

ADAPTIVE ESTIMATION OF RANDOM-EFFECTS DENSITIES IN LINEAR MIXED-EFFECTS MODEL

GWENNAËLLE MABON^{†,‡}

[†]CREST

3 avenue Pierre Larousse
92245 Malakoff, France

[‡]MAP5, Université Paris Descartes

45 rue des Saints-Pères
75006 Paris, France

ABSTRACT. In this paper we consider the problem of adaptive estimation of random-effects densities in linear mixed-effects model. The linear mixed-effects model is defined as $Y_{k,j} = \alpha_k + \beta_k t_j + \varepsilon_{k,j}$ where $Y_{k,j}$ is the observed value for individual k at time t_j for $k = 1, \dots, N$ and $j = 1, \dots, J$. Random variables (α_k, β_k) are known as random effects and stand for the individual random variables of entity k . We denote their densities f_α and f_β and assume that they are independent of the measurement errors $(\varepsilon_{k,j})$. We introduce kernel estimators and present upper risk bounds. We also give rates of convergence. The focus of this work lies on the optimal data driven choice of the smoothing parameter using a penalization strategy in the particular case of fixed interval between times t_j .

Keywords. Adaptive estimation. Nonparametric density estimation. Deconvolution. Linear mixed-effects model. Random effect density. Mean square risk.

1. INTRODUCTION

Mixed models are models which bring together fixed effects and random effects. They allow analysis of repeated measurements or longitudinal data. In this paper, we concentrate on linear mixed-effects models defined as

$$Y_{k,j} = \alpha_k + \beta_k t_j + \varepsilon_{k,j}, \quad k = 1, \dots, N \quad \text{and} \quad j = 1, \dots, J \quad (1)$$

where $Y_{k,j}$ denotes the observed value for individual k at time t_j and (α_k, β_k) represent the individual random variables of entity k . They are known as random effects. $(\varepsilon_{k,j})$ are the measurement errors. We denote their densities f_α , f_β and f_ε . We do not assume that α_k and β_k are independent. We make the following assumptions:

- (A1) Times $(t_j)_{1 \leq j \leq J}$ are known and deterministic and $\Delta_j = \Delta$ for all j , $t_j = j\Delta$ and $J \geq 6$.
- (A2) $(\varepsilon_{k,j})_{k,j}$ are *i.i.d.* with distribution f_ε and the Fourier transform of f_ε does not vanish on the real line.
- (A3) (α_k, β_k) are *i.i.d.* with respective distribution f_α and f_β .
- (A4) (α_k, β_k) are independent of $(\varepsilon_{k,j})_{k,j}$.

This aim of this paper is to recover the densities f_α and f_β from the data $(Y_{k,j})$ in a non-parametric setting.

Mixed models have been widely studied in a parametric context. For example, [Pinheiro and Bates \(2000\)](#) have considered the problem assuming that both random effects and measurement errors are Gaussian, which enables them to use a maximum likelihood approach. Nonetheless the normality assumption can be too strong in some cases. In this way, [Wu and Zhu \(2010\)](#) relaxed the normality assumption estimating the first four moments of the random-effects density. We can also cite previous works of [Shen and Louis \(1999\)](#) who consider a smoothing method without any assumption on the error distribution f_ε , [Zhang and Davidian \(2001\)](#) and [Vock et al. \(2011\)](#) who propose a semi-nonparametric approach based on the approximation of the random-effects density by an Hermite series assuming that the error distribution is a Gaussian.

Here we consider an approach based on deconvolution methods. The convolution model a classical setting in nonparametric statistics which has been widely studied. There exists a large amount of literature on the

E-mail address: gwennaelle.mabon@ensae.fr.

Date: March 13, 2014.

subject assuming first that the noise density is known. We can cite Carroll and Hall (1988), Stefanski (1990), Stefanski and Carroll (1990), Fan (1991), Efromovich (1997) and Delaigle and Gijbels (2004) who study rates of convergence and their optimality for kernel estimators or Butucea (2004), Butucea and Tsybakov (2008a,b) for studies of rate optimality in the minimax sense. Yet the drawback of these methods is that they all work under the assumption that the error distribution is known. However, the main goal of this paper lies in an adaptive choice of a smoothing parameter. For the most part, the adaptive bandwidth selection in deconvolution models has been addressed with a known error distribution, see for example Pensky and Vidakovic (1999) for wavelet strategy, Comte et al. (2006), Butucea and Comte (2009) for projection strategies, or Meister (2009) and references therein. Adaptive estimation in deconvolution problems with unknown error density has been recently studied in a rigorous way. Several papers focus on that matter as those of Comte and Lacour (2011), Johannes and Schwarz (2012), Dattner et al. (2013), Kappus (2014) and Kappus and Mabon (2013). Rates of convergence have been presented in Neumann (1997) and, more recently, in Johannes (2009), or Meister (2009) under the assumption that a preliminary sample of the noise ε is observed.

More precisely, we follow an approach introduced in deconvolution literature but in this particular context we apply a strategy based on repeated measurements in density estimation in deconvolution models. Rates of convergence in a repeated observations model have been presented in Li and Vuong (1998), Neumann (2007), Delaigle et al. (2008) and Comte et al. (2013). For the point of view of adaptive bandwidth selection, we can cite Delaigle et al. (2008). More recently, Kappus and Mabon (2013) study deconvolution model in the context of repeated measurements and achieve a new adaptive procedure. Their method has the advantage of deriving a nearly optimal data driven choice of the smoothing parameter using a penalization strategy under very weak assumptions: in particular no semi parametric assumptions on the shape of the characteristic function of the noise is required. In this paper, we propose to adapt their method in the context of mixed-effects model. For that we are aware of the work of Comte and Samson (2012) who propose nonparametric and nonadaptive estimators of the random effects under several assumptions on the noise, mainly assumed to be known and Dion (2013) who study nonparametric estimators based on Lepski's method and others based on a contrast penalization, but in another asymptotic framework.

This paper is organized as follows. In Section 2, we give the notations, specify the statistical model and estimation procedure for f_α and f_β . In Section 3, we present upper bounds for both densities and rates of convergence for f_β . In Section 4, we introduce adaptive estimators and propose a new adaptive procedure by penalization in the context of linear mixed-effects model under weak assumptions inspired by the work of Kappus and Mabon (2013). Besides the theoretical properties of the adaptive estimators are studied. In Section 5, we lead a study of the adaptive estimators through simulation experiments. Numerical results are then presented. All the proofs are postponed to Section 6.

2. STATISTICAL MODEL AND ESTIMATION PROCEDURE

2.1. Notations. For two real numbers a and b , we denote $a \vee b = \max(a, b)$, $a \wedge b = \min(a, b)$ and $(a)_+ = \max(a, 0)$. For two functions $\varphi, \psi : \mathbb{R} \rightarrow \mathbb{C}$ belonging to $\mathbb{L}^1(\mathbb{R}) \cap \mathbb{L}^2(\mathbb{R})$, we denote $\|\varphi\|$ the \mathbb{L}^2 norm of φ defined by $\|\varphi\|^2 = \int_{\mathbb{R}} |\varphi(x)|^2 dx$, $\langle \varphi, \psi \rangle$ the scalar product between φ and ψ defined by $\langle \varphi, \psi \rangle = \int_{\mathbb{R}} \varphi(x) \overline{\psi(x)} dx$. The Fourier transform φ^* is defined by

$$\varphi^*(x) = \int e^{ixu} \varphi(u) du.$$

Besides, if φ^* belongs to $\mathbb{L}^1(\mathbb{R}) \cap \mathbb{L}^2(\mathbb{R})$, then the function φ is the inverse Fourier transform of φ^* and can be written $\varphi(x) = 1/(2\pi) \int e^{-ixu} \varphi^*(u) du$. Lastly the convolution product $*$ is defined as $(\varphi * \psi)(x) = \int \varphi(x-u) \psi(u) du$.

2.2. Estimation procedure.

2.2.1. Adaptive estimator of f_α . If observations for $t_0 = 0$ are available, then Model (1) writes

$$Y_{k,0} = \alpha_k + \varepsilon_{k,0} \quad \text{and} \quad k = 1, \dots, N. \quad (2)$$

This is a classical deconvolution model. Nonetheless this model has been widely studied when f_ε^* is assumed to be known, see Stefanski and Carroll (1990), Fan (1991), Pensky and Vidakovic (1999) and Comte et al. (2006).

In this paper, we consider the context of unknown measurement errors. This problem has been studied by Comte and Lacour (2011) when a preliminary sample of the error distribution is available. Yet this is not the case we consider. In the context of repeated observations models, we can refer to the works of Li and

Vuong (1998), Neumann (2007), Delaigle et al. (2008), Comte et al. (2013) and the recent work of Kappus and Mabon (2013).

The density distribution of $Y_{k,0}$ is noted f_Y . Under Model (2) and independence assumptions we have clearly that

$$f_Y = f_\alpha * f_\varepsilon$$

which implies that

$$f_\alpha^* = \frac{f_Y^*}{f_\varepsilon^*}.$$

In this case, we have that

$$f_{\alpha,m}(x) = \frac{1}{2\pi} \int_{-\pi m}^{\pi m} e^{-iux} \frac{f_Y^*(u)}{f_\varepsilon^*(u)} du. \quad (3)$$

If f_ε^* were known, we could simply estimate f^* with $\hat{f}_Y^*/f_\varepsilon^*$ where \hat{f}_Y^* is an estimator obtained directly from the data with a simple empirical estimator. We should only apply the inverse Fourier transform to get an estimate of f . Nevertheless, $1/f_\varepsilon^*$ is not integrable over \mathbb{R} because f_ε^* decreases to 0 near infinity. That is why we cannot compute the inverse Fourier transform over \mathbb{R} . We need to regularize the problem, for example, with a spectral cutoff parameter. In this particular case, the estimator of f would be $1/(2\pi) \int_{|u| \leq \pi m} e^{-iux} \hat{f}_Y^*(u)/f_\varepsilon^*(u) du$. We can notice that this estimator corresponds both to a kernel estimator built with a sinc kernel (Butucea (2004)) or to a projection type estimator as in Comte et al. (2006).

In this paper the error distribution is assumed to be unknown. To make the problem identifiable, some additional information on the noise is required. When $t_0 = 0$, linear mixed-effects models can be seen as repeated observation models. Therefore we can recover an estimation of the error distribution from the following data:

$$U_k = Y_{k,4} - Y_{k,3} - (Y_{k,2} - Y_{k,1}) = \varepsilon_{k,4} - \varepsilon_{k,3} - \varepsilon_{k,2} + \varepsilon_{k,1}$$

which imply have the following equality under (A2)

$$f_U^*(x) = \mathbb{E} [e^{ixU}] = |f_\varepsilon^*(x)|^4.$$

For the estimation of f_α , we add the following assumption:

(A5) ε is symmetric.

This latest assumption implies that f_ε^* is real-valued and positive, so that $f_U^*(x) = (f_\varepsilon^*(x))^4$. As a consequence f_ε^{*4} can be estimated as follows

$$\widehat{f_\varepsilon^{*4}}(x) = \left(\frac{1}{N} \sum_{k=1}^N \cos(xU_k) \right)_+. \quad (4)$$

Nevertheless we need to prevent $\widehat{f_\varepsilon^{*4}}$ to become too small. For that we introduce a regularization of the Fourier transform by truncating the estimator following methods presented in Neumann (1997), Comte and Lacour (2011), Kappus (2014) and Kappus and Mabon (2013). We define the following threshold

$$k_N(x) = s_N(x)N^{-1/2} \quad (5)$$

where $s_N(x) \geq 1$ and $s_N(x) = O(N^{1/2})$ when $N \rightarrow \infty$ for all x . Now we can introduce another estimator of f_ε^* .

$$\check{f}_\varepsilon^*(x) = \begin{cases} (\widehat{f_\varepsilon^{*4}}(x))^{1/4} & \text{if } \widehat{f_\varepsilon^{*4}}(x) \geq k_N(x), \\ (k_N(x))^{1/4} & \text{otherwise.} \end{cases} \quad (6)$$

So using the inverse Fourier transform, we can estimate $f_{\alpha,m}$ as follows

$$\hat{f}_{\alpha,m}(x) = \frac{1}{2\pi} \int_{-\pi m}^{\pi m} e^{-ixu} \frac{\hat{f}_Y^*(u)}{\check{f}_\varepsilon^*(u)} du \quad (7)$$

where $\hat{f}_Y^*(u) = 1/N \sum_{k=1}^N e^{iuY_{k,0}}$.

2.2.2. *Adaptive estimator of f_β .* For the estimation of f_β , we use another approach to see the problem as a deconvolution problem. Without loss of generality we assume that J is even. So for $1 \leq j \leq J/2$, we can transform the data as follows

$$Z_{k,j} = \frac{Y_{k,2j} - Y_{k,2j-1}}{\Delta} = \beta_k + \frac{\varepsilon_{k,2j} - \varepsilon_{k,2j-1}}{\Delta} = \beta_k + \frac{\eta_{k,j}}{\Delta}, \quad \eta_{k,j} = \varepsilon_{k,2j} - \varepsilon_{k,2j-1}. \quad (8)$$

Let us notice that for a fixed j , the $(Z_{k,j})$ for $k = 1, \dots, N$ are *i.i.d.* but $Z_{k,j}$ and $Z_{k,l}$ for $j \neq l$ are not independent. It means that we preserve the independence between individuals of the sample.

Since β_k is independent of $\eta_{k,j}$ under **(A4)**, we can write the following equality

$$f_{Z_j} = f_\beta * f_{\frac{\eta_{k,j}}{\Delta}}$$

which implies

$$f_{Z_j}^*(x) = f_\beta^*(x) |f_\varepsilon^*\left(\frac{x}{\Delta}\right)|^2.$$

So under **(A2)** we have

$$f_\beta^*(x) = \frac{f_{Z_j}^*(x)}{|f_\varepsilon^*\left(\frac{x}{\Delta}\right)|^2}.$$

Now using all the observations j we can write that

$$f_\beta^*(x) = \frac{2}{J} \sum_{j=1}^{J/2} \frac{f_{Z_j}^*(x)}{|f_\varepsilon^*\left(\frac{x}{\Delta}\right)|^2}.$$

Unlike in the estimation of f_α , we do not need to estimate the Fourier transform of the error distribution but only $|f_\varepsilon^*|^2$: that is why we do not assume here that the noise is symmetric. Let us notice the following equality

$$\frac{U_k}{\Delta} = Z_{k,2} - Z_{k,1} = \frac{1}{\Delta} (\varepsilon_{k,4} - \varepsilon_{k,3} - \varepsilon_{k,2} + \varepsilon_{k,1}).$$

Then we have

$$f_{\frac{U}{\Delta}}^*(x) = \mathbb{E} \left[e^{ixU/\Delta} \right] = |f_\varepsilon^*\left(\frac{x}{\Delta}\right)|^4. \quad (9)$$

So $|f_\varepsilon^*|^4$ can be estimated as follows

$$\widehat{|f_\varepsilon^{*4}|}\left(\frac{x}{\Delta}\right) = \left(\frac{1}{N} \sum_{k=1}^N \cos(x \frac{U_k}{\Delta}) \right)_+. \quad (10)$$

And to prevent the denominator from becoming too small, we regularize the Fourier transform of the error distribution as follows

$$|\tilde{f}_\varepsilon^*\left(\frac{x}{\Delta}\right)|^2 = \begin{cases} |\hat{f}_\varepsilon^*\left(\frac{x}{\Delta}\right)|^2 = \left(\widehat{|f_\varepsilon^{*4}|}\left(\frac{x}{\Delta}\right)\right)^{1/2} & \text{if } \widehat{|f_\varepsilon^{*4}|}\left(\frac{x}{\Delta}\right) \geq k_N\left(\frac{x}{\Delta}\right), \\ (k_N\left(\frac{x}{\Delta}\right))^{1/2} & \text{otherwise.} \end{cases} \quad (11)$$

Thus we can estimate f_β^* as follows

$$\hat{f}_\beta^*(x) = \frac{2}{J-4} \sum_{j=3}^{J/2} \frac{\hat{f}_{Z_j}^*(x)}{|\tilde{f}_\varepsilon^*\left(\frac{x}{\Delta}\right)|^2} \quad \text{with} \quad \hat{f}_{Z_j}^*(x) = \frac{1}{N} \sum_{k=1}^N e^{ixY_{k,j}}, \quad j = 3, \dots, J/2.$$

We emphasize that the previous definition uses distinct observations for \tilde{f}_ε^* and $\hat{f}_{Z_j}^*$, so that the numerator and the denominator are independent. This is why **(A1)** requires $J \geq 6$. We then define $f_{\beta,m}$ as follows

$$f_{\beta,m}(x) = \frac{1}{2\pi} \int_{-\pi m}^{\pi m} e^{-ixu} \frac{2}{J-4} \sum_{j=3}^{J/2} \frac{f_{Z_j}^*(u)}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} du. \quad (12)$$

Applying the inverse Fourier transform, we get an estimate of $f_{\beta,m}$

$$\begin{aligned} \hat{f}_{\beta,m}(x) &= \frac{1}{2\pi} \int_{-\pi m}^{\pi m} e^{-ixu} \frac{2}{J-4} \sum_{j=3}^{J/2} \frac{\hat{f}_{Z_j}^*(u)}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} du \\ &= \frac{1}{2\pi} \int_{-\pi m}^{\pi m} e^{-ixu} \frac{\hat{f}_Z^*(u)}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} du \end{aligned} \quad (13)$$

where $\hat{f}_Z^*(x) = \frac{2}{J-4} \sum_{j=3}^{J/2} \hat{f}_{Z_j}^*(x)$.

