obtain rates of convergences.

Theorem 3.3 (Rate of convergence for ordinary smooth noise). *Let $(Y_k)_{k \in \{1, \dots, n\}}$ be i.i.d. observations from drawn Model (1) and assumptions (H1) to (H5) hold. Suppose that $\|f\|_\infty < \infty$, f belongs to $W_H^s(L)$, and f_ε is square integrable and ordinary smooth. Consider $\hat{\theta}_m(f)$ be defined in (12). Take $m_{opt} = \lceil n^{\frac{4}{4s+4\gamma+1}} \rceil$. Then, we have*

$$(17) \quad \sup_{f \in W_H^s(L)} \mathbb{E}[(\hat{\theta}_{m_{opt}}(f) - \theta(f))^2] \lesssim \begin{cases} n^{-\frac{8s}{4s+4\gamma+1}} & \text{if } s < \gamma + \frac{1}{4}, \\ n^{-1} & \text{if } s \geq \gamma + \frac{1}{4}. \end{cases}$$

We attain the classical rate in the deconvolution scenario when f_ε exhibits ordinary smoothness, see Butucea (2007). This rate is known to be optimal under additional condition on the derivative of f_ε^* and for densities belonging to the classical Sobolev space. If $\gamma = 0$, corresponding to direct observation case, we achieve the rate as $n^{-8s/(4s+1)} \vee n^{-1}$, see Bickel and Ritov (1988), Laurent (2005).

Theorem 3.4 (Rate of convergence for super smooth noise). *Let $(Y_k)_{k \in \{1, \dots, n\}}$ be n i.i.d. observations drawn from Model (1) and assumptions (H1) to (H5) hold. Suppose that $\|f\|_\infty < \infty$ and f belongs to $W_H^s(L)$, and f_ε is square integrable and super smooth with $0 < \delta \leq 2$. Let $\hat{\theta}_m(f)$ be defined in (12). By selecting $m_{opt} = \lceil \frac{1}{\rho} (\frac{\log n}{4\mu})^{\frac{2}{\delta}} \rceil$, we obtain the following bound*

$$\sup_{f \in W_H^s(L)} \mathbb{E}[(\hat{\theta}_{m_{opt}}(f) - \theta(f))^2] \lesssim (\log n)^{-\frac{4s}{\delta}}.$$

If f_ε is super smooth, we attain the optimal rate, computed over the classical Sobolev spaces $(\log n)^{-\frac{4s}{\delta}}$ (see Butucea (2007)). For the lower bound, Butucea (2007) considers test functions with compact support. Thus, the rate $(\log n)^{-\frac{4s}{\delta}}$ also represents the optimal rate over Sobolev-Hermite spaces, as these spaces coincide with the ones mentioned above (see Section 2.1.3) for compactly supported functions. Let us summarize the previous rates in Table 1. Let us note that the rate obtained for the classes of

| | $\delta = 0$ | $0 < \delta \leq 2$ |
|------|--|---------------------------------|
| | ordinary smooth | super smooth |
| rate | $n^{-\frac{8s}{4s+4\gamma+1}} \vee n^{-1}$ | $(\log n)^{-\frac{4s}{\delta}}$ |

TABLE 1. Rate of convergence for $\mathbb{E}[(\hat{\theta}_{m_{opt}}(f) - \theta(f))^2]$ if $f \in W_H^s(D)$.

mean mixture or variance mixtures of the Gaussian distribution in Sacko (2020) can be extended in our context. Specifically, if f is a Gaussian density and f_ε is ordinary smooth, we deduce from Proposition 3.1 and Lemma 3.1 that:

$$\sup_{f \in W_H^s(L)} \mathbb{E}[(\hat{\theta}_{m_{opt}}(f) - \theta(f))^2] \lesssim \frac{\log^{2\gamma+\frac{1}{2}}(n)}{n^2} \vee \frac{\log^{\gamma-s}(n)}{n} \mathbf{1}_{\gamma>s},$$

where $m_{opt} \asymp \log(n)$.

When f and f_ε are both exponentially decaying, the rate can be better than any power of logarithm. For instance, if $f(x) = \exp(-x^2/(2\sigma^2))$ and $g(x) = \exp(-x^2/(2\zeta^2))$, selecting $m_{opt} = (\sigma^2 + \zeta^2)^{-1} \log(n)$ yields $\sup_{f \in W_H^s(L)} \mathbb{E}[(\hat{\theta}_{m_{opt}}(f) - \theta(f))^2] \lesssim n^{-\frac{\sigma^2}{\sigma^2 + \zeta^2}}$.

4. ADAPTIVE ESTIMATION

In this section, we propose a selection procedure for the estimator $\hat{\theta}_m(f)$. First, let us note that the adaptive procedure from Model (1) has not been explored in the existing literature. We introduce a novel approach inspired by the methods presented in Goldenshluger and Lepski (2011). The method is based on the comparison of estimators of $\hat{\theta}_m(f)$. From now on, let us assume that the constant $\rho \geq 2$ given in (13) is fixed. Consider the following collections of model:

$$\mathcal{M}_n := \left\{ 1 \leq m \leq \lfloor \sqrt{n} \rfloor, \quad \Delta(m)\Delta^{(1)}(m) \leq \frac{n(n-1)}{\log^2(n)}, \quad \Delta^{(1)}(m) \leq \frac{n}{\log(n)} \right\},$$

where $\Delta(m)$ and $\Delta^{(1)}(m)$ are respectively given in (13)-(14). We aim to identify the $m \in \mathcal{M}_n$ that minimizes the bias-variance decomposition of $\hat{\theta}_m(f)$ given in Proposition 3.1(ii). The bias $\|f - f_m\|^4$ is estimated by:

$$(18) \quad \hat{A}(m) = \sup_{m' \in \mathcal{M}_n} \left\{ \left((\hat{\theta}_{m'}(f) - \hat{\theta}_{m \wedge m'}(f))^2 - \kappa_1 V(m') \right)_+ \right\},$$

where $V(m) = V_1(m) + V_2(m)$ is the order of variance term up to $\log(n)$ factor:

$$(19) \quad V_1(m) = \left(\Delta(m)\Delta^{(1)}(m) \right) \frac{\log^2(n)}{n(n-1)},$$

and

$$(20) \quad V_2(m) = \begin{cases} \left(\int_{|u| \leq \sqrt{\rho m}} \frac{|f_m(u)|^2}{|f_\varepsilon^*(u)|^2} du \right) \frac{\log(n)}{n} & \text{if } f_\varepsilon \text{ is ordinary smooth,} \\ \Delta^{(1)}(m) \frac{\log(n)}{n} & \text{if } f_\varepsilon \text{ is super smooth.} \end{cases}$$

Here, κ_1 is a numerical constant to be calibrated.

We shall prove that (see Proof of Theorem 4.1 provided below) for any given noise.

$$\mathbb{E}[\hat{A}(m)] \leq 3\|f - f_m\|^2 + \text{Var}(\hat{\theta}_m(f)) + \mathcal{O}\left(\frac{1}{n}\right).$$

The presence of $\text{Var}(\hat{\theta}_m(f))$ is related to the term $V_2(m)$, which is not necessarily increasing with m in the ordinary smooth case and depends on $|f_m(u)|^2$. Note that $|f_m(u)|^2$ will be replaced later by an estimator (see Section 5). For super-smooth noises, we take an upper bound on $\int_{\sqrt{\rho m}} \frac{|f_m(u)|^2}{|f_\varepsilon^*(u)|^2} du$:

$$\int_{\sqrt{\rho m}} \frac{|f_m(u)|^2}{|f_\varepsilon^*(u)|^2} du \leq \Delta^{(1)}(m) \|f_m\|^2,$$

and this is why $V_2(m) = \Delta^{(1)}(m) \log(n)/n$ in this case. Moreover, in this case, we have

$$\mathbb{E}[\hat{A}(m)] \leq 3\|f - f_m\|^2 + \mathcal{O}\left(\frac{1}{n}\right).$$

Select:

$$(21) \quad \hat{m}_n = \arg \min_{m \in \mathcal{M}_n} \left\{ \hat{A}(m) + \kappa_2 V(m) \right\},$$

where $\kappa_2 \geq \kappa_1 > 0$ is also a numerical constant which must be calibrated.

We can prove the following oracle result.

Theorem 4.1 (Oracle inequality). *Let $(Y_k)_{k \in \{1, \dots, n\}}$ be n i.i.d. observations drawn from Model (1) and assumptions (H1) to (H5) hold. Suppose that $\|f\|_\infty < \infty$, f_ε is square integrable and Condition (7) is fulfilled. Let $\hat{\theta}_{\hat{m}_n}(f)$ be defined in (12) with \hat{m}_n selected in (21). There exists a constant κ_0 such that for any $\kappa_1 \geq \kappa_0$:*

$$\mathbb{E}[(\hat{\theta}_{\hat{m}_n}(f) - \theta(f))^2] \leq C \inf_{m \in \mathcal{M}_n} \left(\|f - f_m\|^4 + V(m) + \frac{\mathbb{E}[|U_{f_m}^*(Y_1)|^2]}{n} \right) + \frac{C'}{n},$$

where $C = \max(36, (6\kappa_2 + 3C_1))$ with C_1 is given Proposition 3.1 and C' is a numerical constant depending on f_ε , $\|f\|_\infty$ and $\|f_Y\|_\infty$.

The term $\mathbb{E}[|U_{f_m}^*(Y_1)|^2]/n$ has the same order as $\|f - f_m\|^4 + V(m)$, see (16) under adequate regularity and choice of m . The bound given in Theorem 4.1 is non asymptotic and shows that the estimator $\hat{\theta}_{\hat{m}_n}(f)$ realizes automatically a bias-variance trade-off up to $\log(n)$ terms in ordinary smooth cases. Unsurprisingly, we recover the optimal rate for super smooth noises. Indeed, it is known that a logarithm factor in the variance does not affect the rate of convergence. Combining Proposition 3.1 and Lemma 3.1, we get

$$\sup_{f \in W_H^s(L)} \mathbb{E}[(\hat{\theta}_{\hat{m}_n}(f) - \theta(f))^2] \lesssim \begin{cases} \left(\frac{n}{\log(n)}\right)^{-\frac{8s}{4s+4\gamma+1}} \vee \left(\frac{n}{\log(n)}\right)^{-1} & \text{if } f_\varepsilon \text{ is ordinary smooth,} \\ (\log n)^{-\frac{4s}{\delta}} & \text{if } f_\varepsilon \text{ is super smooth.} \end{cases}$$

Note that the logarithm factor is due to Goldenschluger and Lepski (2011)'s procedure. In addition, when X_k are available ($\gamma = 0$), we recover the rate $n^{-8s/(4s+1)} \vee n^{-1}$ up to logarithm factor contrary to Laurent (2005)'s where the rate n^{-1} can be achieved for n large enough.

The value κ_0 is not explicit here and depends on the unknown quantities such as $\|f\|_\infty$, $\|f_Y\|$. This is due to the control of our U-statistic of order 2 and to the quantity $\Lambda(m)$, which is bounded by a non explicit constant, see (15). In practice, we set $\kappa_2 = 2\kappa_1$ and calibrate κ_1 through preliminary simulation experiments.

5. NUMERICAL STUDY

5.1. Implementation. In this section, we illustrate the performance of our methodology by implementing the estimator $\hat{\theta}_m(f)$ defined in (12) for both direct and inverse problem scenarios. Let us consider the following distributions for the signal

- (i) $\mathcal{N}(0, 1)$, $\theta(f) = \frac{1}{2\sqrt{\pi}} \approx 0.2820948$,
- (ii) Gamma $\mathcal{G}(\alpha, \beta)$, with $\alpha = 1$ and $\beta = 4$, $\theta(f) = \left(\frac{\beta^\alpha}{\Gamma(\alpha)}\right)^2 \frac{\Gamma(2\alpha-1)}{(2\beta)^{2\alpha-1}} \approx 0.15625$,
- (ii) Laplace $\mathcal{L}(\sqrt{2})$, $\theta(f) = \frac{1}{2\sqrt{2}} \approx 0.3535534$,
- (iv) Cauchy standard, $\theta(f) = \frac{1}{2\pi} \approx 0.1591549$.

Set $\rho = 4$. For the noise, we distinguish three cases: direct case (that is $\varepsilon = 0$ a.s.), ordinary smooth and super-smooth cases.

