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

FABIENNE COMTE[(1)] AND ADELINE SAMSON[(1)]

January 5, 2012

Abstract. We consider a linear mixed-effects model where $Y_{k,j} = \alpha_k + \beta_k t_j + \varepsilon_{k,j}$ is the observed value for individual k at time t_j, $k = 1, \dots, N$, $j = 1, \dots, J$. The random effects α_k, β_k are independent identically distributed random variables with unknown densities f_α and f_β and are independent of the noise. We develop nonparametric estimators of these two densities, which involve a cutoff parameter. We study their mean integrated square risk and propose cutoff-selection strategies, depending on the noise distribution assumptions. Lastly, in the particular case of fixed interval between times t_j, we show that a completely data driven strategy can be implemented without any knowledge on the noise density. Intensive simulation experiments illustrate the method.

Keywords. Linear mixed-effects model, Nonparametric density estimation, Random effect density
MSC 2010 subject classification 62G07

1. Introduction

Longitudinal data and repeated measurements along time of a process are classically analyzed with mixed-effects models. This allows taking into account both the inter-subjects and the intra-subjects variabilities. In this paper, we focus on a simple linear mixed-effects model written as follows. Let $Y_{k,j}$ denote the observed value for individual k at time t_j, for $k = 1, \dots, N$, $j = 1, \dots, J$. The linear mixed-effects model is defined as

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

where (α_k, β_k) represent the individual random variables of subject k, also called random effects, and $(\varepsilon_{k,j})$ are the measurement errors. We assume that:

- [A1] times $(t_j)_{1 \le j \le J}$ are known and deterministic,
- [A2] measurement errors $\varepsilon_{k,j}$ are independent identically distributed (i.i.d.) with a density f_ε, such that $\mathbb{E}(e^{iu\varepsilon}) \neq 0$, for any $u \in \mathbb{R}$,
- [A3] variables (α_k, β_k) are i.i.d. and we denote by f_α and f_β the densities of α_1 and β_1,
- [A4] the sequence $(\alpha_k, \beta_k)_{1 \le k \le N}$ is independent of the sequence $(\varepsilon_{k,j})_{1 \le k \le N, 1 \le j \le J}$.

When the random effects (α_k, β_k) and errors $(\varepsilon_{k,j})$ are normally distributed, the maximum likelihood has been widely studied, the marginal density of Y having a closed form

[(1)]. MAP5, UMR CNRS 8145, Université Paris Descartes, Sorbonne Paris Cité, 45 rue des Saints Pères, 75006 Paris, France.

[see *e.g.* Pinheiro and Bates, 2000, and references therein]. However, the normality assumption of the random effects may be inappropriate in some situations. For example when an important covariate is omitted, a bimodal density may be more pertinent. As recalled by Ghidey et al. [2010], wrongly assuming normality of the random effects can lead to poor estimation results. Some authors propose to relax this assumption by developing estimation of the random effects density first four moments [see *e.g.* Wu and Zhu, 2010, and references therein]. But estimation of the random effects complete density may even be more appropriate, especially if the true density is multimodal. Our aim is therefore to estimate the densities f_α and f_β.

Several approaches have been proposed for this purpose. Shen and Louis [1999] consider a smoothing by roughening method without any assumption on f_ε. Assuming ε Gaussian, Zhang and Davidian [2001], Chen et al. [2002], Vock et al. [2011] propose a semi nonparametric approach based on the approximation of the random effects density by an Hermite series. Verbeke and Lesaffre [1996] develop an heterogeneity model where the random effects have a finite mixture Gaussian density. Ghidey et al. [2004] propose a penalised Gaussian mixture approach. Morris and Carroll [2006] use a wavelet-based approach. Non parametric maximum likelihood has also been studied by Laird [1978], Mallet et al. [1988], Kuhn [2003], Chafaï and Loubes [2006]. Recently Antic et al. [2009] compare several of these approaches with an intensive simulation study.

In this paper, we consider a different approach based on deconvolution tools. Deconvolution has been widely studied in various contexts. First, the noise density was systematically assumed known and different estimators have been proposed: kernel estimators (e.g. Stefanski and Carroll [1990], Fan [1991]), kernel estimators with bandwidth selection strategies (Delaigle and Gijbels [2004]), wavelet estimators (Pensky and Vidakovic [1999]), or projection methods with model selection (Comte et al. [2006]). Then several extensions have been considered to relax the assumption about noise density knowledge. Neumann [1997] first studied the case where the noise density is estimated from a preliminary noise sample. A complete adaptive procedure has been provided by Comte and Lacour [2011]. Recently, several papers focus on repeated observations, which provide another way to estimate the noise density, see Neumann [2007], Delaigle et al. [2008], Meister and Neumann [2010] or Comte et al. [2011].

We propose now a repeated observations strategy applied to mixed-effects models. We describe different estimators whether the noise density is partly or completely unknown. A preliminary estimation of f_ε Fourier transform is used, if needed. We study the risk bounds of these estimators and show how the context of longitudinal data affects the variance terms of the bounds. We also propose cut-off selections depending on the noise distribution assumptions. Especially, a completely data driven estimator is proposed in the case of unknown noise density.

The paper is organized as follows. Sections 2 and 3 present the estimators of f_α and f_β, their risk bounds and the cut-off selection assuming f_ε is known. In section 4, the particular case of a Gaussian noise f_ε is considered. We propose an optimal cut-off selection in the two cases σ_ε known or unknown. Finally, the general case with f_ε completely unknown is considered in section 5 and new estimators and cut-off selections are proposed. Performances of the different estimators are evaluated by simulation in section 6. Concluding remarks are proposed in section 7. Proofs are gathered in Appendix.

2. Definition of the estimators with a known noise density

2.1. **Model and notations.** In this section, we consider model (1) under Assumptions [A1]–[A4] and the additional assumption:

[A5] the noise density f_ε is known.

For the estimation of f_α, we shall distinguish two cases.

- First case, the observation at time 0 is available. For notation simplicity, we will denote by $Y_{k,0}$ the observation associated to time $t_0 = 0$, and we will consider J as the number of other observations with $t_j \neq 0$.
- Second case, no observation is available at time 0, then we have $t_j \neq 0$ for all $j = 1, \dots, J$.

For the estimation of f_β, we do not distinguish the two cases. We shall also consider that the time sequence $(t_j)_{1 \leq j \leq J}$ is in increasing order.

In the following, for notation simplicity and without loss of generality, we will assume that J is even. Then for $j = 1, \dots, J/2$, we denote by

$$\Delta_j = t_{2j} - t_{2j-1}$$

the time step between two successive observations.

Assumptions [A1]–[A4] on model (1) imply that for a given j, $(Y_{k,j})_{k=1,\dots,N}$ are i.i.d. Thus we denote by f_{Y_j} the density of $Y_{k,j}$.

We also need few notations related to Fourier transform theory. If f is an integrable function, then we denote by $f^*(u) = \int e^{iux} f(x)dx$ the Fourier transform of f on $\mathbb{R}$. For two real valued square integrable functions f and g, we denote the convolution product of f and g by $(f \star g)(x) = \int f(x-y)g(y)dy$ and we recall that, if f and g are both integrable and square integrable, then $(f \star g)^* = f^* g^*$. If f is integrable and square integrable we recall that inverse Fourier transform formula yields $f(x) = 1/(2\pi) \int e^{-ixu} f^*(u)du$.

2.2. **Estimator of f_β.** We first remark that by introducing the difference between two successive observations normalized by the length of the time interval, for $j = 1, \dots, J/2$,

$$Z_{k,j} = \frac{Y_{k,2j} - Y_{k,2j-1}}{\Delta_j},$$

we have

$$Z_{k,j} = \beta_k + \frac{\varepsilon_{k,2j} - \varepsilon_{k,2j-1}}{\Delta_j}. \tag{2}$$

For a fixed j, the variables $(Z_{k,j})_{k=1,\dots,N}$ are i.i.d. but the variables $Z_{k,j}$ and $Z_{k,l}$ for $j \neq l$ are not independent. Let us denote f_{Z_j} the density of the variables $Z_{k,j}$. It follows from (2) and the independence of (β_k) and $(\varepsilon_{k,j})$ under [A4], that

$$f_{Z_j} = f_\beta \star f_{(\varepsilon_{k,2j} - \varepsilon_{k,2j-1})/\Delta_j}. \tag{3}$$