3. UPPER BOUND ON THE \mathbb{L}^2 RISK

Before giving upper bounds for $\hat{f}_{\alpha,m}$ and $\hat{f}_{\beta,m}$, we need two key lemmas which are similar to Neumann's lemma but are derived for our particular estimators. These lemmas study the error induced by the truncation.

3.1. Upper bound for f_α .

Lemma 3.1. *Let $q \geq 1$, under assumption (A2) there exists a constant C_q such that*

$$\mathbb{E} \left[\left| \frac{1}{\hat{f}_\varepsilon^*(x)} - \frac{1}{f_\varepsilon^*(x)} \right|^{2q} \right] \leq C_q \left(\frac{1}{|f_\varepsilon^*(x)|^{2q}} \wedge \frac{k_N^{2q}(x)}{|f_\varepsilon^*(x)|^{10q}} \right). \quad (14)$$

Proposition 3.2. *Under assumptions (A2)-(A5), for $k_N(x)$ defined by (5) assume that $s_N(x) = 1$ and for $\hat{f}_{\alpha,m}$ defined by (7) then there is a positive constant C such that*

$$\mathbb{E} \left\| f_\alpha - \hat{f}_{\alpha,m} \right\|^2 \leq \|f_\alpha - f_{\alpha,m}\|^2 + \frac{C}{N} \left(\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{1}{|f_\varepsilon^*(u)|^2} du + \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{|f_\alpha^*(u)|^2}{|f_\varepsilon^*(u)|^8} du \right), \quad (15)$$

where $f_{\alpha,m}$ is defined by (3) and C is a numerical constant.

Remark. The first two terms of the right-hand side of Equation (15) correspond to the usual terms when the error distribution is known (see Comte et al. (2006)): a squared bias term ($\|f_\alpha - f_{\alpha,m}\|^2$) and a bound on the variance depending only on f_ε^* . The last term is due to the estimation of f_ε^* and in addition depends on f_α^* .

For the rates of convergence, we refer to Lacour (2006) for a complete study of the known-error case, Comte and Lacour (2011) for unknown noise. We refer also to Delaigle et al. (2008) who show that in statistical deconvolution problems there is no first-order loss of performance when estimating the error density with repeated data compared to the known-error case under the assumption that " f is smoother than half a derivative of f_ε^* ". Finally, in that model, we also refer to Comte et al. (2013) who complete the theoretical study of Delaigle et al. (2008) by studying this time the integrated risk.

3.2. Upper bound for f_β .

Lemma 3.3. *Let $p \geq 1$, under assumption (A2) there exists a constant C_p such that*

$$\mathbb{E} \left[\left| \frac{1}{|\hat{f}_\varepsilon^*(x)|^2} - \frac{1}{|f_\varepsilon^*(x)|^2} \right|^{2p} \right] \leq C_p \left(\frac{1}{|f_\varepsilon^*(x)|^{4p}} \wedge \frac{k_N^{2p}(x)}{|f_\varepsilon^*(x)|^{12p}} \right). \quad (16)$$

Proposition 3.4. *Under assumptions (A1)-(A4), for $k_N(x)$ defined by (5) assume that $s_N(x) = 1$, then $\hat{f}_{\beta,m}$ defined by (13) satisfies*

$$\mathbb{E} \left\| f_\beta - \hat{f}_{\beta,m} \right\|^2 \leq \|f_\beta - f_{\beta,m}\|^2 + \frac{6m}{N} + \frac{12}{N(J-4)} \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{du}{|f_\varepsilon^*(\frac{u}{\Delta})|^4} + \frac{4C_1}{N} \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{|f_\beta^*(u)|^2}{|f_\varepsilon^*(\frac{u}{\Delta})|^8} du. \quad (17)$$

where $f_{\beta,m}$ is defined by (12) and C_1 is defined in Lemma 3.3.

Remark. The first three terms of the right-hand side of Equation (17) correspond to the terms when the error distribution f_ε^* is known as found in Comte and Samson (2012): a squared bias term ($\|f_\beta - f_{\beta,m}\|^2$) and two terms of variance. The first term of variance with order m/N is the bound we would have if we were in a direct density estimation context. The second term is a classical term appearing in density deconvolution problems when the error distribution is known but the $J-4$ factor is specific to the repeated measurement framework. The last term of variance is due to the estimation of f_ε^* .

3.3. Discussion about resulting rates. In order to derive the corresponding rates of convergence of the estimator of f_β defined by (13) we assume that the density functions f_β and f_ε belong to some nonparametric classes of functions. First, we introduce the following type of smoothness spaces

$$\mathcal{A}(a, r, L) = \left\{ f \in \mathbb{L}^1 \cap \mathbb{L}^2(\mathbb{R}), \int |f^*(u)|^2 e^{2a|u|^r} du \leq L \right\} \quad (18)$$

$$\mathcal{S}(\delta, L) = \left\{ f \in \mathbb{L}^1 \cap \mathbb{L}^2(\mathbb{R}), \int |f^*(u)|^2 (1+u^2)^\delta du \leq L \right\} \quad (19)$$

with $r \geq 0$, $a > 0$, $\delta > 1/2$ and $L > 0$. Then if f_β belongs to $\mathcal{A}(a, r, L)$, the bias term can be bounded as follows

$$\|f_\beta - f_{\beta,m}\|^2 \leq \frac{L}{2\pi} e^{-2a|\pi m|^r},$$

or if f_β belongs to $\mathcal{S}(\delta, L)$

$$\|f_\beta - f_{\beta,m}\|^2 \leq \frac{L}{2\pi} ((\pi m)^2 + 1)^{-\delta}.$$

In order to derive the order of the bound on the variance in Equation (17), we need more information about the regularity of f_ε . We add the following classical assumption:

(A6) There exist positive constants k_0, k'_0, γ, μ , and s such for any real x

$$k_0(x^2 + 1)^{-\gamma/2} e^{-\mu|x|^s} \leq |f_\varepsilon^*(x)| \leq k'_0(x^2 + 1)^{-\gamma/2} e^{-\mu|x|^s}.$$

If $s = 0$, we say that $f_\varepsilon^* \in \text{OS}(\gamma)$ (for *ordinary smooth*) and if $\gamma = 0$ $f_\varepsilon^* \in \text{SS}(s)$ (for *supersmooth*).

We introduce the notation $g(x) \lesssim h(x)$ if there exists a positive constant C such that for all x , $g(x) \leq Ch(x)$ and the notation $g(x) \asymp h(x)$ if $g(x) \lesssim h(x)$ and $h(x) \lesssim g(x)$.

Lemma 3.5. *If f_ε^* satisfies assumption (A6) then*

$$\begin{aligned} \int_{-\pi m}^{\pi m} \frac{du}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4} &\asymp (\pi m)^{4\gamma+1-s} e^{4\mu(\pi m)^s}, \\ \int_{-\pi m}^{\pi m} \frac{|f_\beta^*(u)|^2}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^8} du &\lesssim (\pi m)^{(4\gamma+1-s)\wedge 2(2\gamma-\delta)+} e^{4\mu(\pi m)^s} \mathbf{1}_{\{s>r\}} \\ &\quad + (\pi m)^{2(2\gamma-\delta)+} e^{2(2\mu-a)(\pi m)^s} \mathbf{1}_{\{r=s, 2\mu \geq a\}} + \mathbf{1}_{\{r>s\} \cup \{2\mu \leq a\}}. \end{aligned}$$

A proof can be found in Comte and Lacour (2011).

Proposition 3.6. *Suppose that (A6) is fulfilled, $f_\varepsilon^* \in \text{OS}(a)$ and $f_\beta \in \mathcal{S}(\delta, L)$ defined by (18), then*

$$\mathbb{E} \left\| f_\beta - \hat{f}_{\beta,m} \right\|^2 \leq Cm^{-2\delta} + \frac{6m}{N} + C' \frac{m^{4\gamma+1-s} e^{4\mu(\pi m)^s}}{N(J-4)} + C'' \frac{m^{(4\gamma+1-s)\wedge 2(2\gamma-\delta)+} e^{4\mu(\pi m)^s}}{N} \quad (20)$$

where C, C' and C'' are positive constants independent of N or J .

Proposition 3.7. *Suppose that (A6) is fulfilled, $f_\varepsilon^* \in \text{SS}(s)$ and $f_\beta \in \mathcal{A}(a, r, L)$ defined by (19), then*

$$\mathbb{E} \left\| f_\beta - \hat{f}_{\beta,m} \right\|^2 \leq Cm^{-2\delta} e^{-2a(\pi m)^r} + \frac{6m}{N} + C' \frac{m^{4\gamma+1}}{N(J-4)} + \frac{C''}{N} \quad (21)$$

where C, C' and C'' are positive constants independent of N or J .

The rates are reported in Table 1. We clearly see that the rates of convergence depend on unknown quantities since they describe the regularity of the density function under estimation as well as the error distribution. When J is considered as a constant, we find the usual rates of convergence corresponding to density deconvolution already presented in the literature.

	$f \in \mathcal{A}(a, r, L), f_\varepsilon^* \in \text{OS}(\gamma)$	$f \in \mathcal{S}(\delta, L), f_\varepsilon^* \in \text{SS}(s)$	$f \in \mathcal{A}(a, r, L), f_\varepsilon^* \in \text{OS}(\gamma)$
$r = 0$	$(NJ)^{\frac{-2\delta}{2\delta+4\gamma+1}} + N^{-\left(\frac{2\delta}{2\delta+1} \wedge \frac{\delta}{2\delta}\right)}$	$(\log NJ)^{-2\delta/s}$	$\frac{(\log NJ)^{\frac{4\gamma+1}{r}}}{NJ} + \frac{(\log NJ)^{\frac{1}{r}}}{N}$

TABLE 1. Rates of convergence for the MISE

The idea now for the adaptive estimation is to find a penalty term which would have the same order as the bound on the variance. Thus the adaptive estimator will reach automatically the rates of convergence presented in Table 1. The following section will show that we can obtain an adaptive procedure under weak assumptions.

4. MODEL SELECTION

In this section, we introduce an adaptive estimator of $|f_\varepsilon^*|^2$ which can be uniformly controlled on the real line which brings a model selection procedure with very weak assumptions on the error distribution f_ε^* . For that we need to precise a $s_N(x)$ which will allow us to apply concentration inequalities of Talagrand type. For $\delta > 0$, let us introduce the weight function w defined as

$$\forall x \in \mathbb{R}, w(x) = (\log(e + |x|))^{-\frac{1}{2} - \delta}$$

which has originally been proposed in [Neumann and Reiß \(2009\)](#). In this context, as from now, we set the threshold k_N as follows

$$k_N(x) = \kappa (\log N)^{1/2} w(x)^{-1} N^{-1/2} \quad (22)$$

which means that we take $s_N(x) = \kappa (\log N)^{1/2} w(x)^{-1}$.

We want to propose an estimator $\hat{f}_{\hat{m}}$ of f completely data driven. Following the model selection paradigm, see [Birgé \(1999\)](#), [Birgé and Massart \(1997\)](#) or [Massart \(2003\)](#), we select \hat{m} as the minimizer of a penalized criterion

$$\hat{m} = \operatorname{argmin}_{m \in \mathcal{M}_n} \left\{ -\|\hat{f}_m\|^2 + \widehat{\text{pen}}(m) \right\}$$

where \mathcal{M}_n describes the model collections.

The penalty term should be chosen large enough to counterbalance the fluctuation of \hat{f}_m around \hat{f} , but on the other hand, should ideally not be much larger than the variance terms presented in Equation (17). Here the penalty term is stochastic since the variance terms depend on the error distribution which is supposed unknown.

4.1. Adaptive estimation procedure for f_α . In this section, we adapt the results of [Kappus and Mabon \(2013\)](#) who proposed a completely data driven procedure in the framework of density estimation in deconvolution problems with unknown error distribution. In the present paper, Model (2) can be seen as a repeated observation model, which is studied in their paper. The only difference lies in the fact that in [Kappus and Mabon \(2013\)](#) f_ε^{*2} can be estimated directly from the data with an empirical estimator whereas we can only estimate f_ε^{*4} which is due to our access of the ε -sample. This, in our case, implies that f_ε^* is raised to a greater power in the term specific to the unknown noise density.

We introduce the following notations

$$\begin{aligned} \Theta(m) &= \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w(u)^{-2}}{|f_\varepsilon^*(u)|^2} du \quad \text{and} \quad \Theta^\alpha(m) = \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w(u)^{-2} |f_\alpha^*(u)|^2}{|f_\varepsilon^*(u)|^8} du \\ \hat{\Theta}(m) &= \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w(u)^{-2}}{|\hat{f}_\varepsilon^*(u)|^2} du \quad \text{and} \quad \hat{\Theta}^\alpha(m) = \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w(u)^{-2} |\hat{f}_Y^*(u)|^2}{|\hat{f}_\varepsilon^*(u)|^{10}} du. \end{aligned}$$

These terms correspond to the deterministic and stochastic bounds on the variance appearing in Equation (15). The difference lies in the introduction of the function w essential for the adaptive procedure. Thus, we define empirical penalty as

$$\widehat{\text{pen}}(m) = 12\lambda^2(m, \hat{\Theta}(m)) \frac{\hat{\Theta}(m)}{n} + 16\kappa^2 \log(Nm) \frac{\hat{\Theta}^\alpha(m)}{N}$$

where κ is defined in Equation (22), and a deterministic penalty

$$\text{pen}(m) = 12\lambda^2(m, \Theta(m)) \frac{\Theta(m)}{N} + 16\kappa^2 \log(Nm) \frac{\Theta^\alpha(m)}{N},$$

with

$$\lambda(m, D) = \max \left\{ \sqrt{8 \log(1 + Dm^2)}, \frac{16\sqrt{2}}{3\sqrt{N}} \log(1 + Dm^2) \right\}.$$

Then, we select the cutoff parameter \hat{m} as a minimizer of the following penalized criterion

$$\hat{m} = \operatorname{argmin}_{m \in \mathcal{M}_n} \left\{ -\|\hat{f}_{\alpha, m}\|^2 + \widehat{\text{pen}}(m) \right\} \quad (23)$$

where $\mathcal{M}_n = \{1, \dots, n\}$. We can now state the following oracle inequality:

Theorem 4.1. *Under assumptions (A1)-(A4), let $\hat{f}_{\hat{m},\alpha}$ be defined by (7) and (23). Then there are positive constants C^{ad} and C such that*

$$\mathbb{E}\|f_\alpha - \hat{f}_{\alpha,\hat{m}}\|^2 \leq C^{ad} \inf_{m \in \mathcal{M}_n} \{\|f_\alpha - f_{\alpha,m}\|^2 + \text{pen}(m)\} + \frac{C}{N}. \quad (24)$$

Remark. The latest result is an oracle inequality which means that the bias variance compromise is automatically made and completely data driven in an almost non-asymptotic setting. So rates of convergence are reached by themselves without being specified in the framework. This result is of high interest since in deconvolution problems, rates of convergence are classically intricate and depend on the regularity types of the function f under estimation and the error density f_ε^* (see Section 3.3).

However the penalty term is not exactly the same as the upper bound terms shown in Equation (15). We may wonder if a loss due to adaptation occurs. To answer that question, right-hand side of Equations (15) and (24) have to be compared. More precisely, since the bias term $\|f_\alpha - f_{\alpha,m}\|^2$ is unchanged and N^{-1} is a negligible term, it comes down to compare $\text{pen}(m)$ with

$$\frac{1}{N} \left(\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{1}{|f_\varepsilon^*(u)|^2} du + \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{|f_\alpha^*(u)|^2}{|f_\varepsilon^*(u)|^8} du \right).$$

Clearly the difference lies in the logarithmic terms, and thus, the loss is negligible.