- **The direct and ordinary smooth cases.** We choose the following estimator of the variance $V(m)$:

$$\tilde{V}(m) = \Delta(m)\Delta^{(1)}(m)\frac{\log^2(n)}{n(n-1)} + \tilde{V}_2(m), \quad \rho = 3,$$

where $\tilde{V}_2(m)$ estimates $V_2(m)$ given in (19). We tested two estimators of $\tilde{V}_2(m)$. They are all based on the estimation of $|f_m(u)|^2$. The first is an unbiased estimator of $|f_m(u)|^2$:

$$(22) \quad |\widehat{f_m(u)}|^2 = \sum_{j,j'=0}^{m-1} \widehat{a_j a_{j'}} \varphi_j(u) \varphi_{j'}(u), \quad \widehat{a_j a_{j'}} = \frac{1}{n(n-1)} \sum_{\ell \neq k=1}^n U_{\varphi_j}^*(Y_k) U_{\varphi_{j'}}^*(Y_\ell),$$

where U_t is given (11). The second estimator has a bias but is much faster to compute. It is obtained by replacing $f_m(u)$ by the *Hermite density estimator* studied in Sacko (2020). It is defined as:

$$\hat{f}_m(u) = \sum_{j=0}^{m-1} \hat{a}_j \varphi_j(u), \quad \hat{a}_j = \frac{1}{n} \sum_{i=1}^n U_{\varphi_j}^*(Y_i).$$

Both estimators exhibit equivalent performances. Then, we only implement the following estimator of $V_2(m)$ for risk evaluations:

$$\tilde{V}_2(m) = \left(\int_{|u| \leq \sqrt{\rho m}} \frac{|\hat{f}_m(u)|^2}{|f_\varepsilon^*(u)|^2} du \right) \frac{\log(n)}{n}.$$

This estimator is obtained by substituting $|f_m(u)|^2$ with $|\hat{f}_m(u)|^2$ in (20). Set $\varepsilon = 0$ *a.s.* in the direct case and consider the Laplace noise in the ordinary smooth case with a density given by:

$$f_\varepsilon(x) = \frac{\lambda}{2} e^{-\lambda|x|}; \quad f_\varepsilon^*(x) = \frac{\lambda^2}{\lambda^2 + x^2}; \quad \lambda = 2\sqrt{5}.$$

For the last case, $\Delta(m)$ and $\Delta^{(1)}(m)$ are respectively given by:

$$\Delta(m) = \int_{|u| \leq \sqrt{\rho m}} \left(1 + \frac{u^2}{\lambda^2}\right)^2 du = \left(\sqrt{\rho m} + \frac{2}{3\lambda^2} (\sqrt{\rho m})^3 + \frac{(\sqrt{\rho m})^5}{5\lambda^4}\right)$$

and

$$\Delta^{(1)}(m) = \left(1 + \frac{\rho m}{\lambda^2}\right)^2.$$

- **The super smooth case.** Take a Gaussian noise with density:

$$f_\varepsilon(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-x^2/2\sigma_\varepsilon^2}, \quad f_\varepsilon^*(x) = e^{-\sigma_\varepsilon^2 x^2/2}, \quad \sigma_\varepsilon^2 = 1/10.$$

The variance is defined in (19) with

$$\Delta(m) = 2\sqrt{\rho m} \left(\int_0^1 e^{u^2 \sigma_\varepsilon^2 \rho m} du \right), \quad \Delta^{(1)}(m) = e^{\sigma_\varepsilon^2 \rho m},$$

where $\int_0^1 e^{u^2 \sigma_\varepsilon^2 \rho m} du$ is computed by a Riemann sum discretized in 100 points.

The adaptive method is described in three steps:

- For any $m \in \mathcal{M}_n$, compute $\hat{A}(m) = \sup_{m' \in \mathcal{M}_n} \left\{ \left((\hat{\theta}_{m'}(f) - \hat{\theta}_{m \wedge m'}(f))^2 - \kappa_1 \text{Va}(m') \right)_+ \right\}$, where $\hat{\theta}_m(f)$ is defined in (12) and $\text{Va}(m)$ corresponds to the variance term which is specified above.
- Choose \hat{m}_n via $\hat{m}_n = \arg \min_{m \in \mathcal{M}_n} \left\{ \hat{A}(m) + \kappa_2 \text{Va}(m) \right\}$,
- Compute $\hat{\theta}_{\hat{m}_n}(f) = \sum_{j=0}^{\hat{m}_n-1} \widehat{a_j^2}(f)$, with $\widehat{a_j^2}(f)$ given by (12).

Let us mention that the coefficient $(\widehat{a_j^2(f)})_{0 \leq j \leq m-1}$ is computed by Riemann's approximation in the inverse problem case.

Calibration of κ_1 and κ_2 . Following the idea developed by [Lacour and Massart \(2016\)](#), we set $\kappa_2 = 2\kappa_1$ and $\kappa_1 = 10^{-3}$ in the direct case, $\kappa_1 = 3 \times 10^{-3}$ in the Laplace noise case, and $\kappa_1 = 8 \times 10^{-4}$ in the Gaussian noise case.

5.2. Numerical results. In Table 2, we present simulation results. The first line corresponds to the values of the MSE with standard deviation in parentheses, multiplied by 1000, computed over 200 independent simulations. In the second line, we provide the average of \widehat{m}_n selected by the method, and in the last line, we give the average of $\widehat{\theta}_{\widehat{m}_n}$ for each density. The true unknown parameter $\theta(f)$, represented in bold is provided in the first column of Table 2. We observe that increasing the value of n leads to a smaller MSE, thereby improving the estimation. In addition, we note that the MSE obtained in the case of $\varepsilon = 0$ is smaller than that in the indirect cases. In particular, the MSE obtained in the Laplace noise case is smaller than in the Gaussian noise case. This observation is consistent with theoretical results.

| $\theta(f)$ \backslash Noise | $n = 100$ | | | $n = 500$ | | |
|------------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|--------------------------|
| | $\varepsilon = 0$ | Lap. | Gauss. | $\varepsilon = 0$ | Lap. | Gauss. |
| 0.282... | 0.489 _(0.655) | 1.411 _(2.185) | 2.378 _(6.365) | 0.092 _(0.123) | 0.529 _(0.496) | 0.616 _(0.512) |
| | 6.455 | 3.050 | 4.730 | 2.350 | 1.440 | 1.615 |
| | 0.283 | 0.306 | 0.308 | 0.282 | 0.302 | 0.301 |
| 0.156... | 0.259 _(0.461) | 0.751 _(0.915) | 1.237 _(1.172) | 0.037 _(0.053) | 0.061 _(0.078) | 0.225 _(0.237) |
| | 21.425 | 9.635 | 8.49 | 26.090 | 13.125 | 10.975 |
| | 0.154 | 0.141 | 0.130 | 0.154 | 0.156 | 0.145 |
| 0.354... | 1.700 _(2.343) | 3.322 _(6.592) | 4.009 _(9.443) | 0.314 _(0.475) | 0.633 _(0.833) | 0.780 _(1.189) |
| | 9.005 | 4.095 | 5.085 | 12.295 | 5.260 | 5.54 |
| | 0.351 | 0.375 | 0.367 | 0.351 | 0.366 | 0.368 |
| 0.159... | 0.594 _(0.814) | 0.939 _(1.459) | 0.993 _(1.833) | 0.100 _(0.118) | 0.351 _(0.437) | 0.309 _(0.383) |
| | 4.995 | 2.390 | 3.755 | 6.605 | 1.440 | 3.715 |
| | 0.160 | 0.174 | 0.175 | 0.158 | 0.173 | 0.173 |

TABLE 2. First line: empirical MSE $1000 \times \mathbb{E}[(\widehat{\theta}_{\widehat{m}_n}(f) - \theta(f))^2]$ (with $1000 \times \text{sd}$); second line: mean of \widehat{m}_n ; third line: mean of $\widehat{\theta}_{\widehat{m}_n}(f)$ computed over 200 independent simulations.

6. CONCLUDING REMARKS

In this paper, we propose a new estimation procedure of a quadratic functional of a density based on the expansion of f into Hermite basis from Model (1). We demonstrate that our estimator achieves the optimal rate obtained in [Butucea \(2007\)](#). An adaptive procedure to select the relevant dimension is proposed, and we establish non asymptotic oracle inequality for the resulting estimator. Numerical experiments illustrate the convergence and good performance of our methodology.

To conclude, we outline some perspectives for future work:

- In this study, we consider that $(X_i)_{i \in 1, \dots, n}$ are i.i.d. One possible extension is to investigate the case of dependent X_i . For instance, we can replace Assumption (H1) by: $(X_i)_{i \geq 1}$ is strictly stationary and β -mixing.
- We can also account for uncertainty in the noise distribution, i.e., study the case where the density f_ε is unknown but estimated through an additional sample.

- One of the motivations for this work is the issue of goodness-of-fit testing. We can investigate whether our results extend to this framework.

These points will be studied in future work.

7. PROOFS

7.1. Proof of Proposition 3.1.

7.1.1. *Proof of Part (i).* We have

$$(23) \quad \mathbb{E}[(\widehat{\theta}_m(f) - \theta(f))^2] = \|f - f_m\|^4 + \text{Var}(\widehat{\theta}_m(f)).$$

Let us bound $\text{Var}(\widehat{\theta}_m(f))$. By definition of $\widehat{\theta}_m(f)$ given in (12), it holds

$$(24) \quad \text{Var}(\widehat{\theta}_m(f)) = \sum_{j,k=0}^{m-1} \text{Cov}(\widehat{a_j^2(f)}, \widehat{a_k^2(f)}) = \sum_{j,k=0}^{m-1} \mathbb{E} \left[\left(\widehat{a_j^2(f)} - \mathbb{E}[\widehat{a_j^2(f)}] \right) \left(\widehat{a_k^2(f)} - \mathbb{E}[\widehat{a_k^2(f)}] \right) \right].$$

The following decomposition holds:

$$(25) \quad \begin{aligned} \left(\widehat{a_j^2(f)} - \mathbb{E}[\widehat{a_j^2(f)}] \right) &= \frac{1}{n(n-1)} \sum_{\ell \neq \ell'=1}^n \frac{1}{4\pi^2} \langle \frac{e^{i \cdot Y_\ell}}{f_\varepsilon^*}, \varphi_j^* \rangle \langle \frac{e^{i \cdot Y_{\ell'}}}{f_\varepsilon^*}, \varphi_j^* \rangle - a_j^2(f) \\ &= \frac{1}{n(n-1)} \sum_{\ell \neq \ell'=1}^n \left(U_{\varphi_j}^*(Y_\ell) - \mathbb{E}[U_{\varphi_j}^*(Y_\ell)] \right) \left(U_{\varphi_j}^*(Y_{\ell'}) - \mathbb{E}[U_{\varphi_j}^*(Y_{\ell'})] \right) \\ &\quad + \frac{2}{n} a_j(f) \sum_{\ell=1}^n \left(U_{\varphi_j}^*(Y_\ell) - \mathbb{E}[U_{\varphi_j}^*(Y_\ell)] \right) \\ &= T_{1,j} + T_{2,j}. \end{aligned}$$

It follows that

$$\text{Var}(\widehat{\theta}_m(f)) = \sum_{j,k=0}^{m-1} \text{Cov}(T_{1,j}, T_{1,k}) + 2 \sum_{j,k=0}^{m-1} \text{Cov}(T_{1,j}, T_{2,k}) + \sum_{j,k=0}^{m-1} \text{Cov}(T_{2,j}, T_{2,k}).$$

We state the following Lemma.

Lemma 7.1. *Let j and k be two positive integers. Let U_t be defined in (11). It yields*

$$\text{Cov}(T_{1,j}, T_{1,k}) = \mathbb{E}[T_{1,j}T_{1,k}] = \frac{2}{n(n-1)} \left(\text{Cov} \left(U_{\varphi_j}^*(Y_1), U_{\varphi_k}^*(Y_1) \right) \right)^2,$$

$$\text{Cov}(T_{1,j}, T_{2,k}) = \mathbb{E}[T_{1,j}T_{2,k}] = 0,$$

$$\text{Cov}(T_{2,j}, T_{2,k}) = \mathbb{E}[T_{2,j}T_{2,k}] = \frac{4}{n} a_j(f) a_k(f) \text{Cov}(U_{\varphi_j}^*(Y_1), U_{\varphi_k}^*(Y_1)).$$

Applying Lemma 7.1, we derive that

$$(26) \quad \text{Var}(\widehat{\theta}_m(f)) = \frac{2}{n(n-1)} \sum_{j,k=0}^{m-1} \left(\text{Cov} \left(U_{\varphi_j}^*(Y_1), U_{\varphi_k}^*(Y_1) \right) \right)^2 + \frac{4}{n} \text{Var}(U_{f_m}^*(Y_1)).$$

We bound the $\sum_{j,k=0}^{m-1} \left(\text{Cov} \left(U_{\varphi_j}^*(Y_1), U_{\varphi_k}^*(Y_1) \right) \right)^2$. Let us write

$$\sum_{j,k=0}^{m-1} \left(\text{Cov} \left(U_{\varphi_j}^*(Y_1), U_{\varphi_k}^*(Y_1) \right) \right)^2 \leq 2 \sum_{j,k=0}^{m-1} \left| \mathbb{E} \left[U_{\varphi_j}^*(Y_1) U_{\varphi_k}^*(Y_1) \right] \right|^2 + 2\|f\|^4.$$

Moreover, by definition of U_t^* given (11) and from (5), we have

$$2 \sum_{j,k=0}^{m-1} \left| \mathbb{E} \left[U_{\varphi_j}^*(Y_1) U_{\varphi_k}^*(Y_1) \right] \right|^2 = \frac{1}{\pi^2} \sum_{j,k=0}^{m-1} \left| \iint \frac{\varphi_j(x)}{f_\varepsilon^*(x)} \frac{\varphi_k(y)}{f_\varepsilon^*(y)} f_Y^*(x+y) dx dy \right|^2.$$

We decompose the integrals in 3 pieces:

$$\begin{aligned} 2 \sum_{j,k=0}^{m-1} \left| \mathbb{E} \left[U_{\varphi_j}^*(Y_1) U_{\varphi_k}^*(Y_1) \right] \right|^2 &= \frac{1}{\pi} \sum_{j,k=0}^{m-1} \left| \iint \frac{\varphi_j(x)}{f_\varepsilon^*(x)} \frac{\varphi_k(y)}{f_\varepsilon^*(y)} f_Y^*(x+y) \mathbf{1}_{|x|,|y| \leq \sqrt{\rho m}} dx dy \right|^2 = S_1 \\ &+ \frac{2}{\pi} \sum_{j,k=0}^{m-1} \left| \iint \frac{\varphi_j(x)}{f_\varepsilon^*(x)} \frac{\varphi_k(y)}{f_\varepsilon^*(y)} f_Y^*(x+y) \mathbf{1}_{|x| > \sqrt{\rho m}} \mathbf{1}_{|y| \leq \sqrt{\rho m}} dx dy \right|^2 = S_2 \\ &+ \frac{1}{\pi} \sum_{j,k=0}^{m-1} \left| \iint \frac{\varphi_j(x)}{f_\varepsilon^*(x)} \frac{\varphi_k(y)}{f_\varepsilon^*(y)} f_Y^*(x+y) \mathbf{1}_{|x|,|y| > \sqrt{\rho m}} dx dy \right|^2 = S_3 \\ &= S_1 + S_2 + S_3. \end{aligned}$$

