Série des Documents de Travail

n° 2014-41

Adaptive Estimation of Random-Effects Densities In Linear Mixed-Effects Model

Gwennaëlle. MABON [1]

[1] CREST, Malakoff, France. E-mail: Gwennaelle.Mabon@ensae-paristech.fr

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

G. MABON

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 $k = 1, \ldots, N$ and $j = 1, \ldots, J$. Random variables (α_k, β_k) are called 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 of f_α and f_β and present upper risk bounds. We also compute examples of rates of convergence. The focus of this work lies on the near optimal data driven choice of the smoothing parameter using a penalization strategy in the particular case of fixed interval between times t_j. Risk bounds for the adaptive estimators of f_α and f_β are provided. Simulations illustrate the relevance of the methodology.

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

1. Introduction

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

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

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. The random variables $(\varepsilon_{k,j})$ represent 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 $(t_j)_{1\le j\le J}$ 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 nonparametric 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. Some approaches are based on normal mixtures, see Ghidey et al. (2004), Komárek and Lesaffre (2008). For nonparametric approach, we can cite Claeskens and Hart (2009) who develop a nonparametric goodness-of-fit test in mixed models providing a nonparametric estimator if the normality hypothesis is rejected. Mixed models are also studied in Bayesian litterature, see Ibrahim and Kleinman (1998) who allow the prior to be nonparametric by taking a Dirichlet process.

Here we consider an approach based on deconvolution methods. The convolution model is a classical setting in nonparametric statistics which has been widely studied. There exists a large amount of literature on the subject assuming first that the noise density is known. We can cite Carroll and Hall (1988), Stefanski

CREST-ENSAE, *3 avenue Pierre Larousse, 92245 Malakoff, France*

MAP5, Université Paris Descartes, *45 rue des Saints-Pères, 75006 Paris, France*

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

Date: November 17, 2014.

(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. Concerning studies of rate optimality in the minimax sense, we refer to Butucea (2004) and Butucea and Tsybakov (2008a,b). Yet the drawback of these methods is that they all work under the assumption that the error distribution is known. We can also cite, in the known error case, Dion (2013) who study nonparametric estimators based on Lepski's method. 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 (2013), 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 presence on repeated measurements. 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). More recently, Kappus and Mabon (2013) have achieve a new adaptive procedure in this setting. 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.

Model (1) has been considered by Comte and Samson (2012). Let us emphasize the novelty of our paper. Comte and Samson (2012) derive theoretical properties of the nonparametric estimators of f_α and f_β only when the error distribution is known. In that particular case they establish oracle inequalities which ensure that their method is adaptive. Since assuming that the noise is known is not realistic, they also define an estimator when it is not. They prove an upper bound for its risk for fixed smoothing parameter. Then they propose an adaptive strategy which is implemented but not studied from a theoretical point of view. Here, we modify the procedure, improve the upper bound. Moreover we prove oracle risk bounds for the adaptive estimators of f_α and f_β in presence of unknown noise. These results are difficult and new. We also derive the rates of convergence for f_β.

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 along with upper bounds for both densities and rates of convergence for f_β. In Section 3, 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 4, we lead a study of the adaptive estimators through simulation experiments. Numerical results are then presented. All the proofs are postponed to Section 5.

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 $\varphi, \psi : \mathbb{R} \to \mathbb{C}$ 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 $\|\varphi\|^2 = \int_{\mathbb{R}} |\varphi(x)|^2 \,\mathrm{d}x$, $\langle \varphi, \psi \rangle$ the scalar product between φ and ψ defined by $\langle \varphi, \psi \rangle = \int_{\mathbb{R}} \varphi(x)\overline{\psi(x)}\,\mathrm{d}x$. The Fourier transform φ^* is defined by $\varphi^*(x) = \int e^{ixu}\varphi(u)\,\mathrm{d}u$. 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 $\varphi(x) = 1/(2\pi) \int e^{-ixu}\varphi^*(u)\,\mathrm{d}u$. Lastly the convolution product $*$ is defined as $(\varphi*\psi)(x) = \int \varphi(x-u)\psi(u)\,\mathrm{d}u$.

2.2. **Estimation of f_α.** If observations for $t_0 = 0$ are available, then we can write Model (1) as follows

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

This is a classical deconvolution model in the context of unknown measurement errors (see references above).

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 $f_\alpha^* = f_Y^*/f_\varepsilon^*$. In this case, we have that

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

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 $\mathbb{R}$. 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 $1/(2\pi)\int_{|u|\leq \pi m} e^{-iux}\hat{f}^*_Y(u)/f^*_\varepsilon(u)\,\mathrm{d}u$.

In this paper the error distribution is assumed to be unknown. To make the problem identifiable, some additional information on the noise is required. Model (2) can be seen as a repeated observation model. 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 the following equality under Assumption (**A2**)

$$f^*_U(x) = \mathbb{E}\left[e^{ixU}\right] = |f^*_\varepsilon(x)|^4 .$$

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

(A5) ε is symmetric.

This latest assumption together with (**A2**) implies that f^*_ε is real-valued and positive, so that $f^*_U(x) = (f^*_\varepsilon(x))^4$. As a consequence f^{*4}_ε can be estimated as follows

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

Nevertheless we need to prevent $\widehat{f^{*4}_\varepsilon}$ 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

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

where $s_N(x)\geq 1$ will be defined later. Now we can introduce another estimator of f^*_ε defined by

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

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

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

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

We can state the following upper bound on the $\mathbb{L}^2$ risk for $\hat{f}_{\beta,m}$.

Proposition 2.1. *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*

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

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

The first two terms of the right-hand side of Equation (8) 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. This upper bound is smaller than the upper bound in Comte and Samson (2012).

2.3. **Estimation of f_β.** First let us define the 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

$$Z_{k,j} = \frac{Y_{k,2j} - Y_{k,2j-1}}{\Delta} = \beta_k + \frac{\varepsilon_{k,2j} - \varepsilon_{k,2j-1}}{\Delta} = \beta_k + \frac{\eta_{k,j}}{\Delta}, \quad \eta_{k,j} = \varepsilon_{k,2j} - \varepsilon_{k,2j-1}. \tag{9}$$

Let us notice that for a fixed j, the $(Z_{k,j})_k$ for $k = 1,\ldots,N$ 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 Assumption (**A4**), we can write the following equality $f_{Z_j} = f_\beta * f_{\eta_{k,j}/\Delta}$ which clearly implies $f^*_{Z_j}(x) = f^*_\beta(x)\left|f^*_\varepsilon\left(\frac{x}{\Delta}\right)\right|^2$. Therefore under Assumption **(A2)** we have $f^*_\beta(x) = f^*_{Z_j}(x)/\left|f^*_\varepsilon\left(\frac{x}{\Delta}\right)\right|^2$. 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

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

Then we have

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

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

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

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

$$|\tilde{f}^*_\varepsilon\left(\tfrac{x}{\Delta}\right)|^2 = \begin{cases} |\hat{f}^*_\varepsilon\left(\frac{x}{\Delta}\right)|^2 = \left(\widehat{|f^{*4}_\varepsilon|}\left(\frac{x}{\Delta}\right)\right)^{1/2} & \text{if} \quad \widehat{|f^{*4}_\varepsilon|}\left(\frac{x}{\Delta}\right) \geq k_N\left(\frac{x}{\Delta}\right), \\ \left(k_N\left(\frac{x}{\Delta}\right)\right)^{1/2} & \text{otherwise.} \end{cases} \tag{10}$$

Thus we can estimate f^*_β as follows

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

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

$$f_{\beta,m}(x) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m} e^{-ixu}\frac{2}{J-4}\sum_{j=3}^{J/2}\frac{f^*_{Z_j}(u)}{|f^*_\varepsilon\left(\frac{u}{\Delta}\right)|^2}\,\mathrm{d}u. \tag{11}$$

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

$$\hat{f}_{\beta,m}(x) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m} e^{-ixu}\frac{2}{J-4}\sum_{j=3}^{J/2}\frac{\hat{f}^*_{Z_j}(u)}{|\tilde{f}^*_\varepsilon\left(\frac{u}{\Delta}\right)|^2}\,\mathrm{d}u = \frac{1}{2\pi}\int_{-\pi m}^{\pi m} e^{-ixu}\frac{\hat{f}^*_Z(u)}{|\tilde{f}^*_\varepsilon\left(\frac{u}{\Delta}\right)|^2}\,\mathrm{d}u \tag{12}$$

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

Thus we can state the following upper bound on the $\mathbb{L}^2$ risk for $\hat{f}_{\beta,m}$.