4.2. Adaptive estimation procedure for f_β . As in the previous section, we start by defining the bound appearing in the known-error case then the one appearing in the known-error case and introduce the function w which is principal for the adaptive procedure.

$$\begin{aligned} \Xi(m) &= \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2}}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4} du & \text{and} & \quad \Xi^\beta(m) = \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2} |f_\beta^*(u)|^2}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^8} du. \\ \hat{\Xi}(m) &= \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2}}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4} du & \text{and} & \quad \hat{\Xi}^\beta(m) = \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2} |\hat{f}_Z^*(u)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{12}} du. \end{aligned}$$

We can now define the stochastic penalty associated to the adaptive procedure

$$\begin{aligned} \widehat{\text{qen}}(m) &= \widehat{\text{qen}}_1(m) + \widehat{\text{qen}}_2(m) + \widehat{\text{qen}}_3(m) \\ &= 64 \frac{m}{N} + 16 \frac{\mu^2(m, \hat{\Xi}(m)) \hat{\Xi}(m)}{N(J-4)} + 16\kappa^2 \log(Nm) \frac{\hat{\Xi}^\beta(m)}{N} \end{aligned}$$

and the deterministic penalty

$$\begin{aligned} \text{qen}(m) &= \text{qen}_1(m) + \text{qen}_2(m) + \text{qen}_3(m) \\ &= 64 \frac{m}{N} + 16\mu^2(m, \Xi(m)) \frac{\Xi(m)}{N(J-4)} + 16\kappa^2 \log(Nm) \frac{\Xi^\beta(m)}{N} \end{aligned}$$

with weight

$$\mu(m, D) = \max \left\{ \sqrt{8 \log(1 + Dm^2)}, \frac{16\sqrt{2}}{3\sqrt{N(J-4)}} \log(1 + Dm^2) \right\}.$$

It is worth mentioning that the penalty takes into account the three terms of variance showed in Equation (17) which is consistent with the upper bound given in Proposition 3.4. So the penalty has the same order as the bounds on the variance. As in Kappus and Mabon (2013), the model selection procedure has the advantage of keeping track of the inverse Fourier transform of the target density f_β by estimating it by $\hat{f}_Z^*/\tilde{f}_\varepsilon^*$ in the term $\hat{\Xi}^\beta(m)$ which is a difficulty in deconvolution problems as showed in Section 6. Besides we can also point out that the model collection $\mathcal{M}_n = \{1, \dots, n\}$ is a deterministic one.

Therefore we select the cutoff parameter \hat{m} as a minimizer of the following penalized criterium

$$\hat{m} = \underset{m \in \mathcal{M}_n}{\text{argmin}} \left\{ -\|\hat{f}_{\beta,m}\|^2 + \widehat{\text{qen}}(m) \right\}. \quad (25)$$

Theorem 4.2. *Under assumptions (A1)-(A4), consider $\hat{f}_{\beta,\hat{m}}$ defined by (13) and (25). Then there are positive constants C^{ad} and C such that*

$$\mathbb{E}\|f_\beta - \hat{f}_{\beta,\hat{m}}\|^2 \leq C^{ad} \inf_{m \in \mathcal{M}_n} \{\|f_\beta - f_{\beta,m}\|^2 + \text{qen}(m)\} + \frac{C}{N(J-4)} + \frac{C}{N}. \quad (26)$$

Remark. The same kind of remarks as in Theorem 4.1 hold here. The latest result is an oracle inequality which means that the bias variance compromise is automatically made and completely data driven in an almost non-asymptotic setting. So rates of convergence are reached by themselves without being specified in the framework. As far as we know this result is new in the literature.

5. SIMULATION

In this section, we only concentrate on a simulation study of f_β . Indeed, the proposed method for the estimation of f_α being mainly taken from Kappus and Mabon (2013), we refer to that paper for the performance of the estimator.

The whole implementation is conducted using R software. The integrated squared error $\|f - \hat{f}_{\beta, \hat{m}}\|^2$ is computed via a standard approximation and discretization (over 300 points) of the integral on an interval of \mathbb{R} denoted by I . Then the mean integrated squared error (MISE) $\mathbb{E}\|f - \hat{f}_{\beta, \hat{m}}\|^2$ is computed as the empirical mean of the approximated ISE over 100 simulation samples.

Practical estimation procedure. The adaptive procedure is implemented as follows:

- ▷ For $m \in \mathcal{M}_n = \{m_1, \dots, m_n\}$, compute $-\|\hat{f}_{\beta, m}\|^2 + \widehat{\text{qen}}(m)$.
- ▷ Choose \hat{m} such as $\hat{m} = \underset{m \in \mathcal{M}_n}{\text{argmin}} \left\{ -\|\hat{f}_{\beta, m}\|^2 + \widehat{\text{qen}}(m) \right\}$.
- ▷ And compute $\hat{f}_{\beta, \hat{m}}(x) = \int_{-\pi\hat{m}}^{\pi\hat{m}} e^{-ixu} \frac{\hat{f}_Z^*(u)}{|\hat{f}_\varepsilon^*(\frac{u}{\Delta})|^2} du$.

Riemann's sums are used to approximate all the integrals. The penalties are chosen according to Theorem 4.2 and as in Comte et al. (2007) we consider that m can be fractional by taking the following model collection

$$\mathcal{M}_n = \left\{ m = \frac{k}{10}, \quad 1 \leq k \leq 25 \right\}$$

associated with the following penalty

$$\widehat{\text{qen}}(m) = \kappa_1 \left(\frac{m}{N} + \frac{\log(1 + \hat{\Xi}(m)m^2)\hat{\Xi}(m)}{N(J-4)} \right) + \kappa_2 \log(Nm) \frac{\hat{\Xi}^\beta(m)}{N}$$

Moreover the interval t_j are chosen as $t_j = j\Delta$ with $\Delta = 2$ and $J = 6$ as in Comte and Samson (2012). We consider the four following distributions for β :

- ▷ Standard Gaussian distribution, $I = [-4, 4]$.
- ▷ Cauchy distribution, $f(x) = (\pi(1+x^2))^{-1}$, $I = [-10, 10]$.
- ▷ Gamma distribution : $5 \cdot \Gamma(25, \frac{1}{25})$, $I = [-1, 13]$.
- ▷ Mixed Gamma distribution : $X = W/\sqrt{5.48}$, with $W \sim 0.4\Gamma(5, 1) + 0.6\Gamma(13, 1)$, $I = [-1.5, 26]$.

All the densities are normalized with unit variance except the Cauchy density. Unlike in Comte and Samson (2012), we do not study the influence of the distribution of f_α on the estimation of f_β . In all considered cases, α is a standard gaussian distribution.

The results of the first and third distribution can be compared to those of Comte and Samson (2012) and the first, second and third can be compared to Kappus and Mabon (2013).

We consider the two following noise densities with same variance σ_ε^2 . In the simulation the variance takes the values 1/10 and 1/4. The first one is a gaussian density (*supersmooth* density) which means $f_\varepsilon^* \in \text{SS}(2)$. The second one is a Laplace density (*ordinary smooth* density) which means $f_\varepsilon^* \in \text{OS}(2)$.

Gaussian noise : $f_\varepsilon(x) = \frac{1}{\sigma_\varepsilon\sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma_\varepsilon^2}\right)$, $f_\varepsilon^*(x) = \exp\left(-\frac{\sigma_\varepsilon^2 x^2}{2}\right)$.

Laplace noise : $f_\varepsilon(x) = \frac{1}{2\sigma_\varepsilon} \exp\left(-\frac{|x|}{\sigma_\varepsilon}\right)$, $f_\varepsilon^*(x) = \frac{1}{1 + \sigma_\varepsilon^2 x^2}$.

The calibration of the two constants are done with intensive preliminary simulations with a sample size of 500. In the end, we choose $\kappa_1 = \kappa_2 = 1$. We can notice that in Kappus and Mabon (2013) the constant are bigger. It seems that the greater the power of f_ε^* is in the denominator, the smaller the constants are.

Results. The results of the simulations are given in Tables 2 and 3. For both tables, the MISE is multiplied by 100 and computed from 100 simulated data sets. We also give the medians of the MISE. A first remark: estimating the Fourier transform of the noise f_ε^* reduces the risk compared to knowing the density of the noise, a fact already pointed in Comte and Lacour (2011). This can be explained by the fact that an additional regularization of the characteristic function of the noise comes in. This regularization is not applied in the procedure when the error distribution is known.

Table 2 corresponds to an estimation procedure where the error distribution is a Laplace density while Table 3 corresponds to a Gaussian noise. We notice that increasing the sample size improves the estimation and increasing the variances degrades the estimation but in an acceptable way. Concerning the medians of the MISE, they are always lower than the means of the MISE.

For all the test densities, the results are very good. If we compare to Kappus and Mabon (2013) the results are very close. It is quite remarkable that they have the same order since in this paper we have access to an estimate of $(f_\varepsilon^*)^4$ when they have access to an estimate of f_ε^* or $(f_\varepsilon^*)^2$. Still it proves that their methodology is robust since it does not fail to estimate in that particular case. If we compare the results to those of Comte and Samson (2012), we see that our procedure does not outdo theirs. But the results are very close. Nonetheless our adaptive procedure has the advantage of being based on theoretical properties.

		$\sigma_\varepsilon^2 = \frac{1}{10}$		$\sigma_\varepsilon^2 = \frac{1}{4}$	
		200	2000	200	2000
Gaussian	f_ε known	0.344 (0.273)	0.054 (0.045)	0.514 (0.513)	0.129 (0.105)
	f_ε unknown	0.331 (0.251)	0.042 (0.033)	0.317 (0.236)	0.057 (0.051)
Cauchy	f_ε known	0.625 (0.573)	0.105 (0.097)	0.804 (0.765)	0.216 (0.211)
	f_ε unknown	0.507 (0.427)	0.075 (0.071)	0.657 (0.599)	0.090 (0.079)
Gamma	f_ε known	0.398 (0.360)	0.069 (0.063)	0.620 (0.517)	0.161 (0.140)
	f_ε unknown	0.381 (0.347)	0.051 (0.044)	0.506 (0.430)	0.066 (0.051)
Mixed Gamma	f_ε known	0.545 (0.480)	0.095 (0.084)	0.715 (0.620)	0.150 (0.141)
	f_ε unknown	0.506 (0.453)	0.082 (0.080)	0.518 (0.495)	0.092 (0.084)

TABLE 2. Results of simulation as MISE $\mathbb{E} \left(\|f - \hat{f}_{\beta, \hat{m}}\|^2 \right) \times 100$ averaged over 100 samples. In brackets we give the median of the MISE also averaged over 100 samples with a Laplace noise.

		$\sigma_\varepsilon^2 = \frac{1}{10}$		$\sigma_\varepsilon^2 = \frac{1}{4}$	
		N		N	
Gaussian	f_ε known	200	0.349 (0.296)	2000	0.054 (0.045)
	f_ε unknown	200	0.650 (0.598)	2000	0.142 (0.128)
Cauchy	f_ε known	200	0.285 (0.239)	2000	0.038 (0.029)
	f_ε unknown	200	0.349 (0.273)	2000	0.052 (0.053)
Gamma	f_ε known	200	0.588 (0.532)	2000	0.119 (0.112)
	f_ε unknown	200	0.848 (0.791)	2000	0.272 (0.263)
Mixed Gamma	f_ε known	200	0.481 (0.449)	2000	0.076 (0.070)
	f_ε unknown	200	0.680 (0.607)	2000	0.089 (0.084)
Gamma	f_ε known	200	0.401 (0.332)	2000	0.072 (0.063)
	f_ε unknown	200	0.956 (0.910)	2000	0.207 (0.190)
Mixed Gamma	f_ε known	200	0.402 (0.316)	2000	0.049 (0.041)
	f_ε unknown	200	0.461 (0.418)	2000	0.067 (0.055)
Mixed Gamma	f_ε known	200	0.504 (0.454)	2000	0.089 (0.084)
	f_ε unknown	200	0.704 (0.639)	2000	0.163 (0.146)
Mixed Gamma	f_ε known	200	0.504 (0.446)	2000	0.079 (0.072)
	f_ε unknown	200	0.552 (0.484)	2000	0.101 (0.093)

TABLE 3. Results of simulation as MISE $\mathbb{E} \left(\|f - \hat{f}_{\beta, \hat{m}}\|^2 \right) \times 100$ averaged over 100 samples. In brackets we give the median of the MISE also averaged over 100 samples with a Gaussian noise.

6. PROOFS

6.1. **Proof of Lemma 3.3.** We start by proving Lemma 3.3 since Lemma 3.1 is obtained as a consequence of it. Let $p \geq 1$ be. Using that $1/|\tilde{f}_\varepsilon^*(x)|^2 + |f_\varepsilon^*(x)|^2 \leq 1/|\tilde{f}_\varepsilon^*(x)|^4$, we have

$$\begin{aligned}
\mathbb{E} \left| \frac{1}{|\tilde{f}_\varepsilon^*(x)|^2} - \frac{1}{|f_\varepsilon^*(x)|^2} \right|^{2p} &= \mathbb{E} \left[\mathbf{1} \left\{ |\tilde{f}_\varepsilon^*(x)|^4 < k_N(x) \right\} \left| \frac{1}{\sqrt{k_N(x)}} - \frac{1}{|f_\varepsilon^*(x)|^2} \right|^{2p} \right] \\
&\quad + \mathbb{E} \left[\mathbf{1} \left\{ |\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x) \right\} \left| \frac{1}{|\tilde{f}_\varepsilon^*(x)|^2} - \frac{1}{|f_\varepsilon^*(x)|^2} \right|^{2p} \right] \\
&\leq |f_\varepsilon^*(x)|^{-4p} \mathbb{P} \left[|\tilde{f}_\varepsilon^*(x)|^4 < k_N(x) \right] \frac{\left| |f_\varepsilon^*(x)|^2 - \sqrt{k_N(x)} \right|^{2p}}{k_N(x)^p} \\
&\quad + \mathbb{E} \left[\mathbf{1} \left\{ |\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x) \right\} \frac{\left| |\tilde{f}_\varepsilon^*(x)|^4 - |f_\varepsilon^*(x)|^4 \right|^{2p}}{|f_\varepsilon^*(x)|^{4p} |\tilde{f}_\varepsilon^*(x)|^{4p} \left| |\tilde{f}_\varepsilon^*(x)|^2 + |f_\varepsilon^*(x)|^2 \right|^{2p}} \right] \\
&\leq |f_\varepsilon^*(x)|^{-4p} \frac{\left| |f_\varepsilon^*(x)|^2 - \sqrt{k_N(x)} \right|^{2p}}{k_N(x)^p} \\
&\quad + \mathbb{E} \left[\mathbf{1} \left\{ |\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x) \right\} \frac{\left| |\tilde{f}_\varepsilon^*(x)|^4 - |f_\varepsilon^*(x)|^4 \right|^{2p}}{|f_\varepsilon^*(x)|^{4p} |\tilde{f}_\varepsilon^*(x)|^{8p}} \right] \tag{27}
\end{aligned}$$

• 1st case : $|f_\varepsilon^*(x)|^4 < 2k_N(x)$. In this case we have $\frac{1}{|f_\varepsilon^*(x)|^{4p}} \wedge \frac{k_N^{2p}(x)}{|f_\varepsilon^*(x)|^{12p}} = |f_\varepsilon^*(x)|^{-4p}$. Then starting from (27), we get

$$\begin{aligned} \mathbb{E} \left| \frac{1}{|\tilde{f}_\varepsilon^*(x)|^2} - \frac{1}{|f_\varepsilon^*(x)|^2} \right|^{2p} &\leq 9^p |f_\varepsilon^*(x)|^{-4p} + |f_\varepsilon^*(x)|^{-4p} \mathbb{E} \left[\mathbb{1} \left\{ |\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x) \right\} \frac{||\tilde{f}_\varepsilon^*(x)|^4 - |f_\varepsilon^*(x)|^4|^{2p}}{|\tilde{f}_\varepsilon^*(x)|^{8p}} \right] \\ &\leq 9^p |f_\varepsilon^*(x)|^{-4p} + |f_\varepsilon^*(x)|^{-4p} k_N(x)^{-2p} \mathbb{E} \left[\left| \widehat{|f_\varepsilon^{*4}|}(x) - |f_\varepsilon^*(x)|^4 \right|^{2p} \right] \\ &\leq 9^p |f_\varepsilon^*(x)|^{-4p} + |f_\varepsilon^*(x)|^{-4p} k_N(x)^{-2p} N^{-p} \\ &\leq 9^p |f_\varepsilon^*(x)|^{-4p} + |f_\varepsilon^*(x)|^{-4p} (\log N)^{-p} w(x)^{2p} N^p N^{-p} \\ &\leq O(|f_\varepsilon^*(x)|^{-4p}). \end{aligned}$$