Thus, by noting that, for all $j = 1, \dots, J/2$,

$$\begin{aligned} f^*_{(\varepsilon_{k,2j} - \varepsilon_{k,2j-1})/\Delta_j}(u) &= \mathbb{E}\left(e^{iu\frac{\varepsilon_{k,2j} - \varepsilon_{k,2j-1}}{\Delta_j}}\right) = \mathbb{E}\left(e^{i\frac{u}{\Delta_j}\varepsilon_{k,2j}} e^{-i\frac{u}{\Delta_j}\varepsilon_{k,2j-1}}\right) \\ &= \mathbb{E}\left(e^{i\frac{u}{\Delta_j}\varepsilon}\right) \mathbb{E}\left(e^{-i\frac{u}{\Delta_j}\varepsilon}\right) = \left|f^*_\varepsilon(\frac{u}{\Delta_j})\right|^2, \end{aligned}$$

we get, by taking the Fourier transform of equality (3),

$$f^*_{Z_j}(u) = f^*_\beta(u)\,|f^*_\varepsilon(u/\Delta_j)|^2. \tag{4}$$

It follows from (4) that, for all $j = 1,\dots,J/2$,

$$f^*_\beta(u) = \frac{f^*_{Z_j}(u)}{|f^*_\varepsilon(u/\Delta_j)|^2}.$$

In order to exploit all the available observations, we can also write:

$$f^*_\beta(u) = \frac{2}{J}\sum_{j=1}^{J/2}\frac{f^*_{Z_j}(u)}{|f^*_\varepsilon(u/\Delta_j)|^2}.$$

Now, Fourier inversion implies

$$f_\beta(x) = \frac{1}{2\pi}\int e^{-iux} f^*_\beta(u)du = \frac{1}{2\pi}\int e^{-iux}\frac{2}{J}\sum_{j=1}^{J/2}\frac{f^*_{Z_j}(u)}{|f^*_\varepsilon(u/\Delta_j)|^2}du. \tag{5}$$

This formula allows us to define the estimator of f_β based on the natural estimator of $f^*_{Z_j}(u)$

$$\widehat{f^*_{Z_j}}(u) = \frac{1}{N}\sum_{k=1}^{N} e^{iuZ_{k,j}} = \frac{1}{N}\sum_{k=1}^{N} e^{iu\frac{Y_{k,2j}-Y_{k,2j-1}}{\Delta_j}}, \tag{6}$$

Plugging (6) in (5) would give a proposal but may induce convergence problems of the integral. Thus, we introduce a cutoff πm in the integral defining the estimator of f_β:

$$\widehat{f_{\beta,m}}(x) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m} e^{-ixu}\frac{2}{J}\sum_{j=1}^{J/2}\frac{\widehat{f^*_{Z_j}}(u)}{|f^*_\varepsilon(u/\Delta_j)|^2}du$$

where $\widehat{f^*_{Z_j}}(u)$ is given by (6). To summarize, our proposal to estimate f_β is:

$$\widehat{f_{\beta,m}}(x) = \frac{2}{NJ}\sum_{k=1}^{N}\sum_{j=1}^{J/2}\frac{1}{2\pi}\int_{-\pi m}^{\pi m} e^{-ixu}\frac{e^{iuZ_{k,j}}}{|f^*_\varepsilon(u/\Delta_j)|^2}du \tag{7}$$

Note that under [A5], the estimator can indeed be computed.

2.3. **Estimator of f_α.** First, we can notice that, if observations for $t_0 = 0$ are available, then we have

$$Y_{k,0} = \alpha_k + \varepsilon_{k,0},\ k = 1,\dots,N.$$

This model is a classical deconvolution model. Thus, we propose to estimate f_α with the deconvolution estimator proposed by Fan [1991] with specific kernel as in Comte et al. [2006]:

$$\widehat{f^0_{\alpha,m}}(x) = \frac{1}{2\pi N}\sum_{k=1}^{N}\int_{-\pi m}^{\pi m} e^{-ixu}\frac{e^{iuY_{k,0}}}{f^*_\varepsilon(u)}du. \tag{8}$$

Its theoretical properties have been first studied by Stefanski and Carroll [1990], Fan [1991], and then by Comte et al. [2006] for cutoff selection and in more general context of functional regularities (see also Pensky and Vidakovic [1999] for such ideas in wavelet framework).

Now, we also provide an estimator when observations at $t_0 = 0$ are not available. In that case, we follow a construction similar to the one used for $\widehat{f_{\beta,m}}$. Set, for $j = 1, \ldots, J/2$,

$$V_{k,j} = \frac{Y_{k,2j}}{t_{2j}} - \frac{Y_{k,2j-1}}{t_{2j-1}}.$$

By definition of $Y_{k,j}$, we have

$$V_{k,j} = \left(\frac{1}{t_{2j}} - \frac{1}{t_{2j-1}}\right) \alpha_k + \left(\frac{\varepsilon_{k,2j}}{t_{2j}} - \frac{\varepsilon_{k,2j-1}}{t_{2j-1}}\right)$$

Remark that for a fixed j, the variables $(V_{k,j})_{k=1,\ldots,N}$ are i.i.d. but the variables $V_{k,j}$ and $V_{k,l}$ for $j \neq l$ are not independent. Let us denote by $f^*_{V_j}$ the Fourier transform of the density of the variables $V_{k,j}$ and

$$p_j = \frac{1}{t_{2j}} - \frac{1}{t_{2j-1}}.$$

We have, for all $j = 1, \ldots, J/2$,

$$f^*_\alpha(u) = \frac{f^*_{V_j}(u/p_j)}{f^*_\varepsilon(u/(p_j t_{2j}))\, f^*_\varepsilon(-u/(p_j t_{2j-1}))}. \tag{9}$$

A natural estimator of $f^*_\alpha(u)$ would be to compute the mean of the estimators of (9) for $j = 1, \ldots, J/2$. However, this choice can lead to numerical instability because the quantity $1/(p_j t_{2j})$ involved in the denominator of (9) can be large for large values of j. Indeed, when $\Delta_j = \Delta$ is fixed, then $t_j = j\Delta$, $1/(p_j t_{2j}) = -(2j-1)$ and $f^*_\varepsilon(u/(p_j t_{2j})) = f^*_\varepsilon(-(2j-1)u)$. Since f^*_ε tends to zero near infinity, $f^*_\varepsilon(u/(p_j t_{2j}))$ decreases when j increases. Therefore the estimator of f^*_α based on (9) may artificially take large values for large j. It shall be noted this is not the case for the estimator of $f^*_\beta(u)$ which only involves the step size Δ_j.

Thus, for numerical reasons, we propose an estimator of $f^*_\alpha(u)$ which is only based on the first observation $V_{k,1}$

$$\widehat{f^*_\alpha}(u) = \frac{1}{N}\sum_{k=1}^{N} \frac{e^{iV_{k,1}u/p_1}}{f^*_\varepsilon(u/(p_1 t_2))\, f^*_\varepsilon(-u/(p_1 t_1))}$$

Finally, the estimator of f_α is defined by

$$\widehat{f_{\alpha,m}}(x) = \frac{1}{2\pi N}\sum_{k=1}^{N}\int_{-\pi m}^{\pi m} e^{-iux}\frac{e^{iV_{k,1}u/p_1}}{f^*_\varepsilon(u/(p_1 t_2))\, f^*_\varepsilon(-u/(p_1 t_1))}du \tag{10}$$

3. Risk bounds and cutoff selection with known noise density

3.1. Risk bound for the estimator of f_β.

Let us define $f_{\beta,m}$ such that $f^*_{\beta,m} = f^*_\beta \mathbf{1}_{[-\pi m;\pi m]}$. The function $f_{\beta,m}$ is the function which is in fact estimated by $\widehat{f_{\beta,m}}$. We wish to bound the mean integrated squared error (MISE) defined by $\mathbb{E}\left(||f_\beta - \widehat{f_{\beta,m}}||^2\right)$. We first remark that the MISE is the sum of the integrated bias and the integrated variance:

$$\mathbb{E}\left\|\widehat{f_{\beta,m}}-f_\beta\right\|^2 = \left\|\mathbb{E}\left(\widehat{f_{\beta,m}}\right)-f_\beta\right\|^2+\mathbb{E}\left\|\widehat{f_{\beta,m}}-\mathbb{E}\left(\widehat{f_{\beta,m}}\right)\right\|^2$$

We can easily calculate the expectation $\mathbb{E}\left(\widehat{f_{\beta,m}}\right)$ of the estimator $\widehat{f_{\beta,m}}$. We have

$$\begin{aligned}\mathbb{E}\left(\widehat{f_{\beta,m}}(x)\right) &= \frac{2}{NJ}\sum_{k=1}^{N}\sum_{j=1}^{J/2}\frac{1}{2\pi}\int_{-\pi m}^{\pi m}e^{-iux}\frac{\mathbb{E}\left(e^{iuZ_{k,j}}\right)}{|f_\varepsilon^*(u/\Delta_j)|^2}\,du\\ &= \frac{2}{NJ}\sum_{k=1}^{N}\sum_{j=1}^{J/2}\frac{1}{2\pi}\int_{-\pi m}^{\pi m}e^{-iux}\frac{f_{Z_j}^*(u)}{|f_\varepsilon^*(u/\Delta_j)|^2}\,du\\ &= \frac{1}{2\pi}\int_{-\pi m}^{\pi m}e^{-iux}f_\beta^*(u)\,du = f_{\beta,m}(x)\end{aligned}$$

where the last line follows from (4). Therefore the pointwise bias is

$$f_\beta(x)-\mathbb{E}\left(\widehat{f_{\beta,m}}(x)\right) = \frac{1}{2\pi}\int e^{-iux}(f_\beta^*(u)-f_{\beta,m}^*(u))du = \frac{1}{2\pi}\int_{|u|\geq\pi m}e^{-iux}f_\beta^*(u)du,$$

and the integrated bias is equal to

$$\|f_\beta-f_{\beta,m}\|^2=\frac{1}{2\pi}\int_{|u|\geq\pi m}|f_\beta^*(u)|^2\,du.$$