Proposition 2.2. *Under Assumptions **(A1)**-**(A4)**, for $k_N(x)$ defined by (5) assume that $s_N(x) = 1$, then $\hat{f}_{\beta,m}$ defined by (12) satisfies*

$$\mathbb{E}\left\|f_\beta - \hat{f}_{\beta,m}\right\|^2 \leq \|f_\beta - f_{\beta,m}\|^2 + \frac{6m}{N} + \frac{12}{N(J-4)}\frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{\mathrm{d}u}{\left|f^*_\varepsilon\left(\frac{u}{\Delta}\right)\right|^4} + \frac{4C_1}{N}\frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{|f^*_\beta(u)|^2}{\left|f^*_\varepsilon\left(\frac{u}{\Delta}\right)\right|^8}\,\mathrm{d}u. \tag{13}$$

where $f_{\beta,m}$ is defined by (11) and C_1 is defined in Lemma B.2.

The terms of the right-hand side of Equation (13) correspond to a squared bias ($\|f_\beta - f_{\beta,m}\|^2$) variance decomposition. Compared to Comte and Samson (2012), this inequality differs from the last term which is smaller than theirs. 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.

2.3.1. *Discussion about resulting rates.* In order to derive the corresponding rates of convergence of the estimator of f_β defined by (12), 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

$$\mathcal{A}(a,r,L) = \left\{ f \in \mathbb{L}^1 \cap \mathbb{L}^2(\mathbb{R}),\ \int |f^*(u)|^2 e^{2a|u|^r}\,\mathrm{d}u \le L \right\} \tag{14}$$

$$\mathcal{S}(\delta,L) = \left\{ f \in \mathbb{L}^1 \cap \mathbb{L}^2(\mathbb{R}),\ \int |f^*(u)|^2 (1+u^2)^\delta\,\mathrm{d}u \le L \right\} \tag{15}$$

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

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

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

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

To derive the order of the bound on the variance in Equation (13), we need more information about the regularity of f_ε. We add the following classical assumption:

There exist positive constants k_0, k_0', γ, μ, and s such that for any real x

$$k_0(x^2+1)^{-\gamma/2} e^{-\mu|x|^s} \le |f_\varepsilon^*(x)| \le k_0'(x^2+1)^{-\gamma/2} e^{-\mu|x|^s}. \tag{16}$$

- We say that $f_\varepsilon^* \in \mathrm{OS}(\gamma)$ (for *ordinary smooth*), if f_ε^* satisfies (16) with $\gamma = 0$, $s > 0$ and $\mu > 0$.
- We say that $\gamma = 0$ $f_\varepsilon^* \in \mathrm{SS}(s)$ (for *supersmooth*, if f_ε^* satisfies (16) with $s = 0$.

Proposition 2.3. *If f_ε^* satisfies (16) and $f_\beta \in \mathcal{S}(\delta,L)$ defined by (14), then*

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

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

Proposition 2.4. *Suppose that $f_\varepsilon^* \in \mathrm{OS}(\gamma)$ and $f_\beta \in \mathcal{A}(a,r,L)$ defined by (15), then*

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

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. Increasing J may improve the rates.

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

TABLE 1. Rates of convergence for the MISE

The idea now for the adaptive estimation is to find a penalty term which 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. In particular, we will not assume that the characteristic function of the error distribution has a particular shape (ordinary smooth or supersmooth). In other words, Assumption **(A6)** is not necessary to derive an adaptive procedure..

3. Model selection

In this section, we introduce an adaptive estimator of $|f_\varepsilon^*|^2$ which can be uniformly controlled on the real line. This brings a model selection procedure with very weak assumptions on the error distribution f_ε^*. For that we need to choose a new and adequate $s_N(x)$, larger than previously, 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). This function is of high importance because empirical estimators of the Fourier transform will appear in the penalty terms more precisely in the denominator. Thus oracle bounds depend on a control of the deviation of $\widehat{f_\varepsilon^{*4}}$ from f_ε^{*4} which is possible when weighted by w as shown in Neumann and Reiß (2009). As from now, we set the threshold k_N defined by (5) as follows

$$s_N(x) = \kappa\,(\log N)^{1/2}\,w(x)^{-1} \tag{17}$$

where κ is a positive universal constant.

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

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

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 (13). Here the penalty term is stochastic since the variance terms depend on the error distribution which is supposed unknown.

3.1. 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 is their paper. The main 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 here f_ε^{*4} can be estimated. 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

$$\Theta(m) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m} \frac{w(u)^{-2}}{|f_\varepsilon^*(u)|^2}\,\mathrm{d}u \quad \text{and} \quad \Theta^\alpha(m) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m} \frac{w(u)^{-2}|f_\alpha^*(u)|^2}{|f_\varepsilon^*(u)|^8}\,\mathrm{d}u$$

$$\hat{\Theta}(m) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m} \frac{w(u)^{-2}}{|\check{f}_\varepsilon^*(u)|^2}\,\mathrm{d}u \quad \text{and} \quad \hat{\Theta}^\alpha(m) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m} \frac{w(u)^{-2}|\hat{f}_Y^*(u)|^2}{|\check{f}_\varepsilon^*(u)|^{10}}\,\mathrm{d}u.$$

These terms correspond to the deterministic and stochastic bounds on the variance appearing in Equation (8). The difference lies in the introduction of the function w essential for the adaptive procedure. Thus, we define an empirical penalty as

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

where κ is the same as in Equation (17), and a deterministic penalty

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

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

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

$$\hat{m} = \underset{m\in\mathcal{M}_N}{\operatorname{argmin}} \left\{ -\|\hat{f}_{\alpha,m}\|^2 + \widehat{\mathrm{pen}}(m) \right\} \tag{18}$$

where $\mathcal{M}_N = \{1,\ldots,N\}$. We can now state the following oracle inequality:

Theorem 3.1. *Under Assumptions **(A1)-(A4)**, let $\hat{f}_{\hat{m},\alpha}$ be defined by (7) and (18). Then there are positive constants C^{ad} and C such that*

$$\mathbb{E}\|f_\alpha - \hat{f}_{\alpha,\hat{m}}\|^2 \leq C^{ad} \inf_{m\in\mathcal{M}_N} \left\{\|f_\alpha - f_{\alpha,m}\|^2 + \mathrm{pen}(m)\right\} + \frac{C}{N}. \tag{19}$$

The latest result is an oracle inequality which means that the squared bias variance compromise is automatically made and completely data driven in a non-asymptotic setting. So rates of convergence are reached 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 2.3.1).

However the penalty term is not exactly the same as the upper bound terms shown in Equation (8). We may wonder if a loss due to adaptation occurs. To answer that question, right-hand side of Equations (8) and (19) have to be compared. More precisely, since the squared bias term $\|f_\alpha - f_{\alpha,m}\|^2$ is unchanged and N^{-1} is a negligible term, it comes down to compare $\mathrm{pen}(m)$ with

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

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

3.2. **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 unknown-error case, with additional w function.

$$\Xi(m) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{w\left(\frac{u}{\Delta}\right)^{-2}}{\left|f_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^4}\,\mathrm{d}u \quad \text{and} \quad \Xi^\beta(m) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{w\left(\frac{u}{\Delta}\right)^{-2}|f_\beta^*(u)|^2}{\left|f_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^8}\,\mathrm{d}u.$$

$$\hat{\Xi}(m) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{w\left(\frac{u}{\Delta}\right)^{-2}}{\left|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^4}\,\mathrm{d}u \quad \text{and} \quad \hat{\Xi}^\beta(m) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{w\left(\frac{u}{\Delta}\right)^{-2}|\hat{f}_Z^*(u)|^2}{\left|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^{12}}\,\mathrm{d}u.$$

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

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

and the deterministic penalty

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

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

It is worth mentioning that the penalty takes into account the three terms of variance showed in Equation (13). So the penalty has the same order as the bounds on the variance. Therefore we select the cutoff parameter $\hat{m}$ as a minimizer of the following penalized criterion

$$\hat{m} = \underset{m\in\mathcal{M}_N}{\mathrm{argmin}}\left\{-\left\|\hat{f}_{\beta,m}\right\|^2 + \widehat{\mathrm{qen}}(m)\right\}. \tag{20}$$

Theorem 3.2. *Under Assumptions **(A1)-(A4)**, consider $\hat{f}_{\beta,\hat{m}}$ defined by (12) and (20). Then there are positive constants C^{ad} and C such that*

$$\mathbb{E}\left\|f_\beta - \hat{f}_{\beta,\hat{m}}\right\|^2 \leq C^{ad}\inf_{m\in\mathcal{M}_N}\left\{\|f_\beta - f_{\beta,m}\|^2 + \mathrm{qen}(m)\right\} + \frac{C}{N(J-4)} + \frac{C}{N}. \tag{21}$$

The same kind of remarks as in Theorem 3.1 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 automatically without being specified in the framework. As far as we know this result is new in the literature.