• 2nd case : $|f_\varepsilon^*(x)|^4 \geq 2k_N(x)$. In this case we have $\frac{1}{|f_\varepsilon^*(x)|^{4p}} \wedge \frac{k_N^{2p}(x)}{|f_\varepsilon^*(x)|^{12p}} = \frac{k_N^{2p}(x)}{|f_\varepsilon^*(x)|^{12p}}$. Now using the Markov and Rosenthal inequalities.

$$\begin{aligned} \mathbb{P} \left[|\tilde{f}_\varepsilon^*(x)|^4 \leq k_N(x) \right] &= \mathbb{P} \left[\widehat{|f_\varepsilon^{*4}|}(x) \leq k_N(x) \right] \\ &\leq \mathbb{P} \left[\left| \widehat{|f_\varepsilon^{*4}|}(x) - |f_\varepsilon^*(x)|^4 \right| > |f_\varepsilon^*(x)|^4 - k_N(x) \right] \\ &\leq \mathbb{P} \left[\left| \widehat{|f_\varepsilon^{*4}|}(x) - |f_\varepsilon^*(x)|^4 \right| > |f_\varepsilon^*(x)|^4 / 2 \right] \\ &\leq \frac{\mathbb{E} \left| \widehat{|f_\varepsilon^{*4}|}(x) - |f_\varepsilon^*(x)|^4 \right|^{2p}}{(|f_\varepsilon^*(x)|^4 / 2)^{2p}} \\ &\leq \frac{c_p N^{-p}}{|f_\varepsilon^*(x)|^{8p}} \end{aligned}$$

Then we can bound the first term of Equation (27) as follows

$$\begin{aligned} |f_\varepsilon^*(x)|^{-4p} \mathbb{P} \left[|\tilde{f}_\varepsilon^*(x)|^4 < k_N(x) \right] \frac{||f_\varepsilon^*(x)|^2 - \sqrt{k_N(x)}|^{2p}}{k_N(x)^p} \\ \leq |f_\varepsilon^*(x)|^{-4p} \mathbb{P} \left[|\tilde{f}_\varepsilon^*(x)|^4 < k_N(x) \right] \frac{(1 + \sqrt{k_N(x)})^{2p}}{k_N(x)^p} \\ \leq |f_\varepsilon^*(x)|^{-4p} \mathbb{P} \left[|\tilde{f}_\varepsilon^*(x)|^4 < k_N(x) \right] \frac{C k_N(x)^p}{k_N(x)^p} \\ \leq O(N^{-p} |f_\varepsilon^*(x)|^{-12p}) \\ \leq O(k_N(x)^{2p} |f_\varepsilon^*(x)|^{-12p}). \end{aligned} \tag{28}$$

Moreover using this time that $1/|\tilde{f}_\varepsilon^*(x)|^2 + |f_\varepsilon^*(x)|^2 \leq 1/|f_\varepsilon^*(x)|^4$, we can bound the second term of Equation (27) as follows

$$\begin{aligned} &\mathbb{E} \left[\mathbb{1} \left\{ |\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x) \right\} \left| \frac{1}{|\tilde{f}_\varepsilon^*(x)|^2} - \frac{1}{|f_\varepsilon^*(x)|^2} \right|^{2p} \right] \\ &= \mathbb{E} \left[\mathbb{1} \left\{ |\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x) \right\} \left| \frac{1}{|f_\varepsilon^*(x)|^2} \right|^{2p} \left| \frac{1}{|\tilde{f}_\varepsilon^*(x)|^2} \right|^{2p} \left| \frac{\widehat{|f_\varepsilon^{*4}|}(x) - |f_\varepsilon^*(x)|^4}{|\tilde{f}_\varepsilon^*(x)|^2 + |f_\varepsilon^*(x)|^2} \right|^{2p} \right] \\ &= |f_\varepsilon^*(x)|^{-4p} \mathbb{E} \left[\mathbb{1} \left\{ |\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x) \right\} \left| \frac{1}{|\tilde{f}_\varepsilon^*(x)|^2} - \frac{1}{|f_\varepsilon^*(x)|^2} + \frac{1}{|f_\varepsilon^*(x)|^2} \right|^{2p} \left| \frac{\widehat{|f_\varepsilon^{*4}|}(x) |f_\varepsilon^*(x)|^4}{|\tilde{f}_\varepsilon^*(x)|^2 + |f_\varepsilon^*(x)|^2} \right|^{2p} \right] \end{aligned}$$

We then deduce the following bounds

$$\begin{aligned}
& \mathbb{E} \left[\mathbb{1} \left\{ |\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x) \right\} \left| \frac{1}{|\tilde{f}_\varepsilon^*(x)|^2} - \frac{1}{|f_\varepsilon^*(x)|^2} \right|^{2p} \right] \\
& \leq 2^{2p-1} |f_\varepsilon^*(x)|^{-8p} \mathbb{E} \left[\mathbb{1} \left\{ |\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x) \right\} \left| \frac{|\widehat{f_\varepsilon^{*4}}(x) - |f_\varepsilon^*(x)|^4|}{|\tilde{f}_\varepsilon^*(x)|^2 + |f_\varepsilon^*(x)|^2} \right|^{2p} \right] \\
& \quad + 2 |f_\varepsilon^*(x)|^{-8p} \mathbb{E} \left[\mathbb{1} \left\{ |\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x) \right\} \frac{|\widehat{f_\varepsilon^{*4}}(x) - |f_\varepsilon^*(x)|^4|^{4p}}{|\tilde{f}_\varepsilon^*(x)|^{4p}} \right] \\
& \leq 2^{2p-1} |f_\varepsilon^*(x)|^{-12p} \mathbb{E} \left[\left| \widehat{f_\varepsilon^{*4}}(x) - |f_\varepsilon^*(x)|^4 \right|^{2p} \right] \\
& \quad + 2 |f_\varepsilon^*(x)|^{-12p} \mathbb{E} \left[\mathbb{1} \left\{ |\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x) \right\} \frac{|\widehat{f_\varepsilon^{*4}}(x)|^4 - |f_\varepsilon^*(x)|^4}{|\tilde{f}_\varepsilon^*(x)|^{4p}} \right] \\
& \leq 2^{2p-1} |f_\varepsilon^*(x)|^{-12p} N^{-p} + 2 |f_\varepsilon^*(x)|^{-12p} k_N(x)^{-p} N^{-2p} \\
& \leq 2^{2p-1} |f_\varepsilon^*(x)|^{-12p} N^{-p} + 2 |f_\varepsilon^*(x)|^{-12p} N^p N^{-2p} \\
& \leq O \left(|f_\varepsilon^*(x)|^{-12p} N^{-p} \right) \\
& \leq O(k_N(x)^{2p} |f_\varepsilon^*(x)|^{-12p}). \tag{29}
\end{aligned}$$

Then gathering Equations (28) and (29), we just proved that if $|f_\varepsilon^*(x)|^4 \geq 2k_N(x)$ then

$$\mathbb{E} \left| \frac{1}{|\tilde{f}_\varepsilon^*(x)|^2} - \frac{1}{|f_\varepsilon^*(x)|^2} \right|^{2p} \leq O(k_N(x)^{2p} |f_\varepsilon^*(x)|^{-12p}).$$

In the end

$$\mathbb{E} \left| \frac{1}{|\tilde{f}_\varepsilon^*(x)|^2} - \frac{1}{|f_\varepsilon^*(x)|^2} \right|^{2p} \leq C_p \left(\frac{1}{|f_\varepsilon^*(x)|^{4p}} \wedge \frac{k_N^{2p}(x)}{|f_\varepsilon^*(x)|^{12p}} \right).$$

□

6.2. Proof of Lemma 3.1. Under (A5) and applying Lemma 3.3, we have

$$\begin{aligned}
\mathbb{E} \left[\left| \frac{1}{|\tilde{f}_\varepsilon^*(x)} - \frac{1}{f_\varepsilon^*(x)} \right|^{2p} \right] &= \mathbb{E} \left[\left| \frac{1}{|\tilde{f}_\varepsilon^*(x)|^2} - \frac{1}{|f_\varepsilon^*(x)|^2} \right|^{2p} \middle/ \left| \frac{1}{\tilde{f}_\varepsilon^*(x)} + \frac{1}{f_\varepsilon^*(x)} \right|^{2p} \right] \\
&= \mathbb{E} \left[\left| \frac{1}{(\tilde{f}_\varepsilon^*(x))^2} - \frac{1}{(f_\varepsilon^*(x))^2} \right|^{2p} \middle/ \left| \frac{1}{\tilde{f}_\varepsilon^*(x)} + \frac{1}{f_\varepsilon^*(x)} \right|^{2p} \right] \\
&= |f_\varepsilon^*(x)|^{2p} \mathbb{E} \left[\left| \frac{1}{(\tilde{f}_\varepsilon^*(x))^2} - \frac{1}{(f_\varepsilon^*(x))^2} \right|^{2p} \right] \\
&= |f_\varepsilon^*(x)|^{2p} C_p \left(\frac{1}{|f_\varepsilon^*(x)|^{4p}} \wedge \frac{k_N^{2p}(x)}{|f_\varepsilon^*(x)|^{12p}} \right) \\
&= C_p \left(\frac{1}{|f_\varepsilon^*(x)|^{2p}} \wedge \frac{k_N^{2p}(x)}{|f_\varepsilon^*(x)|^{10p}} \right)
\end{aligned}$$

□

We just proved the desired result.

6.3. Proof of Proposition 3.4. As aforementioned $\hat{f}_{\beta,m}$ can be seen as a projection estimator. We can then write the following equality using Pythagoras' theorem

$$\|f_\beta - \hat{f}_{\beta,m}\|^2 = \|f_\beta - f_{\beta,m}\|^2 + \|f_{\beta,m} - \hat{f}_{\beta,m}\|^2.$$

Now using Plancherel's formula, we can write

$$\begin{aligned}
\|f_{\beta,m} - \hat{f}_{\beta,m}\|^2 &= \frac{1}{2\pi} \int |f_{\beta,m}^*(u) - \hat{f}_{\beta,m}^*(u)|^2 du \\
&= \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \left| \frac{\hat{f}_Z^*(u)}{\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)^2} - \frac{f_Z^*(u)}{f_\varepsilon^*\left(\frac{u}{\Delta}\right)^2} \right|^2 du \\
&= \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \left| \frac{\hat{f}_Z^*(u)}{\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)^2} - \frac{\hat{f}_Z^*(u)}{f_\varepsilon^*\left(\frac{u}{\Delta}\right)^2} + \frac{\hat{f}_Z^*(u)}{f_\varepsilon^*\left(\frac{u}{\Delta}\right)^2} - \frac{f_Z^*(u)}{f_\varepsilon^*\left(\frac{u}{\Delta}\right)^2} \right|^2 du \\
&= \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \left| \hat{f}_Z^*(u) R\left(\frac{u}{\Delta}\right) + \frac{\hat{f}_Z^*(u) - f_Z^*(u)}{f_\varepsilon^*\left(\frac{u}{\Delta}\right)^2} \right|^2 du \\
&\leq \frac{1}{\pi} \int_{-\pi m}^{\pi m} \left| \hat{f}_Z^*(u) R\left(\frac{u}{\Delta}\right) \right|^2 du + \frac{1}{\pi} \int_{-\pi m}^{\pi m} \left| \frac{\hat{f}_Z^*(u) - f_Z^*(u)}{f_\varepsilon^*\left(\frac{u}{\Delta}\right)^2} \right|^2 du \tag{30}
\end{aligned}$$

with $R\left(\frac{u}{\Delta}\right) = \frac{1}{\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)^2} - \frac{1}{f_\varepsilon^*\left(\frac{u}{\Delta}\right)^2}$. Taking the expectation, we get:

$$\mathbb{E} \|f_{\beta,m} - \hat{f}_{\beta,m}\|^2 \leq \frac{1}{\pi} \int_{-\pi m}^{\pi m} \mathbb{E} \left[\left| \hat{f}_Z^*(u) R\left(\frac{u}{\Delta}\right) \right|^2 \right] du + \frac{1}{\pi} \int_{-\pi m}^{\pi m} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{-4} \mathbb{E} \left[\left| \hat{f}_Z^*(u) - f_Z^*(u) \right|^2 \right] du.$$

Yet

$$\begin{aligned}
\mathbb{E} \left[\left| \hat{f}_Z^*(u) - f_Z^*(u) \right|^2 \right] &= \frac{4}{N(J-4)^2} \sum_{j=3}^{J/2} \text{Var} (e^{iuZ_{1,j}}) + \frac{4}{N(J-4)^2} \sum_{\substack{3 \leq j, j' \leq J/2 \\ j \neq j'}} \text{Cov} (e^{iuZ_{1,j}}, e^{iuZ_{1,j'}}) \\
&\leq \frac{4}{N(J-4)^2} \left(\frac{J-4}{2} + \frac{(J-4)^2}{4} (1 - |f_\beta^*(u)|^2) |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4 \right) \\
&\leq \frac{2}{N(J-4)} \left(1 + \frac{J-4}{2} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4 \right),
\end{aligned}$$

hence

$$\frac{1}{\pi} \int_{-\pi m}^{\pi m} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{-4} \mathbb{E} \left[\left| \hat{f}_Z^*(u) - f_Z^*(u) \right|^2 \right] du \leq \frac{1}{\pi} \frac{2}{N(J-4)} \int_{-\pi m}^{\pi m} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{-4} du + \frac{2m}{N}. \tag{31}$$

Now noticing that $|\hat{f}_Z^*(u) - f_Z^*(u)|$ and $|R\left(\frac{u}{\Delta}\right)|$ are independent and applying Lemma 3.3 for $p = 1$, we get

$$\begin{aligned}
\int_{-\pi m}^{\pi m} \mathbb{E} \left[\left| \hat{f}_Z^*(u) R\left(\frac{u}{\Delta}\right) \right|^2 \right] du &= \int_{-\pi m}^{\pi m} \mathbb{E} \left[\left| \hat{f}_Z^*(u) - f_Z^*(u) + f_Z^*(u) \right|^2 |R\left(\frac{u}{\Delta}\right)|^2 \right] du \\
&\leq 2 \int_{-\pi m}^{\pi m} |f_Z^*(u)|^2 \mathbb{E} |R\left(\frac{u}{\Delta}\right)|^2 du + 2 \int_{-\pi m}^{\pi m} \mathbb{E} \left[\left| \hat{f}_Z^*(u) - f_Z^*(u) \right|^2 |R\left(\frac{u}{\Delta}\right)|^2 \right] du \\
&\leq 2 \int_{-\pi m}^{\pi m} |f_Z^*(u)|^2 \mathbb{E} |R\left(\frac{u}{\Delta}\right)|^2 du + 2 \int_{-\pi m}^{\pi m} \mathbb{E} \left| \hat{f}_Z^*(u) - f_Z^*(u) \right|^2 \mathbb{E} |R\left(\frac{u}{\Delta}\right)|^2 du \\
&\leq 2 \int_{-\pi m}^{\pi m} |f_Z^*(u)|^2 \frac{k_N^2(u)}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{12}} du \\
&\quad + \frac{2}{\pi} \int_{-\pi m}^{\pi m} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{-4} \frac{2}{N(J-4)} \left(1 + \frac{J-4}{2} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4 \right) du \\
&\leq 2C_1 \int_{-\pi m}^{\pi m} |f_Z^*(u)|^2 \frac{k_N^2(u)}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{12}} du + \frac{4m\pi}{N} + \frac{2}{N(J-4)} \int_{-\pi m}^{\pi m} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{-4} du \tag{32}
\end{aligned}$$