To compute the integrated variance we first write, using Parseval formula,

$$\begin{aligned}\left\|\widehat{f_{\beta,m}}-f_{\beta,m}\right\|^2 &= \left\|\frac{2}{J}\frac{1}{2\pi}\sum_{j=1}^{J/2}\int e^{-iux}\frac{\widehat{f_{Z_j}^*}(u)-f_{Z_j}^*(u)}{|f_\varepsilon^*(u/\Delta_j)|^2}\mathbf{1}_{[-\pi m,\pi m]}(u)du\right\|^2\\ &= \frac{1}{2\pi}\int_{-\pi m}^{\pi m}\left|\frac{2}{J}\sum_{j=1}^{J/2}\frac{\widehat{f_{Z_j}}(u)-f_{Z_j}(u)}{|f_\varepsilon^*(u/\Delta_j)|^2}\right|^2\,du\\ &= \frac{1}{2\pi}\int_{-\pi m}^{\pi m}\left|\frac{2}{NJ}\sum_{j=1}^{J/2}\sum_{k=1}^{N}\frac{e^{iuZ_{k,j}}-\mathbb{E}(e^{iuZ_{k,j}})}{|f_\varepsilon^*(u/\Delta_j)|^2}\right|^2\,du \qquad (11)\end{aligned}$$

Thus, we get

$$\begin{aligned}
\mathbb{E}\left(\left\|\widehat{f_{\beta,m}}-f_{\beta,m}\right\|^2\right) &= \frac{2}{\pi N J^2}\int_{-\pi m}^{\pi m}\mathrm{Var}\left(\sum_{j=1}^{J/2}\frac{e^{iuZ_{1,j}}}{|f_\varepsilon^*(u/\Delta_j)|^2}\right)du\\
&= \frac{2}{\pi N J^2}\int_{-\pi m}^{\pi m}\left(\sum_{j=1}^{J/2}\frac{\mathrm{Var}(e^{iuZ_{1,j}})}{|f_\varepsilon^*(u/\Delta_j)|^4}+\sum_{j,j'=1,j\neq j'}^{J/2}\frac{\mathrm{cov}(e^{iuZ_{1,j}},e^{iuZ_{1,j'}})}{|f_\varepsilon^*(u/\Delta_j)|^2\left|f_\varepsilon^*(u/\Delta_{j'})\right|^2}\right)du\\
&\le \frac{2}{\pi N J^2}\int_{-\pi m}^{\pi m}\left(\sum_{j=1}^{J/2}\frac{1}{|f_\varepsilon^*(u/\Delta_j)|^4}+\sum_{j,j'=1}^{J/2}(1-|f_\beta^*(u)|^2)\right)du\\
&\le \frac{4}{NJ^2}\sum_{j=1}^{J/2}\left(\frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{du}{|f_\varepsilon^*(u/\Delta_j)|^4}\right)+\frac{m}{N}
\end{aligned}$$

Note that when the observation times are equally spaced ($\Delta_j=\Delta$), this reduces to

$$\mathbb{E}\left\|\widehat{f_{\beta,m}}-f_{\beta,m}\right\|^2 \le \frac{1}{\pi NJ}\int_{-\pi m}^{\pi m}\frac{du}{|f_\varepsilon^*(u/\Delta)|^4}+\frac{m}{N}$$

This result shows that we reduce the variance of a factor $1/J$ by taking the mean of all J available values of j in (7).

To summarize, the following result holds.

Proposition 1. *Consider Model (1) under Assumptions [A1]–[A5] and* $\widehat{f_{\beta,m}}$ *the estimator given by (7). If* f_β *is integrable and square-integrable, then*

$$\mathbb{E}\left\|\widehat{f_{\beta,m}}-f_\beta\right\|^2 \le \frac{1}{2\pi}\int_{|u|\ge\pi m}|f_\beta^*(u)|^2\,du+\frac{4}{NJ^2}\sum_{j=1}^{J/2}\left(\frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{du}{|f_\varepsilon^*(u/\Delta_j)|^4}\right)+\frac{m}{N} \tag{12}$$

Inequality (12) requires few comments. First, the term

$$\frac{1}{2\pi}\int_{|u|\ge\pi m}|f_\beta^*(u)|^2\,du$$

is a squared bias term due to the truncation of the integral. It decreases when m increases, and the rate of decrease is faster when the function f_β is more regular. Indeed, classical regularity spaces considered for density f_β on $\mathbb{R}$ are described by:

[A6] $f_\beta\in\mathcal{A}_b(L)=\{f_\beta\in\mathbb{L}^1\cap\mathbb{L}^2,\ \int|f_\beta^*(x)|^2(x^2+1)^b dx\le L\}$ with $b>1/2$, $L>0$.

Then, under [A6], we have the following bias order:

$$\|f_\beta-f_{\beta,m}\|^2\le CL(\pi m)^{-2b}. \tag{13}$$

Obviously, the larger the regularity index b of f_β, the faster the bias decreases.

The two other terms of inequality (12) are variance terms. They clearly increase when m increases. Moreover, the first of these two terms is dominating: since $|f^*(u)|\le 1, \forall u\in\mathbb{R}$, we have $\int_{-\pi m}^{\pi m}du/|f_\varepsilon^*(u/\Delta_j)|^4\ge 2\pi m$ and it is usually much larger. For instance, if the noise is Gaussian, $\int_{-\pi m}^{\pi m}du/|f_\varepsilon^*(u/\Delta_j)|^4$ is larger than $e^{2\sigma_\varepsilon^2(\pi m/\Delta_j)^2}\Delta_j^2/(2\sigma_\varepsilon^2\pi m)$.

We have therefore to find how to realize a compromise between the bias and the variance terms. This is the purpose of section 3.3 in which a cutoff selection is proposed. Note that we can also provide pointwise risk bounds for the estimator $\widehat{f_{\beta,m}}$; details about this are gathered in Appendix A.

3.2. **Risk bound for the estimator of f_α.** Let us define $f_{\alpha,m}$ such that $f^*_{\alpha,m} = f^*_\alpha \mathbf{1}_{[-\pi m;\,\pi m]}$. The function $f_{\alpha,m}$ is the function which is in fact estimated by $\widehat{f_{\alpha,m}}$. Using similar calculations to those used for $\widehat{f_{\beta,m}}$, we get the following global MISE bound.

Proposition 2. *Consider Model (1) under Assumptions [A1]–[A5] and estimator $\widehat{f_{\alpha,m}}$ given by (10). If f_α is integrable and square integrable, then we have*

$$\text{(14)}\quad \mathbb{E}\left\|\widehat{f_{\alpha,m}} - f_\alpha\right\|^2 \le \frac{1}{2\pi}\int_{|u|\ge \pi m} |f^*_\alpha(u)|^2\,du + \frac{1}{2\pi N}\int_{-\pi m}^{\pi m} \frac{du}{\left|f^*_\varepsilon\left(\frac{u}{p_1 t_2}\right)\, f^*_\varepsilon\left(\frac{u}{p_1 t_1}\right)\right|^2} + \frac{m}{N}.$$

The same comments as for bound (12) apply here.

3.3. **Cutoff selection.** We propose the following model selection procedure for choosing a relevant cutoff m for the estimators $\widehat{f_{\beta,m}}$ and $\widehat{f_{\alpha,m}}$. We define for $\omega = \alpha$ or $\omega = \beta$

$$\hat m_\omega = \arg\min_{m\in\mathcal{M}_{\omega,N}} \left\{-\|\widehat{f_{\omega,m}}\|^2 + \mathrm{pen}_\omega(m)\right\},$$

where for $\omega = \beta$:

$$\text{(15)}\qquad \mathrm{pen}_\beta(m) = \kappa_\beta\left(\frac{4}{NJ^2}\sum_{j=1}^{J/2}\left(\frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{du}{|f^*_\varepsilon(u/\Delta_j)|^4}\right) + \frac{m}{N}\right),$$

and for $\omega = \alpha$,

$$\text{(16)}\qquad \mathrm{pen}_\alpha(m) = \kappa_\alpha\left(\frac{1}{2\pi N}\int_{-\pi m}^{\pi m}\frac{du}{\left|f^*_\varepsilon\left(\frac{u}{p_1 t_2}\right)\, f^*_\varepsilon\left(\frac{u}{p_1 t_1}\right)\right|^2} + \frac{m}{N}.\right).$$

Here κ_β and κ_α are constants which are calibrated once for all on preliminary simulation experiments. Moreover, we set

$$\mathcal{M}_{\omega,N} = \{m \in \{1,\dots N\},\ \text{such that } \mathrm{pen}_\omega(m) \le 1\}.$$

We can prove the following result.