We bound successively these above terms. Let

$$\Psi_k(x) = \int \frac{f_Y^*(x+y)}{f_\varepsilon^*(y)} \varphi_k(y) \mathbf{1}_{|y| \leq \sqrt{\rho m}} dy = \left\langle \frac{\tau_{-x} f_Y^*}{f_\varepsilon^*} \mathbf{1}_{|\cdot| \leq \sqrt{\rho m}}, \varphi_k \right\rangle,$$

where $(\tau_x h)(y) = h(y-x)$. Through Bessel's inequality (applied twice), it holds for the first term

$$\begin{aligned} S_1 &= \frac{1}{\pi} \sum_{j,k=0}^{m-1} \left| \left\langle \frac{\Psi_k}{f_\varepsilon^*} \mathbf{1}_{|\cdot| \leq \sqrt{\rho m}}, \varphi_j \right\rangle \right|^2 \leq \sum_{k=0}^{m-1} \int \frac{|\Psi_k(x)|^2}{|f_\varepsilon^*(x)|^2} \mathbf{1}_{|x| \leq \sqrt{\rho m}} dx \\ &\leq \frac{1}{\pi} \iint \frac{|f_Y^*(x+y)|^2}{|f_\varepsilon^*(x)|^2 |f_\varepsilon^*(y)|^2} \mathbf{1}_{|x|,|y| \leq \sqrt{\rho m}} dx dy. \end{aligned}$$

Let us remark that by the Cauchy-Schwarz inequality and Parseval's theorem, we have $\int |f_Y^*(z)| dz \leq 2\pi \|f\| \|f_\varepsilon\| \leq 2\pi \|f\|_\infty^{\frac{1}{2}} \|f_\varepsilon\|$. Thus, with the substitution $u = x+y$ and $v = x$, we obtain for S_1 :

$$S_1 \leq \frac{\Delta^{(1)}(m)}{\pi} \iint \frac{|f_Y^*(u)|^2}{|f_\varepsilon^*(v)|^2} \mathbf{1}_{|v|,|u-v| \leq \sqrt{\rho m}} du dv \leq 2\|f\|_\infty^{\frac{1}{2}} \|f_\varepsilon\| \Delta^{(1)}(m) \Delta(m).$$

Let $\Phi_j(x) = \left\langle \frac{\tau-x f_Y^*}{f_\varepsilon^*} \mathbf{1}_{|\cdot| > \sqrt{\rho m}}, \varphi_k \right\rangle$. Using the Bessel and Cauchy-Schwarz inequalities, it yields for S_2

$$\begin{aligned} S_2 &= \frac{2}{\pi} \sum_{j,k=0}^{m-1} \left| \left\langle \frac{\Phi_j}{f_\varepsilon^*} \mathbf{1}_{|\cdot| \leq \sqrt{\rho m}}, \varphi_j \right\rangle \right|^2 \\ &\leq \frac{2}{\pi} \sum_{j=0}^{m-1} \int_{|y| \leq \sqrt{\rho m}} \frac{1}{|f_\varepsilon^*(y)|^2} \left| \int_{|x| > \sqrt{\rho m}} f_Y^*(x+y) \frac{\varphi_j(x)}{f_\varepsilon^*(x)} dx \right|^2 dy \\ &\leq \frac{2}{\pi} \sum_{j=0}^{m-1} \int_{|y| \leq \sqrt{\rho m}} \frac{1}{|f_\varepsilon^*(y)|^2} \left(\int |f_Y^*(x+y)|^2 dx \int_{|x| > \sqrt{\rho m}} \frac{|\varphi_j(x)|^2}{|f_\varepsilon^*(x)|^2} dx \right) dy \\ &\leq 4 \|f\|_\infty^{\frac{1}{2}} \|f_\varepsilon\| \Delta(m) \Lambda(m), \end{aligned}$$

where we use in the last line $\int |f_Y^*(x+y)|^2 dx \leq \int |f_Y^*(z)| dz \leq 2\pi \|f\|_\infty^{\frac{1}{2}} \|f_\varepsilon\|$. Finally, by making the substitution $u = x+y$ and $v = x$, and as $\int |f_Y^*(z)| dz \leq 2\pi \|f\|_\infty^{\frac{1}{2}} \|f_\varepsilon\|$, we have from the Cauchy-Schwarz inequality,

$$\begin{aligned} S_3 &\leq \frac{1}{\pi} \sum_{j,k=0}^{m-1} \left(\iint |f_Y^*(x+y)| \frac{|\varphi_j(x)|^2}{|f_\varepsilon^*(x)|^2} \mathbf{1}_{|x|,|y| > \sqrt{\rho m}} dx dy \right) \times \left(\iint |f_Y^*(x+y)| \frac{|\varphi_j(y)|^2}{|f_\varepsilon^*(y)|^2} \mathbf{1}_{|x|,|y| > \sqrt{\rho m}} dx dy \right) \\ &= \frac{1}{\pi} \left(\sum_{j=0}^{m-1} \iint |f_Y^*(x+y)| \frac{|\varphi_j(x)|^2}{|f_\varepsilon^*(x)|^2} dx dy \mathbf{1}_{|x|,|y| > \sqrt{\rho m}} \right)^2 \\ &\leq 4\pi \|f\|_\infty \|f_\varepsilon\|^2 \Lambda^2(m). \end{aligned}$$

Therefore, we get

$$\begin{aligned} \frac{2}{n(n-1)} \sum_{j,k=0}^{m-1} \left(\text{Cov} \left(U_{\varphi_j}^*(Y_1), U_{\varphi_k}^*(Y_1) \right) \right)^2 &\leq \frac{C_0}{n(n-1)} \Delta(m) \left(\Delta^{(1)}(m) + \Lambda(m) \right) + \frac{C_0^2}{n(n-1)} \Lambda^2(m) \\ (27) \qquad \qquad \qquad &+ \frac{4\|f\|_\infty^2}{n(n-1)}. \end{aligned}$$

Injecting (27) in (26), we deduce

$$\begin{aligned} \text{Var}(\hat{\theta}_m(f)) &\leq \frac{C_0}{n(n-1)} \Delta(m) \left(\Delta^{(1)}(m) + \Lambda(m) \right) + \frac{C_0^2}{n(n-1)} \Lambda^2(m) \\ (28) \qquad \qquad \qquad &+ \frac{4\|f\|_\infty^2}{n(n-1)} + \frac{4}{n} \text{Var}(U_{f_m}^*(Y_1)). \end{aligned}$$

Plugging (28) in (23) completes the proof of Part (i).

7.1.2. Proof of Part (ii). It is a consequence of Part (i). It is based on the fact that $\Lambda(m)$ is bounded uniformly in m . Indeed, we have from (4)-(7) and $\rho \geq 2$ that (see also Proposition 3.1 (ii) given in Sacko (2020))

$$\Lambda(m) = \sum_{j=0}^{m-1} \int_{|u| > \sqrt{\rho m}} \frac{|\varphi_j(u)|^2}{|f_\varepsilon^*(u)|^2} du \leq C^2 e^{-\xi \rho m} \sum_{j=0}^{m-1} \int_{|u| > \sqrt{\rho m}} (1+u^2)^\gamma e^{\mu|u|^\delta} e^{-\xi u^2} du.$$

Since $\int_{|u|>\sqrt{\rho m}}(1+u^2)^\gamma e^{\mu|u|^\delta} e^{-\xi u^2} du < \infty$ if $\delta < 2$ or $\delta = 2$ and $\xi \leq \mu$, it follows that

$$\Lambda(m) \lesssim m e^{-\xi m} \lesssim 1.$$

This implies Part (ii) and concludes the proof.

7.1.3. *Proof of Lemma 7.1.* Let $\tilde{Y}_k^{(j)}$ be a real random variable defined as:

$$\tilde{Y}_k^{(j)} = U_{\varphi_j}^*(Y_k) - \mathbb{E}[U_{\varphi_j}^*(Y_k)],$$

where U_t is defined in (11). The variables $(\tilde{Y}_k^{(j)})_{k \in \{1, \dots, n\}}$ are i.i.d with $\mathbb{E}[\tilde{Y}_k^{(j)}] = 0$ and $\mathbb{E}[(\tilde{Y}_k^{(j)})^2] = \text{Var}(\tilde{Y}_k^{(j)})$. By definition of $T_{1,j}$ given (25), we have

$$\begin{aligned} \text{Cov}(T_{1,j}, T_{1,j'}) &= \mathbb{E}[T_{1,j} T_{1,j'}] = \frac{1}{n^2(n-1)^2} \sum_{k \neq \ell=1, k' \neq \ell'=1}^n \mathbb{E}[\tilde{Y}_k^{(j)} \tilde{Y}_\ell^{(j)} \tilde{Y}_{k'}^{(j')} \tilde{Y}_{\ell'}^{(j')}] \\ &= \frac{1}{n^2(n-1)^2} \left(\sum_{k \neq \ell=1, k' \neq \ell'=1, k=k', k \neq \ell'}^n \mathbb{E}[\tilde{Y}_k^{(j)} \tilde{Y}_\ell^{(j)} \tilde{Y}_{k'}^{(j')} \tilde{Y}_{\ell'}^{(j')}] + \sum_{k \neq \ell=1, k' \neq \ell'=1, k \neq k', k=\ell'}^n \mathbb{E}[\tilde{Y}_k^{(j)} \tilde{Y}_\ell^{(j)} \tilde{Y}_{k'}^{(j')} \tilde{Y}_{\ell'}^{(j')}] \right) \\ &= \frac{2}{n^2(n-1)^2} \sum_{k \neq \ell=1, k' \neq \ell'=1, k \neq \ell'}^n \mathbb{E}[\tilde{Y}_k^{(j)} \tilde{Y}_\ell^{(j)} \tilde{Y}_{k'}^{(j')} \tilde{Y}_{\ell'}^{(j')}] \\ &= \frac{2}{n^2(n-1)^2} \sum_{k \neq \ell=1, k' \neq \ell'=1, k \neq \ell', \ell=\ell'}^n \mathbb{E}[\tilde{Y}_k^{(j)} \tilde{Y}_\ell^{(j)} \tilde{Y}_{k'}^{(j')} \tilde{Y}_{\ell'}^{(j')}] \\ &= \frac{2}{n^2(n-1)^2} \sum_{k \neq \ell=1}^n \mathbb{E}[\tilde{Y}_k^{(j)} \tilde{Y}_k^{(j')}] \mathbb{E}[\tilde{Y}_\ell^{(j)} \tilde{Y}_\ell^{(j')}] = \frac{2}{n^2(n-1)^2} \sum_{k \neq \ell=1}^n \mathbb{E}[\tilde{Y}_1^{(j)} \tilde{Y}_1^{(j')}] \mathbb{E}[\tilde{Y}_1^{(j)} \tilde{Y}_1^{(j')}] \\ &= \frac{2}{n(n-1)} \left(\mathbb{E} \left[\left(U_{\varphi_j}^*(Y_1) - \mathbb{E}[U_{\varphi_j}^*(Y_1)] \right) \left(U_{\varphi_{j'}}^*(Y_1) - \mathbb{E}[U_{\varphi_{j'}}^*(Y_1)] \right) \right] \right)^2 \\ &= \frac{2}{n(n-1)} \left(\text{Cov}(U_{\varphi_j}^*(Y_1), U_{\varphi_{j'}}^*(Y_1)) \right)^2 \end{aligned}$$

By remembering that $\mathbb{E}[\tilde{Y}_k^{(j)}] = a_j(f)$ by construction, we get

$$\begin{aligned} \mathbb{E}[T_{1,j} T_{2,j'}] &= \frac{2}{n^2(n-1)} a_{j'}(f) \sum_{k \neq \ell=1, \ell'=1}^n \mathbb{E}[\tilde{Y}_k^{(j)} \tilde{Y}_\ell^{(j)} \tilde{Y}_{\ell'}^{(j')}] \\ &= \frac{2}{n^2(n-1)} a_{j'}(f) \left(\sum_{k \neq \ell=1, \ell'=1, \ell=\ell'}^n \mathbb{E}[\tilde{Y}_k^{(j)} \tilde{Y}_\ell^{(j)} \tilde{Y}_{\ell'}^{(j')}] + \sum_{k \neq \ell=1, \ell'=1, \ell \neq \ell'}^n \mathbb{E}[\tilde{Y}_k^{(j)} \tilde{Y}_\ell^{(j)} \tilde{Y}_{\ell'}^{(j')}] \right) = 0. \end{aligned}$$

Besides, we have

$$\begin{aligned} \mathbb{E}[T_{2,j} T_{2,j'}] &= \frac{4}{n^2} a_j(f) a_{j'}(f) \sum_{k=1, \ell=1}^n \mathbb{E}[\tilde{Y}_k^{(j)} \tilde{Y}_\ell^{(j')}] = \frac{4}{n^2} a_j(f) a_{j'}(f) \sum_{k=1}^n \mathbb{E}[\tilde{Y}_k^{(j)} \tilde{Y}_k^{(j')}] \\ &= \frac{4}{n} a_j(f) a_{j'}(f) \mathbb{E} \left[\left(U_{\varphi_j}^*(Y_1) - \mathbb{E}[U_{\varphi_j}^*(Y_1)] \right) \left(U_{\varphi_{j'}}^*(Y_1) - \mathbb{E}[U_{\varphi_{j'}}^*(Y_1)] \right) \right] \\ &= \frac{4}{n} a_j(f) a_{j'}(f) \text{Cov}(U_{\varphi_j}^*(Y_1), U_{\varphi_{j'}}^*(Y_1)). \end{aligned}$$

Lemma 7.1 is therefore proved.