Plugging (31) and (32) into Equation (30) yields

$$\begin{aligned}
\mathbb{E} \left\| f_{\beta,m} - \hat{f}_{\beta,m} \right\|^2 &\leq \frac{6m}{N} + \frac{6}{N(J-4)} \int_{-\pi m}^{\pi m} |f_{\varepsilon}^* \left(\frac{u}{\Delta} \right)|^{-4} du + \frac{2C_1}{\pi} \int_{-\pi m}^{\pi m} |f_Z^*(u)|^2 \frac{k_N^2(u)}{|f_{\varepsilon}^* \left(\frac{u}{\Delta} \right)|^{12}} du \\
&\leq \frac{6m}{N} + \frac{12}{N(J-4)} \frac{1}{2\pi} \int_{-\pi m}^{\pi m} |f_{\varepsilon}^* \left(\frac{u}{\Delta} \right)|^{-4} du + \frac{2C_1}{\pi} \int_{-\pi m}^{\pi m} |f_{\beta}^*(u)|^2 |f_{\varepsilon}^* \left(\frac{u}{\Delta} \right)|^4 \frac{k_N^2(u)}{|f_{\varepsilon}^* \left(\frac{u}{\Delta} \right)|^{12}} du \\
&\leq \frac{6m}{N} + \frac{12}{N(J-4)} \frac{1}{2\pi} \int_{-\pi m}^{\pi m} |f_{\varepsilon}^* \left(\frac{u}{\Delta} \right)|^{-4} du + \frac{2C_1}{\pi} \int_{-\pi m}^{\pi m} |f_{\beta}^*(u)|^2 \frac{k_N^2(u)}{|f_{\varepsilon}^* \left(\frac{u}{\Delta} \right)|^8} du \\
&\leq \frac{6m}{N} + \frac{12}{N(J-4)} \frac{1}{2\pi} \int_{-\pi m}^{\pi m} |f_{\varepsilon}^* \left(\frac{u}{\Delta} \right)|^{-4} du + \frac{2C_1}{\pi} \frac{1}{N} \int_{-\pi m}^{\pi m} \frac{|f_{\beta}^*(u)|^2}{|f_{\varepsilon}^* \left(\frac{u}{\Delta} \right)|^8} du
\end{aligned}$$

In the end

$$\mathbb{E} \left\| f_{\beta} - \hat{f}_{\beta,m} \right\|^2 \leq \|f_{\beta} - f_{\beta,m}\|^2 + \frac{6m}{N} + \frac{12}{N(J-4)} \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{1}{|f_{\varepsilon}^* \left(\frac{u}{\Delta} \right)|^4} du + \frac{4C_1}{N} \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{|f_{\beta}^*(u)|^2}{|f_{\varepsilon}^* \left(\frac{u}{\Delta} \right)|^8} du.$$

□

6.4. Proof of Theorem 4.1. The proof is similar to the proof of Theorem 4.2 and particularly of Theorem 3.1 in Kappus and Mabon (2013). For the last one, the only difference lies in $\hat{\Theta}^{\alpha}$ since f_{ε}^* is raised to a greater power. Then the proof can be easily adapted.

6.5. Proof of Theorem 4.2. Before proving any result, let us introduce some notations: for $k > m$,

$$\begin{aligned}
\hat{\Xi}(m, k) &= \hat{\Xi}(k) - \hat{\Xi}(m), \\
\hat{\Xi}^{\beta}(m, k) &= \hat{\Xi}^{\beta}(k) - \hat{\Xi}^{\beta}(m).
\end{aligned}$$

Moreover,

$$\begin{aligned}
\widehat{\text{qen}}(m, k) &= \text{qen}_1(m, k) + \text{qen}_2(m, k) + \text{qen}_3(m, k) \\
&= \frac{64(k-m)}{N} + 16 \frac{\hat{\mu}^2(m, k) \hat{\Xi}(m, k)}{N(J-4)} + 16\kappa^2 \log(N(k-m)) \frac{\hat{\Xi}^{\beta}(m, k)}{N}
\end{aligned}$$

with

$$\hat{\mu}(m, k) = \max \left\{ \sqrt{8 \log \left(1 + \hat{\Xi}(m, k) m^2 \right)}, \frac{16\sqrt{2}}{3\sqrt{N(J-4)}} \log \left(1 + \hat{\Xi}(m, k) (k-m)^2 \right) \right\}.$$

Now we can start the proof of Theorem 4.2. We denote by m^* the oracle cutoff defined by

$$m^* = \underset{m \in \mathcal{M}_n}{\text{argmin}} \left\{ -\|f_{\beta,m}\|^2 + \widehat{\text{qen}}(m) \right\}.$$

We have

$$\left\| f_{\beta} - \hat{f}_{\beta, \hat{m}} \right\|^2 \leq 2 \left\| f_{\beta} - \hat{f}_{\beta, m^*} \right\|^2 + 2 \left\| \hat{f}_{\beta, m^*} - \hat{f}_{\beta, \hat{m}} \right\|^2.$$

• Let us notice on the set $G = \{\hat{m} \leq m^*\}$:

$$\left\| \hat{f}_{\beta, m^*} - \hat{f}_{\beta, \hat{m}} \right\|^2 \mathbf{1}_G = \left(\left\| \hat{f}_{\beta, m^*} \right\|^2 - \left\| \hat{f}_{\beta, \hat{m}} \right\|^2 \right) \mathbf{1}_G.$$

Besides according to the definition of \hat{m} , one has the following inequalities:

$$-\left\| \hat{f}_{\beta, \hat{m}} \right\|^2 + \widehat{\text{qen}}(\hat{m}) \leq -\left\| \hat{f}_{\beta, m^*} \right\|^2 + \widehat{\text{qen}}(m^*) \quad (33)$$

which implies

$$-\left\| \hat{f}_{\beta, \hat{m}} \right\|^2 \leq -\left\| \hat{f}_{\beta, m^*} \right\|^2 + \widehat{\text{qen}}(m^*).$$

Thus

$$\left\| \hat{f}_{\beta, m^*} - \hat{f}_{\beta, \hat{m}} \right\|^2 \mathbf{1}_G = \left(\left\| \hat{f}_{\beta, m^*} \right\|^2 - \left\| \hat{f}_{\beta, \hat{m}} \right\|^2 \right) \mathbf{1}_G \leq \widehat{\text{qen}}(m^*).$$

Taking expectation, we apply the following Lemma

Lemma 6.1. *There is a positive constant C such that for any arbitrary $m \in \mathcal{M}_n$*

$$\mathbb{E} [\widehat{\text{qen}}(m)] \leq C \text{qen}(m), \quad (34)$$

It yields for some positive constant C

$$\begin{aligned} \mathbb{E} \left[\left\| f_\beta - \hat{f}_{\beta, \hat{m}} \right\|^2 \mathbf{1}_G \right] &\leq 2\mathbb{E} \left[\left\| f_\beta - \hat{f}_{\beta, m^*} \right\|^2 \right] + 2\mathbb{E} [\widehat{\text{qen}}(m^*)] \\ &\leq 2 \|f_\beta - f_{\beta, m^*}\|^2 + 2C \text{qen}(m^*). \end{aligned}$$

We just proved the desired result on G

$$\mathbb{E} \left[\left\| f_\beta - \hat{f}_{\beta, \hat{m}} \right\|^2 \mathbf{1}_G \right] \leq C \inf_{m \in \mathcal{M}_n} \left\{ \|f_\beta - f_{\beta, m}\|^2 + \text{pen}(m) \right\}. \quad (35)$$

• We now consider the set $G^c = \{\hat{m} > m^*\}$.

$$\begin{aligned} \left\| \hat{f}_{\beta, \hat{m}} - \hat{f}_{\beta, m^*} \right\|^2 \mathbf{1}_{G^c} &= \left(\left\| \hat{f}_{\beta, \hat{m}} - \hat{f}_{\beta, m^*} \right\|^2 - 4 \|f_{\beta, \hat{m}} - f_{\beta, m^*}\|^2 - \frac{1}{2} \widehat{\text{qen}}(m^*, \hat{m}) \right) \mathbf{1}_{G^c} \\ &\quad + \left(4 \|f_{\beta, \hat{m}} - f_{\beta, m^*}\|^2 + \frac{1}{2} \widehat{\text{qen}}(m^*, \hat{m}) \right) \mathbf{1}_{G^c} \\ &\leq \sup_{\substack{k \geq m^* \\ k \in \mathcal{M}_n}} \left\{ \left\| \hat{f}_{\beta, k} - \hat{f}_{\beta, m^*} \right\|^2 - 4 \|f_{\beta, k} - f_{\beta, m^*}\|^2 - \frac{1}{2} \widehat{\text{qen}}(m^*, k) \right\}_+ \\ &\quad + 4 \|f_{\beta, \hat{m}} - f_{\beta, m^*}\|^2 + \frac{1}{2} \sum_{\substack{k \geq m^* \\ k \in \mathcal{M}_n}} \widehat{\text{qen}}(m^*, k) \mathbf{1}\{\hat{m} = k\}. \end{aligned} \quad (36)$$

Let us first notice the following inequality

$$\forall k > m, \quad \widehat{\text{qen}}(m, k) \leq \widehat{\text{qen}}(k). \quad (37)$$

Besides by definition of \hat{m} (see Equation (25)), on the set $\{\hat{m} = k\} \cap G^c$ and applying Equation (33), one has

$$\begin{aligned} \frac{1}{2} (\widehat{\text{qen}}(k) - \widehat{\text{qen}}(m^*)) &\leq \left\| \hat{f}_{\beta, \hat{m}} - \hat{f}_{\beta, m^*} \right\|^2 - \frac{1}{2} \widehat{\text{qen}}(k) + \frac{1}{2} \widehat{\text{qen}}(m^*) \\ \frac{1}{2} \widehat{\text{qen}}(k) &\leq \left\| \hat{f}_{\beta, \hat{m}} - \hat{f}_{\beta, m^*} \right\|^2 - \frac{1}{2} \widehat{\text{qen}}(m^*, k) + \frac{1}{2} \widehat{\text{qen}}(m^*) \\ &\leq \left(\left\| \hat{f}_{\beta, \hat{m}} - \hat{f}_{\beta, m^*} \right\|^2 - 4 \|f_{\beta, \hat{m}} - f_{\beta, m^*}\|^2 - \frac{1}{2} \widehat{\text{qen}}(m^*, k) \right) + 4 \|f_{\beta, \hat{m}} - f_{\beta, m^*}\|^2 + \frac{1}{2} \widehat{\text{qen}}(m^*). \end{aligned} \quad (38)$$

Now using Equations (37) and (38)

$$\begin{aligned} \frac{1}{2} \sum_{\substack{k \geq m^* \\ k \in \mathcal{M}_n}} \widehat{\text{qen}}(m^*, k) &\leq \sup_{\substack{k \geq m^* \\ k \in \mathcal{M}_n}} \left\{ \left\| \hat{f}_{\beta, \hat{m}} - \hat{f}_{\beta, m^*} \right\|^2 - 4 \|f_{\beta, \hat{m}} - f_{\beta, m^*}\|^2 - \frac{1}{2} \widehat{\text{qen}}(m^*, k) \right\}_+ \\ &\quad + 4 \|f_{\beta, \hat{m}} - f_{\beta, m^*}\|^2 + \frac{1}{2} \widehat{\text{qen}}(m^*). \end{aligned}$$

From Equation (36), we now have

$$\begin{aligned} \left\| \hat{f}_{\beta, \hat{m}} - \hat{f}_{\beta, m^*} \right\|^2 \mathbf{1}_{G^c} &\leq 2 \sup_{\substack{k \geq m^* \\ k \in \mathcal{M}_n}} \left\{ \left\| \hat{f}_{\beta, k} - \hat{f}_{\beta, m^*} \right\|^2 - 4 \|f_{\beta, k} - f_{\beta, m^*}\|^2 - \frac{1}{2} \widehat{\text{qen}}(m^*, k) \right\}_+ \\ &\quad + 8 \|f_{\beta, \hat{m}} - f_{\beta, m^*}\|^2 + \frac{1}{2} \widehat{\text{qen}}(m^*). \end{aligned} \quad (39)$$

Taking expectation the first summand is negligible by applying the following Proposition.

Proposition 6.2. *There is a positive constant C such that for any arbitrary $m \in \mathcal{M}_n$*

$$\mathbb{E} \left[\sup_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \left\{ \left\| \hat{f}_{\beta, k} - \hat{f}_{\beta, m} \right\|^2 - 4 \|f_{\beta, k} - f_{\beta, m}\|^2 - \frac{1}{2} \widehat{\text{qen}}(m, k) \right\}_+ \right] \leq \frac{C}{N}. \quad (40)$$

Finally we have

$$\mathbb{E} \left[\left\| f_\beta - \hat{f}_{\beta, \hat{m}} \right\|^2 \mathbf{1}_{G^c} \right] \leq C \left(\|f_\beta - f_{\beta, m^*}\|^2 + \text{qen}(m^*) \right) + \frac{C'}{N(J-4)} + \frac{C'}{N}.$$

This combining with (35) complete the proof. \square

6.6. Proof of Lemma 6.1. • First consider $\widehat{\text{qen}}_2(m)$. For $q = 1/2$ or 1 , using Cauchy-Schwarz's inequality, we have

$$\begin{aligned} \mathbb{E} \left[\log^q \left(1 + \hat{\Xi}(m)m^2 \right) \hat{\Xi}(m) \right] &\leq \sqrt{\mathbb{E} \left[\log^{2q} \left(1 + \hat{\Xi}(m)m^2 \right) \right]} \mathbb{E} \left[\hat{\Xi}^2(m) \right]. \\ \hat{\Xi}^2(m) &= \left(\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2}}{\left| \tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^4} du \right)^2 \\ &\leq \frac{1}{4\pi^2} \left(\int_{-\pi m}^{\pi m} w\left(\frac{u}{\Delta}\right)^{-2} \left| \frac{1}{\left| \tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} - \frac{1}{\left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} + \frac{1}{\left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} \right|^2 du \right)^2 \\ &\leq \frac{1}{\pi^2} \left(\int_{-\pi m}^{\pi m} w\left(\frac{u}{\Delta}\right)^{-2} \left| \frac{1}{\left| \tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} - \frac{1}{\left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} \right|^2 + w\left(\frac{u}{\Delta}\right)^{-2} \left| \frac{1}{\left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} \right|^2 du \right)^2 \\ &\leq \frac{2}{\pi^2} \left(\int_{-\pi m}^{\pi m} w\left(\frac{u}{\Delta}\right)^{-2} \left| \frac{1}{\left| \tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} - \frac{1}{\left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} \right|^2 du \right)^2 + 8\Xi^2(m). \end{aligned}$$

Now noticing that we can write the first term of the lastest inequality, we have

$$\begin{aligned} &\mathbb{E} \left(\int_{-\pi m}^{\pi m} w\left(\frac{u}{\Delta}\right)^{-2} \left| \frac{1}{\left| \tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} - \frac{1}{\left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} \right|^2 du \right)^2 \\ &= \mathbb{E} \int_{-\pi m}^{\pi m} \int_{-\pi m}^{\pi m} w\left(\frac{u}{\Delta}\right)^{-2} w\left(\frac{v}{\Delta}\right)^{-2} \left| \frac{1}{\left| \tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} - \frac{1}{\left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} \right|^2 \left| \frac{1}{\left| \tilde{f}_\varepsilon^*\left(\frac{v}{\Delta}\right) \right|^2} - \frac{1}{\left| f_\varepsilon^*\left(\frac{v}{\Delta}\right) \right|^2} \right|^2 du dv \\ &= \int_{-\pi m}^{\pi m} \int_{-\pi m}^{\pi m} w\left(\frac{u}{\Delta}\right)^{-2} w\left(\frac{v}{\Delta}\right)^{-2} \mathbb{E} \left[\left| \frac{1}{\left| \tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} - \frac{1}{\left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} \right|^2 \left| \frac{1}{\left| \tilde{f}_\varepsilon^*\left(\frac{v}{\Delta}\right) \right|^2} - \frac{1}{\left| f_\varepsilon^*\left(\frac{v}{\Delta}\right) \right|^2} \right|^2 \right] du dv. \end{aligned}$$

Now applying Cauchy-Schwarz's inequality and Lemma 3.3 for $p = 2$

$$\begin{aligned} &\mathbb{E} \left[\left| \frac{1}{\left| \tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} - \frac{1}{\left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} \right|^2 \left| \frac{1}{\left| \tilde{f}_\varepsilon^*\left(\frac{v}{\Delta}\right) \right|^2} - \frac{1}{\left| f_\varepsilon^*\left(\frac{v}{\Delta}\right) \right|^2} \right|^2 \right] \\ &\leq \sqrt{\mathbb{E} \left| \frac{1}{\left| \tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} - \frac{1}{\left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2} \right|^4} \sqrt{\mathbb{E} \left| \frac{1}{\left| \tilde{f}_\varepsilon^*\left(\frac{v}{\Delta}\right) \right|^2} - \frac{1}{\left| f_\varepsilon^*\left(\frac{v}{\Delta}\right) \right|^2} \right|^4} \\ &\leq \sqrt{C_2 |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{-8}} \sqrt{C_2 |f_\varepsilon^*\left(\frac{v}{\Delta}\right)|^{-8}} \\ &\leq C_2 |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{-4} |f_\varepsilon^*\left(\frac{v}{\Delta}\right)|^{-4} \end{aligned}$$