Theorem 1. *Consider Model (1) under Assumptions [A1]–[A5] with f_ω integrable and square integrable. Assume that the noise is ordinary smooth, i.e. that there exist two constants c_ϵ, C_ϵ such that, $\forall x \in \mathbb{R}$,*

$$\text{(17)}\qquad c_\epsilon(1+x^2)^\delta \le 1/|f^*_\varepsilon(x)|^2 \le C_\epsilon(1+x^2)^\delta.$$

Then, for $\omega = \alpha, \beta$,

$$\text{(18)}\qquad \mathbb{E}\left(\|\widehat{f_{\omega,\hat m_\omega}} - f_\omega\|^2\right) \le C \inf_{m\in\mathcal{M}_{\omega,N}}\left(\|f_\omega - f_{\omega,m}\|^2 + \mathrm{pen}_\omega(m)\right) + \frac{C'}{N},$$

where C and C' are constants depending on the problem.

Inequality (18) shows that the resulting estimator automatically realizes the squared-bias/variance trade-off, up to a multiplicative constant C. For the sake of simplicity, the result is given only in the ordinary smooth case. It can be obtained in general setting of ordinary or super smooth noise, provided that a factor is added in the penalty for the super smooth case, as detailed in Comte et al. [2006]. In practice, in the case of super smooth noise, we multiply the penalty pen_β by the factor

$$\log\left(\frac{4}{NJ^2}\sum_{j=1}^{J/2}\left(\frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{du}{|f_\varepsilon^*(u/\Delta_j)|^4}\right)\right)/\log(m+1) \tag{19}$$

and the penalty pen_α by the factor

$$\log\left(\frac{1}{2\pi N}\int_{-\pi m}^{\pi m}\frac{du}{\left|f_\varepsilon^*\left(\frac{u}{p_1t_2}\right)\, f_\varepsilon^*\left(\frac{u}{p_1t_1}\right)\right|^2}\right)/\log(m+1). \tag{20}$$

Indeed, it is shown in Comte and Lacour [2011] that this is a slight over-penalization which has the advantage of being easy to generalize to the unknown noise case.

4. Special case of Gaussian noise

In this section, we assume that f_ε is Gaussian and centered, and that [A6] holds. We distinguish two cases whether σ_ε^2 is known or not.

4.1. Cutoff choice for $\widehat{f_{\beta,m}}$ when σ_ε^2 is known.

The Fourier transform of f_ε is $f_\varepsilon^*(u) = \exp(-\sigma_\varepsilon^2 u^2/2)$. Let Δ_{min} be a constant such that for all $j = 1, \ldots, J$, $\Delta_{min} \leq \Delta_j$. Then we get

$$\begin{aligned}\int_{-\pi m}^{\pi m}\frac{du}{|f_\varepsilon^*(u/\Delta_j)|^4} &\leq \int_{-\pi m}^{\pi m}\exp(2\sigma_\varepsilon^2(u/\Delta_{min})^2)du\\ &\leq 2\pi m\exp(2\sigma_\varepsilon^2\Delta_{min}^{-2}\pi^2 m^2)\end{aligned}$$

which, associated to [A6] and inequality (12), gives the following result.

Proposition 3. *Consider Model (1) under Assumptions [A1]–[A6] with f_β integrable and square integrable, and assume that ε is Gaussian and $\Delta_j \geq \Delta_{min}$, $\forall j = 1, \ldots, J$. Then the choice*

$$m_{0,\beta} = m_{0,\beta}(\sigma_\varepsilon) = \left(\frac{\kappa'_\beta\log(NJ)}{4\pi^2\sigma_\varepsilon^2\Delta_{min}^{-2}}\right)^{1/2} \tag{21}$$

gives the bound $\mathbb{E}(\|\widehat{f_{\beta,m_{0,\beta}}} - f_\beta\|^2) = O(1/[\log(NJ)]^b)$ provided that $\kappa'_\beta < 1$.

The consequence is that the convergence rate of the estimator is logarithmic, which is rather slow. Nevertheless, simulation experiments show that deconvolution estimators behave well also in this setting. It is easy to see that the rate will be much improved if f_β is also Gaussian or more generally super-smooth.

4.2. **Cutoff choice for $\widehat{f_{\alpha,m}}$ when σ_ε^2 is known.** Now, we assume that $\Delta_j = \Delta$ is fixed. The main variance term of Inequality (14) with $p_1 t_1 = -1/2$ and $p_1 t_2 = -1$ is of order:

$$\frac{2}{\pi N}\int_{-\pi m}^{\pi m} \exp(5\sigma_\varepsilon^2 u^2)du \quad \leq \quad \frac{4m}{N}\exp(5\pi^2\sigma_\varepsilon^2 m^2).$$

Under [A6], the bias order is given by (13). We deduce that a good choice of m is

$$m_{0,\alpha} = m_{0,\alpha}(\sigma_\varepsilon) = \left(\frac{\kappa'_\alpha \log(N)}{5\pi^2\sigma_\varepsilon^2}\right)^{1/2} \tag{22}$$

with $\kappa'_\alpha < 1$. It is worth noticing that here and contrary to $m_{0,\beta}$, the choice $m_{0,\alpha}$ does not depend on the step Δ.

4.3. **Cutoff selection when σ_ε^2 is unknown.** The optimal choices $m_{0,\beta}$ and $m_{0,\alpha}$ provided in the two previous sections depend on σ_ε^2. When this variance is unknown, we propose to replace it by an estimator. A natural estimator can be obtained based on the following relations

$$Var(Y_{k,j}) \quad = \sigma_j^2 = \quad \sigma_\alpha^2 + t_j^2\sigma_\beta^2 + 2t_j\sigma_{\alpha,\beta} + \sigma_\varepsilon^2 \tag{23}$$

$$cov(Y_{k,j}, Y_{k,j'}) \quad = \sigma_{jj'}^2 = \quad \sigma_\alpha^2 + t_j t_{j'}\sigma_\beta^2 + (t_j + t_{j'})\sigma_{\alpha,\beta} \tag{24}$$

which hold for all k, j, j' and where σ_α^2, σ_β^2 and $\sigma_{\alpha,\beta}$ are the variances of α_k, β_k and the covariance of (α_1, β_1), respectively. Set $Y_{.j} = \frac{1}{N}\sum_{k=1}^N Y_{k,j}$. The following quantities

$$\hat\sigma_{Y.j}^2 \quad = \quad \frac{1}{N}\sum_{k=1}^N (Y_{k,j} - Y_{.j})^2, \quad \text{and} \quad \hat\sigma_{Y.j,Y.j'}^2 = \frac{1}{N}\sum_{k=1}^N (Y_{k,j} - Y_{.j})(Y_{k,j'} - Y_{.j'})$$

are natural estimators of σ_j^2 and $\sigma_{jj'}^2$, respectively. From equations (23) and (24), let us define the 4×4 - matrix A

$$A = \begin{pmatrix} 1 & \frac{3}{J}\sum_{j=1}^{J/3} t_{3j}^2 & \frac{3}{J}\sum_{j=1}^{J/3} t_{3j} & 1 \\ 1 & \frac{3}{J}\sum_{j=1}^{J/3} t_{3j-1}^2 & \frac{3}{J}\sum_{j=1}^{J/3} t_{3j-1} & 1 \\ 1 & \frac{3}{J}\sum_{j=1}^{J/3} t_{3j-2}^2 & \frac{3}{J}\sum_{j=1}^{J/3} t_{3j-2} & 1 \\ 1 & \frac{2}{J(J-1)}\sum_{1\leq j<j'}^{J/2} t_j t_{j'} & \frac{2}{J(J-1)}\sum_{1\leq j<j'}^{J/2}(t_j + t_{j'}) & 0 \end{pmatrix},$$

and the vector

$$\widehat\Sigma_Y = \begin{pmatrix} \frac{3}{J}\sum_{j=1}^{J/3} \hat\sigma_{Y.3j}^2 \\ \frac{3}{J}\sum_{j=1}^{J/3} \hat\sigma_{Y.3j-1}^2 \\ \frac{3}{J}\sum_{j=1}^{J/3} \hat\sigma_{Y.3j-2}^2 \\ \frac{2}{J(J-1)}\sum_{1\leq j<j'}^{J/2} \hat\sigma_{Y.j,Y.j}^2 \end{pmatrix}.$$