7.2. Proof of Lemma 3.1. Let us begin by Part (ii). By Plancherel-Parseval Equality, we have

$$(29) \quad \begin{aligned} \mathbb{E}[|U_{f_m}^*(Y_1)|^2] &= \int |U_{f_m}^*(z)|^2 f_Y(z) dy \leq \|f_Y\|_\infty \frac{1}{2\pi} \int \frac{|f_m^*(-u)|^2}{|f_\varepsilon^*(u)|^2} du \\ &= \|f_Y\|_\infty \left(\frac{1}{2\pi} \int_{|u| \leq \sqrt{\rho m}} \frac{|f_m^*(-u)|^2}{|f_\varepsilon^*(u)|^2} du + \frac{1}{2\pi} \int_{|u| > \sqrt{\rho m}} \frac{|f_m^*(-u)|^2}{|f_\varepsilon^*(u)|^2} du \right). \end{aligned}$$

Using the Cauchy Schwarz inequality on the sum, (5) and as $\|f\|^2 \leq \|f\|_\infty$, we have for the second integral

$$(30) \quad \int_{|u| > \sqrt{\rho m}} \frac{|f_m^*(-u)|^2}{|f_\varepsilon^*(u)|^2} du \leq 2\pi \sum_{j=0}^{m-1} a_j^2(f) \sum_{j=0}^{m-1} \int_{|u| > \sqrt{\rho m}} \frac{|\varphi_j(u)|^2}{|f_\varepsilon^*(u)|^2} du \leq 2\pi \|f\|_\infty \Lambda(m).$$

Parseval's equality and $\|f\|^2 \leq \|f\|_\infty$ imply that

$$\int_{|u| \leq \sqrt{\rho m}} \frac{|f_m^*(-u)|^2}{|f_\varepsilon^*(u)|^2} du \leq 2\pi \Delta^{(1)}(m) \|f\|_\infty.$$

It follows that

$$(31) \quad \mathbb{E}[|U_{f_m}^*(Y_1)|^2] \leq \|f\|_\infty \|f_Y\|_\infty \left(\Delta^{(1)}(m) + \Lambda(m) \right) = C'_0 \left(\Delta^{(1)}(m) + \Lambda(m) \right),$$

where $C'_0 = \|f\|_\infty \|f_Y\|_\infty$. This gives Part (ii).

Now, we demonstrate Part (i). The difference with Part (i) is related to the study of the first integral given in (29). We have

$$\begin{aligned} \int_{|u| \leq \sqrt{\rho m}} \frac{|f_m^*(-u)|^2}{|f_\varepsilon^*(u)|^2} du &\leq 2 \int_{|u| \leq \sqrt{\rho m}} \frac{|(f_m - f)^*(-u)|^2}{|f_\varepsilon^*(u)|^2} du + 2 \int_{|u| \leq \sqrt{\rho m}} \frac{|f^*(-u)|^2}{|f_\varepsilon^*(u)|^2} du \\ &\leq 4\pi \Delta^{(1)}(m) \|f - f_m\|^2 + 2 \int_{|u| \leq \sqrt{\rho m}} \frac{|f^*(-u)|^2}{|f_\varepsilon^*(u)|^2} du. \end{aligned}$$

Injecting this and (30) in (29), we get Part (i) and then the announced result.

7.3. Proof of Proposition 3.2. We only prove the ordinary smooth case. The super smooth is handled in the same way using Part (ii) of Lemma 3.1. By Proposition 3.1 (ii) and (16), it yields

$$\begin{aligned} \mathbb{E}[(\hat{\theta}_m(f) - \theta(f))^2] &\leq 2\|f - f_m\|^4 + \frac{C_1 + 16\|f_Y\|_\infty^2}{n(n-1)} \Delta(m) \Delta^{(1)}(m) \\ &\quad + \frac{4\|f_Y\|_\infty}{\pi n} \int_{|u| \leq \sqrt{\rho m}} \frac{|f^*(-u)|^2}{|f_\varepsilon^*(u)|^2} du + \frac{C}{n}, \end{aligned}$$

where C is an absolute constant. We evaluate the order of each term. For $f \in W_H^s(L)$, we recall that $\|f - f_m\|^4 = \left(\sum_{j \geq m} a_j^2(f) \right)^2 \leq L^2 m^{-2s}$. From Lemma 1 in Comte and Lacour (2011) p.586 and elementary computation, it yields under (7):

$$(32) \quad \Delta(m) \asymp m^{\gamma + \frac{1-\delta}{2}} e^{\mu \rho^{\frac{\delta}{2}} m^{\frac{\delta}{2}}}, \quad \Delta^{(1)}(m) \leq c_1 (1 + \rho m)^\gamma e^{\mu(\rho m)^{\frac{\delta}{2}}}.$$

As $f \in W_H^s(L) \subset W^s(L')$ (see Section 2.1.3) and f_ε is ordinary smooth, it holds

$$\int_{|u| \leq \sqrt{\rho m}} \left| \frac{f^*(-u)}{f_\varepsilon^*(u)} \right|^2 du \leq c_1 \int_{|u| \leq \sqrt{\rho m}} (1 + u^2)^s |f^*(-u)|^2 (1 + u^2)^{-(s-\gamma)} du \leq L' c_1 \left(\mathbf{1}_{s \geq \gamma} + (1 + \rho m)^{-(s-\gamma)} \mathbf{1}_{s < \gamma} \right),$$

where c_1 is given in (7). Denote by $\mathfrak{C} > 0$ the maximum among constants that appear in the previous upper bound of $\mathbb{E}[(\hat{\theta}_m(f) - \theta(f))^2]$, obtained by computing the order of each term. Setting $\delta = 0$ in the above bounds, we get $f \in W_H^s(L)$ and (7)

$$\mathbb{E}[(\hat{\theta}_m(f) - \theta(f))^2] \leq 2L^2 m^{-2s} + \mathfrak{C} \left(\frac{1}{n(n-1)} m^{2\gamma + \frac{1}{2}} + \frac{1}{n} m^{-(s-\gamma)} \mathbf{1}_{s < \gamma} + \frac{1}{n} \right).$$

Hence the result for the ordinary smooth case.

7.4. Proof of Theorem 3.3 and 3.4. The idea is to replace m by m_{opt} in Proposition 3.2. We distinguish two cases: ordinary smooth and super smooth cases.

- **The ordinary smooth case.** Plugging the value $m_{opt} = \lceil n^{\frac{4}{4s+4\gamma+1}} \rceil$ in the bound given in Proposition 3.2 (i), it yields

$$\mathbb{E}[(\hat{\theta}_{m_{opt}}(f) - \theta(f))^2] \lesssim \left(n^{-\frac{8s}{4s+4\gamma+1}} + n^{-\frac{8s+1}{4s+4\gamma+1}} + n^{-1} \right) \lesssim \left(n^{-\frac{8s}{4s+4\gamma+1}} \vee n^{-1} \right).$$

The result is proven for ordinary smooth noise.

- **The super smooth case.** Injecting $m_{opt} = \frac{1}{\rho} \left(\frac{\log n}{4\mu} \right)^{\frac{2}{\delta}}$ in Proposition 3.2 (ii), one obtains

$$\mathbb{E}[(\hat{\theta}_m(f) - \theta(f))^2] \lesssim (\log(n))^{-\frac{4s}{\delta}}.$$

This concludes the proof of Theorem 3.3 and 3.4.

7.5. Proof of Theorem 4.1. For the sake of simplicity, we set $\hat{\theta}_m(f) = \hat{\theta}_m$. By definition of $\hat{A}(m)$ given in (18) and \hat{m}_n in (21) and $\forall m \in \mathcal{M}_n$, we have the following decomposition:

$$\begin{aligned} (\hat{\theta}_{\hat{m}_n} - \theta)^2 &\leq 3(\hat{\theta}_{\hat{m}_n} - \hat{\theta}_{\hat{m}_n \wedge m})^2 + 3(\hat{\theta}_{\hat{m}_n \wedge m} - \hat{\theta}_m)^2 + 3(\hat{\theta}_m - \theta)^2 \\ &\leq 3(\hat{A}(\hat{m}_n) + \kappa_1 V(\hat{m}_n)) + 3(\hat{A}(m) + \kappa_1 V(m)) + 3(\hat{\theta}_m - \theta)^2 \\ &\leq 6\hat{A}(m) + 6\kappa_2 V(m) + 3(\hat{\theta}_m - \theta)^2. \end{aligned}$$

By taking the expectation, one obtains

$$(33) \quad \mathbb{E} \left[(\hat{\theta}_{\hat{m}_n}(f) - \theta(f))^2 \right] \leq 6 \mathbb{E}[\hat{A}(m)] + 6\kappa_2 V(m) + 3 \mathbb{E}[(\hat{\theta}_m(f) - \theta(f))^2].$$

We bound $\mathbb{E}[\hat{A}(m)]$. Decompose $(\hat{\theta}_{m' \wedge m} - \hat{\theta}_{m'})^2$ as follows:

$$(34) \quad (\hat{\theta}_{m' \wedge m} - \hat{\theta}_{m'})^2 \leq 3(\hat{\theta}_{m' \wedge m} - \mathbb{E}[\hat{\theta}_{m' \wedge m}])^2 + 3(\mathbb{E}[\hat{\theta}_{m' \wedge m}] - \mathbb{E}[\hat{\theta}_{m'}])^2 + 3(\hat{\theta}_{m'} - \mathbb{E}[\hat{\theta}_{m'}])^2.$$

Consequently, it yields

$$\begin{aligned} \hat{A}(m) &\leq 3 \sup_{m' \in \mathcal{M}_n} \left\{ \left((\hat{\theta}_{m' \wedge m} - \mathbb{E}[\hat{\theta}_{m' \wedge m}])^2 - \frac{\kappa_1}{6} V(m') \right)_+ \right\} + 3 \sup_{m' \in \mathcal{M}_n} \left\{ \left((\hat{\theta}_{m'} - \mathbb{E}[\hat{\theta}_{m'}])^2 - \frac{\kappa_1}{6} V(m') \right)_+ \right\} \\ &\quad + 3 \sup_{m' \in \mathcal{M}_n} \left\{ (\mathbb{E}[\hat{\theta}_{m' \wedge m}] - \mathbb{E}[\hat{\theta}_{m'}])^2 \right\}. \end{aligned}$$

Observe that, if $m' \leq m$, the last term is equal to zero. So, by definition of $\hat{\theta}_m$ given in (12), we have

$$\begin{aligned} \sup_{m' \in \mathcal{M}_n} \left\{ (\mathbb{E}[\hat{\theta}_{m' \wedge m}] - \mathbb{E}[\hat{\theta}_{m'}])^2 \right\} &= \sup_{m' \in \mathcal{M}_n, m < m'} \left\{ (\mathbb{E}[\hat{\theta}_{m'}] - \mathbb{E}[\hat{\theta}_m])^2 \right\} = \sup_{m' \in \mathcal{M}_n, m < m'} \left\{ (\|f_{m'}\|^2 - \|f_m\|^2)^2 \right\} \\ &\leq \sup_{m' \in \mathcal{M}_n, m < m'} \left\{ (\|f\|^2 - \|f_m\|^2)^2 \right\} = \|f - f_m\|^4, \end{aligned}$$

where the last inequality is obtained by using the Pythagorean theorem. It comes that:

$$\hat{A}(m) \leq 3 \sup_{m' \in \mathcal{M}_n} \left\{ \left((\hat{\theta}_{m'} - \mathbb{E}[\hat{\theta}_{m'}])^2 - \frac{\kappa_1}{6} V(m') \right)_+ \right\} + 3 \sup_{m' \in \mathcal{M}_n} \left\{ \left((\hat{\theta}_{m' \wedge m} - \mathbb{E}[\hat{\theta}_{m' \wedge m}])^2 - \frac{\kappa_1}{6} V(m') \right)_+ \right\} + 3 \|f - f_m\|^4.$$

As

$$\begin{aligned} & \sup_{m' \in \mathcal{M}_n} \left\{ \left((\hat{\theta}_{m' \wedge m} - \mathbb{E}[\hat{\theta}_{m' \wedge m}])^2 - \frac{\kappa_1}{6} V(m') \right)_+ \right\} \\ & \leq \max \left(\sup_{m' \in \mathcal{M}_n, m' \geq m} \left\{ \left((\hat{\theta}_m - \mathbb{E}[\hat{\theta}_m])^2 - \frac{\kappa_1}{6} V(m') \right)_+ \right\}, \sup_{m' \in \mathcal{M}_n, m' < m} \left\{ \left((\hat{\theta}_{m'} - \mathbb{E}[\hat{\theta}_{m'}])^2 - \frac{\kappa_1}{6} V(m') \right)_+ \right\} \right) \\ & \leq \sup_{m' \in \mathcal{M}_n, m' < m} \left\{ \left((\hat{\theta}_{m'} - \mathbb{E}[\hat{\theta}_{m'}])^2 - \frac{\kappa_1}{6} V(m') \right)_+ \right\} + (\hat{\theta}_m - \mathbb{E}[\hat{\theta}_m])^2, \end{aligned}$$

we deduce that

$$\mathbb{E}[\hat{A}(m)] \leq 3 \|f - f_m\|^4 + 6 \mathbb{E} \left[\sup_{m' \in \mathcal{M}_n, m' < m} \left\{ \left((\hat{\theta}_{m'} - \mathbb{E}[\hat{\theta}_{m'}])^2 - \frac{\kappa_1}{6} V(m') \right)_+ \right\} \right] + 3 \text{Var}(\hat{\theta}_m).$$

By (28), (15), it holds

$$\text{Var}(\hat{\theta}_m) \leq C_1 V(m) + \frac{4}{n} \mathbb{E}[|U_{f_m}^*(Y_1)|^2] + \frac{C_2}{n},$$

where C_1 and C_2 are constant appearing in Proposition 3.1 (ii) and U_t is given in (11). This implies

$$\begin{aligned} \mathbb{E}[\hat{A}(m)] & \leq 3 \|f - f_m\|^4 + 6C_1 V(m) + \frac{24}{n} \mathbb{E}[|U_{f_m}^*(Y_1)|^2] + \frac{6C_2}{n} \\ & \quad + 6 \mathbb{E} \left[\sup_{m' \in \mathcal{M}_n} \left\{ \left((\hat{\theta}_{m'} - \mathbb{E}[\hat{\theta}_{m'}])^2 - \frac{\kappa_1}{6} V(m') \right)_+ \right\} \right]. \end{aligned}$$

Now, we need the following result which leads to control the last term of the above bound.