So we have

$$\mathbb{E} \left[\hat{\Xi}^2(m) \right] \leq (8C_2 + 8)\Xi^2(m).$$

Besides we have

$$\begin{aligned}\hat{\Xi}(m) &= \frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2}}{\left|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^4} du \\ &\leq \frac{1}{\pi} \int_{-\pi m}^{\pi m} w\left(\frac{u}{\Delta}\right)^{-2} \left| \frac{1}{\left|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^2} - \frac{1}{\left|f_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^2} \right|^2 du + 2\Xi(m).\end{aligned}$$

Once again applying Lemma 3.3 for $p = 1$

$$\begin{aligned}\mathbb{E} \left[\int_{-\pi m}^{\pi m} w\left(\frac{u}{\Delta}\right)^{-2} \left| \frac{1}{\left|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^2} - \frac{1}{\left|f_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^2} \right|^2 du \right] &\leq \int_{-\pi m}^{\pi m} w\left(\frac{u}{\Delta}\right)^{-2} \mathbb{E} \left[\left| \frac{1}{\left|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^2} - \frac{1}{\left|f_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^2} \right|^2 \right] du \\ &\leq C_1 \int_{-\pi m}^{\pi m} w\left(\frac{u}{\Delta}\right)^{-2} \left|f_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^{-4} du.\end{aligned}$$

So we have

$$\mathbb{E} \left[\hat{\Xi}(m) \right] \leq (2C_1 + 2)\Xi(m).$$

Now using Jensen's inequality (since log is concave)

$$\begin{aligned}\mathbb{E} \left[\log^{2q} \left(1 + \hat{\Xi}(m)m^2 \right) \right] &\leq \log^{2q} \left(\mathbb{E} \left[1 + \hat{\Xi}(m)m^2 \right] \right) \\ &\leq \log^{2q} \left(1 + \mathbb{E} \left[\hat{\Xi}(m) \right] m^2 \right) \\ &\leq \log^{2q} \left(1 + C\Xi(m)m^2 \right) \\ &\leq C \log^{2q} \left(1 + \Xi(m)m^2 \right).\end{aligned}$$

So

$$\mathbb{E} \left[\log^q \left(1 + \hat{\Xi}(m)m^2 \right) \hat{\Xi}(m) \right] \leq C \log^q \left(1 + \Xi(m)m^2 \right) \Xi(m)$$

which means $\mathbb{E}[\widehat{\text{qen}}_2(m)] \leq C\text{qen}_2(m)$.

• Consider now $\widehat{\text{qen}}_3(m)$. Let $A_p\left(\frac{u}{\Delta}\right) = \left| \frac{1}{\left|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^2} - \frac{1}{\left|f_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^2} \right|^{2p} \left/ \left(\frac{1}{\left|f_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^{4p}} \wedge \frac{k_N^{2p}\left(\frac{u}{\Delta}\right)}{\left|f_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^{12p}} \right) \right.$. Another application of Lemma 3.3 yields

$$\begin{aligned}&\frac{1}{N} \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2} \left|\hat{f}_Z^*(u)\right|^2}{\left|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^{12}} du \right] \\ &\leq \frac{2}{N} \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2} \left|f_Z^*(u)\right|^2}{\left|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^{12}} du \right] + \frac{2}{N} \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2} \left|\hat{f}_Z^*(u) - f_Z^*(u)\right|^2}{\left|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^{12}} du \right] \\ &\leq \frac{4}{N} \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2} \left|f_Z^*(u)\right|^2}{\left|f_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^{12}} du \right] \mathbb{E} \left[1 + \sup_{u \in \mathbb{R}} A_3\left(\frac{u}{\Delta}\right) \right] + \frac{2}{N} \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2} \left|\hat{f}_Z^*(u) - f_Z^*(u)\right|^2}{\left|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^{12}} du \right] \\ &\leq \frac{C}{N} \Xi^\beta(m) + \frac{2}{N} \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2} \left|\hat{f}_Z^*(u) - f_Z^*(u)\right|^2}{\left|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^{12}} du \right].\end{aligned}$$

Let us notice that

$$\begin{aligned}e^{iuZ_{k,j}} - \mathbb{E} \left[e^{iuZ_{k,j}} \right] &= e^{iuZ_{k,j}} - f_\beta^*(u) \left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2 \\ &= e^{iu(\beta_k + \eta_{k,j}/\Delta)} - f_\beta^*(u) \left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2 \\ &= e^{iu\beta_k} \left(e^{iu\eta_{k,j}/\Delta} - \left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2 \right) + (e^{iu\beta_k} - f_\beta^*(u)) \left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2,\end{aligned}$$

hence

$$\hat{f}_Z^*(u) - f_Z^*(u) = \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2 + \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} \left(e^{iu\beta_k} \left(e^{iu\eta_{k,j}/\Delta} - \left| f_\varepsilon^*\left(\frac{u}{\Delta}\right) \right|^2 \right) \right). \quad (41)$$

Then

$$\frac{1}{N} \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2} |\hat{f}_Z^*(u) - f_Z^*(u)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{12}} du \right] \quad (42)$$

$$\begin{aligned} &\leq \frac{2}{N} \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2} |\hat{f}_\beta^*(u) - f_\beta^*(u)|^2 |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{12}} du \right] \\ &+ \frac{2}{N} \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2} \left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} \left(e^{iu\eta_k} \left(e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2 \right) \right) \right|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{12}} du \right]. \quad (43) \end{aligned}$$

Let us consider the first term on the right-hand side of Equation (43). We use the fact that $|\tilde{f}_\varepsilon^*(u)|^4 \geq N^{-1/2}(\log N)^{1/2}w(u)^{-1}$ as well as the independence of \hat{f}_β^* and \tilde{f}_ε^* to find

$$\begin{aligned} &\frac{1}{N} \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2} |\hat{f}_\beta^*(u) - f_\beta^*(u)|^2 |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{12}} du \right] \\ &\leq \frac{1}{N} \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2} w(u)^{-2} w(u)^2 |\hat{f}_\beta^*(u) - f_\beta^*(u)|^2 |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4 N^{-1} (\log N) w\left(\frac{u}{\Delta}\right)^{-2}} du \right] \\ &\leq \mathbb{E} \left[\sup_{u \in \mathbb{R}} |\hat{f}_\beta^*(u) - f_\beta^*(u)|^2 w(u)^2 \right] \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w(u)^{-2} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}{|f_\varepsilon^*(u)|^4} du \right]. \end{aligned}$$

Thanks to Theorem 5.1 in Neumann and Reiß (2009), for some positive constant C ,

$$\mathbb{E} \left[\sup_{u \in \mathbb{R}} |\hat{f}_\beta^*(u) - f_\beta^*(u)|^2 w(u)^2 \right] \leq \frac{C}{N}.$$

Applying Lemma 3.3 for $p = 1$, we get

$$\begin{aligned} \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w(u)^{-2} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4} du \right] &\leq \frac{C}{2\pi} \int_{-\pi m}^{\pi m} \frac{w(u)^{-2} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4} du \\ &\leq Cm. \end{aligned}$$

which means that there exists C such that

$$\frac{1}{N} \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w\left(\frac{u}{\Delta}\right)^{-2} |\hat{f}_\beta^*(u) - f_\beta^*(u)|^2 |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{12}} du \right] \leq C \frac{m}{N}.$$

Now let us consider the second term on the right-hand side of Equation (43) and let us notice that

$$\begin{aligned} &\mathbb{E} \left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} \left(e^{iu\beta_k} \left(e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2 \right) \right) \right|^2 \\ &= \frac{4}{N^2(J-4)^2} \mathbb{E} \left[\sum_{k,k'=1}^N \sum_{j,j'=3}^{J/2} e^{iu(\beta_k - \beta_{k'})} \left(e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2 \right) \left(e^{iu\eta_{k',j'}/\Delta} - |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2 \right) \right] \\ &= \frac{4}{N^2(J-4)^2} \sum_{k,k'=1}^N \sum_{j,j'=3}^{J/2} \mathbb{E} \left[e^{iu(\beta_k - \beta_{k'})} \right] \mathbb{E} \left[\left(e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2 \right) \left(e^{iu\eta_{k',j'}/\Delta} - |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2 \right) \right] \end{aligned}$$

The term $e^{iu(\beta_k - \beta_{k'})}$ is equal to 1 if $k = k'$ otherwise $\mathbb{E}[e^{iu(\beta_k - \beta_{k'})}] = |f_\beta^*(u)|^2 \leq 1$. Moreover if $k \neq k'$ and $j \neq j'$, $\mathbb{E}[(e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*(\frac{u}{\Delta})|^2)(e^{iu\eta_{k',j'}/\Delta} - |f_\varepsilon^*(\frac{u}{\Delta})|^2)] = 0$, then

$$\begin{aligned} & \mathbb{E} \left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} \left(e^{iu\beta_k} \left(e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*(\frac{u}{\Delta})|^2 \right) \right) \right|^2 \\ & \leq \frac{4}{N(J-4)^2} \sum_{j=3}^{J/2} \mathbb{E} \left[\left(e^{iu\eta_{1,j}/\Delta} - |f_\varepsilon^*(\frac{u}{\Delta})|^2 \right) \left(e^{iu\eta_{1,j}/\Delta} - |f_\varepsilon^*(\frac{u}{\Delta})|^2 \right) \right] \\ & \leq \frac{4}{N(J-4)} (1 - |f_\varepsilon^*(u)|^4) \\ & \leq \frac{4}{N(J-4)}. \end{aligned}$$

Now noticing the independence between the numerator and the denominator, using that $|\tilde{f}_\varepsilon^*(u)|^4 \geq N^{-1/2}(\log N)^{1/2}w(u)^{-1}$ and applying Lemma 3.3 for $p = 1$, we have

$$\begin{aligned} & \frac{1}{N} \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w(\frac{u}{\Delta})^{-2} \left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} \left(e^{iu\beta_k} \left(e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*(\frac{u}{\Delta})|^2 \right) \right) \right|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^{12}} du \right] \\ & \leq \frac{4}{N(J-4)} \mathbb{E} \left[\frac{1}{2\pi} \int_{-\pi m}^{\pi m} \frac{w(\frac{u}{\Delta})^{-2}}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} du \right] \\ & \leq \frac{C}{N(J-4)} \Xi(m). \end{aligned}$$

This completes the proof. \square

6.7. Proof of Proposition 6.2. Before proving Proposition 6.2, we first to need prove two auxiliary lemmas. In the sequel, C will always denote some universal positive constant, but the value may vary from line to line. For $k > m$, let us introduce the following notation : $A(m, k) := \{u \in \mathbb{R}, |u| \in [\pi m, \pi k]\}$.

Lemma 6.3. *For an estimator of \tilde{f}_ε^* defined by (11), assume $\kappa > \sqrt{c_1 p}$. Let $\tau \geq 2\kappa$ and $x \geq 1$. Then for some positive constant C*

$$\mathbb{P} \left[\exists u \in \mathbb{R} : |(\tilde{f}_\varepsilon^*)^4(u) - (f_\varepsilon^*)^4(u)| > \tau (\log(Nx))^{1/2} w(u)^{-1} N^{-1/2} \right] \leq Cx^{-p} N^{-p}$$

Proof.

$$\left| (\tilde{f}_\varepsilon^*)^4(u) - (f_\varepsilon^*)^4(u) \right| \leq \left| (\tilde{f}_\varepsilon^*)^4(u) - (\hat{f}_\varepsilon^*)^4(u) \right| + \left| (\hat{f}_\varepsilon^*)^4(u) - (f_\varepsilon^*)^4(u) \right| \leq 2k_N(u) + \left| (\hat{f}_\varepsilon^*)^4(u) - (f_\varepsilon^*)^4(u) \right|$$

By Lemma A.4, we have

$$\begin{aligned} & \mathbb{P} \left[\exists u \in \mathbb{R} : |(\tilde{f}_\varepsilon^*)^4(u) - (f_\varepsilon^*)^4(u)| > \tau (\log(Nx))^{1/2} w(u)^{-1} N^{-1/2} \right] \\ & \leq \mathbb{P} \left[\exists u \in \mathbb{R} : \left| (\hat{f}_\varepsilon^*)^4(u) - (f_\varepsilon^*)^4(u) \right| + 2k_N(u) > \tau (\log(Nx))^{1/2} w(u)^{-1} N^{-1/2} \right] \\ & \leq \mathbb{P} \left[\exists u \in \mathbb{R} : |(\tilde{f}_\varepsilon^*)^4(u) - (f_\varepsilon^*)^4(u)| > (\tau - 2\kappa) (\log(Nx))^{1/2} w(u)^{-1} N^{-1/2} \right] \\ & \leq Cx^{-p} N^{-p}. \end{aligned}$$

\square

Lemma 6.4. *In the situation of the preceding Lemma*

$$\mathbb{P} \left[\exists u \in \mathbb{R} : \left| |\tilde{f}_\varepsilon^*(u)|^2 - |f_\varepsilon^*(u)|^2 \right| \mathbf{1}_{\left\{ |\tilde{f}_\varepsilon^*(u)| < |f_\varepsilon^*(u)| \right\}} > \frac{\tau (\log(Nx))^{1/2} w(u)^{-1} N^{-1/2}}{|\tilde{f}_\varepsilon^*(u)|^2} \right] \leq Cx^{-p} N^{-p}$$

Proof. This is a direct consequence of Lemma 6.3 using the fact that for $x, y \geq 0$, $|\sqrt{x} - \sqrt{y}| \leq \frac{|x-y|}{2\sqrt{x \wedge y}}$ holds. \square

Lemma 6.5. *There is a positive constant C such that for any arbitrary $m \in \mathcal{M}_n$*

$$\mathbb{E} \left[\sup_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \left\{ \int_{A(m,k)} \left| \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \right|^2 du - \frac{1}{24} \widehat{\text{qen}}_1(m, k) \right\}_+ \right] \leq \frac{C}{N} \quad (44)$$

Proof.

$$\begin{aligned} & \mathbb{E} \left[\sup_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \left\{ \frac{1}{2\pi} \int_{A(m,k)} \left| \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \right|^2 du - \frac{1}{24} \widehat{\text{qen}}_1(m, k) \right\}_+ \right] \\ & \leq \mathbb{E} \left[\sum_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \left\{ \frac{1}{2\pi} \int_{A(m,k)} \left| \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \right|^2 du - \frac{1}{24} \widehat{\text{qen}}_1(m, k) \right\}_+ \right] \\ & \leq \sum_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \mathbb{E} \left[\left\{ \frac{1}{2\pi} \int_{A(m,k)} \left| \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \right|^2 du - \frac{1}{24} \widehat{\text{qen}}_1(m, k) \right\}_+ \right] \\ & \leq \sum_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \mathbb{E} \left[\left\{ \sup_{t \in S(m,k)} \left| \frac{1}{2\pi} \int_{A(m,k)} t^*(u) \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) du \right|^2 - \frac{\pi(k-m)}{N} \right\}_+ \right]. \end{aligned}$$

We then study the following empirical process

$$\nu_N(t) = \frac{1}{N} \sum_{k=1}^N \left(\frac{1}{2\pi} \int_{A(m,k)} \overline{t^*(u)} e^{iu\beta_k} du - \mathbb{E} \left[\frac{1}{2\pi} \int_{A(m,k)} \overline{t^*(u)} e^{iu\beta_k} du \right] \right)$$

and define the following space: $S(m, k) = \{t \in A(m, k), \|t\| = 1\}$. Then we can write

$$\begin{aligned} \sup_{t \in S(m,k)} |\nu_N(t)|^2 &= \left| \int_{A(m,k)} \frac{1}{2\pi} \overline{t^*(u)} \left(\frac{1}{N} \sum_{k=1}^N e^{iu\beta_k} - f_\beta^*(u) \right) du \right|^2 \\ &\leq \sup_{t \in S(m,k)} \frac{1}{2\pi} \int_{A(m,k)} |t^*(u)|^2 du \frac{1}{2\pi} \int_{A(m,k)} \left| \frac{1}{N} \sum_{k=1}^N e^{iu\beta_k} - f_\beta^*(u) \right|^2 du \\ &\leq \frac{1}{2\pi} \int_{A(m,k)} \left| \frac{1}{N} \sum_{k=1}^N e^{iu\beta_k} - f_\beta^*(u) \right|^2 du, \end{aligned}$$

hence

$$\begin{aligned} \mathbb{E} \left[\sup_{t \in S(m,k)} |\nu_N(t)|^2 \right] &\leq \frac{1}{2\pi} \int_{A(m,k)} \mathbb{E} \left| \frac{1}{N} \sum_{k=1}^N e^{iu\beta_k} - f_\beta^*(u) \right|^2 du \\ &\leq \frac{(k-m)}{N} := H^2. \end{aligned}$$

$$\text{Var} \left[\left| \frac{1}{2\pi} \int_{A(m,k)} \overline{t^*(u)} e^{iu\beta_1} du \right| \right] \leq \mathbb{E} \left[\left| \frac{1}{2\pi} \int_{A(m,k)} \overline{t^*(u)} e^{iu\beta_1} du \right|^2 \right]$$