We assume that A is invertible. In this case, we deduce from equations (23)-(24) estimators of $\sigma^2 = (\sigma_\alpha^2, \sigma_\beta^2, \sigma_{\alpha,\beta}, \sigma_\varepsilon^2)$, which are defined as

$$\widehat{\sigma}^2 = \left(\hat\sigma_\alpha^2, \hat\sigma_\beta^2, \hat\sigma_{\alpha\beta}, \hat\sigma_\varepsilon^2\right)' = A^{-1}\widehat\Sigma_Y$$

As $\widehat{\Sigma}_Y$ is a M-estimator of parameters $(\frac{3}{J}\sum_{j=1}^{J/3}\sigma^2_{3j}, \frac{3}{J}\sum_{j=1}^{J/3}\sigma^2_{3j-1}, \frac{3}{J}\sum_{j=1}^{J/3}\sigma^2_{3j-2}, \frac{2}{J(J-1)}\sum_{1\leq j<j'}^{J/2}\sigma^2_{jj'})$, if $\mathbb{E}(|Y_{k,j}|^4)<\infty$, there exists an explicit matrix $\mathcal{I}$ such that

$$\sqrt{N}\left(\widehat{\Sigma}_Y - \left(\frac{3}{J}\sum_{j=1}^{J/3}\sigma^2_{3j}, \frac{3}{J}\sum_{j=1}^{J/3}\sigma^2_{3j-1}, \frac{3}{J}\sum_{j=1}^{J/3}\sigma^2_{3j-2}, \frac{2}{J(J-1)}\sum_{1\leq j<j'}^{J/2}\sigma^2_{jj'}\right)'\right) \xrightarrow[N\to\infty]{\mathcal{L}} \mathcal{N}(0,\mathcal{I})$$

We deduce that

$$\sqrt{N}\left(\widehat{\sigma}^2-\sigma^2\right) \xrightarrow[N\to\infty]{\mathcal{L}} \mathcal{N}(0, A^{-1}\mathcal{I}A'^{-1}) \tag{25}$$

We assume that there exists a known upper bound for the unknown value of σ_ε, denoted by $\sigma_{\varepsilon,max}$. By plugging $\hat{\sigma}^2_\varepsilon$ in the definition (21) of $m_{0,\beta}$, we obtain a random cutoff m:

$$\widehat{m_{0,\beta}} = m_{0,\beta}(\hat{\sigma}_\varepsilon)/\sqrt{2}\wedge m_n = \frac{\sqrt{\log(NJ)}}{2\sqrt{2}\hat{\sigma}_\varepsilon\Delta^{-1}_{min}\pi}\wedge m_n, \tag{26}$$

where $m_n = m_{0,\beta}(\sigma_{\varepsilon,max})/\sqrt{2}$.

The study of $\widehat{f_{\beta,\widehat{m_{0,\beta}}}}$ is complex in that case. We can prove the following upper bound of the integrated risk.

Proposition 4. *Consider Model (1) under Assumptions [A1]–[A6], and assume that ε is Gaussian $\mathcal{N}(0,\sigma^2_\varepsilon)$ with unknown $\sigma_\varepsilon\leq\sigma_{\varepsilon,max}$. Then the estimator $\widehat{f_{\beta,\widehat{m_{0,\beta}}}}$ with $\widehat{f_{\beta,m}}$ defined by (7) and $\widehat{m_{0,\beta}}$ defined by (26) is such that*

$$\mathbb{E}(\|\widehat{f_{\beta,\widehat{m_{0,\beta}}}}-f_\beta\|^2)\leq C\left([\log(NJ)]^{-b}+\frac{1}{N}\right). \tag{27}$$

It follows from Inequality (27) that the estimator $\widehat{f_{\beta,\widehat{m_{0,\beta}}}}$ automatically reaches its best possible rate, without requiring any information on the unknown function.

A similar proposal can be done for $\widehat{f^0_{\alpha,m}}$.

5. Estimators with unknown noise density

In this section, we consider Model (1) under Assumptions [A1], [A3], [A4] and Assumption [A2'] replacing Assumption [A2]:

[A2'] Assumption [A2] holds and the measurement errors $\varepsilon_{k,j}$ are symmetric.

Note that the symmetry of the noise together with the condition $f^*_\varepsilon(u)\neq 0$ imply that f^*_ε takes values in $\mathbb{R}^+/\{0\}$, i.e. $|f^*(u)| = f^*(u)>0$, $\forall u\in\mathbb{R}$.

Furthermore, we restrict to the following assumption on the observations design:

[A7] $\Delta_j=\Delta$ for all j, such that $t_j = j\Delta$ and $J\geq 6$.

5.1. Estimator of f^*_ε.

Under Assumptions [A2'] and [A7], we can propose an estimator of the density f_ε. Indeed, let us introduce

$$\begin{aligned} W_k &= Z_{k,2}-Z_{k,1} \\ &= \beta_k+\frac{\varepsilon_{k,4}-\varepsilon_{k,3}}{\Delta}-\beta_k-\frac{\varepsilon_{k,2}-\varepsilon_{k,1}}{\Delta} = \frac{1}{\Delta}(\varepsilon_{k,4}-\varepsilon_{k,3}-\varepsilon_{k,2}+\varepsilon_{k,1}) \end{aligned}$$

which implies that

$$f_W^*(u) = \mathbb{E}(e^{iuW_k}) = \mathbb{E}(\cos(uW_k)) = |f_\varepsilon^*\left(\frac{u}{\Delta}\right)|^4 = \left(f_\varepsilon^*(\frac{u}{\Delta})\right)^4.$$

Under Assumption [A2'], we can estimate $(f_\varepsilon^*)^4$ via

$$\widehat{(f_\varepsilon^*)^4}(u/\Delta) = \frac{1}{N}\sum_{k=1}^N \cos(uW_k).$$

Let us define estimators of f_β and f_α when f_ε is unknown. For numerical reasons, we can not directly plug estimators of $1/(f_\varepsilon^*)^2$ and $1/f_\varepsilon^*$ in (4) and (9) by considering $1/(\widehat{(f_\varepsilon^*)^4})^{1/2}$ and $1/(\widehat{(f_\varepsilon^*)^4})^{1/4}$. Therefore, following Neumann [1997], Comte and Lacour [2011] or Comte et al. [2011], we define a truncated estimator of $1/(f_\varepsilon^*)^2$

$$\frac{1}{\widetilde{(f_\varepsilon^*)^2}(u)} = \frac{\mathbb{1}_{\widehat{(f_\varepsilon^*)^4}(u)\geq N^{-1/2}}}{\left[\widehat{(f_\varepsilon^*)^4}(u)\right]^{1/2}}$$

to be plugged in $\widehat{f_{\beta,m}}$ and a truncated estimator of $1/f_\varepsilon^*$

$$\frac{1}{\widetilde{f_\varepsilon^*}(u)} = \frac{\mathbb{1}_{\widehat{(f_\varepsilon^*)^4}(u)\geq N^{-1/2}}}{\left[\widehat{(f_\varepsilon^*)^4}(u)\right]^{1/4}}$$

to be plugged in $\widehat{f_{\alpha,m}^0}$ and $\widehat{f_{\alpha,m}}$.

The error induced by the truncation is studied in the following lemma, which is an extension of Neumann [1997]'s lemma for the case we study here.

Lemma 1. *Assume that Assumption [A2'] holds.*

(1) *There exists a constant C_0 such that*

$$\mathbb{E}\left(\left|\frac{1}{\widetilde{(f_\varepsilon^*)^2}(u)} - \frac{1}{(f_\varepsilon^*)^2(u)}\right|^2\right) \leq \frac{2}{|f_\varepsilon^*(u)|^4} \wedge \frac{C_0 N^{-1/2}}{|f_\varepsilon^*(u)|^8} \wedge \frac{C_0 N^{-1}}{|f_\varepsilon^*(u)|^{12}}. \tag{28}$$

(2) *There exists a constant C_1 such that*

$$\mathbb{E}\left(\left|\frac{1}{\widetilde{f_\varepsilon^*}(u)} - \frac{1}{f_\varepsilon^*(u)}\right|^2\right) \leq \frac{1}{|f_\varepsilon^*(u)|^2} \wedge \left(C_1 \min_{p\in\{1,2,3,4\}} \frac{N^{-p/4}}{|f_\varepsilon^*(u)|^{2+2p}}\right). \tag{29}$$

5.2. **Estimator of f_β with unknown f_ε.** We can now define an estimator for f_β by plugging $\widetilde{(f_\varepsilon^*)^2}$ in (4). Under Assumption [A7], $J \geq 6$ and the observations used to estimate f_ε^* can be different from those used to estimate f_β. The estimator of f_β when f_ε is unknown is denoted by $\widetilde{f_{\beta,m}}$ and defined by

$$\widetilde{f_{\beta,m}}(x) = \frac{2}{N(J-4)}\sum_{k=1}^N\sum_{j=3}^{J/2}\frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{e^{-iu(x-Z_{k,j})}}{\widetilde{(f_\varepsilon^*)^2}(u/\Delta)}du. \tag{30}$$

Applying Lemma 1 shows that the risk bound in the case of estimated noise density is getting more complicated. We denote for any function f integrable and square integrable

$$D_k(m,f) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m} \frac{|f^*(u)|^2}{|f_\varepsilon^*(u/\Delta)|^{2k}}du.$$

Then we obtain the following risk bound

Proposition 5. *Consider Model (1) under Assumptions [A1], [A2'], [A3], [A4], [A7]. Assume moreover that f_β is integrable and square-integrable, then the estimator defined by (30) satisfies*

$$\mathbb{E}(\|\widetilde{f_{\beta,m}} - f_\beta\|^2) \le \|f_{\beta,m} - f_\beta\|^2 + 16\frac{D_2(m,1)}{N(J-4)} + 4C_0\left(\frac{D_2(m,f_\beta)}{\sqrt{N}}\right)\wedge\left(\frac{D_4(m,f_\beta)}{N}\right) + 6\frac{m}{N},$$

where C_0 is defined in Lemma 1.