Proposition 7.1. *Under the assumptions of Theorem 4.1, there exists a constant $\kappa_1 \geq \kappa_0$ such that*

$$\mathbb{E} \left[\sup_{m' \in \mathcal{M}_n} \left\{ \left((\hat{\theta}_{m'} - \mathbb{E}[\hat{\theta}_{m'}])^2 - \frac{\kappa_1}{6} V(m') \right)_+ \right\} \right] \leq \frac{C^{(0)}}{n},$$

where $C^{(0)}$ depends on $f_\varepsilon, \|f\|_\infty$ and $\|f_Y\|_\infty$.

Mainly, the constant κ_0 depends on $\|f\|_\infty, \|f_Y\|_\infty$, and the density f_ε . Proposition 7.1 and 3.1 (ii) and (33) imply that

$$\mathbb{E} \left[(\hat{\theta}_{\hat{m}_n}(f) - \theta(f))^2 \right] \leq 21 \|f - f_m\|^2 + (6\kappa_2 + 9C_1) V(m) + \frac{36}{n} \mathbb{E}[|U_{f_m}^*(Y_1)|^2] + \frac{36C^{(0)} + 9C_2}{n},$$

where U_t is defined in (11).

Choosing $C = \max(36, (6\kappa_2 + 3C_1))$ and $C' = 36C^{(0)} + 3C_2$ and then taking the infimum on \mathcal{M}_n completes the proof.

7.6. **Proof of Proposition 7.1.** Define a degenerate U-statistic of order 2 by:

$$(35) \quad U_n(H_m) = \frac{1}{n(n-1)} \sum_{k \neq \ell=1}^n H_m(Y_k, Y_\ell), \quad H_m(x, y) = \sum_{j=0}^{m-1} (U_{\varphi_j}^*(x) - a_j)(U_{\varphi_j}^*(y) - a_j)$$

where $a_j = a_j(f) = \int f \varphi_j$ and U_t is defined in (11) and the empirical measure by

$$(36) \quad P_n(t) := \frac{1}{n} \sum_{i=1}^n U_t^*(Y_i) - \int_{\mathbb{R}} t f.$$

When X_k are available, that is $\varepsilon_k = 0$ almost surely $P_n(t)$ is equal to the classical process $\frac{1}{n} \sum_{i=1}^n t(X_i) - \int_{\mathbb{R}} t f$. From (25), we have:

$$\hat{\theta}_m - \mathbb{E}[\hat{\theta}_m] = U_n(H_m) + P_n(2f_m).$$

This implies that

$$(37) \quad \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left((\hat{\theta}_m - \mathbb{E}[\hat{\theta}_m])^2 - \frac{\kappa_1}{6} V(m) \right)_+ \right\} \right] \leq 2 \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(U_n^2(H_m) - \frac{\kappa_1}{12} V(m) \right)_+ \right\} \right] \\ + 2 \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(P_n^2(2f_m) - \frac{\kappa_1}{12} V(m) \right)_+ \right\} \right].$$

We bound successively the two terms in the right hand side of (37).

Bounding of $\mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(U_n^2(H_m) - \frac{\kappa_1}{12} V(m) \right)_+ \right\} \right]$. First, we decompose the $U_n(H_m)$ given by (35) in three processes by splitting the integrals. By Plancherel-Parseval's formula, it holds

$$H_m(Y_k, Y_\ell) = \sum_{j=0}^{m-1} \left(\frac{1}{2\pi} \int_{\mathbb{R}} e^{iuY_k} \frac{\varphi_j^*(-u)}{f_\varepsilon^*(u)} du - \mathbb{E} \left[\frac{1}{2\pi} \int_{\mathbb{R}} e^{iuY_k} \frac{\varphi_j^*(-u)}{f_\varepsilon^*(u)} du \right] \right) \\ \times \left(\frac{1}{2\pi} \int_{\mathbb{R}} e^{iuY_\ell} \frac{\varphi_j^*(-u)}{f_\varepsilon^*(u)} du - \mathbb{E} \left[\frac{1}{2\pi} \int_{\mathbb{R}} e^{iuY_\ell} \frac{\varphi_j^*(-u)}{f_\varepsilon^*(u)} du \right] \right) \\ = \sum_{j=0}^{m-1} \left\{ \left(\begin{aligned} & \frac{1}{2\pi} \int_{|u| \leq \sqrt{\rho m}} e^{iuY_k} \frac{\varphi_j^*(-u)}{f_\varepsilon^*(u)} du - \mathbb{E} \left[\frac{1}{2\pi} \int_{|u| \leq \sqrt{\rho m}} e^{iuY_k} \frac{\varphi_j^*(-u)}{f_\varepsilon^*(u)} du \right] \\ & + \frac{1}{2\pi} \int_{|u| \geq \sqrt{\rho m}} e^{iuY_k} \frac{\varphi_j^*(-u)}{f_\varepsilon^*(u)} du - \mathbb{E} \left[\frac{1}{2\pi} \int_{|u| \geq \sqrt{\rho m}} e^{iuY_k} \frac{\varphi_j^*(-u)}{f_\varepsilon^*(u)} du \right] \end{aligned} \right) \right\} \\ \times \left\{ \left(\begin{aligned} & \frac{1}{2\pi} \int_{|u| \leq \sqrt{\rho m}} e^{iuY_\ell} \frac{\varphi_j^*(-u)}{f_\varepsilon^*(u)} du - \mathbb{E} \left[\frac{1}{2\pi} \int_{|u| \leq \sqrt{\rho m}} e^{iuY_\ell} \frac{\varphi_j^*(-u)}{f_\varepsilon^*(u)} du \right] \\ & + \frac{1}{2\pi} \int_{|u| \geq \sqrt{\rho m}} e^{iuY_\ell} \frac{\varphi_j^*(-u)}{f_\varepsilon^*(u)} du - \mathbb{E} \left[\frac{1}{2\pi} \int_{|u| \geq \sqrt{\rho m}} e^{iuY_\ell} \frac{\varphi_j^*(-u)}{f_\varepsilon^*(u)} du \right] \end{aligned} \right) \right\}.$$

Define

$$(38) \quad v_{\varphi_j}(x) = \frac{1}{2\pi} \int_{|u| \leq \sqrt{\rho m}} e^{iux} \frac{\varphi_j^*(-u)}{f_\varepsilon^*(u)} du.$$

and

$$(39) \quad \Psi_{\varphi_j}(Y_\ell) = \frac{1}{2\pi} \int_{|u| \geq \sqrt{\rho m}} (e^{iuY_\ell} - \mathbb{E}[e^{iuY_\ell}]) \frac{\varphi_j^*(-u)}{f_\varepsilon^*(u)} du.$$

Thanks to (38) and (39), we get

$$\begin{aligned}
H_m(Y_k, Y_\ell) &= \sum_{j=0}^{m-1} (v_{\varphi_j}(Y_k) - \mathbb{E}[v_{\varphi_j}(Y_k)])(v_{\varphi_j}(Y_\ell) - \mathbb{E}[v_{\varphi_j}(Y_\ell)]) + \sum_{j=0}^{m-1} (v_{\varphi_j}(Y_\ell) - \mathbb{E}[v_{\varphi_j}(Y_\ell)])\Psi_{\varphi_j}(Y_k) \\
&\quad + \sum_{j=0}^{m-1} (v_{\varphi_j}(Y_k) - \mathbb{E}[v_{\varphi_j}(Y_k)])\Psi_{\varphi_j}(Y_\ell) + \sum_{j=0}^{m-1} \Psi_{\varphi_j}(Y_k)\Psi_{\varphi_j}(Y_\ell).
\end{aligned}$$

Introduce:

$$(40) \quad U_n(I_m) = \frac{1}{n(n-1)} \sum_{k \neq \ell=1}^n I_m(Y_k, Y_\ell),$$

$$I_m(x, y) = \sum_{j=0}^{m-1} (v_{\varphi_j}(x) - \mathbb{E}[v_{\varphi_j}(Y_k)])(v_{\varphi_j}(y) - \mathbb{E}[v_{\varphi_j}(Y_\ell)]),$$

$$(41) \quad \nu_{n,1}(\varphi_j) = \frac{2}{n(n-1)} \sum_{k \neq \ell=1}^n \sum_{j=0}^{m-1} (v_{\varphi_j}(Y_k) - \mathbb{E}[v_{\varphi_j}(Y_k)])\Psi_{\varphi_j}(Y_\ell),$$

and

$$(42) \quad \nu_{n,2}(\varphi_j) = \frac{1}{n(n-1)} \sum_{k \neq \ell=1}^n \sum_{j=0}^{m-1} \Psi_{\varphi_j}(Y_k)\Psi_{\varphi_j}(Y_\ell).$$

We rewrite

$$U_n(H_m) = U_n(I_m) + \nu_{n,1}(\varphi_j) + \nu_{n,2}(\varphi_j),$$

and

$$\begin{aligned}
\mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(U_n^2(H_m) - \frac{\kappa_1}{12} V(m) \right)_+ \right\} \right] &\leq 3 \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(U_n^2(I_m) - \frac{\kappa_1}{36} V(m) \right)_+ \right\} \right] \\
&\quad + 3 \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \nu_{n,1}^2(\varphi_j) \right] + 3 \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \nu_{n,2}^2(\varphi_j) \right].
\end{aligned}$$

We establish the following result.

Proposition 7.2. *Let $U_n(I_m)$ be defined in (40), $\nu_{n,1}(\varphi_j)$ in (41) and $\nu_{n,2}(\varphi_j)$ in (42).*

(i) *Then, we have for κ_1 large enough, depending only on f_ε .*

$$\mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(U_n^2(I_m) - \frac{\kappa_1}{36} V(m) \right)_+ \right\} \right] \leq 4\rho \frac{(1 \vee 4\|f\|_\infty)^2}{n}.$$

(ii) *There exist two constants $C_{f_\varepsilon,1}$ and $C_{f_\varepsilon,2}$ depending on f_ε such that*

$$\mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \nu_{n,1}^2(\varphi_j) \right] \leq \frac{C_{f_\varepsilon,1}}{n}, \quad \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \nu_{n,2}^2(\varphi_j) \right] \leq \frac{C_{f_\varepsilon,2}}{n}.$$

Proposition 7.2 implies

$$(43) \quad \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(U_n^2(H_m) - \frac{\kappa_1}{12} V(m) \right)_+ \right\} \right] \leq 12\rho \frac{(1 \vee 4\|f\|_\infty)^2}{n} + 3 \frac{C_{f_\varepsilon,1} + C_{f_\varepsilon,2}}{n},$$

Bounding of $\mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(P_n^2(2f_m) - \frac{\kappa_1}{12} V(m) \right)_+ \right\} \right]$. To control the deviation of $P_n(2f_m)$, we apply Bernstein inequality. Let us first decompose the process $P_n(t)$ defined in (36) as follows: for any $t \in S_m$, we have using the Plancherel Parseval identity

$$\begin{aligned}
P_n(t) &= \frac{1}{n} \sum_{k=1}^n (U_t^*(Y_k) - \langle t, f \rangle) \\
&= \frac{1}{n} \sum_{k=1}^n \left(\frac{1}{2\pi} \int_{|u| \leq \sqrt{\rho m}} \frac{t^*(-u)}{f_\varepsilon^*(u)} e^{iuY_k} du - \mathbb{E} \left[\frac{1}{2\pi} \int_{|u| \leq \sqrt{\rho m}} \frac{t^*(-u)}{f_\varepsilon^*(u)} e^{iuY_k} du \right] \right) \\
&\quad + \frac{1}{n} \sum_{k=1}^n \left(\frac{1}{2\pi} \int_{|u| > \sqrt{\rho m}} \frac{t^*(-u)}{f_\varepsilon^*(u)} e^{iuY_k} du - \mathbb{E} \left[\frac{1}{2\pi} \int_{|u| > \sqrt{\rho m}} \frac{t^*(-u)}{f_\varepsilon^*(u)} e^{iuY_k} du \right] \right) \\
(44) \quad &= \frac{1}{n} \sum_{k=1}^n (v_t(Y_k) - \mathbb{E}[v_t(Y_k)]) + \frac{1}{2\pi} \int_{|u| > \sqrt{\rho m}} \frac{t^*(-u)}{f_\varepsilon^*(u)} (\hat{f}_Y^*(u) - f_Y^*(u)) du,
\end{aligned}$$

with v_t is given (38) and \hat{f}_Y^* is defined in (10). Therefore, we write $P_n(t) = P_{n,1}(t) + P_{n,2}(t)$ where

$$P_{n,1}(t) = \frac{1}{n} \sum_{k=1}^n (v_t(Y_k) - \mathbb{E}[v_t(Y_k)])$$

and

$$P_{n,2}(t) = \frac{1}{2\pi} \int_{|u| > \sqrt{\rho m}} \frac{t^*(-u)}{f_\varepsilon^*(u)} (\hat{f}_Y^*(u) - f_Y^*(u)) du.$$

It follows that

$$\begin{aligned}
\mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(P_n^2(2f_m) - \frac{\kappa_1}{12} V(m) \right)_+ \right\} \right] &\leq 2 \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(P_{n,1}^2(2f_m) - \frac{\kappa_1}{24} V(m) \right)_+ \right\} \right] \\
(45) \quad &\quad + 2 \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} P_{n,2}^2(2f_m) \right].
\end{aligned}$$

Proposition 7.3. *Under the assumptions of Theorem 4.1, the following hold.*

- (i) *There exists a constant $C_{f_\varepsilon,3}$ such that for $\kappa_1 \geq \max(384 \|f_Y\|_{\infty p}, 384 \|f\|_{\infty} p^2 / (\alpha^2 c_1'))$ with p an integer ($p = 4$ suits) and $\alpha \in (0, 1)$*

$$\mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(P_{n,1}^2(2f_m) - \frac{\kappa_1}{24} V(m) \right)_+ \right\} \right] \leq \frac{C_{f_\varepsilon,3}}{n}.$$

- (ii) *There exists a constant $C_{f_\varepsilon,4}$ such that*

$$\mathbb{E} \left[\sup_{m \in \mathcal{M}_n} P_{n,2}^2(2f_m) \right] \leq \frac{C_{f_\varepsilon,4}}{n}.$$

From Proposition 7.3 and (45), we obtain

$$\mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(P_n^2(2f_m) - \frac{\kappa_1}{12} V(m) \right)_+ \right\} \right] \leq 2 \frac{C_{f_\varepsilon,4} + C_{f_\varepsilon,3}}{n}.$$

Finally, injecting this and (43) in (37), we derive that for κ_1 large enough

$$\mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left((\hat{\theta}_m - \mathbb{E}[\hat{\theta}_m])^2 - \frac{\kappa_1}{6} V(m) \right)_+ \right\} \right] \leq \frac{C^{(0)}}{n},$$

where $C^{(0)} = 24\rho(1 \vee 4\|f\|_\infty)^2 + 6(C_{f_\varepsilon,1} + C_{f_\varepsilon,2} + C_{f_\varepsilon,3} + C_{f_\varepsilon,4})$. Hence the announced result.