Let us notice that the expectation can be rewritten as follows

$$\begin{aligned} \mathbb{E} \left[\left| \frac{1}{2\pi} \int_{A(m,k)} \overline{t^*(u)} e^{iu\beta_1} du \right|^2 \right] &= \mathbb{E} \left[\frac{1}{4\pi^2} \int_{A(m,k)} \overline{t^*(u)} e^{iu\beta_1} du \int_{A(m,k)} \overline{t^*(v)} e^{-iv\beta_1} du dv \right] \\ &= \frac{1}{4\pi^2} \iiint_{\mathbb{R} \times A(m,k) \times A(m,k)} f_\beta(x) e^{i(u-v)x} \overline{t^*(u)} t^*(v) du dv dx \\ &= \frac{1}{4\pi^2} \iint_{A(m,k) \times A(m,k)} f_\beta^*(u-v) \overline{t^*(u)} t^*(v) du dv. \end{aligned}$$

Now applying Cauchy-Schwarz's inequality

$$\begin{aligned}
\mathbb{V}\text{ar} \left[\left| \frac{1}{2\pi} \int_{A(m,k)} \overline{t^*(u)} e^{iu\beta_1} du \right| \right] &\leq \frac{1}{4\pi^2} \iint_{A(m,k) \times A(m,k)} |f_\beta^*(u-v) t^*(u) t^*(v)| du dv \\
&\leq \frac{1}{4\pi^2} \sqrt{\iint_{A(m,k) \times A(m,k)} |f_\beta^*(u-v)| |t^*(u)|^2 du dv} \\
&\quad \cdot \sqrt{\iint_{A(m,k) \times A(m,k)} |f_\beta^*(u-v)| |t^*(v)|^2 du dv} \\
&\leq \frac{1}{4\pi^2} \sqrt{\iint_{A(m,k) \times \mathbb{R}} |f_\beta^*(w)| |t^*(u)|^2 du dw} \\
&\quad \cdot \sqrt{\iint_{\mathbb{R} \times A(m,k)} |f_\beta^*(w)| |t^*(v)|^2 dw dv} \\
&\leq \frac{1}{4\pi^2} \iint_{A(m,k) \times A(m,k)} |f_\beta^*(w)| |t^*(u)|^2 du dw \\
&\leq \frac{\|f_\beta^*\|_1}{2\pi},
\end{aligned}$$

hence

$$\sup_{t \in S(m,k)} \mathbb{V}\text{ar} \left[\left| \frac{1}{2\pi} \int_{A(m,k)} \overline{t^*(u)} e^{iu\beta_1} du \right| \right] \leq \frac{\|f_\beta^*\|_1}{2\pi} := \nu.$$

And

$$\left| \frac{1}{2\pi} \int_{A(m,k)} \overline{t^*(u)} e^{iux} du \right|^2 \leq k - m,$$

hence

$$\sup_{x \in \mathbb{R}} \left| \frac{1}{2\pi} \int_{A(m,k)} \overline{t^*(u)} e^{iux} du \right| \leq \sqrt{k - m} := M_1.$$

We can now apply Talagrand's inequality

$$\begin{aligned}
&\mathbb{E} \left[\sup_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \left\{ \int_{A(m,k)} \left| \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \right|^2 du - \frac{1}{24} \widehat{\text{qen}}_1(m, k) \right\}_+ \right] \\
&\leq \sum_{\substack{k \geq m \\ k \in \mathcal{M}_n}} K_1 \left(\frac{\|f_\beta^*\|_1}{N} e^{-K_2 \frac{k-m}{\|f_\beta^*\|_1}} + K_2 \frac{k-m}{N^2} e^{-K_3 \frac{\sqrt{N(k-m)}}{\sqrt{k-m}}} \right).
\end{aligned}$$

Finally

$$\mathbb{E} \left[\sup_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \left\{ \frac{1}{2\pi} \int_{A(m,k)} \left| \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \right|^2 du - \frac{1}{24} \widehat{\text{qen}}_1(m, k) \right\}_+ \right] \leq \frac{C}{N}$$

□

Lemma 6.6. *There is a positive constant C such that for any arbitrary $m \in \mathcal{M}_n$*

$$\mathbb{E} \left[\sup_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \left\{ \frac{1}{2\pi} \int_{A(m,k)} \frac{\left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} e^{iu\beta_k} (e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*(\frac{u}{\Delta})|^2) \right|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} du - \frac{1}{16} \widehat{\text{qen}}_2(m, k) \right\}_+ \right] \leq \frac{C}{N(J-4)} \quad (45)$$

Proof. We introduce the notation $\mathbb{E} \left[X | \tilde{f}_\varepsilon^*, \beta \right]$ which corresponds to the conditional expectation of a random variable X given β_1, \dots, β_N and $\varepsilon_{k,j}$ for $j = 1, 2$ and $k = 1, \dots, N$.

$$\begin{aligned} & \mathbb{E} \left[\sup_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \left\{ \frac{1}{2\pi} \int_{A(m,k)} \frac{\left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} e^{iu\beta_k} (e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*(\frac{u}{\Delta})|^2) \right|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} du - \frac{1}{16} \widehat{\text{qen}}_2(m, k) \right\}_+ \right] \\ & \leq \mathbb{E} \left[\sum_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \left\{ \frac{1}{2\pi} \int_{A(m,k)} \frac{\left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} e^{iu\beta_k} (e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*(\frac{u}{\Delta})|^2) \right|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} du - \frac{\hat{\mu}^2(m, k) \hat{\Xi}(m, k)}{N(J-4)} \right\}_+ \right] \\ & \leq \mathbb{E} \left[\sum_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \frac{1}{2\pi} \int_{A(m,k)} \mathbb{E} \left[\left\{ \frac{\left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} e^{iu\beta_k} (e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*(\frac{u}{\Delta})|^2) \right|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} - \frac{\hat{\mu}^2(m, k)}{N(J-4) |\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} \right\}_+ \middle| \tilde{f}_\varepsilon^*, \beta \right] du \right] \end{aligned}$$

Now $\left(\frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} e^{iu\beta_k} e^{iu\eta_{k,j}/\Delta} \right) / |\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^2$ (conditional on $\tilde{f}_\varepsilon^*(\frac{u}{\Delta})$ and β_1, \dots, β_N) is the sum of $N(J-4)$ independent and identically distributed random variables with variance $v^2 \leq 1/|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4$ which are surely bounded by $1/|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^2$. Thus Lemma A.1 gives

$$\begin{aligned} & \mathbb{E} \left[\left\{ \frac{\left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} e^{iu\beta_k} (e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*(\frac{u}{\Delta})|^2) \right|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} - \frac{\hat{\mu}^2(m, k)}{N(J-4) |\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} \right\}_+ \middle| \tilde{f}_\varepsilon^*, \beta \right] \\ & \leq \frac{32}{N(J-4) |\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} \exp\left(-\frac{\hat{\mu}^2(m, k)}{8}\right) + \frac{128\sqrt{2}}{N^2(J-4)^2 |\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} \exp\left(-\frac{3}{16\sqrt{2}} \sqrt{N(J-4)} \hat{\mu}(m, k)\right) \\ & \leq \frac{32}{N(J-4) |\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} (k-m)^{-2} \hat{\Xi}(m, k)^{-1} + \frac{128\sqrt{2}}{N^2(J-4)^2 |\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} (k-m)^{-2} \hat{\Xi}(m, k)^{-1} \end{aligned}$$

where we used the fact that

$$\hat{\mu}(m, k) \leq \max \left\{ \sqrt{8 \log \left(1 + \hat{\Xi}(m, k)(k-m)^2 \right)}, \frac{16\sqrt{2}}{3\sqrt{N(J-4)}} \log \left(1 + \hat{\Xi}(m, k)(k-m)^2 \right) \right\}.$$

We have thus shown for a universal positive constant C that for any $m, k \in \mathcal{M}_n$

$$\begin{aligned} & \int_{A(m,k)} \mathbb{E} \left[\left\{ \frac{\left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} e^{iu\beta_k} (e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*(\frac{u}{\Delta})|^2) \right|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} - \frac{\hat{\mu}^2(m, k)}{N(J-4) |\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} \right\}_+ \middle| \tilde{f}_\varepsilon^*, \beta \right] du \\ & \leq \frac{C}{N(J-4)} (k-m)^{-2} \hat{\Xi}(m, k)^{-1} \int_{A(m,k)} \frac{du}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} \\ & \leq \frac{C}{N(J-4)} (k-m)^{-2}. \end{aligned}$$

Finally

$$\mathbb{E} \left[\sup_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \left\{ \frac{1}{2\pi} \int_{A(m,k)} \frac{\left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} e^{iu\beta_k} (e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*(\frac{u}{\Delta})|^2) \right|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} du - \frac{1}{16} \widehat{\text{qen}}_2(m, k) \right\}_+ \right] \leq \frac{C}{N(J-4)}.$$

□

Proof of Proposition 6.2. Using Plancherel's formula, we get

$$\begin{aligned}
\|\hat{f}_{\beta,k} - \hat{f}_{\beta,m}\|^2 &= \frac{1}{2\pi} \int_{A(m,k)} \frac{|\hat{f}_Z^*(u)|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} du \\
&= \frac{1}{2\pi} \int_{A(m,k)} \frac{|\hat{f}_Z^*(u)|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} \mathbb{1}_{\{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})| > |f_\varepsilon^*(\frac{u}{\Delta})|\}} du + \frac{1}{2\pi} \int_{A(m,k)} \frac{|\hat{f}_Z^*(u)|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} \mathbb{1}_{\{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})| \leq |f_\varepsilon^*(\frac{u}{\Delta})|\}} du \\
&\leq \frac{1}{\pi} \int_{A(m,k)} \left(\frac{|\hat{f}_Z^*(u) - f_Z^*(u)|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} + \frac{|f_Z^*(u)|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} \right) \mathbb{1}_{\{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})| > |f_\varepsilon^*(\frac{u}{\Delta})|\}} du \\
&\quad + \frac{1}{\pi} \int_{A(m,k)} \left(|\hat{f}_Z^*(u)|^2 \left| \frac{1}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^2} - \frac{1}{|f_\varepsilon^*(\frac{u}{\Delta})|^2} \right|^2 + \frac{|f_Z^*(u)|^2}{|f_\varepsilon^*(\frac{u}{\Delta})|^4} \right) \mathbb{1}_{\{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})| \leq |f_\varepsilon^*(\frac{u}{\Delta})|\}} du.
\end{aligned}$$

Then it follows that

$$\begin{aligned}
\|\hat{f}_{\beta,k} - \hat{f}_{\beta,m}\|^2 &\leq \frac{1}{\pi} \int_{A(m,k)} \frac{|\hat{f}_Z^*(u) - f_Z^*(u)|^2}{|f_\varepsilon^*(\frac{u}{\Delta})|^4} \mathbb{1}_{\{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})| > |f_\varepsilon^*(\frac{u}{\Delta})|\}} du + \frac{1}{\pi} \int_{A(m,k)} \frac{|f_Z^*(u)|^2}{|f_\varepsilon^*(\frac{u}{\Delta})|^4} \mathbb{1}_{\{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})| > |f_\varepsilon^*(\frac{u}{\Delta})|\}} du \\
&\quad + \frac{1}{\pi} \int_{A(m,k)} |\hat{f}_Z^*(u)|^2 \frac{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})^2 - |f_\varepsilon^*(\frac{u}{\Delta})|^2|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4 |f_\varepsilon^*(\frac{u}{\Delta})|^4} \mathbb{1}_{\{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})| \leq |f_\varepsilon^*(\frac{u}{\Delta})|\}} du \\
&\quad + \frac{2}{\pi} \int_{A(m,k)} \frac{|\hat{f}_Z^*(u) - f_Z^*(u)|^2}{|f_\varepsilon^*(\frac{u}{\Delta})|^4} \mathbb{1}_{\{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})| \leq |f_\varepsilon^*(\frac{u}{\Delta})|\}} du + \frac{2}{\pi} \int_{A(m,k)} \frac{|f_Z^*(u)|^2}{|f_\varepsilon^*(\frac{u}{\Delta})|^4} \mathbb{1}_{\{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})| \leq |f_\varepsilon^*(\frac{u}{\Delta})|\}} du \\
&\leq \frac{2}{\pi} \int_{A(m,k)} \frac{|\hat{f}_Z^*(u) - f_Z^*(u)|^2}{|f_\varepsilon^*(\frac{u}{\Delta})|^4} du + \frac{2}{\pi} \int_{A(m,k)} \frac{|f_Z^*(u)|^2}{|f_\varepsilon^*(\frac{u}{\Delta})|^4} du \\
&\quad + \frac{1}{\pi} \int_{A(m,k)} |\hat{f}_Z^*(u)|^2 \frac{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})^2 - |f_\varepsilon^*(\frac{u}{\Delta})|^2|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4 |f_\varepsilon^*(\frac{u}{\Delta})|^4} \mathbb{1}_{\{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})| \leq |f_\varepsilon^*(\frac{u}{\Delta})|\}} du \\
&\leq \frac{2}{\pi} \int_{A(m,k)} \frac{|\hat{f}_Z^*(u) - f_Z^*(u)|^2}{|f_\varepsilon^*(\frac{u}{\Delta})|^4} du + 4 \|\beta_{\beta,k} - \beta_{\beta,m}\|^2 \\
&\quad + \frac{1}{\pi} \int_{A(m,k)} |\hat{f}_Z^*(u)|^2 \frac{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^2 - |f_\varepsilon^*(\frac{u}{\Delta})|^2|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^8} \mathbb{1}_{\{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})| \leq |f_\varepsilon^*(\frac{u}{\Delta})|\}} du. \tag{46}
\end{aligned}$$

To bound the second term on the right-hand-side of inequality (46), we introduce the following set

$$C(m,k) = \left\{ \forall u \in \mathbb{R} : \left| |\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^2 - |f_\varepsilon^*(\frac{u}{\Delta})|^2 \right|^2 \mathbb{1}_{\{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})| \leq |f_\varepsilon^*(\frac{u}{\Delta})|\}} \leq \frac{4\kappa^2 \log(N(k-m)) w(\frac{u}{\Delta})^{-2} N^{-1}}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^4} \right\}.$$

On $C(m,k)$, the following inequalities can be deduced

$$\begin{aligned}
&\frac{1}{\pi} \int_{A(m,k)} |\hat{f}_Z^*(u)|^2 \frac{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})^2 - |f_\varepsilon^*(\frac{u}{\Delta})|^2|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^8} \mathbb{1}_{\{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})| \leq |f_\varepsilon^*(\frac{u}{\Delta})|\}} du \\
&\leq 8\kappa^2 \log(N(k-m)) N^{-1} \frac{1}{2\pi} \int_{A(m,k)} \frac{w(\frac{u}{\Delta})^{-2} |\hat{f}_Z^*(u)|^2}{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|^{12}} \mathbb{1}_{\{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})| \leq |f_\varepsilon^*(\frac{u}{\Delta})|\}} du \\
&\leq 8\kappa^2 \log(N(k-m)) N^{-1} \hat{\Xi}^\beta(m,k) := \frac{1}{2} \widehat{\text{qen}}_3(m,k).
\end{aligned}$$