This risk bound implies the usual bias term $\|f_{\beta,m}-f_\beta\|^2$ and three terms of variance. The terms $16D_2(m,1)/[N(J-4)]$ and $6m/N$ correspond to the terms obtained in Inequality (12) of Proposition 1 for known f_ε^*. The additional term, $4C_0(D_2(m,f_\beta)/\sqrt{N})\wedge(D_4(m,f_\beta)/N)$ comes from estimating f_ε^*. If $|f_\beta^*|^2$ decreases faster than $|f_\varepsilon^*|^8$, it can happen that $D_4(m,f_\beta)$ is bounded by a fixed constant. Then this term is negligible compared to $D_2(m,1)/[N(J-4)]$. Moreover, in Comte et al. [2011], in a context of repeated measurements, similar variance terms are also obtained and their simulation experiments show that the first variance term remains the dominating one. We conjecture that the same thing happens here, even if the estimation may be more difficult.

For the adaptive procedure, we customize the proposals of Comte and Lacour [2011] to the present case. More precisely, we replace f_ε^* by its estimate in the penalty and in the definition of the collection of cutoffs. For super smooth noise, an additional multiplicative factor has to be added. Therefore, when the noise density is unknown, this multiplicative factor must be systematically added "blindly". Practically, our estimator is $\widetilde{f_{\beta,\tilde m_\beta}}$ defined by (30) with

$$\tilde m_\beta = \arg\min_{m\in\widetilde{\mathcal{M}}_{\beta,N}} \left\{-\|\widetilde{f_{\beta,m}}\|^2 + \widetilde{\mathrm{pen}}_\beta(m)\right\}$$

with

$$\widetilde{\mathrm{pen}}_\beta(m) = \tilde\kappa_\beta \frac{\log(\frac{4\pi\hat D_2(m,1)}{J-4})}{\log(m+1)}\left(\frac{4\pi\hat D_2(m,1)}{N(J-4)} + \frac{m}{N}\right),\quad \hat D_2(m,1) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{du}{\widetilde{(f_\varepsilon^*)^4}(\frac{u}{\Delta})}$$

and $\widetilde{\mathcal{M}}_{\beta,N} = \{m\in\{1,\dots,N\}, \widetilde{\mathrm{pen}}_\beta(m)\le 1\}$.

For ordinary smooth noise, this multiplicative factor behaves roughly like a constant. For super smooth noise, it provides a slight overpenalization in the Gaussian case, as required by the theory, see Comte et al. [2006] and Comte and Lacour [2011].

5.3. **Estimator for f_α with unknown f_ε.** Analogously, we define for the estimation of f_α, either, if observations for $j=0$, $t_0=0$ are available

$$\widetilde{f_{\alpha,m}^0}(x) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m} e^{-iux}\frac{\frac{1}{N}\sum_{k=1}^N e^{iuY_{k,0}}}{\widetilde{f_\varepsilon^*}(u)}du, \tag{31}$$

or when observations at time 0 are not available

$$\widetilde{f_{\alpha,m}}(x) = \frac{1}{2\pi}\int_{-\pi m}^{\pi m} e^{-iux}\frac{\frac{1}{N}\sum_{k=1}^{N} e^{-2i\Delta u V_{k,3}}}{\widetilde{f_\varepsilon^*}(-u)\widetilde{f_\varepsilon^*}(2u)}du, \tag{32}$$

otherwise. In that second case, observations used for the estimation of f_ε^* are not used for the estimation of the numerator.

Now we detail the risk bound of the estimator $\widetilde{f^0_{\alpha,m}}$ deduced from Lemma 1.

Proposition 6. *Consider Model (1) under Assumptions [A1], [A2'], [A3], [A4] and [A7]. Assume moreover that f_α is integrable and square-integrable and observations for $t_0 = 0$ are available, then the estimator defined by (31) satisfies*

(33)

$$\mathbb{E}(\|\widetilde{f^0_{\alpha,m}} - f_\alpha\|^2) \le \|f_{\alpha,m} - f_\alpha\|^2 + \frac{3}{\pi N}\int_{-\pi m}^{\pi m}\frac{du}{|f_\varepsilon^*(u)|^2} + 4C_1 \min_{p\in\{1,\dots,4\}}\left(\frac{N^{p/4}}{2\pi}\int_{-\pi m}^{\pi m}\frac{|f_\alpha^*(u)|^2}{|f_\varepsilon^*(u)|^{2p}}du\right),$$

where C_1 is defined in Lemma 1.

The first two terms of the right-hand-side of inequality (33) are the standard terms, which are also obtained for known f_ε^*. The last term comes from the substitution of f_ε^* by its estimate. The main difference with the deconvolution estimators studied in Comte and Lacour [2011] comes from the fact that we replace f_ε^* by a truncated estimator based on $\widehat{(f_\varepsilon^*)^4}$, while in Comte and Lacour [2011], the truncated estimator is based on an estimator of f_ε^*. We notice that a similar (but still different) phenomenon happens for $\widetilde{f_{\beta,m}}$ where we replace $(f_\varepsilon^*)^2$ by a truncated estimator based on $\widehat{(f_\varepsilon^*)^4}$. The risk bound for $\widetilde{f_{\beta,m}}$ is in fact similar to the one found in Comte et al. [2011].

If f_α is very smooth, and in particular much smoother than f_ε, then the integrals $\int_{-\pi m}^{\pi m}|f_\alpha^*(u)|^2/(f_\varepsilon^*(u))^{2p}du$ may be convergent and the last term negligible. For instance, if α is Gaussian and ε is Laplace i.e. $\varepsilon = \sigma_\varepsilon\eta/\sqrt{2}$ where η has density $f_\eta(x) = e^{-|x|}/2$, then, for $p = 1, 2, 3, 4$,

$$\int_{-\pi m}^{\pi m}\frac{|f_\alpha^*(u)|^2}{(f_\varepsilon^*(u))^{2p}}du \le \kappa_p$$

and the last term is less than $(2C_1\kappa_4/\pi)/N$. If α is Gaussian with variance σ_α^2 and ε is Gaussian with variance σ_ε^2, then the same behavior happens provided that $\sigma_\alpha^2 > 4\sigma_\varepsilon^2$.

Practically, our estimator is $\widetilde{f^0_{\alpha,\tilde m_\alpha}}$ defined by (31) with

$$\tilde m_\alpha = \arg\min_{m\in\widetilde{\mathcal{M}}_{\alpha,N}}\left\{-\|\widetilde{f^0_{\alpha,m}}\|^2 + \widetilde{\mathrm{pen}}^0_\alpha(m)\right\}$$

with

$$\widetilde{\mathrm{pen}}^0_\alpha(m) = \tilde\kappa^0_\alpha\frac{\log(2\pi\hat I)}{\log(m+1)}\frac{2\pi\hat I}{N}, \quad \hat I = \frac{1}{2\pi}\int_{-\pi m}^{\pi m}\frac{du}{\widetilde{(f_\varepsilon^*)^2}(u)} \tag{34}$$

and $\widetilde{\mathcal{M}}_{\alpha,N} = \{m\in\{1,\dots,N\}, \widetilde{\mathrm{pen}}^0_\alpha(m)\le 1\}$.

A similar study can be performed for the estimator $\widetilde{f_{\alpha,m}}$.