4. 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 $-\|\hat{f}_{\beta,m}\|^2 + \widehat{\mathrm{qen}}(m)$.
- ▷ Choose $\hat{m}$ such as $\hat{m} = \underset{m \in \mathcal{M}_N}{\operatorname{argmin}} \left\{ -\|\hat{f}_{\beta,m}\|^2 + \widehat{\mathrm{qen}}(m) \right\}$.
- ▷ And compute $\hat{f}_{\beta,\hat{m}}(x) = \displaystyle\int_{-\pi\hat{m}}^{\pi\hat{m}} e^{-ixu} \frac{\hat{f}_Z^*(u)}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} du$.

Riemann's sums are used to approximate all the integrals. The penalties are chosen according to Theorem 3.2 and as in Comte et al. (2007) we consider that m can be fractional by taking the following model collection $\mathcal{M}_N = \{m = k/10, \quad 1 \le k \le 25\}$ associated with the following penalty

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

Moreover the times 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, f_α is a standard Gaussian distribution.

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 $f_\varepsilon^* \in \mathrm{SS}(2)$. The second one is a Laplace density (*ordinary smooth* density) which means $f_\varepsilon^* \in \mathrm{OS}(2)$.

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

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

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 larger. 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 ISE. 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.

Comparison with existing results. The results of the standard Gaussian distribution, Cauchy distribution and Gamma distribution can be compared to those of Kappus and Mabon (2013). We estimate the same functions but in a more difficult context. In fact we have access to an estimate of $(f_\varepsilon^*)^4$ when they have access to an estimate of f_ε^* or $(f_\varepsilon^*)^2$. That makes the problem tougher since the information about f_ε^* is less precise. Nevertheless, if we compare to Kappus and Mabon (2013) the results are very close which is quite remarkable.

Now if we compare our results to those of Comte and Samson (2012), we must compare: standard Gaussian distribution and Gamma distribution. We have exactly the same model. We can point out that our results are twice better when the error distribution is Gaussian. When the error distribution is a Laplace the results are similar. So our methodology can handle more cases (supersmooth error) than Comte and Samson (2012) and give very good results of estimation.

		$\sigma_\varepsilon^2 = \frac{1}{10}$		$\sigma_\varepsilon^2 = \frac{1}{4}$	
	N	200	2000	200	2000
Gaussian	f_ε known	0.344	0.054	0.514	0.129
		(0.273)	(0.045)	(0.513)	(0.105)
	f_ε unknown	0.331	0.042	0.317	0.057
		(0.251)	(0.033)	(0.236)	(0.051)
Cauchy	f_ε known	0.625	0.105	0.804	0.216
		(0.573)	(0.097)	(0.765)	(0.211)
	f_ε unknown	0.507	0.075	0.657	0.090
		(0.427)	(0.071)	(0.599)	(0.079)
Gamma	f_ε known	0.398	0.069	0.620	0.161
		(0.360)	(0.063)	(0.517)	(0.140)
	f_ε unknown	0.381	0.051	0.506	0.066
		(0.347)	(0.044)	(0.430)	(0.051)
Mixed Gamma	f_ε known	0.545	0.095	0.715	0.150
		(0.480)	(0.084)	(0.620)	(0.141)
	f_ε unknown	0.506	0.082	0.518	0.092
		(0.453)	(0.080)	(0.495)	(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 ISE also averaged over 100 samples with a Laplace noise.

5. Proofs

5.1. **Sketch of the proof of Proposition 2.2.** The proof follows the same lines as Proposition 5.2 in Comte and Samson (2012). The difference is that we apply our Lemma B.2 instead.

5.2. **Proof of Theorem 3.1.** The proof is similar to the proof of Theorem 3.2. See also Kappus and Mabon (2013). Thus the proof is omitted.

5.3. **Proof of Theorem 3.2.** Let us introduce some notations: for $k > m$,

$$\hat{\Xi}(m,k) = \hat{\Xi}(k) - \hat{\Xi}(m), \quad \hat{\Xi}^\beta(m,k) = \hat{\Xi}^\beta(k) - \hat{\Xi}^\beta(m).$$

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 3.2. We denote by $m^\star$ the oracle cutoff defined by

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

		$\sigma_\varepsilon^2 = \frac{1}{10}$		$\sigma_\varepsilon^2 = \frac{1}{4}$	
	N	200	2000	200	2000
Gaussian	f_ε known	0.349	0.054	0.650	0.142
		(0.296)	(0.045)	(0.598)	(0.128)
	f_ε unknown	0.285	0.038	0.349	0.052
		(0.239)	(0.029)	(0.273)	(0.053)
Cauchy	f_ε known	0.588	0.119	0.848	0.272
		(0.532)	(0.112)	(0.791)	(0.263)
	f_ε unknown	0.481	0.076	0.680	0.089
		(0.449)	(0.070)	(0.607)	(0.084)
Gamma	f_ε known	0.401	0.072	0.956	0.207
		(0.332)	(0.063)	(0.910)	(0.190)
	f_ε unknown	0.402	0.049	0.461	0.067
		(0.316)	(0.041)	(0.418)	(0.055)
Mixed Gamma	f_ε known	0.504	0.089	0.704	0.163
		(0.454)	(0.084)	(0.639)	(0.146)
	f_ε unknown	0.504	0.079	0.552	0.101
		(0.446)	(0.072)	(0.484)	(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 ISE also averaged over 100 samples with a Gaussian noise.

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

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

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

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

$$-\left\|\hat{f}_{\beta,\hat{m}}\right\|^2 + \widehat{\text{qen}}(\hat{m}) \leq -\left\|\hat{f}_{\beta,m^\star}\right\|^2 + \widehat{\text{qen}}(m^\star) \tag{22}$$

which implies

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

Thus

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

Taking expectation, we apply the following Lemma proved in Section 5.4.

Lemma 5.1. *There is a positive constant C such that for any arbitrary $m \in \mathcal{M}_N$*

$$\mathbb{E}\left[\widehat{\text{qen}}(m)\right] \leq C\text{qen}(m). \tag{23}$$

It yields for some positive constant C

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

We just proved the desired result on G as $\|f_\beta - f_{\beta,m^\star}\|^2 = \|f_\beta\|^2 - \|f_{\beta,m^\star}\|^2$ and using the definition of $m^\star$

$$\mathbb{E}\left[\|f_\beta - \hat{f}_{\beta,\hat{m}}\|^2 \mathbb{1}_G\right] \leq C \inf_{m \in \mathcal{M}_N}\left\{\|f_\beta - f_{\beta,m}\|^2 + \text{qen}(m)\right\}. \tag{24}$$

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

$$\begin{aligned}
\left\|\hat{f}_{\beta,\hat{m}} - \hat{f}_{\beta,m^\star}\right\|^2 \mathbb{1}_{G^c} &= \left(\left\|\hat{f}_{\beta,\hat{m}} - \hat{f}_{\beta,m^\star}\right\|^2 - 4\left\|f_{\beta,\hat{m}} - f_{\beta,m^\star}\right\|^2 - \frac{1}{2}\widehat{\text{qen}}(m^\star,\hat{m})\right)\mathbb{1}_{G^c} \\
&\quad + \left(4\left\|f_{\beta,\hat{m}} - f_{\beta,m^\star}\right\|^2 + \frac{1}{2}\widehat{\text{qen}}(m^\star,\hat{m})\right)\mathbb{1}_{G^c} \\
&\leq \sup_{\substack{k\geq m^\star \\ k\in\mathcal{M}_N}} \left\{\left\|\hat{f}_{\beta,k} - \hat{f}_{\beta,m^\star}\right\|^2 - 4\left\|f_{\beta,k} - f_{\beta,m^\star}\right\|^2 - \frac{1}{2}\widehat{\text{qen}}(m^\star,k)\right\}_+ \\
&\quad + 4\|f_{\beta,\hat{m}} - f_{\beta,m^\star}\|^2 + \frac{1}{2}\sum_{\substack{k\geq m^\star \\ k\in\mathcal{M}_N}} \widehat{\text{qen}}(m^*,k)\mathbb{1}_{\{\hat{m}=k\}}. \qquad (25)
\end{aligned}$$