7.7. Proof of Proposition 7.2.

7.7.1. *Proof of Part (i).* We use the following Lemma, which is proved in the sequel.

Lemma 7.2. *Let $U_n(I_m)$ be defined in (40). Assume that $\|f\|_\infty < \infty$, (H5) and f_ε is square integrable. There exists a constant $C_\star > 0$ depending on $\|f\|_\infty, \|f_Y\|_\infty$, such that for any $x > 0$,*

$$(46) \quad \mathbb{P} \left\{ |U_n(I_m)| > \frac{C_\star}{n-1} \left(\sqrt{\Delta^{(1)}(m)\Delta(m)x} + \left(\sqrt{\Delta^{(1)}(m)\Delta(m) + \Delta^{(1)}(m)} \right) x + \frac{\Delta(m)x^2}{n} \right) \right\} \leq 5.54 \exp(-x),$$

where $\Delta(m)$ is defined in (13) and $\Delta^{(1)}(m)$ in (14).

In the sequel, we choose $x = 4 \log(n)$ and set

$$\Omega_1(m) := \left\{ |U_n(I_m)| \leq \frac{C_\star}{n-1} \left(\sqrt{\Delta^{(1)}(m)\Delta(m)4 \log(n)} + \left(\sqrt{\Delta^{(1)}(m)\Delta(m) + \Delta^{(1)}(m)} \right) 4 \log(n) + \frac{\Delta(m)16 \log^2(n)}{n} \right) \right\}.$$

Let us write that

$$\begin{aligned} \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(U_n^2(I_m) - \frac{\kappa_1}{36} V(m) \right)_+ \right\} \right] &\leq \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(U_n^2(I_m) - \frac{\kappa_1}{36} V_1(m) \right)_+ \right\} \mathbf{1}_{\Omega_1(m)} \right] = \mathbb{E}_1 \\ &+ \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(U_n^2(I_m) - \frac{\kappa_1}{36} V_1(m) \right)_+ \right\} \mathbf{1}_{\Omega_1^c(m)} \right] = \mathbb{E}_2. \end{aligned}$$

On $\Omega_1(m)$ and as $m \in \mathcal{M}_n$, we have for $C'_\star > 0$ a numerical constant $U_n^2(I_m) \leq C'_\star V_1(m)$, where $V_1(m)$ is defined in (19). Indeed, it is easy to check $\forall m \in \mathcal{M}_n$

$$\frac{\Delta(m)^2 \log^4(n)}{n^2(n-1)^2} \leq 4\rho m \frac{\Delta^{(1)}(m)^2 \log^4(n)}{n^2(n-1)^2} \lesssim \frac{\log^2(n)}{n(n-1)},$$

and

$$\begin{aligned} &\left(\sqrt{\Delta^{(1)}(m)\Delta(m)4 \log(n)} + \left(\sqrt{\Delta^{(1)}(m)\Delta(m) + \Delta^{(1)}(m)} \right) 4 \log(n) \right)^2 \frac{1}{(n-1)^2} \\ &\lesssim \Delta^{(1)}(m)\Delta(m) \frac{\log^2(n)}{n(n-1)}. \end{aligned}$$

Thus, for $\kappa_1 \geq 36C'_\star$, it yields $(U_n^2(I_m) - \frac{\kappa_1}{36} V_1(m)) \leq 0$ and \mathbb{E}_1 is equal to zero.

For the other term, it holds

$$(47) \quad \mathbb{E}_2 \leq \sum_{m \in \mathcal{M}_n} \mathbb{E} \left[U_n^2(I_m) \mathbf{1}_{\Omega_1^c(m)} \right].$$

By definition of $U_n(I_m)$ given in (40) and the Cauchy-Schwarz inequality, we have

$$|U_n(I_m)| \leq \frac{1}{n(n-1)} \sum_{k \neq \ell=1}^n \sqrt{\sum_{j=0}^{m-1} |v_{\varphi_j}(Y_k) - \mathbb{E}[v_{\varphi_j}(Y_k)]|^2} \sqrt{\sum_{j=0}^{m-1} |v_{\varphi_j}(Y_\ell) - \mathbb{E}[v_{\varphi_j}(Y_\ell)]|^2},$$

where v_t is given in (38) for $t \in S_m$. Furthermore, it yields by Bessel's inequality

$$(48) \quad \sum_{j=0}^{m-1} |v_{\varphi_j}(Y_k)|^2 \leq \frac{\Delta(m)}{2\pi}$$

and by the orthonormality of $(\varphi_j)_{j \geq 0}$ and $\|f\|_\infty < \infty$: $|\mathbb{E}[v_{\varphi_j}(Y_k)]|^2 = \frac{1}{4\pi^2} \int |f^*(u)|^2 du \int |\varphi_j^*(u)|^2 du \leq \|f\|_\infty$. It comes that

$$|U_n(I_m)| \leq 2 \left(\frac{\Delta(m)}{2\pi} + \|f\|_\infty \right) \leq (1 \vee 4\|f\|_\infty) \Delta(m).$$

Injecting this in (47), using $\Delta(m) \leq 2\sqrt{\rho m} \Delta^{(1)}(m)$ and by taking $x = 4 \log(n)$ in (46), we deduce that

$$\mathbb{E}_2 \leq 4\rho(1 \vee 4\|f\|_\infty)^2 \frac{n^2}{\log^2(n)} \sum_{m \in \mathcal{M}_n} \mathbb{P}(\Omega_1^c) \leq 4\rho \frac{(1 \vee 4\|f\|_\infty)^2}{n},$$

since $m \leq \sqrt{n}$ by assumption on \mathcal{M}_n . Part (i) of Lemma 7.2 is therefore proven.

7.7.2. *Proof of Part (ii).* Let us start by the centered process $\nu_{n,1}(\varphi_j)$ defined in (36). By the Cauchy-Schwarz inequality, it holds

$$\begin{aligned} \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \nu_{n,1}^2(\varphi_j) \right] &\leq \sum_{m \in \mathcal{M}_n} \text{Var}(\nu_{n,1}(\varphi_j)) \\ &= \frac{4}{n^2(n-1)^2} \sum_{m \in \mathcal{M}_n} \sum_{k \neq \ell=1}^n \text{Var} \left(\sum_{j=0}^{m-1} (v_{\varphi_j}(Y_k) - \mathbb{E}[v_{\varphi_j}(Y_k)]) \Psi_{\varphi_j}(Y_\ell) \right) \\ &= \frac{1}{n^2(n-1)^2} \sum_{m \in \mathcal{M}_n} \sum_{k \neq \ell=1}^n \mathbb{E} \left[\left| \sum_{j=0}^{m-1} (v_{\varphi_j}(Y_k) - \mathbb{E}[v_{\varphi_j}(Y_k)]) \Psi_{\varphi_j}(Y_\ell) \right|^2 \right] \\ &\leq \frac{1}{n^2(n-1)^2} \sum_{m \in \mathcal{M}_n} \sum_{k \neq \ell=1}^n \sum_{j=0}^{m-1} \mathbb{E} \left[|v_{\varphi_j}(Y_k) - \mathbb{E}[v_{\varphi_j}(Y_k)]|^2 \right] \sum_{j=0}^{m-1} \mathbb{E} \left[|\Psi_{\varphi_j}(Y_\ell)|^2 \right]. \end{aligned}$$

Thanks to (48), it yields

$$\sum_{j=0}^{m-1} \mathbb{E} \left[|v_{\varphi_j}(Y_k) - \mathbb{E}[v_{\varphi_j}(Y_k)]|^2 \right] \leq \sum_{j=0}^{m-1} \mathbb{E} \left[|v_{\varphi_j}(Y_k)|^2 \right] \leq \frac{\Delta(m)}{2\pi}.$$

As the integral $\int_{|u| > \sqrt{\rho m}} \frac{e^{-\xi u^2}}{|f_\varepsilon^*(u)|^2} du$ is convergent if f_ε is ordinary smooth or super smooth if $\delta < 2$ or $\delta = 2$ with $\mu \leq \xi$ and, by definition of $\Psi_{\varphi_j}(Y_\ell)$ given (39) and (4), it follows that

$$(49) \quad \begin{aligned} \mathbb{E} \left[|\Psi_{\varphi_j}(Y_\ell)|^2 \right] &\leq \mathbb{E} \left[\left| \int_{|u| > \sqrt{\rho m}} e^{iuY_\ell} \frac{\varphi_j(u)}{f_\varepsilon^*(u)} du \right|^2 \right] \\ &\leq C_\infty \mathbb{E} \left[\int_{|u| > \sqrt{\rho m}} \frac{e^{-\xi u^2}}{|f_\varepsilon^*(u)|^2} du \int_{|u| > \sqrt{\rho m}} e^{-\xi u^2} du \right] \\ &\lesssim e^{-\xi m} \int_{|u| > \sqrt{\rho m}} \frac{e^{-\xi u^2}}{|f_\varepsilon^*(u)|^2} du \lesssim e^{-\xi m}. \end{aligned}$$

Consequently, we get,

$$\mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \nu_{n,1}^2(\varphi_j) \right] \lesssim \frac{1}{n(n-1)} \sum_{m \in \mathcal{M}_n} m^2 \Delta(m) e^{-\xi m} \lesssim \frac{1}{n} \sum_{m=1}^{+\infty} e^{-\frac{\xi m}{2}} \lesssim \frac{1}{n},$$

because $\Delta(m) \leq 2\sqrt{\rho m} \Delta^{(1)}(m) \leq \frac{2\sqrt{\rho m}}{n}$ by definition of \mathcal{M}_n .

Let us now turn to the control of $\mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \nu_{n,2}^2(\varphi_j) \right] \leq \sum_{m \in \mathcal{M}_n} \text{Var}(\nu_{n,2}(\varphi_j))$ with $\nu_{n,2}(\varphi_j)$ defined in (42). Similarly to the study of $\nu_{n,1}(\varphi_j)$, we have

$$\mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \nu_{n,1}^2(\varphi_j) \right] \leq \frac{1}{n^2(n-1)^2} \sum_{m \in \mathcal{M}_n} \sum_{k \neq \ell=1}^n \sum_{j=0}^{m-1} \mathbb{E} \left[|\Psi_{\varphi_j}(Y_k)|^2 \right] \sum_{j=0}^{m-1} \mathbb{E} \left[|\Psi_{\varphi_j}(Y_\ell)|^2 \right].$$

Equation (49) imply that

$$\mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \nu_{n,1}^2(\varphi_j) \right] \leq \frac{1}{n^2(n-1)^2} \lesssim \frac{1}{n(n-1)} \sum_{m=1}^{+\infty} e^{-\xi m} \lesssim \frac{1}{n(n-1)}.$$

Hence the proof of Part (ii) and therefore Proposition 7.2 is proven.

7.8. Proof of Lemma 7.2. Let $x > 0$. Applying Theorem 3.4.8 given in Giné and Nickl (2016) (see Section A.2), we have

$$\mathbb{P} \left\{ n(n-1) |U_n(I_m)| > C \left(A_1 \sqrt{x} + A_2 x + A_3 x^{\frac{3}{2}} + A_4 x^2 \right) \right\} \leq 5.54 e^{-x},$$

where the constants $(A_i)_{i \in 1, \dots, 4}$ are given by:

$$\begin{aligned} A_1^2 &= \sum_{j=2}^n \sum_{i=1}^{j-1} \mathbb{E} [I_m^2(Y_i, Y_j)] \\ A_2 &= \sup \left\{ \mathbb{E} \left[\sum_{j=2}^n \sum_{i=1}^{j-1} I_m(Y_i, Y_j) \alpha_i(Y_i) \beta_j(Y_j) \right], \mathbb{E} \left[\sum_{i=1}^{n-1} \alpha_i^2(Y_i) \right] \leq 1, \mathbb{E} \left[\sum_{j=2}^n \beta_j^2(Y_j) \right] \leq 1 \right\}, \\ A_3^2 &= \max \left(\max_j \sup_x \left\{ \sum_{i=1}^{j-1} \mathbb{E}_{Y_i} [I_m^2(Y_i, x)] \right\}, \max_i \sup_x \left\{ \sum_{j=i+1}^n \mathbb{E}_{Y_j} [I_m^2(x, Y_j)] \right\} \right), \\ A_4 &= \sup_{x, y} |I_m(x, y)|, \end{aligned}$$

with I_m is given in (40). As the map $(x, y) \mapsto I_m(x, y)$ is symmetric in its entries (that is $I_m(x, y) = I_m(y, x)$) and the random variables Y_1, \dots, Y_n are i.i.d, the first three constants becomes (see also Giné and Nickl (2016), Section 3.4.3, pp. 176):

$$\begin{aligned} A_1^2 &= \frac{n(n-1)}{2} \mathbb{E} [|I_m(Y_1, Y_2)|^2], \\ A_2 &= \frac{n}{2} \sup \left\{ \mathbb{E} [I_m(Y_1, Y_2) \alpha(Y_1) \beta(Y_2)], \mathbb{E} [\alpha^2(Y_1)] \leq 1, \mathbb{E} [\beta^2(Y_2)] \leq 1 \right\} \\ A_3^2 &= (n-1) \sup_x \left\{ \mathbb{E}_{Y_1} [I_m^2(Y_1, x)] \right\} = (n-1) \sup_x \left\{ \mathbb{E}_{Y_2} [I_m^2(x, Y_2)] \right\}. \end{aligned}$$

Let us now compute these terms.