		$\sigma_\varepsilon = 1/10$		$\sigma_\varepsilon = 1/4$		$\sigma_\varepsilon = 1/2$	
	Estimator	$N = 200$	2000	200	2000	200	2000
α Gaussian	$\widehat{f^0_{\alpha,m}}$	0.324	0.037	0.402	0.048	0.745	0.105
	$\widetilde{f^0_{\alpha,m}}$	0.243	0.082	0.266	0.103	0.544	0.113
	$\widehat{f_{\alpha,m}}$	0.469	0.055	0.520	0.094	2.022	1.014
	$\widetilde{f_{\alpha,m}}$	0.700	0.180	0.736	0.122	1.170	0.228
β Gaussian	$\widehat{f_{\beta,m}}$	0.427	0.044	0.313	0.032	0.402	0.041
	$\widetilde{f_{\beta,m}}$	0.301	0.032	0.285	0.042	0.535	0.151
α Gaussian	$\widehat{f^0_{\alpha,m}}$	0.335	0.036	0.417	0.048	0.744	0.111
	$\widetilde{f^0_{\alpha,m}}$	0.238	0.074	0.271	0.105	0.434	0.118
	$\widehat{f_{\alpha,m}}$	0.469	0.050	0.538	0.097	2.080	1.046
	$\widetilde{f_{\alpha,m}}$	0.448	0.092	0.770	0.162	1.151	0.234
β Mixed	$\widehat{f_{\beta,m}}$	1.500	0.312	2.288	0.425	6.648	1.311
Gaussian	$\widetilde{f_{\beta,m}}$	1.407	0.211	6.578	2.107	15.074	10.091
α Mixed	$\widehat{f^0_{\alpha,m}}$	1.690	0.339	6.116	0.684	10.166	5.203
Gaussian	$\widetilde{f^0_{\alpha,m}}$	2.088	0.353	6.168	0.703	14.302	6.014
	$\widehat{f_{\alpha,m}}$	5.673	2.269	22.945	9.926	36.288	34.551
	$\widetilde{f_{\alpha,m}}$	3.328	0.700	14.279	6.555	34.705	31.040
β Gaussian	$\widehat{f_{\beta,m}}$	0.424	0.053	0.283	0.034	0.428	0.043
	$\widetilde{f_{\beta,m}}$	0.267	0.036	0.297	0.047	0.522	0.156
α Gamma	$\widehat{f^0_{\alpha,m}}$	0.362	0.047	0.483	0.053	0.896	0.125
	$\widetilde{f^0_{\alpha,m}}$	0.288	0.057	0.348	0.112	0.601	0.173
	$\widehat{f_{\alpha,m}}$	0.476	0.053	0.618	0.126	2.438	1.221
	$\widetilde{f_{\alpha,m}}$	0.612	0.118	0.730	0.179	1.477	0.307
β Gamma	$\widehat{f_{\beta,m}}$	0.395	0.053	0.351	0.047	0.410	0.080
	$\widetilde{f_{\beta,m}}$	0.315	0.047	0.335	0.054	0.517	0.183
α Gamma	$\widehat{f^0_{\alpha,m}}$	0.402	0.047	0.412	0.047	0.821	0.113
	$\widetilde{f^0_{\alpha,m}}$	0.356	0.050	0.313	0.108	0.514	0.164
	$\widehat{f_{\alpha,m}}$	0.653	0.060	0.586	0.115	2.368	1.232
	$\widetilde{f_{\alpha,m}}$	0.859	0.106	0.876	0.136	1.295	0.289
β Mixed	$\widehat{f_{\beta,m}}$	1.247	0.360	1.228	0.465	2.848	0.639
Gamma	$\widetilde{f_{\beta,m}}$	1.194	0.331	1.457	0.689	6.769	1.790

Table 1. Empirical MISE $\times 100$ computed from 100 simulated datasets with ε Laplace. Estimators of f_α are $\widehat{f^0_{\alpha,m}}$, $\widetilde{f^0_{\alpha,m}}$, $\widehat{f_{\alpha,m}}$ and $\widetilde{f_{\alpha,m}}$. Estimators of f_β are $\widehat{f_{\beta,m}}$ and $\widetilde{f_{\beta,m}}$.

6. Simulation study

We consider simulations of Model (1) with noise having either Laplace density ($\sigma_\varepsilon\eta$ with $f_\eta(x) = e^{-\sqrt{2}|x|}/\sqrt{2}$) or Gaussian $\mathcal{N}(0, \sigma_\varepsilon^2)$ density and $t_j = j\Delta$, $\Delta = 2$, $J = 6$. For both α and β distributions, we experiment four possibilities:

- Gaussian, $\mathcal{N}(0, 1)$,

- Mixed Gaussian, with $0.3\mathcal{N}(-1,(1/4)^2)+0.7\mathcal{N}(1,(1/4)^2)$,
- Gamma distribution, $\gamma(25,1/25)/5$,
- Mixed Gamma distribution, $[0.3\gamma(2,1/2)+0.7\gamma(20,1/5)]/\sqrt{3}$.

All these densities are calibrated so that their variance is approximately 1. We consider three values of σ_ε: $1/10$, $1/4$, $1/2$. This means that the ratios of standard deviations of signal over noise are equal to 10, 4 and 2. In the last case, there is a lot of noise in the model, and the robustness of the procedure is really tested. Moreover, to see the improvement due to sample sizes, we take two values for N, $N=200$ and $N=2000$.

We compare the performances in term of MISE computed over 100 samples of the estimators

- $\widehat{f_{\beta,m}}$ as given by (7) (known noise) with model selection as described in Section 3.3 with constant κ_β of penalty (15) equal to 0.5 for Laplace errors and 0.001 for Gaussian errors,
- $\widetilde{f_{\beta,m}}$ as given by (30) (unknown noise) with model selection described in Section 5.2 and constant $\tilde{\kappa}_\beta$ equal to 0.5,
- $\widehat{f_{\beta,m_{opt}}}$ given by (7) (known Gaussian noise) with $m_{opt}=m_{0,\beta}(\sigma_\varepsilon)$ given by (21) with $\kappa'_\beta=1$. Note that the value of κ'_β does not fulfill the constraint $\kappa'_\beta<1$ but other values seemed too small.

We also compare

- $\widehat{f^0_{\alpha,m}}$ given by (8) (known noise) with model selection as described in Section 3.3 with constant κ^0_α equal to 2,
- $\widehat{f_{\alpha,m}}$ given by (10) (known noise) and penalization device with constant κ_α equal to 0.5 in the Laplace case and 0.0001 in the Gaussian case,
- $\widetilde{f^0_{\alpha,m}}$ given by (31) (unknown noise) with model selection described in Section 5.3 and constant $\tilde{\kappa}^0_\alpha$ equal to 2,
- $\widetilde{f_{\alpha,m}}$ given by (32) (unknown noise) with constant $\tilde{\kappa}_\alpha$ equal to 10,
- $\widehat{f_{\alpha,m_{opt}}}$ (known Gaussian noise) with $m_{opt}=m_{0,\alpha}(\sigma_\varepsilon)$ given by (22) with $\kappa'_\alpha=0.5$.

The results are gathered in Tables 1 and 2.

Clearly, for the estimation of f_α, the estimator $\widehat{f^0_{\alpha,m}}$ based on the observations $Y_{k,0}$ at $t_0=0$ has better performances than the estimator $\widehat{f_{\alpha,m}}$ based on the other observations. For simple problems, $\widehat{f_{\alpha,m}}$ performs well but fails to recover the two bumps of the bimodal distributions, unless the noise level is very low ($\sigma_\varepsilon=1/10$). It is worth mentioning that estimating the noise characteristic function often improves the estimation: this has been already observed in Comte and Lacour [2011] and in Comte et al. [2011] recently. This may be due to the truncation of $\widehat{f_\varepsilon}$ which attenuates the small values of f_ε involved in the denominator of the estimators of f_β and f_α. When the true f_ε is used, these small values appearing in the denominator are not truncated.

The estimation of f_β is very satisfactory and quite stable even for bimodal densities. Globally, increasing the noise level does not degrade too much the results.

As expected, for both functions, increasing the sample size improves the estimation.

In the Gaussian case, we experiment the specific proposals of Section 4. The results are always really interesting, and rather convincing for both sample sizes 200 and 2000; the main exception corresponds to the case where α is bimodal, where the estimator fails to correctly estimate f_α even for small noise. This method can be used for estimating f_β and more cautiously f_α, if one is convinced that the noise is Gaussian.

7. Conclusion.

In this paper, we consider a linear mixed-effects model with random i.i.d. coefficients α_k and β_k and we study how to estimate their unknown distributions. We propose several solutions, depending on the available information about noise density. Since it is often assumed to be Gaussian, we show that specific strategies can be developed in this case. In the more realistic case where it is unknown, we also propose general solutions based on deconvolution strategies. All this material is tested on simulation experiments which show the relevance of the methods. These proposals are all new and very different from existing strategies for such models.

Several extensions may be considered in future works. First, we may wish to estimate the joint distribution of α and β, with known or unknown noise density. Secondly, remaining in a linear setting, we may add fixed regressors with constant coefficients to be estimated. Lastly, finding out if such strategies may be successful for more general nonlinear mixed-effects model remains an open question.