Let us first notice the following inequality

$$\forall k > m, \quad \widehat{\text{qen}}(m,k) \leq \widehat{\text{qen}}(k). \qquad (26)$$

Besides by definition of $\hat{m}$ (see Equation (20)), on the set $\{\hat{m} = k\} \cap G^c$ and applying Equation (22), one has

$$\begin{aligned}
&\frac{1}{2}\left(\widehat{\text{qen}}(k) - \widehat{\text{qen}}(m^\star)\right) \leq \left\|\hat{f}_{\beta,\hat{m}} - \hat{f}_{m^\star}\right\|^2 - \frac{1}{2}\widehat{\text{qen}}(k) + \frac{1}{2}\widehat{\text{qen}}(m^\star) \\
&\frac{1}{2}\widehat{\text{qen}}(k) \leq \left\|\hat{f}_{\beta,\hat{m}} - \hat{f}_{\beta,m^\star}\right\|^2 - \frac{1}{2}\widehat{\text{qen}}(m^\star,k) + \frac{1}{2}\widehat{\text{qen}}(m^\star) \\
&\qquad \leq \left(\left\|\hat{f}_{\beta,\hat{m}} - \hat{f}_{\beta,m^\star}\right\|^2 - 4\left\|f_{\beta,\hat{m}} - f_{\beta,m^\star}\right\|^2 - \frac{1}{2}\widehat{\text{qen}}(m^\star,k)\right) + 4\left\|f_{\beta,\hat{m}} - f_{\beta,m^\star}\right\|^2 + \frac{1}{2}\widehat{\text{qen}}(m^\star). \qquad (27)
\end{aligned}$$

Now using Equations (26) and (27)

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

From Equation (25), we now have

$$\begin{aligned}
\left\|\hat{f}_{\beta,\hat{m}} - \hat{f}_{\beta,m^\star}\right\|^2 \mathbb{1}_{G^c} &\leq 2\sup_{\substack{k\geq m^\star \\ k\in\mathcal{M}_N}} \left\{\left\|\hat{f}_{\beta,k} - \hat{f}_{\beta,m^\star}\right\|^2 - 4\left\|f_{\beta,k} - f_{\beta,m^\star}\right\|^2 - \frac{1}{2}\widehat{\text{qen}}(m^\star,k)\right\}_+ \\
&\quad + 8\left\|f_{\beta,\hat{m}} - f_{\beta,m^\star}\right\|^2 + \frac{1}{2}\widehat{\text{qen}}(m^\star). \qquad (28)
\end{aligned}$$

Taking expectation the first summand is negligible by applying the following Proposition proved in Section 5.5.

Proposition 5.2. *There is a positive constant C such that for any arbitrary $m \in \mathcal{M}_N$*

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

Finally we have $\mathbb{E}\left[\left\|f_\beta - \hat{f}_{\beta,\hat{m}}\right\|^2 \mathbb{1}_{G^c}\right] \leq C\left(\left\|f_\beta - f_{\beta,m^\star}\right\|^2 + \text{qen}(m^\star)\right) + \dfrac{C'}{N(J-4)} + \dfrac{C'}{N}$. This combining with (24) complete the proof. □

5.4. **Proof of Lemma 5.1.** There is nothing to prove for $\widehat{\text{qen}}_1(m)$.

• Consider $\widehat{\text{qen}}_2(m)$. For $q = 1/2$ or 1, using Cauchy-Schwarz's inequality, we have

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

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

$$\leq \frac{2}{\pi^2}\left(\int_{-\pi m}^{\pi m} w\left(\frac{u}{\Delta}\right)^{-2}\left|\frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} - \frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2}\right|^2 \mathrm{d}u\right)^2 + 8\Xi^2(m).$$

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

$$\mathbb{E}\left(\int_{-\pi m}^{\pi m} w\left(\frac{u}{\Delta}\right)^{-2}\left|\frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} - \frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2}\right|^2 \mathrm{d}u\right)^2$$

$$= \int_{-\pi m}^{\pi m}\int_{-\pi m}^{\pi m} w\left(\frac{u}{\Delta}\right)^{-2} w\left(\frac{v}{\Delta}\right)^{-2}\mathbb{E}\left[\left|\frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} - \frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2}\right|^2\left|\frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{v}{\Delta}\right)|^2} - \frac{1}{|f_\varepsilon^*\left(\frac{v}{\Delta}\right)|^2}\right|^2\right] \mathrm{d}u\,\mathrm{d}v.$$

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

$$\mathbb{E}\left[\left|\frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} - \frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2}\right|^2\left|\frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{v}{\Delta}\right)|^2} - \frac{1}{|f_\varepsilon^*\left(\frac{v}{\Delta}\right)|^2}\right|^2\right]$$

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

So we have

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

Besides we have

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

Once again applying Lemma B.2 for $p = 1$

$$\mathbb{E}\left[\int_{-\pi m}^{\pi m} w\left(\frac{u}{\Delta}\right)^{-2}\left|\frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2} - \frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2}\right|^2 \mathrm{d}u\right]$$

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

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)

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

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

Finally $\mathbb{E}[\widehat{\mathrm{qen}}_2(m)] \leq C\mathrm{qen}_2(m)$.

• Consider now $\widehat{\mathrm{qen}}_3(m)$. Another application of Lemma B.2 yields

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

Let us notice that

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

hence

$$\hat{f}_Z^*(u)-f_Z^*(u)=\frac{1}{N}\sum_{k=1}^N\left(e^{iu\beta_k}-f_\beta^*(u)\right)|f_\varepsilon^*\left(\tfrac{u}{\Delta}\right)|^2+\frac{2}{N(J-4)}\sum_{k=1}^N\sum_{j=3}^{J/2}\left(e^{iu\beta_k}\left(e^{iu\eta_{k,j}/\Delta}-|f_\varepsilon^*\left(\tfrac{u}{\Delta}\right)|^2\right)\right). \tag{29}$$

Then

$$\begin{aligned}
&\frac{1}{N}\mathbb{E}\left[\frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{w\left(\frac{u}{\Delta}\right)^{-2}|\hat{f}_Z^*(u)-f_Z^*(u)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{12}}\mathrm{d}u\right]\\
&\leq \frac{2}{N}\mathbb{E}\left[\frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{w\left(\frac{u}{\Delta}\right)^{-2}|\hat{f}_\beta^*(u)-f_\beta^*(u)|^2|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{12}}\mathrm{d}u\right]\\
&\quad+\frac{2}{N}\mathbb{E}\left[\frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{w\left(\frac{u}{\Delta}\right)^{-2}|\frac{2}{N(J-4)}\sum_{k=1}^N\sum_{j=3}^{J/2}\left(e^{iu\beta_k}\left(e^{iu\eta_{k,j}/\Delta}-|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2\right)\right)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{12}}\mathrm{d}u\right].
\end{aligned}\tag{30}$$

Let us consider the first term on the right-hand side of Equation (30). 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 $\hat{f}_\beta^*$ and $\tilde{f}_\varepsilon^*$ to find

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

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

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

Applying Lemma B.2 for $p=1$, we get

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

which means that there exists C such that

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

Now let us consider the second term on the right-hand side of Equation (30) and let us notice that

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

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 $j\neq j'$, $\mathbb{E}\left[\left(e^{iu\eta_{k,j}/\Delta}-|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2\right)\left(e^{iu\eta_{k',j'}/\Delta}-|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2\right)\right]=0$, then

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

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 B.2 for $p=1$, we have

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

This completes the proof. □

5.5. **Proof of Proposition 5.2.** Before proving Proposition 5.2, 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 5.3. *For an estimator of $\tilde{f}_\varepsilon^*$ defined by (10), assume $\kappa>\sqrt{c_1p}$. Let $\tau\geq 2\kappa$ and $x\geq 1$. Then for some positive constant C*

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

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.3, we have

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

□

Lemma 5.4. *In the situation of the preceding Lemma*

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

Proof. This is a direct consequence of Lemma 5.3 using the fact that for $x, y \geq 0$, $|\sqrt{x} - \sqrt{y}| \leq \frac{|x-y|}{2\sqrt{x \wedge y}}$ holds. □

Lemma 5.5. *There is a positive constant C such that for any arbitrary $m \in \mathcal{M}_N$*

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

Proof.