To bound the first term on the right-hand-side of inequality (46), we use Equation (41) and so we have

$$\begin{aligned}
& \frac{2}{\pi} \int_{A(m,k)} \frac{|\hat{f}_Z^*(u) - f_Z^*(u)|^2}{|\tilde{f}_\varepsilon^*(u)|^2} du \\
& \leq \frac{4}{\pi} \int_{A(m,k)} \frac{\left| \frac{1}{N} \sum_{k=1}^N |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2 (e^{iu\beta_k} - f_\beta^*(u)) \right|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4} du \\
& \quad + \frac{4}{\pi} \int_{A(m,k)} \frac{\left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} e^{iu\beta_k} (e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2) \right|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4} du \\
& \leq \frac{8}{\pi} \int_{A(m,k)} \frac{\left| \frac{1}{N} \sum_{k=1}^N |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2 (e^{iu\beta_k} - f_\beta^*(u)) \right|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4} du \\
& \quad + \frac{4}{\pi} \int_{A(m,k)} \frac{\left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} e^{iu\beta_k} (e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2) \right|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4} du \\
& \quad + \frac{8}{\pi} \int_{A(m,k)} \left| \frac{1}{N} \sum_{k=1}^N |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2 (e^{iu\beta_k} - f_\beta^*(u)) \right|^2 \left| \frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} - \frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} \right|^2 du \\
& \leq \frac{8}{\pi} \int_{A(m,k)} \left| \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \right|^2 du \\
& \quad + \frac{4}{\pi} \int_{A(m,k)} \frac{\left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} e^{iu\beta_k} (e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2) \right|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4} du \\
& \quad + \frac{8}{\pi} \int_{A(m,k)} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4 \left| \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \right|^2 \left| \frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} - \frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} \right|^2 du.
\end{aligned}$$

We can now write the following inequalities, using (46) and the above remarks

$$\begin{aligned}
& \|\hat{f}_k - \hat{f}_m\|^2 - 4\|f_k - f_m\|^2 - \frac{1}{2}\widehat{\text{qen}}(m, k) \\
& \leq \frac{8}{\pi} \int_{A(m,k)} \left| \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \right|^2 du - \frac{1}{4}\widehat{\text{qen}}_1(m, k) \\
& \quad + \frac{4}{\pi} \int_{A(m,k)} \frac{\left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} e^{iu\beta_k} (e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2) \right|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4} du - \frac{1}{2}\widehat{\text{qen}}_2(m, k) \\
& \quad + \frac{8}{\pi} \int_{A(m,k)} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4 \left| \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \right|^2 \left| \frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} - \frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} \right|^2 du - \frac{1}{4}\widehat{\text{qen}}_1(m, k) \\
& \quad + \frac{1}{2}\widehat{\text{qen}}_3(m, k) - \frac{1}{2}\widehat{\text{qen}}_3(m, k) \\
& \quad + \frac{1}{\pi} \int_{A(m,k)} |\hat{f}_Z^*(u)|^2 \frac{\left| |\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2 - |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2 \right|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^8} \mathbf{1}_{\left\{ |\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)| \leq |f_\varepsilon^*\left(\frac{u}{\Delta}\right)| \right\}} du \mathbf{1}_{\{C(m,k)^c\}}
\end{aligned}$$

Taking expectation, we get

$$\begin{aligned}
& \mathbb{E} \left[\sup_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \left\{ \left\| \hat{f}_k - \hat{f}_m \right\|^2 - 4 \|f_k - f_m\|^2 - \frac{1}{2} \widehat{\text{qen}}(m, k) \right\}_+ \right] \\
& \leq \sum_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \mathbb{E} \left[\left\{ \left\| \hat{f}_k - \hat{f}_m \right\|^2 - 4 \|f_k - f_m\|^2 - \frac{1}{2} \widehat{\text{qen}}(m, k) \right\}_+ \right] \\
& \leq 16 \sum_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \mathbb{E} \left[\left\{ \frac{1}{2\pi} \int_{A(m,k)} \left| \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \right|^2 du - \frac{1}{64} \text{qen}_1(m, k) \right\}_+ \right] \\
& \quad + \sum_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \mathbb{E} \left[\left\{ \frac{8}{\pi} \int_{A(m,k)} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4 \left| \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \right|^2 \left| \frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} - \frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} \right|^2 du - \frac{1}{4} \text{qen}_1(m, k) \right\}_+ \right] \\
& \quad + 8 \sum_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \mathbb{E} \left[\left\{ \frac{1}{2\pi} \int_{A(m,k)} \frac{\left| \frac{2}{N(J-4)} \sum_{k=1}^N \sum_{j=3}^{J/2} e^{iu\beta_k} (e^{iu\eta_{k,j}/\Delta} - |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2) \right|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4} du - \frac{1}{16} \widehat{\text{qen}}_2(m, k) \right\}_+ \right] \\
& \quad + \frac{1}{\pi} \sum_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \mathbb{E} \left[\int_{A(m,k)} |\hat{f}_Z^*(u)|^2 \frac{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2 - |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^8} \mathbb{1}_{\{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)| \leq |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|\}} du \mathbb{1}_{\{C(m,k)^c\}} \right].
\end{aligned}$$

Now noticing that $\left| \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \right|^2$ and $\left| \frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} - \frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} \right|^2$ are independent and applying Lemma 3.3, we have

$$\begin{aligned}
& \frac{8}{\pi} \mathbb{E} \left[\int_{A(m,k)} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4 \left| \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \right|^2 \left| \frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} - \frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} \right|^2 du \right] \\
& = \frac{8}{\pi} \int_{A(m,k)} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4 \mathbb{E} \left[\left| \frac{1}{N} \sum_{k=1}^N (e^{iu\beta_k} - f_\beta^*(u)) \right|^2 \right] \mathbb{E} \left[\left| \frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} - \frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} \right|^2 \right] du \\
& \leq \frac{8}{\pi} \int_{A(m,k)} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4 \frac{1}{N} |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{-4} du \\
& \leq \frac{16(k-m)}{N} := \frac{1}{4} \text{qen}_1(m, k).
\end{aligned}$$

Lemma 6.3 implies that $\mathbb{P}[C(m, k)^c] \leq N^{-3}(k-m)^{-3}$, we then get

$$\begin{aligned}
& \mathbb{E} \left[\int_{A(m,k)} |\hat{f}_Z^*(u)|^2 \frac{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2 - |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^8} \mathbb{1}_{\{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)| \leq |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|\}} du \mathbb{1}_{\{C(m,k)^c\}} \right] \\
& \leq 4 \mathbb{E} \left[\int_{A(m,k)} |\hat{f}_Z^*(u)|^2 \frac{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}{k_N^4\left(\frac{u}{\Delta}\right)} \mathbb{1}_{\{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)| \leq |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|\}} du \mathbb{1}_{\{C(m,k)^c\}} \right] \\
& \leq 4 \mathbb{E} \left[\int_{A(m,k)} \kappa^{-4} (\log N)^{-2} w\left(\frac{u}{\Delta}\right)^4 N^2 du \mathbb{1}_{\{C(m,k)^c\}} \right] \\
& \leq 4\kappa^{-4} (\log N)^{-2} N^2 (k-m) \mathbb{P}[C(m, k)^c] \\
& \leq 4\kappa^{-4} (\log N)^{-2} N^2 (k-m) N^{-3} (k-m)^{-3} \\
& \leq 4\kappa^{-4} N^{-1} (k-m)^{-2}.
\end{aligned}$$

Finally applying Lemma 6.5 and 6.6, we have

$$\mathbb{E} \left[\sup_{\substack{k \geq m \\ k \in \mathcal{M}_n}} \left\{ \left\| \hat{f}_{\beta,k} - \hat{f}_{\beta,m} \right\|^2 - 4 \|f_{\beta,k} - f_{\beta,m}\|^2 - \frac{1}{2} \widehat{\text{qen}}(m, k) \right\}_+ \right] \leq C \left(\frac{1}{N} + \frac{2}{N(J-4)} \right).$$

This completes the proof of Proposition 6.2. \square

ACKNOWLEDGEMENTS

The author would like to thank Charlotte Dion for useful discussions on the subject and Fabienne Comte for advice and suggestions.

APPENDIX A.

We remind, for the readers convenience, some useful results.

Lemma A.1. *Let X_1, \dots, X_n be i.i.d. random variables with $\text{Var}[X_1] \leq v^2$ and suppose that almost surely $\|X_1\|_\infty \leq B$. Let $S_n = 1/n \sum_{j=1}^n (X_j - \mathbb{E}[X_1])$. Let $\mathbb{E}|S_n| \leq H$. Then*

$$\mathbb{E} \left[\left\{ |S_n|^2 - H^2 \right\}_+ \right] \leq 32 \frac{v^2}{n} \exp \left(-n \frac{H^2}{8v^2} \right) + 128 \sqrt{2} \frac{B^2}{n^2} \exp \left(-n \frac{H}{\frac{16\sqrt{2}}{3} B} \right).$$

Lemma A.2. *(Talagrand's inequality). Let I be some countable index set. For each $i \in I$, let $X_1^{(i)}, \dots, X_n^{(i)}$ be centered i.i.d. random variables, defined on the same probability space, with $\|X_1^{(i)}\| \leq B$ for some $B < \infty$. Let $v^2 := \sup_{i \in I} \text{Var} X_1$. Then for arbitrary $\epsilon > 0$, there are positive constants c_1 and $c_2 = c_2(\epsilon)$ depending only on ϵ such that for any $\kappa > 0$:*

$$\mathbb{P} \left[\left\{ \sup_{i \in I} |S_n^{(i)}| \leq (1 + \epsilon) \mathbb{E} \left[\sup_{i \in I} |S_n^{(i)}| \right] + \kappa \right\} \right] \leq 2 \exp \left(-n \left(\frac{\kappa^2}{c_1 v^2} \wedge \frac{\kappa}{c_2 B} \right) \right).$$

A proof can be found, for example, on page 170 in Massart (2003).

Next we give some technical results which will be essential for the proofs.

Lemma A.3 (Lemma 2 p.35 (Butucea and Tsybakov (2008a))). *Let γ, μ , and s be positive constants then for any $m > 0$*

$$\int_0^m (x^2 + 1)^\gamma e^{2\mu x^s} dx \approx m^{2\gamma+1-s} e^{2\mu m^s}.$$

Lemma A.4. *In the definition of \tilde{f}_ε^* , assume $\kappa > \sqrt{c_1 p}$. Let $\tau \geq 2\kappa$ and $x \geq 1$. Then for some positive constant C*

$$\mathbb{P} \left[\exists u \in \mathbb{R} : |\hat{f}_\varepsilon^*(u)^4 - f_\varepsilon^*(u)^4| > \tau (\log(Nx))^{1/2} w(u)^{-1} N^{-1/2} \right] \leq C x^{-p} N^{-p}.$$

See Lemma 5.5 in Kappus (2014) for the proof.

REFERENCES

- Birgé, L. (1999). An alternative point of view on Lepski's method. In *State of the art in probability and statistics: Festschrift for Willem R. van Zwet*, pages 113–133. IMS Lecture Notes-Monograph Series.
- Birgé, L. and Massart, P. (1997). From model selection to adaptive estimation. In *Festschrift for Lucien Le Cam*, pages 55–87. Springer.
- Butucea, C. (2004). Deconvolution of supersmooth densities with smooth noise. *The Canadian Journal of Statistics*, 32(2):181–192.
- Butucea, C. and Comte, F. (2009). Adaptive estimation of linear functionals in the convolution model and applications. *Bernoulli*, 15(1):69–98.
- Butucea, C. and Tsybakov, A. (2008a). Sharp optimality in density deconvolution with dominating bias I. *Theory Proba. Appl.*, 52(1):24–39.
- Butucea, C. and Tsybakov, A. (2008b). Sharp optimality in density deconvolution with dominating bias II. *Theory Proba. Appl.*, 52(2):237–249.
- Carroll, R. J. and Hall, P. (1988). Optimal rates of convergence for deconvolving a density. *Journal of the American Statistical Association*, 83(404):1184–1186.
- Comte, F. and Lacour, C. (2011). Data-driven density estimation in the presence of additive noise with unknown distribution. *Journal of the Royal Statistical Society: Series B*, 73:601–627.

- Comte, F., Rozenholc, Y., and Taupin, M.-L. (2006). Penalized contrast estimator for adaptive density deconvolution. *The Canadian Journal of Statistics*, 3(34):431–452.
- Comte, F., Rozenholc, Y., and Taupin, M.-L. (2007). Finite sample penalization in adaptive density deconvolution. *Journal of Statistical Computation and Simulation*, 77(11):977–1000.
- Comte, F. and Samson, A. (2012). Nonparametric estimation of random-effects densities in linear mixed-effects model. *Journal of Nonparametric Statistics*, 24(4):951–975.
- Comte, F., Samson, A., and Stirnemann, J. (2013). Deconvolution estimation of onset of pregnancy with replicate observations. To appear in *Scandinavian Journal of Statistics*.
- Dattner, I., Reiß, M., and Trabs, M. (2013). Adaptive quantile estimation in deconvolution with unknown error distribution. Preprint, arXiv:1303.1698.
- Delaigle, A. and Gijbels, I. (2004). Bootstrap bandwidth selection in kernel density estimation from a contaminated sample. *Annals of the Institute of Statistical Mathematics*, 56(1):19–47.
- Delaigle, A., Hall, P., and Meister, A. (2008). On deconvolution with repeated measurements. *The Annals of Statistics*, 36(2):665–685.
- Dion, C. (2013). New strategies for nonparametric estimation in a linear mixed model. Preprint, hal-00906379.
- Efromovich, S. (1997). Density estimation for the case of supersmooth measurement errors. *Journal of the American Statistical Association*, 92:526–535.
- Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problems. *The Annals of Statistics*, 19(3):1257–1272.
- Johannes, J. (2009). Deconvolution with unknown error distribution. *The Annals of Statistics*, 37(5a):2301–2323.
- Johannes, J. and Schwarz, M. (2012). Adaptive circular deconvolution by model selection under unknown error distribution. arXiv:0912.1207v2. To appear in *Bernoulli*.
- Kappus, J. (2014). Adaptive nonparametric estimation for Lévy processes observed at low frequency. *Stochastic Processes and their Applications*, 124:730–758.
- Kappus, J. and Mabon, G. (2013). Adaptive density estimation in deconvolution problems with unknown error distribution. Preprint, hal-00915982.
- Lacour, C. (2006). Rates of convergence for nonparametric deconvolution. *C. R. Acad. Sci. Paris, Ser. I* 342(11):877–882.
- Li, T. and Vuong, Q. (1998). Nonparametric estimation of the measurement error model using multiple indicators. *Journal of Multivariate Analysis*, 65:139–165.
- Massart, P. (2003). *Concentration Inequalities and Model selection*, volume 1896 of *Lecture Notes in Mathematics*. Springer, Berlin. Lectures from the 33rd Summer School on Probability Theory held in Saint-Flour, July 6–23, 2003, With a foreword by Jean Picard.
- Meister, A. (2009). *Deconvolution Problems in Nonparametric Statistics*, volume 193 of *Lecture Notes in Statistics*. Springer-Verlag, Berlin.
- Neumann, M. (2007). Deconvolution from panel data with unknown error distribution. *Journal of Multivariate Analysis*, 98(10):1955–1968.
- Neumann, M. and Reiß, M. (2009). Nonparametric estimation for Lévy processes from low frequency observations. *Bernoulli*, 15(1):223–248.
- Neumann, M. H. (1997). On the effect of estimating the error density in nonparametric deconvolution. *Journal of Nonparametric Statistics*, 7(4):307–330.
- Pensky, M. and Vidakovic, B. (1999). Adaptive wavelet estimator for nonparametric density deconvolution. *The Annals of Statistics*, 27(6):2033–2053.
- Pinheiro, J. and Bates, D. (2000). *Mixed-effect models in S and Splus*. Statistics and Computing Series. Springer-Verlag, New York, NY.
- Shen, W. and Louis, T. (1999). Empirical bayes estimation via the smoothing by roughening approach. *J. Comput. Graph. Stat.*, 8(4):800–823.
- Stefanski, L. (1990). Rates of convergence of some estimators in a class of deconvolution problems. *Statistics and Probability Letters*, 9:229–235.
- Stefanski, S. and Carroll, R. (1990). Deconvoluting kernel density estimators. *Statistics*, 21:169–184.
- Vock, D., Davidian, M., Tsiatis, A., and Muir, A. (2011). Mixed model analysis of censored longitudinal data with flexible random-effects density. *Biostatistics*, 13:61–73.
- Wu, P. and Zhu, L. (2010). An orthogonality-based estimation of moments for linear mixed models. *Scand. J. Stat.*, 37:253–263.
- Zhang, D. and Davidian, M. (2001). Linear mixed models with flexible distributions of random effects for longitudinal data. *Biometrics*, 57(3):795–802.