- **Computing of A_1 .** Notice that for any real-valued function t , the function $x \mapsto v_t(x)$ (defined in (38)) is also real-valued. By definition, we have

$$\begin{aligned} \mathbb{E}[|I_m(Y_1, Y_2)|^2] &= \sum_{j,k=0}^{m-1} \left\{ \mathbb{E} \left[(v_{\varphi_j}(Y_1) - \mathbb{E}[v_{\varphi_j}(Y_1)]) \left(\overline{v_{\varphi_k}(Y_1)} - \overline{\mathbb{E}[v_{\varphi_k}(Y_1)]} \right) \right] \right\}^2 \\ &\leq 2 \sum_{j,k=0}^{m-1} |\mathbb{E}[v_{\varphi_j}(Y_1)v_{\varphi_k}(Y_1)]|^2 + 2 \sum_{j,k=0}^{m-1} |\mathbb{E}[v_{\varphi_j}(Y_1)] \mathbb{E}[v_{\varphi_k}(Y_1)]|^2 \end{aligned}$$

Using the definition of v_t given in (38), as $\varphi_j^* = \sqrt{2\pi}(i^j)\varphi_j$ (see (5)) and Bessel's inequality (applied twice), it yields

$$\begin{aligned} \sum_{j,k=0}^{m-1} |\mathbb{E}[v_{\varphi_j}(Y_1)v_{\varphi_k}(Y_1)]|^2 &= \frac{1}{2\pi} \sum_{j,k=0}^{m-1} \left| \int \int \mathbb{E}[e^{iuY_1+ivY_1}] \frac{\varphi_j(u)}{f_\varepsilon^*(u)} \frac{\varphi_k(v)}{f_\varepsilon^*(v)} \mathbb{1}_{|u,v| \leq \sqrt{\rho m}} dudv \right|^2 \\ &\leq \frac{1}{2\pi} \int \int \frac{|f_Y^*(u+v)|^2}{|f_\varepsilon^*(u)f_\varepsilon^*(v)|^2} \mathbb{1}_{|u,v| \leq \sqrt{\rho m}} dudv. \end{aligned}$$

Since $\int |f_Y^*(u)| du \leq 2\pi \|f\| \|f_\varepsilon\|$ and $|f_Y^*(u)| \leq 1$, we get

$$\sum_{j,k=0}^{m-1} |\mathbb{E}[v_{\varphi_j}(Y_1)v_{\varphi_k}(Y_1)]|^2 \leq \|f\|_{\infty}^{\frac{1}{2}} \|f_\varepsilon\| \Delta(m) \Delta^{(1)}(m).$$

Again, with the Bessel inequality and (5), we have

$$\sum_{j=0}^{m-1} |\mathbb{E}[v_{\varphi_j}(Y_1)]|^2 = \frac{1}{2\pi} \sum_{j=0}^{m-1} \left| \int_{|u| \leq \sqrt{\rho m}} f^*(-u) \varphi_j(u) du \right|^2 \leq \frac{\|f^*\|^2}{2\pi} \leq \|f\|_{\infty}$$

Therefore, we deduce $\mathbb{E}[|I_m(Y_1, Y_2)|^2] \leq 2\|f\|_{\infty}^{\frac{1}{2}} \|f_\varepsilon\| \Delta(m) \Delta^{(1)}(m) + 2\|f\|_{\infty}^2$ and

$$A_1 \leq 2n(\|f\|_{\infty} \vee \|f\|_{\infty}^{\frac{1}{4}} \|f_\varepsilon\|^{\frac{1}{2}}) \sqrt{\Delta(m) \Delta^{(1)}(m)}.$$

- **Computing of A_2 .** We have by the Cauchy Schwarz inequately, $\mathbb{E}[\alpha_i^2(Y_1)] \leq 1$ and $\mathbb{E}[\beta_j^2(Y_2)] \leq 1$

$$\begin{aligned} \mathbb{E}[I_m(Y_1, Y_2) \alpha_i(Y_1) \beta_j(Y_2)] &= \iint I_m(u, v) \alpha_i(u) \beta_j(v) f_Y(u) f_Y(v) dudv \\ &\leq \sqrt{\iint |I_m(u, v)|^2 f_Y(u) f_Y(v) dudv} \sqrt{\iint \alpha_i^2(u) \beta_j^2(v) f_Y(u) f_Y(v) dudv} \\ &\leq \sqrt{\mathbb{E}[|I_m(Y_1, Y_2)|^2]}. \end{aligned}$$

Following the computation of A_1 , we derive

$$A_2 \leq n(\|f\|_{\infty} \vee \|f\|_{\infty}^{\frac{1}{4}} \|f_\varepsilon\|^{\frac{1}{2}}) \sqrt{\Delta(m) \Delta^{(1)}(m)}.$$

- **Computing of A_3 .** Let fixed x . Set $\tilde{a}_j = \frac{1}{2\pi} \int_{|u| \leq \sqrt{\rho m}} f^*(u) \varphi_j^*(-u) du$. By remarking that v_t is real-value and \tilde{a}_j is a real number, we write

$$\begin{aligned} \mathbb{E}_{Y_2}[|I_m(x, Y_2)|^2] &= \sum_{j,k=0}^{m-1} (v_{\varphi_j}(x) - \tilde{a}_j)(\overline{v_{\varphi_k}(x) - \tilde{a}_k}) \mathbb{E}_{Y_2} \left[(v_{\varphi_j}(Y_2) - \tilde{a}_j)(\overline{v_{\varphi_k}(Y_2) - \tilde{a}_k}) \right] \\ &= \sum_{j,k=0}^{m-1} (v_{\varphi_j}(x) - \tilde{a}_j)(v_{\varphi_k}(x) - \tilde{a}_k) \left(\int v_{\varphi_j} v_{\varphi_k} f_Y - \tilde{a}_k \int v_{\varphi_j} f_Y - \tilde{a}_j \int v_{\varphi_k} f_Y + \tilde{a}_j \tilde{a}_k \right). \end{aligned}$$

Let us recall that $\mathbb{E}[v_{\varphi_j}(Y_1)] = \tilde{a}_j$ by definition of v_t given in (38). Thus, Plancherel Equality and the orthogonality of (φ_j) imply

$$\begin{aligned} \mathbb{E}_{Y_2}[|I_m(x, Y_2)|^2] &\leq \sum_{j,k=0}^{m-1} (v_{\varphi_j}(x) - \tilde{a}_j)(v_{\varphi_k}(x) - \tilde{a}_k) \int v_{\varphi_j} v_{\varphi_k} f_Y = \int \left| \sum_{j=0}^{m-1} (v_{\varphi_j}(x) - \tilde{a}_j) v_{\varphi_j}(y) \right|^2 f_Y(y) dy \\ &\leq \|f_Y\|_\infty \int \left| \sum_{j=0}^{m-1} (v_{\varphi_j}(x) - \tilde{a}_j) v_{\varphi_j}(y) \right|^2 dy \\ &= \|f_Y\|_\infty \int \left| \left(z \mapsto \sum_{j=0}^{m-1} (v_{\varphi_j}(x) - \tilde{a}_j) \frac{1}{2\pi} \frac{\varphi_j^*(-z)}{f_\varepsilon^*(z)} \mathbf{1}_{|z| \leq \sqrt{\rho m}} \right)^* (y) \right|^2 dy \\ &= \|f_Y\|_\infty \int_{|z| \leq \sqrt{\rho m}} \left| \sum_{j=0}^{m-1} (v_{\varphi_j}(x) - \tilde{a}_j) \frac{\varphi_j(z)}{f_\varepsilon^*(z)} \right|^2 dz = \|f_Y\|_\infty \Delta^{(1)}(m) \sum_{j=0}^{m-1} (v_{\varphi_j}(x) - \tilde{a}_j)^2. \end{aligned}$$

It comes that

$$\mathbb{E}_{Y_2}[|I_m(x, Y_2)|^2] \leq \|f_Y\|_\infty \Delta^{(1)}(m) \sum_{j=0}^{m-1} (v_{\varphi_j}(x) - \tilde{a}_j)^2.$$

The Bessel and Plancherel equalities give (see (48))

$$\begin{aligned} \sum_{j=0}^{m-1} |v_{\varphi_j}(x) - \tilde{a}_j|^2 &\leq 2 \left(\sum_{j=0}^{m-1} |v_{\varphi_j}(x)|^2 + \sum_{j=0}^{m-1} \tilde{a}_j^2 \right) \\ &\leq 4(\|f\|_\infty \vee 1) \Delta(m). \end{aligned}$$

We derive that

$$\mathbb{E}_{Y_2}[|I_m(x, Y_2)|^2] \leq 4(\|f\|_\infty \vee 1) \|f_Y\|_\infty \Delta(m) \Delta^{(1)}(m),$$

and

$$A_3 \leq 2\sqrt{(\|f\|_\infty \vee 1) \|f_Y\|_\infty} \sqrt{n \Delta(m) \Delta^{(1)}(m)}.$$

- **Computing of A_4 .** Let us write

$$|I_m(x, y)| \leq \sum_{j=0}^{m-1} |(v_{\varphi_j}(x) - \tilde{a}_j)(v_{\varphi_j}(y) - \tilde{a}_j)|.$$

We bound successively the terms. By the Cauchy Schwarz inequality and (48), it holds

$$\sum_{j=0}^{m-1} |v_{\varphi_j}(x)v_{\varphi_j}(y)| \leq \sqrt{\sum_{j=0}^{m-1} |v_{\varphi_j}(x)|^2} \sqrt{\sum_{j=0}^{m-1} |v_{\varphi_j}(y)|^2} \leq \frac{\Delta(m)}{2\pi},$$

and

$$\sum_{j=0}^{m-1} |v_{\varphi_j}(x)\tilde{a}_j| \leq \sqrt{\sum_{j=0}^{m-1} |v_{\varphi_j}(x)|^2} \sqrt{\sum_{j=0}^{m-1} \tilde{a}_j^2} \leq \sqrt{\frac{\|f\|_\infty}{2}} \Delta(m),$$

since $\sum_{j=0}^{m-1} \tilde{a}_j^2 \leq \|f\|_\infty$. This implies $|I_m(x, y)| \leq 4\left(\frac{1}{2\pi} \vee \sqrt{\frac{\|f\|_\infty}{2}} \vee \|f\|_\infty\right) \Delta(m)$ and

$$A_4 \leq 12 \left(\frac{1}{2\pi} \vee \sqrt{\frac{\|f\|_\infty}{2}} \vee \|f\|_\infty \right) \Delta(m).$$

Collecting the above evaluations, it yields $\forall x > 0$

$$\begin{aligned} \mathbb{P} \left\{ |U_n(I_m)| > \frac{C_\star}{n-1} \left(\sqrt{\Delta^{(1)}(m)\Delta(m)x} + \sqrt{\Delta^{(1)}(m)\Delta(m)x} + \frac{\sqrt{\Delta(m)\Delta^{(1)}(m)x^{\frac{3}{2}}}}{\sqrt{n}} + \frac{\Delta(m)x^2}{n} \right) \right\} \\ \leq 5.54 \exp(-x), \end{aligned}$$

where $C_\star > 0$ depending on f_ε , $\|f\|_\infty$ and $\|f_Y\|_\infty$. Since $2\frac{\sqrt{\Delta(m)\Delta^{(1)}(m)x^{\frac{3}{2}}}}{\sqrt{n}} \leq x\Delta^{(1)}(m) + \frac{\Delta(m)x^2}{n}$, it follows that

$$\begin{aligned} \mathbb{P} \left\{ |U_n(I_m)| > \frac{C_\star}{n-1} \left(\sqrt{\Delta^{(1)}(m)\Delta(m)x} + \left(\sqrt{\Delta^{(1)}(m)\Delta(m)} + \Delta^{(1)}(m) \right) x + \frac{\Delta(m)x^2}{n} \right) \right\} \\ \leq 5.54 \exp(-x). \end{aligned}$$

Hence the announced result.

7.9. Proof of Proposition 7.3. We only proof Part (i). The proof of Part (ii) can be found in [Sacko \(2020\)](#) (see Proof of Lemma 7.3, p. 23). Notice that

$$\begin{aligned} \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(P_{n,1}^2(2f_m) - \frac{\kappa_1}{24} V(m) \right)_+ \right\} \right] &\leq \sum_{m \in \mathcal{M}_n} \int_{\mathbb{R}^+} \mathbb{P} \left\{ \left(P_{n,1}^2(2f_m) - \frac{\kappa_1}{24} V(m) \right)_+ > x \right\} dx \\ &\leq \sum_{m \in \mathcal{M}_n} \int_{\mathbb{R}^+} \mathbb{P} \left\{ |P_{n,1}(2f_m)| \geq \sqrt{\frac{\kappa_1}{24} V(m) + x} \right\} dx. \end{aligned}$$

Now, we apply Bernstein inequality given in Appendix A.1. We need to evaluate two quantities b and v . We distinguish two cases.