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

We then study the following empirical process

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

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

$$\begin{aligned}
&\sup_{t \in S(m,k)} |\nu_N(t)|^2 = \left| \int_{A(m,k)} \frac{1}{2\pi} \overline{t^*(u)} \left(\frac{1}{N} \sum_{k=1}^N e^{iu\beta_k} - f_\beta^*(u) \right) \mathrm{d}u \right|^2 \\
&\leq \sup_{t \in S(m,k)} \frac{1}{4\pi^2} \int_{A(m,k)} |t^*(u)|^2 \, \mathrm{d}u \int_{A(m,k)} \left| \frac{1}{N} \sum_{k=1}^N e^{iu\beta_k} - f_\beta^*(u) \right|^2 \mathrm{d}u \leq \frac{1}{2\pi} \int_{A(m,k)} \left| \frac{1}{N} \sum_{k=1}^N e^{iu\beta_k} - f_\beta^*(u) \right|^2 \mathrm{d}u,
\end{aligned}$$

hence

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

For the variance term, we have $\mathbb{V}\mathrm{ar}\left[\left| \frac{1}{2\pi} \int_{A(m,k)} \overline{t^*(u)} e^{iu\beta_1} \, \mathrm{d}u \right| \right] \leq \mathbb{E}\left[\left| \frac{1}{2\pi} \int_{A(m,k)} \overline{t^*(u)} e^{iu\beta_1} \, \mathrm{d}u \right|^2 \right]$. Let us notice that the expectation can be rewritten as follows

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

Now applying Cauchy-Schwarz's inequality

$$\mathbb{V}\mathrm{ar}\left[\left|\frac{1}{2\pi}\int_{A(m,k)}\overline{t^*(u)}e^{iu\beta_1}\,\mathrm{d}u\right|\right] \le \frac{1}{4\pi^2}\iint_{A(m,k)\times A(m,k)}\left|f^*_\beta(u-v)t^*(u)t^*(v)\right|\,\mathrm{d}u\,\mathrm{d}v$$

$$\le \frac{1}{4\pi^2}\sqrt{\iint_{A(m,k)\times A(m,k)}|f^*_\beta(u-v)||t^*(u)|^2\,\mathrm{d}u\,\mathrm{d}v}\sqrt{\iint_{A(m,k)\times A(m,k)}|f^*_\beta(u-v)||t^*(v)|^2\,\mathrm{d}u\,\mathrm{d}v}$$

$$\le \frac{1}{4\pi^2}\sqrt{\iint_{A(m,k)\times\mathbb{R}}|f^*_\beta(w)||t^*(u)|^2\,\mathrm{d}u\,\mathrm{d}w}\sqrt{\iint_{\mathbb{R}\times A(m,k)}|f^*_\beta(w)||t^*(v)|^2\,\mathrm{d}w\,\mathrm{d}v}$$

$$\le \frac{1}{4\pi^2}\iint_{A(m,k)\times A(m,k)}|f^*_\beta(w)||t^*(u)|^2\,\mathrm{d}u\,\mathrm{d}w \le \frac{\|f^*_\beta\|_1}{2\pi}.$$

So we choose ν as follows

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

And noticing that $\left|\frac{1}{2\pi}\int_{A(m,k)}\overline{t^*(u)}e^{iux}\,\mathrm{d}u\right|^2 \le k-m,$, we choose M_1 as

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

We can now apply Talagrand's inequality

$$\mathbb{E}\left[\sup_{\substack{k\ge m\\ k\in\mathcal{M}_N}}\left\{\int_{A(m,k)}\left|\frac{1}{N}\sum_{k=1}^{N}(e^{iu\beta_k}-f^*_\beta(u))\right|^2\mathrm{d}u - \frac{1}{24}\widehat{\mathrm{qen}}_1(m,k)\right\}_+\right]$$

$$\le \sum_{\substack{k\ge m\\ k\in\mathcal{M}_N}} K_1\left(\frac{\|f^*_\beta\|_1}{N}e^{-K_2\frac{k-m}{\|f^*_\beta\|_1}} + K_2\frac{k-m}{N^2}e^{-K_3\frac{\sqrt{N(k-m)}}{\sqrt{k-m}}}\right).$$

Finally $\mathbb{E}\left[\sup_{\substack{k\ge m\\ k\in\mathcal{M}_N}}\left\{\frac{1}{2\pi}\int_{A(m,k)}\left|\frac{1}{N}\sum_{k=1}^{N}(e^{iu\beta_k}-f^*_\beta(u))\right|^2\mathrm{d}u - \frac{1}{24}\widehat{\mathrm{qen}}_1(m,k)\right\}_+\right] \le \frac{C}{N}.$ □

Lemma 5.6. *There is a positive constant C such that for any arbitrary $m\in\mathcal{M}_N$*

$$\mathbb{E}\left[\sup_{\substack{k\ge m\\ k\in\mathcal{M}_N}}\left\{\frac{1}{2\pi}\int_{A(m,k)}\frac{|S_{NJ}(u)|^2}{|\tilde f^*_\varepsilon\left(\frac{u}{\Delta}\right)|^4}\mathrm{d}u - \frac{1}{16}\widehat{\mathrm{qen}}_2(m,k)\right\}_+\right] \le \frac{C}{N(J-4)}$$

with $S_{NJ}(u) = \frac{2}{N(J-4)}\sum_{k=1}^{N}\sum_{j=3}^{J/2}e^{iu\beta_k}\left(e^{iu\eta_{k,j}/\Delta} - |f^*_\varepsilon\left(\frac{u}{\Delta}\right)|^2\right)$.

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 $\beta_1,\ldots,\beta_N$ and $\varepsilon_{k,j}$ for $j=1,2$ and $k=1,\ldots,N$.

$$\mathbb{E}\left[\sup_{\substack{k\ge m\\ k\in\mathcal{M}_N}}\left\{\frac{1}{2\pi}\int_{A(m,k)}\frac{|S_{NJ}(u)|^2}{|\tilde f^*_\varepsilon\left(\frac{u}{\Delta}\right)|^4}\mathrm{d}u - \frac{1}{16}\widehat{\mathrm{qen}}_2(m,k)\right\}_+\right]$$

$$\le \mathbb{E}\left[\sum_{\substack{k\ge m\\ k\in\mathcal{M}_N}}\left\{\frac{1}{2\pi}\int_{A(m,k)}\frac{|S_{NJ}(u)|^2}{|\tilde f^*_\varepsilon\left(\frac{u}{\Delta}\right)|^4}\mathrm{d}u - \frac{\hat\mu^2(m,k)\hat\Xi(m,k)}{N(J-4)}\right\}_+\right]$$

$$\le \mathbb{E}\left[\sum_{\substack{k\ge m\\ k\in\mathcal{M}_N}}\frac{1}{2\pi}\int_{A(m,k)}\mathbb{E}\left[\left\{\frac{|S_{NJ}(u)|^2}{|\tilde f^*_\varepsilon\left(\frac{u}{\Delta}\right)|^4} - \frac{\hat\mu^2(m,k)}{N(J-4)|\tilde f^*_\varepsilon\left(\frac{u}{\Delta}\right)|^4}\right\}_+\Big|\tilde f^*_\varepsilon,\beta\right]\mathrm{d}u\right]$$

Now $\left(\frac{2}{N(J-4)}\sum_{k=1}^{N}\sum_{j=3}^{J/2} e^{iu\beta_k}e^{iu\eta_{k,j}/\Delta}\right)/|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2$ (conditional on $\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)$ and $\beta_1,\dots,\beta_N$) is the sum of $N(J-4)$ independent and identically distributed random variables with variance $v^2 \le 1/|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4$ which are surely bounded by $1/|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2$. Thus Lemma A.1 gives

$$\begin{aligned}
&\mathbb{E}\left[\left\{\frac{|S_{NJ}(u)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4} - \frac{\hat{\mu}^2(m,k)}{N(J-4)|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\right\}_+ \middle| \tilde{f}_\varepsilon^*, \beta\right]\\
&\le \frac{32}{N(J-4)|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\exp\left(-\frac{\hat{\mu}^2(m,k)}{8}\right) + \frac{128\sqrt{2}}{N^2(J-4)^2|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\exp\left(-\frac{3}{16\sqrt{2}}\sqrt{N(J-4)}\hat{\mu}(m,k)\right)\\
&\le \frac{32}{N(J-4)|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}(k-m)^{-2}\hat{\Xi}(m,k)^{-1} + \frac{128\sqrt{2}}{N^2(J-4)^2|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}(k-m)^{-2}\hat{\Xi}(m,k)^{-1}
\end{aligned}$$

where we used the fact that

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

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

$$\begin{aligned}
&\int_{A(m,k)}\mathbb{E}\left[\left\{\frac{|S_{NJ}(u)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4} - \frac{\hat{\mu}^2(m,k)}{N(J-4)|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\right\}_+ \middle| \tilde{f}_\varepsilon^*, \beta\right]\mathrm{d}u\\
&\le \frac{C}{N(J-4)}(k-m)^{-2}\hat{\Xi}(m,k)^{-1}\int_{A(m,k)}\frac{\mathrm{d}u}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4} \le \frac{C}{N(J-4)}(k-m)^{-2}.
\end{aligned}$$

Finally, $\mathbb{E}\left[\sup_{\substack{k\ge m\\ k\in\mathcal{M}_N}}\left\{\frac{1}{2\pi}\int_{A(m,k)}\frac{|S_{NJ}(u)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\mathrm{d}u - \frac{1}{16}\widehat{\mathrm{qen}}_2(m,k)\right\}_+\right] \le \frac{C}{N(J-4)}$. □

Proof of Proposition 5.2. Using Plancherel's formula, we get

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

Then it follows that

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

Therefore we can write

$$\begin{aligned}
\left\|\hat{f}_{\beta,k}-\hat{f}_{\beta,m}\right\|^2 &\le \frac{2}{\pi}\int_{A(m,k)}\frac{|\hat{f}_Z^*(u)-f_Z^*(u)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\,\mathrm{d}u+\frac{2}{\pi}\int_{A(m,k)}\frac{|f_Z^*(u)|^2}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\,\mathrm{d}u\\
&\quad+\frac{1}{\pi}\int_{A(m,k)}|\hat{f}_Z^*(u)|^2\frac{\left||\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)^2-|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2\right|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\mathbb{1}_{\left\{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|\le|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|\right\}}\,\mathrm{d}u\\
&\le \frac{2}{\pi}\int_{A(m,k)}\frac{|\hat{f}_Z^*(u)-f_Z^*(u)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\,\mathrm{d}u+4\left\|f_{\beta,k}-f_{\beta,m}\right\|^2\\
&\quad+\frac{1}{\pi}\int_{A(m,k)}|\hat{f}_Z^*(u)|^2\frac{\left||\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2-|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2\right|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^8}\mathbb{1}_{\left\{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|\le|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|\right\}}\,\mathrm{d}u\\
&:= I_1(m,k)+4\left\|f_{\beta,k}-f_{\beta,m}\right\|^2+I_2(m,k). \qquad (31)
\end{aligned}$$

To bound $I_2(m,k)$, we introduce the following set

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

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

$$\begin{aligned}
I_2(m,k)&\le 8\kappa^2\log\left(N(k-m)\right)N^{-1}\frac{1}{2\pi}\int_{A(m,k)}\frac{w\left(\frac{u}{\Delta}\right)^{-2}|\hat{f}_Z^*(u)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^{12}}\mathbb{1}\left\{|\tilde{f}_\varepsilon^*\left(\tfrac{u}{\Delta}\right)|\le|f_\varepsilon^*\left(\tfrac{u}{\Delta}\right)|\right\}\mathrm{d}u\\
&\le 8\kappa^2\log\left(N(k-m)\right)N^{-1}\hat{\Xi}^\beta(m,k):=\frac{1}{2}\widehat{\mathrm{qen}}_3(m,k).
\end{aligned}$$

To bound $I_2(m,k)$, we use Equation (29) and the notation S_{NJ} defined in Lemma 5.6. Thus we have

$$\begin{aligned}
I_2(m,k)&\le\frac{4}{\pi}\int_{A(m,k)}\frac{\left|\frac{1}{N}\sum_{k=1}^N|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2(e^{iu\beta_k}-f_\beta^*(u))\right|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\,\mathrm{d}u+\frac{4}{\pi}\int_{A(m,k)}\frac{|S_{NJ}(u)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\,\mathrm{d}u\\
&\le\frac{8}{\pi}\int_{A(m,k)}\frac{\left|\frac{1}{N}\sum_{k=1}^N|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2(e^{iu\beta_k}-f_\beta^*(u))\right|^2}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\,\mathrm{d}u+\frac{4}{\pi}\int_{A(m,k)}\frac{|S_{NJ}(u)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\,\mathrm{d}u\\
&\quad+\frac{8}{\pi}\int_{A(m,k)}\left|\frac{1}{N}\sum_{k=1}^N|f_\varepsilon^*\left(\tfrac{u}{\Delta}\right)|^2(e^{iu\beta_k}-f_\beta^*(u))\right|^2\left|\frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2}-\frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2}\right|^2\mathrm{d}u\\
&\le\frac{8}{\pi}\int_{A(m,k)}\left|\frac{1}{N}\sum_{k=1}^N(e^{iu\beta_k}-f_\beta^*(u))\right|^2\mathrm{d}u+\frac{4}{\pi}\int_{A(m,k)}\frac{|S_{NJ}(u)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\,\mathrm{d}u\\
&\quad+\frac{8}{\pi}\int_{A(m,k)}|f_\varepsilon^*\left(\tfrac{u}{\Delta}\right)|^4\left|\frac{1}{N}\sum_{k=1}^N(e^{iu\beta_k}-f_\beta^*(u))\right|^2\left|\frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2}-\frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2}\right|^2\mathrm{d}u.
\end{aligned}$$

We can now write the following inequalities, using (31) and the above remarks

$$\begin{aligned}
&\left\|\hat{f}_{\beta,k}-\hat{f}_{\beta,m}\right\|^2-4\|f_{\beta,k}-f_{\beta,m}\|^2-\frac{1}{2}\widehat{\mathrm{qen}}(m,k)\\
&\le\frac{8}{\pi}\int_{A(m,k)}\left|\frac{1}{N}\sum_{k=1}^N(e^{iu\beta_k}-f_\beta^*(u))\right|^2\mathrm{d}u-\frac{1}{4}\mathrm{qen}_1(m,k)+\frac{4}{\pi}\int_{A(m,k)}\frac{|S_{NJ}(u)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\,\mathrm{d}u-\frac{1}{2}\widehat{\mathrm{qen}}_2(m,k)\\
&+\frac{8}{\pi}\int_{A(m,k)}|f_\varepsilon^*\left(\tfrac{u}{\Delta}\right)|^4\left|\frac{1}{N}\sum_{k=1}^N(e^{iu\beta_k}-f_\beta^*(u))\right|^2\left|\frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2}-\frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2}\right|^2\mathrm{d}u-\frac{1}{4}\mathrm{qen}_1(m,k)\\
&+\frac{1}{2}\widehat{\mathrm{qen}}_3(m,k)-\frac{1}{2}\widehat{\mathrm{qen}}_3(m,k)\\
&+\frac{1}{\pi}\int_{A(m,k)}|\hat{f}_Z^*(u)|^2\frac{\left||\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2-|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2\right|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^8}\mathbb{1}_{\left\{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|\le|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|\right\}}\,\mathrm{d}u\,\mathbb{1}_{\{C(m,k)^c\}}
\end{aligned}$$

Taking expectation, we get

$$\begin{aligned}
&\mathbb{E}\left[\sup_{\substack{k\geq m\\ k\in\mathcal{M}_N}}\left\{\left\|\hat{f}_{\beta,k}-\hat{f}_{\beta,m}\right\|^2-4\|f_{\beta,k}-f_{\beta,m}\|^2-\frac{1}{2}\widehat{\mathrm{qen}}(m,k)\right\}_+\right]\\
&\leq\sum_{\substack{k\geq m\\ k\in\mathcal{M}_N}}\mathbb{E}\left[\left\{\left\|\hat{f}_{\beta,k}-\hat{f}_{\beta,m}\right\|^2-4\|f_{\beta,k}-f_{\beta,m}\|^2-\frac{1}{2}\widehat{\mathrm{qen}}(m,k)\right\}_+\right]\\
&\leq 16\sum_{\substack{k\geq m\\ k\in\mathcal{M}_N}}\mathbb{E}\left[\left\{\frac{1}{2\pi}\int_{A(m,k)}\left|\frac{1}{N}\sum_{k=1}^N(e^{iu\beta_k}-f_\beta^*(u))\right|^2\mathrm{d}u-\frac{1}{64}\mathrm{qen}_1(m,k)\right\}_+\right]\\
&\quad+\sum_{\substack{k\geq m\\ k\in\mathcal{M}_N}}\mathbb{E}\left[\left\{\frac{8}{\pi}\int_{A(m,k)}|f_\varepsilon^*\left(\tfrac{u}{\Delta}\right)|^4\left|\frac{1}{N}\sum_{k=1}^N(e^{iu\beta_k}-f_\beta^*(u))\right|^2\left|\frac{1}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2}-\frac{1}{|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2}\right|^2\mathrm{d}u-\frac{1}{4}\mathrm{qen}_1(m,k)\right\}_+\right]\\
&\quad+8\sum_{\substack{k\geq m\\ k\in\mathcal{M}_N}}\mathbb{E}\left[\left\{\frac{1}{2\pi}\int_{A(m,k)}\frac{|S_{NJ}(u)|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4}\mathrm{d}u-\frac{1}{16}\widehat{\mathrm{qen}}_2(m,k)\right\}_+\right]\\
&\quad+\frac{1}{\pi}\sum_{\substack{k\geq m\\ k\in\mathcal{M}_N}}\mathbb{E}\left[\int_{A(m,k)}|\hat{f}_Z^*(u)|^2\frac{\left||\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2-|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^2\right|^2}{|\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)|^8}\mathbb{1}_{\{|\tilde{f}_\varepsilon^*(\frac{u}{\Delta})|\leq|f_\varepsilon^*(\frac{u}{\Delta})|\}}\,\mathrm{d}u\,\mathbb{1}_{\{C(m,k)^c\}}\right].
\end{aligned}$$