The ordinary smooth case. By the Cauchy-Schwarz inequality, it holds

$$\begin{aligned} |v_{2f_m}(Y_k)| &= \frac{1}{2\pi} \left| \int_{|u| \leq \sqrt{\rho m}} e^{iuY_k} \frac{(2f_m)^*(-u)}{f_\varepsilon^*(u)} du \right| \leq \frac{1}{\pi} \sqrt{\|f_m^*\|^2 \int_{|u| \leq \sqrt{\rho m}} \frac{1}{|f_\varepsilon(u)|^2} du} \\ &\leq \sqrt{2\|f\|_\infty \Delta(m)} = b. \end{aligned}$$

By the Plancherel theorem and under (H5), we take (see (20))

$$\mathbb{E} [|v_{2f_m}(Y_k)|^2] \leq 4\|f_Y\|_\infty \int_{|u| \leq \sqrt{\rho m}} \frac{|f_m(u)|^2}{|f_\varepsilon^*(u)|^2} du = 4\|f_Y\|_\infty V_2(m) \frac{n}{\log(n)} = v.$$

Applying Bernstein inequality leads to have for any $\alpha \in]0, 1[$

$$\begin{aligned} & \mathbb{P} \left\{ |P_{n,1}(2f_m)| \geq \sqrt{\frac{\kappa_1}{24} V(m) + x} \right\} \\ & \leq 2 \left\{ \exp \left(-\frac{n}{4v} \left(\frac{\kappa_1}{24} V(m) + x \right) \right) \vee \exp \left(-\frac{n}{4b} \sqrt{\frac{\kappa_1}{24} V(m) + x} \right) \right\} \\ & \leq 2 \left\{ \exp \left(-\frac{n}{4v} \left(\frac{\kappa_1}{24} V(m) + x \right) \right) \vee \exp \left(-\frac{n\alpha}{4b} \sqrt{\frac{\kappa_1}{24} V(m)} \right) \exp \left(-\frac{n(1-\alpha)}{4b} \sqrt{x} \right) \right\}. \end{aligned}$$

Moreover, for any $\kappa_1 \geq 384 \|f_Y\|_\infty p$, we have

$$\frac{n}{4v} \frac{\kappa_1}{24} V(m) \geq \frac{V_2(m)}{384 \|f_Y\|_\infty V_2(m)} \log(n) \geq p \log(n),$$

where p is an integer which shall be specified in the sequel. It also holds for any $m \in \mathcal{M}_n$ and $\kappa_1 \geq 384 \|f\|_\infty p^2 / (\alpha^2 c'_1)$

$$\frac{n\alpha}{4b} \sqrt{\frac{\kappa_1}{24} V(m)} \geq \frac{n\alpha}{4b} \sqrt{\frac{\kappa_1}{24} V_1(m)} \frac{n\alpha}{4\sqrt{2} \|f\|_\infty \Delta(m)} \sqrt{\frac{\kappa_1}{24} \Delta(m) \Delta^{(1)}(m) \frac{\log^2(n)}{n^2}} \geq p \log(n),$$

since $V(m) = V_1(m) + V_2(m) \geq V_1(m) = \Delta^{(1)}(m) \Delta(m) \frac{\log^2(n)}{n^2}$ and $\Delta^{(1)}(m) \geq c'_1$ (see (7)). Therefore, we obtain for κ_1 large enough

$$(50) \quad \mathbb{P} \left\{ |P_{n,1}(2f_m)| \geq \sqrt{\frac{\kappa_1}{24} V(m) + x} \right\} \leq 2n^{-p} \left\{ \exp \left(-\frac{n}{4v} x \right) \vee \exp \left(-\frac{n(1-\alpha)}{4b} \sqrt{x} \right) \right\}.$$

The super smooth case. We take an upper bound of $\int_{|u| \leq \sqrt{\rho m}} \frac{|f_m(u)|^2}{|f_\varepsilon^*(u)|^2} du$ in such a way as to set $v = 4 \|f\|_\infty \|f_Y\|_\infty \Delta^{(1)}(m)$ and b does not change. Again with $V(m) = V_1(m) + V_2(m) \geq V_2(m) = \Delta^{(1)}(m) \log(n)/n$, it holds for $\kappa_1 \geq 384 \|f\|_\infty \|f_Y\|_\infty$ that $\frac{n}{4v} \frac{\kappa_1}{24} V(m) \geq p \log(n)$. Thanks to $V(m) \geq \Delta^{(1)}(m) \Delta(m) \frac{\log^2(n)}{n^2}$ and $\kappa_1 \geq 384 \|f\|_\infty p^2 / (\alpha^2 c'_1)$, we get $\frac{n\alpha}{4b} \sqrt{\frac{\kappa_1}{24} V(m)} \geq p \log(n)$. We derive that (50) holds for the super smooth case.

By choosing $p = 3$ with some changes variables, we deduce

$$\begin{aligned} \mathbb{E} \left[\sup_{m \in \mathcal{M}_n} \left\{ \left(P_{n,1}^2(2f_m) - \frac{\kappa_1}{24} V(m) \right)_+ \right\} \right] & \leq 2n^{-p} \sum_{m \in \mathcal{M}_n} \int_{\mathbb{R}^+} \left\{ \exp \left(-\frac{n}{4v} x \right) \vee \exp \left(-\frac{n(1-\alpha)}{4b} \sqrt{x} \right) \right\} dx \\ & \leq 2n^{-p} \sum_{m \in \mathcal{M}_n} \left(\frac{4 \|f_Y\|_\infty \|f\|_\infty \Delta^{(1)}(m)}{n} \vee 32 \|f\|_\infty \frac{\Delta(m)}{n^2 (1-\alpha)^2} \right) \leq \frac{C_1^*}{n}, \end{aligned}$$

where C_1^* is a constant depending on $\|f\|_\infty$, $\|f_Y\|_\infty$ and f_ε via (7). This concludes the proof.

APPENDIX A. CONCENTRATION INEQUALITIES

A.1. Bernstein inequality. Let Y_1, \dots, Y_n be n independent real random variables and $S_n = \sum_{i=1}^n (Y_i - \mathbb{E}[Y_i])$. Assume there exist two constants s^2 and b , such that $\text{Var}(Y_i) \leq s^2$ and $|Y_i| \leq b$. Then, for all $x > 0$, we have

$$\mathbb{P}(|S_n| \geq nx) \leq 2 \max \left(\exp \left(-\frac{nx}{4v^2} \right), \exp \left(-\frac{nx^2}{4b} \right) \right).$$

Moreover, it holds

$$\mathbb{P} \left(|S_n| \geq \sqrt{2ns^2x} + \frac{bx}{3} \right) \leq 2e^{-x}.$$

A proof of Bernstein inequality can be found in [Birgé and Massart \(1998\)](#), p. 366.

A.2. Exponential inequalities for U-statistics of order two. Let Y_1, \dots, Y_n be n independent random variables defined on a Polish space S and let us consider the degenerate (or canonical) U-statistic of order 2: $U_n(h) = \sum_{i \neq j=1}^n h(Y_i, Y_j)$, where the function $h : S^2 \mapsto \mathbb{R}$ is such that $\mathbb{E}[|h(Y_i, Y_j)|] < \infty$ for any $1 \leq i, j \leq n$. Let A_1, A_2, A_3 and A_4 be defined as follows:

$$\begin{aligned} A_1^2 &= \sum_{j=2}^n \sum_{i=1}^{j-1} \mathbb{E} [h^2(Y_i, Y_j)] \\ A_2 &= \sup \left\{ \mathbb{E} \left[\sum_{j=2}^n \sum_{i=1}^{j-1} h(Y_i, Y_j) \alpha_i(Y_i) \beta_j(Y_j) \right], \mathbb{E} \left[\sum_{i=1}^{n-1} \alpha_i^2(Y_i) \right] \leq 1, \mathbb{E} \left[\sum_{j=2}^n \beta_j^2(Y_j) \right] \leq 1 \right\}, \\ A_3^2 &= \max \left(\max_j \sup_x \left\{ \sum_{i=1}^{j-1} \mathbb{E}_{Y_i} [h^2(Y_i, x)] \right\}, \max_i \sup_x \left\{ \sum_{j=i+1}^n \mathbb{E}_{Y_j} [h^2(x, Y_j)] \right\} \right), \\ A_4 &= \sup_{x,y} |h(x, y)|, \end{aligned}$$

where $\mathbb{E}_Y[X]$ denotes the conditional expectation of X given Y . Then, there exists an absolute constant $C > 0$ such that for any $x > 0$,

$$\mathbb{P} \left\{ |U_n(h)| \geq C \left(A_1 \sqrt{x} + A_2 x + A_3 x^{\frac{3}{2}} + A_4 x^2 \right) \right\} \leq 5.54e^{-x}.$$

The Bernstein-type inequality for canonical U-statistics of order 2 is proven in [Houdré and Reynaud-Bouret \(2003\)](#) for real variables. The above result is a simplified version of Theorem 3.4.8 given in [Giné and Nickl \(2016\)](#).

Acknowledgements: The author thanks F. Comte for helpful advices and kind proofreadings.

REFERENCES

- Abramowitz, M. and Stegun, I. A. (1964). *Handbook of mathematical functions with formulas, graphs, and mathematical tables*, volume 55 of *National Bureau of Standards Applied Mathematics Series*. For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, D.C.
- Askey, R. and Wainger, S. (1965). Mean convergence of expansions in Laguerre and Hermite series. *Amer. J. Math.*, 87:695–708. MR182834.
- Belomestny, D., Comte, F., and Genon-Catalot, V. (2019). Sobolev-Hermite versus Sobolev nonparametric density estimation on \mathbb{R} . *Ann. Inst. Statist. Math.*, 71(1):29–62.
- Bickel, P. J. and Ritov, Y. (1988). Estimating integrated squared density derivatives: sharp best order of convergence estimates. *Sankhyā Ser. A*, 50(3):381–393.
- Birgé, L. and Massart, P. (1995). Estimation of integral functionals of a density. *Ann. Statist.*, 23(1):11–29.
- Birgé, L. and Massart, P. (1998). Minimum contrast estimators on sieves: exponential bounds and rates of convergence. *Bernoulli*, 4(3):329–375.
- Bongioanni, B. and Torrea, J. L. (2006). Sobolev spaces associated to the harmonic oscillator. *Proc. Indian Acad. Sci. Math. Sci.*, 116(3):337–360.

- Butucea, C. (2007). Goodness-of-fit testing and quadratic functional estimation from indirect observations. *Ann. Statist.*, 35(5):1907–1930.
- Butucea, C., Rohde, A., and Steinberger, L. (2023). Interactive versus noninteractive locally differentially private estimation: two elbows for the quadratic functional. *Ann. Statist.*, 51(2):464–486.
- Chesneau, C. (2011). On adaptive wavelet estimation of a quadratic functional from a deconvolution problem. *Ann. Inst. Statist. Math.*, 63(2):405–429.
- Comte, F. and Lacour, C. (2011). Data-driven density estimation in the presence of additive noise with unknown distribution. *J. R. Stat. Soc. Ser. B Stat. Methodol.*, 73(4):601–627.
- Efromovich, S. and Low, M. (1996). On Bickel and Ritov’s conjecture about adaptive estimation of the integral of the square of density derivative. *Ann. Statist.*, 24(2):682–686.
- Giné, E. and Nickl, R. (2016). *Mathematical foundations of infinite-dimensional statistical models*. Cambridge Series in Statistical and Probabilistic Mathematics, [40]. Cambridge University Press, New York.
- Goldenshluger, A. and Lepski, O. (2011). Bandwidth selection in kernel density estimation: oracle inequalities and adaptive minimax optimality. *Ann. Statist.*, 39(3):1608–1632.
- Goldenshluger, A. and Lepski, O. V. (2022a). Minimax estimation of norms of a probability density: I. Lower bounds. *Bernoulli*, 28(2):1120–1154.
- Goldenshluger, A. and Lepski, O. V. (2022b). Minimax estimation of norms of a probability density: II. Rate-optimal estimation procedures. *Bernoulli*, 28(2):1155–1178.
- Houdré, C. and Reynaud-Bouret, P. (2003). Exponential inequalities, with constants, for u-statistics of order two. In *Stochastic inequalities and applications*, pages 55–69. Springer.
- Indritz, J. (1961). An inequality for Hermite polynomials. *Proc. Amer. Math. Soc.*, 12:981–983.
- Kerkycharian, G. and Picard, D. (1996). Estimating nonquadratic functionals of a density using Haar wavelets. *Ann. Statist.*, 24(2):485–507.
- Lacour, C. and Massart, P. (2016). Minimal penalty for Goldenshluger-Lepski method. *Stochastic Process. Appl.*, 126(12):3774–3789.
- Laurent, B. (1996). Efficient estimation of integral functionals of a density. *Ann. Statist.*, 24(2):659–681.
- Laurent, B. (2005). Adaptive estimation of a quadratic functional of a density by model selection. *ESAIM Probab. Stat.*, 9:1–18.
- Loubes, J. M. and Marteau, C. (2014). Goodness-of-fit testing strategies from indirect observations. *J. Nonparametr. Stat.*, 26(1):85–99.
- Sacko, O. (2020). Hermite density deconvolution. *ALEA Lat. Am. J. Probab. Math. Stat.*, 17(1):419–443.
- Schluttenhofer, S. and Johannes, J. (2020). Minimax testing and quadratic functional estimation for circular convolution. *arXiv preprint arXiv:2004.12714*.
- Tribouley, K. (2000). Adaptive estimation of integrated functionals. *Math. Methods Statist.*, 9(1):19–38.