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

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

Lemma 5.3 implies that $\mathbb{P}\left[C(m,k)^c\right]\leq N^{-3}(k-m)^{-3}$, we then get

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

Finally applying Lemma 5.5 and 5.6, we have

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

This completes the proof of Proposition 5.2. ☐

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 $\mathbb{V}\mathrm{ar}[X_1] \le v^2$ and suppose that almost surely $\|X_1\|_\infty \le B$. Let $S_n = 1/n \sum_{j=1}^n (X_j - \mathbb{E}[X_1])$. Let $\mathbb{E}|S_n| \le H$. Then*

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

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 $\|X_1^{(i)}\| \le B$ for some $B < \infty$. Let $v^2 := \sup_{i\in I} \mathbb{V}\mathrm{ar}X_1$. 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 $\kappa > 0$:*

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

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

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

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

Lemma A.4 (Lemma 2 p.35 (Butucea and Tsybakov (2008a)))**.** *Let γ, μ, and s be positive constants then for any $m > 0$*

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

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

Lemma A.5. *If f_ε^* satisfies assumption **(A6)** then*

$$\int_{-\pi m}^{\pi m} \frac{\mathrm{d}u}{\left|f_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^4} \asymp (\pi m)^{4\gamma+1-s} e^{4\mu(\pi m)^s},$$

$$\int_{-\pi m}^{\pi m} \frac{|f_\beta^*(u)|^2}{\left|f_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^8} \mathrm{d}u \lesssim (\pi m)^{(4\gamma+1-s)\wedge 2(2\gamma-\delta)_+} e^{4\mu(\pi m)^s} \mathbb{1}_{\{s>r\}}$$
$$+ (\pi m)^{2(2\gamma-\delta)_+} e^{2(2\mu-a)(\pi m)^s} \mathbb{1}_{\{r=s, 2\mu\ge a\}} + \mathbb{1}_{\{r>s\}\cup\{2\mu\le a\}}.$$

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

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.

Claeskens, G. and Hart, J. D. (2009). Goodness-of-fit tests in mixed models. *TEST*, 18(2):213–239.

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. *Scandinavian Journal of Statistics*, 41(2):325–345.

Dattner, I., Reiß, M., and Trabs, M. (2013). Adaptive quantile estimation in deconvolution with unknown error distribution. *To appear in Bernoulli, 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.

Ghidey, W., Lesaffre, E., and Eilers, P. (2004). Smooth random effects distribution in a linear mixed model. *Biometrics*, 60(4):945–953.

Ibrahim, J. G. and Kleinman, K. P. (1998). Semiparametric Bayesian methods for random effects models. In *Practical nonparametric and semiparametric Bayesian statistics*, volume 133 of *Lecture Notes in Statist.*, pages 89–114. Springer, New York.

Johannes, J. (2009). Deconvolution with unknown error distribution. *The Annals of Statistics*, 37(5a):2301–2323.

Johannes, J. and Schwarz, M. (2013). Adaptive circular deconvolution by model selection under unknown error distribution. *Bernoulli*, 19(5A):1576–1611.

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.

Komárek, A. and Lesaffre, E. (2008). Generalized linear mixed model with a penalized Gaussian mixture as a random effects distribution. *Comput. Statist. Data Anal.*, 52(7):3441–3458.

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.

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-effetcs 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.

Appendix B. Supplementary material

To prove Propositions 2.1 and 2.2, we need the two following technical lemma.

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

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

Lemma B.2. *Let $p \geq 1$, under assumption* ***(A2)*** *there exists a constant C_p such that*

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

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

$$\begin{aligned}
&\mathbb{E}\left|\frac{1}{|\tilde{f}_\varepsilon^*(x)|^2} - \frac{1}{|f_\varepsilon^*(x)|^2}\right|^{2p} \\
&= \mathbb{E}\left[\mathbb{1}_{\{|\tilde{f}_\varepsilon^*(x)|^4 < k_N(x)\}}\left|\frac{1}{\sqrt{k_N(x)}} - \frac{1}{|f_\varepsilon^*(x)|^2}\right|^{2p}\right] + \mathbb{E}\left[\mathbb{1}_{\{|\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x)\}}\left|\frac{1}{|\tilde{f}_\varepsilon^*(x)|^2} - \frac{1}{|f_\varepsilon^*(x)|^2}\right|^{2p}\right] \\
&\leq |f_\varepsilon^*(x)|^{-4p}\mathbb{P}\left[|\tilde{f}_\varepsilon^*(x)|^4 < k_N(x)\right]\frac{\left||f_\varepsilon^*(x)|^2 - \sqrt{k_N(x)}\right|^{2p}}{k_N(x)^p} + \mathbb{E}\left[\frac{\mathbb{1}_{\{|\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x)\}}\left||\tilde{f}_\varepsilon^*(x)|^4 - |f_\varepsilon^*(x)|^4\right|^{2p}}{|f_\varepsilon^*(x)|^{4p}|\tilde{f}_\varepsilon^*(x)|^{4p}\left||\tilde{f}_\varepsilon^*(x)|^2 + |f_\varepsilon^*(x)|^2\right|^{2p}}\right] \\
&\leq |f_\varepsilon^*(x)|^{-4p}]\frac{\left||f_\varepsilon^*(x)|^2 - \sqrt{k_N(x)}\right|^{2p}}{k_N(x)^p} + \mathbb{E}\left[\frac{\mathbb{1}_{\{|\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x)\}}\left||\tilde{f}_\varepsilon^*(x)|^4 - |f_\varepsilon^*(x)|^4\right|^{2p}}{|f_\varepsilon^*(x)|^{4p}|\tilde{f}_\varepsilon^*(x)|^{8p}}\right] \qquad (32)
\end{aligned}$$

• 1^{st} 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 (32), we get

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

• 2^{nd} 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.

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

Then we can bound the first term of Equation (32) as follows

$$|f_\varepsilon^*(x)|^{-4p}\mathbb{P}\left[|\tilde{f}_\varepsilon^*(x)|^4 < k_N(x)\right] \frac{\left||f_\varepsilon^*(x)|^2 - \sqrt{k_N(x)}\right|^{2p}}{k_N(x)^p} \leq |f_\varepsilon^*(x)|^{-4p}\mathbb{P}\left[|\tilde{f}_\varepsilon^*(x)|^4 < k_N(x)\right] \frac{\left(1 + \sqrt{k_N(x)}\right)^{2p}}{k_N(x)^p}$$
$$\leq |f_\varepsilon^*(x)|^{-4p}\mathbb{P}\left[|\tilde{f}_\varepsilon^*(x)|^4 < k_N(x)\right] \frac{Ck_N(x)^p}{k_N(x)^p} \leq \mathrm{O}(N^{-p}|f_\varepsilon^*(x)|^{-12p}) \leq \mathrm{O}(k_N(x)^{2p}|f_\varepsilon^*(x)|^{-12p}). \tag{33}$$

Moreover using this time that $1/\left||\tilde{f}_\varepsilon^*(x)|^2 + |f_\varepsilon^*(x)|^2\right|^2 \leq 1/|f_\varepsilon^*(x)|^4$, we can bound the second term of Equation (32) as follows

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

We then deduce the following bounds

$$\mathbb{E}\left[\mathbb{1}_{\{|\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x)\}}\left|\frac{1}{|\tilde{f}_\varepsilon^*(x)|^2} - \frac{1}{|f_\varepsilon^*(x)|^2}\right|^{2p}\right]$$
$$\leq 2^{2p-1}\,|f_\varepsilon^*(x)|^{-8p}\,\mathbb{E}\left[\mathbb{1}_{\{|\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x)\}}\left|\frac{\widehat{|f_\varepsilon^{*4}|}(x) - |f_\varepsilon^*(x)|^4}{|\tilde{f}_\varepsilon^*(x)|^2 + |f_\varepsilon^*(x)|^2}\right|^{2p}\right]$$
$$+ 2\,|f_\varepsilon^*(x)|^{-8p}\,\mathbb{E}\left[\mathbb{1}_{\{|\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x)\}}\frac{\left|\widehat{|f_\varepsilon^{*4}|}(x) - |f_\varepsilon^*(x)|^4\right|^{4p}}{|\tilde{f}_\varepsilon^*(x)}\right]$$
$$\leq 2^{2p-1}\,|f_\varepsilon^*(x)|^{-12p}\,\mathbb{E}\left[\left|\widehat{|f_\varepsilon^{*4}|}(x) - |f_\varepsilon^*(x)|^4\right|^{2p}\right]$$
$$+ 2\,|f_\varepsilon^*(x)|^{-12p}\,\mathbb{E}\left[\mathbb{1}_{\{|\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x)\}}\frac{\left|\widehat{|f_\varepsilon^{*4}|}(x)|^4 - |f_\varepsilon^*(x)|^4\right|^{4p}}{|\tilde{f}_\varepsilon^*(x)|^{4p}}\right]$$

which implies

$$\mathbb{E}\left[\mathbb{1}_{\{|\tilde{f}_\varepsilon^*(x)|^4 \geq k_N(x)\}}\left|\frac{1}{|\tilde{f}_\varepsilon^*(x)|^2} - \frac{1}{|f_\varepsilon^*(x)|^2}\right|^{2p}\right] \leq 2^{2p-1}\,|f_\varepsilon^*(x)|^{-12p}\,N^{-p} + 2\,|f_\varepsilon^*(x)|^{-12p}\,k_N(x)^{-p}N^{-2p}$$
$$\leq 2^{2p-1}\,|f_\varepsilon^*(x)|^{-12p}\,N^{-p} + 2\,|f_\varepsilon^*(x)|^{-12p}\,N^pN^{-2p} \leq \mathrm{O}\left(|f_\varepsilon^*(x)|^{-12p}\,N^{-p}\right) \leq \mathrm{O}(k_N(x)^{2p}|f_\varepsilon^*(x)|^{-12p}). \tag{34}$$

Then gathering Equations (33) and (34), we just proved that if $|f_\varepsilon^*(x)|^4 \geq 2k_N(x)$ then

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

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

B.2. **Proof of Lemma B.1.** Under (**A5**) and applying Lemma B.2, we have

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

□

B.3. **Proof of Proposition 2.2.** As aforementioned $\hat{f}_{\beta,m}$ can be seen as a projection estimator. We can then write the following equality using Pythagoras' theorem

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

Now using Plancherel's formula, we can write

$$
\begin{aligned}
\left\|f_{\beta,m}-\hat{f}_{\beta,m}\right\|^2&=\frac{1}{2\pi}\int |f_{\beta,m}^*(u)-\hat{f}_{\beta,m}^*(u)|^2\,\mathrm{d}u=\frac{1}{2\pi}\int_{-\pi m}^{\pi m}\left|\frac{\hat{f}_Z^*(u)}{\tilde{f}_\varepsilon^*\left(\frac{u}{\Delta}\right)^2}-\frac{f_Z^*(u)}{f_\varepsilon^*\left(\frac{u}{\Delta}\right)^2}\right|^2\mathrm{d}u \\
&\leq \frac{1}{\pi}\int_{-\pi m}^{\pi m}\left|\hat{f}_Z^*(u)R\left(\tfrac{u}{\Delta}\right)\right|^2\mathrm{d}u+\frac{1}{\pi}\int_{-\pi m}^{\pi m}\left|\frac{\hat{f}_Z^*(u)-f_Z^*(u)}{f_\varepsilon^*\left(\frac{u}{\Delta}\right)^2}\right|^2\mathrm{d}u \qquad (35)
\end{aligned}
$$

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

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

Yet we can write

$$
\mathbb{E}\left[\left|\hat{f}_Z^*(u)-f_Z^*(u)\right|^2\right]=\frac{4}{N(J-4)^2}\sum_{j=3}^{J/2}\mathbb{V}\mathrm{ar}\left(e^{iuZ_{1,j}}\right)+\frac{4}{N(J-4)^2}\sum_{\substack{3\leq j,j'\leq J/2\\ j\neq j'}}\mathbb{C}\mathrm{ov}\left(e^{iuZ_{1,j}},e^{iuZ_{1,j'}}\right),
$$

which implies that

$$
\mathbb{E}\left|\hat{f}_Z^*(u)-f_Z^*(u)\right|^2 \leq \frac{4}{N(J-4)^2}\left(\frac{J-4}{2}+\frac{(J-4)^2}{4}\left(1-|f_\beta^*(u)|^2\right)\left|f_\varepsilon^*\left(\tfrac{u}{\Delta}\right)\right|^4\right) \leq \frac{2}{N(J-4)}\left(1+\frac{J-4}{2}\left|f_\varepsilon^*\left(\tfrac{u}{\Delta}\right)\right|^4\right),
$$

hence

$$
\frac{1}{\pi}\int_{-\pi m}^{\pi m}\left|f_\varepsilon^*\left(\tfrac{u}{\Delta}\right)\right|^{-4}\mathbb{E}\left[\left|\hat{f}_Z^*(u)-f_Z^*(u)\right|^2\right]\mathrm{d}u \leq \frac{1}{\pi}\frac{2}{N(J-4)}\int_{-\pi m}^{\pi m}\left|f_\varepsilon^*\left(\tfrac{u}{\Delta}\right)\right|^{-4}\mathrm{d}u+\frac{2m}{N}. \qquad (36)
$$

Now noticing that $|\hat{f}_Z^*(u)-f_Z^*(u)|$ and $|R\left(\frac{u}{\Delta}\right)|$ are independent and applying Lemma B.2 for $p=1$, we get

$$
\begin{aligned}
\int_{-\pi m}^{\pi m}\mathbb{E}\left[\left|\hat{f}_Z^*(u)R\left(\tfrac{u}{\Delta}\right)\right|^2\right]\mathrm{d}u&=\int_{-\pi m}^{\pi m}\mathbb{E}\left[\left|\hat{f}_Z^*(u)-f_Z^*(u)+f_Z^*(u)\right|^2\left|R\left(\tfrac{u}{\Delta}\right)\right|^2\right]\mathrm{d}u \\
&\leq 2\int_{-\pi m}^{\pi m}|f_Z^*(u)|^2\,\mathbb{E}\left|R\left(\tfrac{u}{\Delta}\right)\right|^2\mathrm{d}u+2\int_{-\pi m}^{\pi m}\mathbb{E}\left|\hat{f}_Z^*(u)-f_Z^*(u)\right|^2\mathbb{E}\left|R\left(\tfrac{u}{\Delta}\right)\right|^2\mathrm{d}u \\
&\leq 2\int_{-\pi m}^{\pi m}|f_Z^*(u)|^2\frac{k_N^2(u)}{\left|f_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^{12}}\mathrm{d}u \\
&\quad+2\int_{-\pi m}^{\pi m}\left|f_\varepsilon^*\left(\tfrac{u}{\Delta}\right)\right|^{-4}\frac{2}{N(J-4)}\left(1+\frac{J-4}{2}\left|f_\varepsilon^*\left(\tfrac{u}{\Delta}\right)\right|^2\right)\mathrm{d}u \\
&\leq 2C_1\int_{-\pi m}^{\pi m}|f_Z^*(u)|^2\frac{k_N^2(u)}{\left|f_\varepsilon^*\left(\frac{u}{\Delta}\right)\right|^{12}}\mathrm{d}u+\frac{4m\pi}{N}+\frac{4}{N(J-4)}\int_{-\pi m}^{\pi m}\left|f_\varepsilon^*\left(\tfrac{u}{\Delta}\right)\right|^{-4}\mathrm{d}u
\end{aligned}
\qquad (37)
$$

Plugging (36) and (37) into Equation (35) yields

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

as $|f_Z^*(u)|^2=|f_\beta^*(u)|^2|f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4$. In the end

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

□