Appendix A. Pointwise risk bound of f_β

The variance of the estimator is

$$\begin{aligned}\mathrm{Var}\left(\widehat{f_{\beta,m}}(x)\right) &= \frac{4}{N^2J^2}\frac{1}{4\pi^2}\sum_{k=1}^{N}\mathrm{Var}\left(\sum_{j=1}^{J/2}\int_{-\pi m}^{\pi m} e^{-ixu}\frac{e^{iuZ_{k,j}}}{|f_\varepsilon^*(u/\Delta_j)|^2}\,du\right)\\ &= \frac{4}{NJ^2}\frac{1}{4\pi^2}\sum_{j=1}^{J/2}\mathrm{Var}\left(\int_{-\pi m}^{\pi m} e^{-ixu}\frac{e^{iuZ_{k,j}}}{|f_\varepsilon^*(u/\Delta_j)|^2}\,du\right)\\ &+\frac{4}{NJ^2}\frac{1}{4\pi^2}\sum_{j,j'=1,j\neq j'}^{J/2}\mathrm{cov}\left(\int_{-\pi m}^{\pi m} e^{-ixu}\frac{e^{iuZ_{k,j}}}{|f_\varepsilon^*(u/\Delta_j)|^2}\,du,\int_{-\pi m}^{\pi m} e^{-ixz}\frac{e^{izZ_{k,j'}}}{\left|f_\varepsilon^*(z/\Delta_{j'})\right|^2}\,dz\right).\end{aligned}$$

For the first part of the decomposition, we have

$$\begin{aligned}&\mathrm{Var}\left(\int_{-\pi m}^{\pi m} e^{-ixu}\frac{e^{iuZ_{k,j}}}{|f_\varepsilon^*(u/\Delta_j)|^2}\,du\right)\leq\mathbb{E}\left(\left|\int_{-\pi m}^{\pi m} e^{-ixu}\,\frac{e^{iuZ_{k,j}}}{|f_\varepsilon^*(u/\Delta_j)|^2}\,du\right|^2\right)\\ \leq\ &\mathbb{E}\left(\int_{-\pi m}^{\pi m}\int_{-\pi m}^{\pi m} e^{-ix(u-z)}\,\frac{e^{i(u-z)Z_{k,j}}}{|f_\varepsilon^*(u/\Delta_j)|^2\,|f_\varepsilon^*(z/\Delta_j)|^2}\,du\,dz\right)\\ \leq\ &\int_{-\pi m}^{\pi m}\int_{-\pi m}^{\pi m} e^{-ix(u-z)}\,\frac{f_{Z_j}^*(u-z)}{|f_\varepsilon^*(u/\Delta_j)|^2\,|f_\varepsilon^*(z/\Delta_j)|^2}\,du\,dz\\ \leq\ &\int_{-\pi m}^{\pi m}\frac{1}{|f_\varepsilon^*(u/\Delta_j)|^4}\,du\int|f_{Z_j}^*(w)|dw\end{aligned}$$

	Estimator	$\sigma_\varepsilon = 1/10$		$\sigma_\varepsilon = 1/4$		$\sigma_\varepsilon = 1/2$	
		$N = 200$	2000	200	2000	200	2000
α Gaussian	$\widehat{f^0_{\alpha,m}}$	0.378	0.032	0.475	0.041	0.839	0.115
	$\widetilde{f^0_{\alpha,m}}$	0.273	0.078	0.281	0.102	0.871	0.115
	$\widehat{f_{\alpha,m}}$	0.999	0.100	0.463	0.119	3.448	3.357
	$\widetilde{f_{\alpha,m}}$	0.597	0.060	0.878	0.124	2.891	0.855
	$\widehat{f_{\alpha,m_{opt}}}$	0.519	0.066	1.219	0.467	7.883	5.750
β Gaussian	$\widehat{f_{\beta,m}}$	1.542	0.171	0.448	0.058	0.245	0.044
	$\widetilde{f_{\beta,m}}$	1.081	0.137	1.939	0.576	1.279	0.401
	$\widehat{f_{\beta,m_{opt}}}$	1.201	0.154	0.339	0.051	0.449	0.126
α Gaussian	$\widehat{f^0_{\alpha,m}}$	0.339	0.034	0.405	0.049	0.769	0.116
	$\widetilde{f^0_{\alpha,m}}$	0.251	0.073	0.276	0.103	0.771	0.119
	$\widehat{f_{\alpha,m}}$	1.057	0.106	0.464	0.125	3.418	3.358
	$\widetilde{f_{\alpha,m}}$	0.519	0.062	0.845	0.128	2.770	0.834
	$\widehat{f_{\alpha,m_{opt}}}$	0.552	0.071	1.193	0.470	7.883	5.750
β Mixed	$\widehat{f_{\beta,m}}$	1.423	0.150	5.705	3.745	0.225	17.173
Gaussian	$\widetilde{f_{\beta,m}}$	1.945	0.189	3.910	0.666	1.279	2.659
	$\widehat{f_{\beta,m_{opt}}}$	1.212	0.146	11.329	6.843	0.449	32.819
α Mixed	$\widehat{f^0_{\alpha,m}}$	1.758	0.354	6.854	1.298	16.623	7.780
Gaussian	$\widetilde{f^0_{\alpha,m}}$	2.270	0.391	6.785	2.300	21.627	13.979
	$\widehat{f_{\alpha,m}}$	5.586	3.605	30.503	30.099	37.200	37.163
	$\widetilde{f_{\alpha,m}}$	6.528	0.917	22.195	13.291	36.105	34.323
	$\widehat{f_{\alpha,m_{opt}}}$	10.692	6.270	35.626	34.765	40.127	38.672
β Gaussian	$\widehat{f_{\beta,m}}$	1.464	0.163	0.534	0.052	0.253	0.040
	$\widetilde{f_{\beta,m}}$	0.941	0.086	2.821	0.795	1.949	0.552
	$\widehat{f_{\beta,m_{opt}}}$	1.186	0.145	0.386	0.047	0.463	0.126
α Gamma	$\widehat{f^0_{\alpha,m}}$	0.403	0.048	0.481	0.056	0.912	0.162
	$\widetilde{f^0_{\alpha,m}}$	0.330	0.052	0.336	0.117	0.783	0.177
	$\widehat{f_{\alpha,m}}$	1.106	0.102	0.519	0.180	3.852	3.765
	$\widetilde{f_{\alpha,m}}$	0.648	0.054	1.020	0.175	3.005	1.015
	$\widehat{f_{\alpha,m_{opt}}}$	0.572	0.070	1.418	0.617	8.208	6.129
β Gamma	$\widehat{f_{\beta,m}}$	1.578	0.155	0.497	0.056	0.339	0.053
	$\widetilde{f_{\beta,m}}$	0.832	0.114	1.927	0.574	1.573	0.484
	$\widehat{f_{\beta,m_{opt}}}$	1.224	0.143	0.390	0.051	0.591	0.197
α Gamma	$\widehat{f^0_{\alpha,m}}$	0.364	0.045	0.462	0.050	0.948	0.158
	$\widetilde{f^0_{\alpha,m}}$	0.310	0.057	0.330	0.128	0.826	0.171
	$\widehat{f_{\alpha,m}}$	0.968	0.085	0.483	0.176	3.839	3.757
	$\widetilde{f_{\alpha,m}}$	0.574	0.056	1.008	0.173	3.002	0.977
	$\widehat{f_{\alpha,m_{opt}}}$	0.534	0.064	1.404	0.611	8.199	6.128
β Mixed	$\widehat{f_{\beta,m}}$	1.743	0.305	1.120	0.551	5.719	4.999
Gamma	$\widetilde{f_{\beta,m}}$	2.032	0.342	2.971	1.054	2.959	1.016
	$\widehat{f_{\beta,m_{opt}}}$	1.418	0.308	2.229	0.977	7.094	6.656

TABLE 2. Empirical MISE $\times 100$ computed from 100 simulated datasets with ε Gaussian. Estimators of f_α are $\widehat{f^0_{\alpha,m}}$, $\widetilde{f^0_{\alpha,m}}$, $\widetilde{f_{\alpha,m}}$ and $\widehat{f_{\alpha,m_{opt}}}$. Estimators of f_β are $\widehat{f_{\beta,m}}$, $\widetilde{f_{\beta,m}}$ and $\widehat{f_{\beta,m_{opt}}}$.

by applying bidimensional Cauchy Schwarz inequality with respect to the measure $|f^*_{Z_j}(u-z)|dudz$ and by using then Fubini. Thus

$$\begin{aligned}\mathrm{Var}\left(\int_{-\pi m}^{\pi m} e^{-ixu}\frac{e^{iuZ_{k,j}}}{|f^*_\varepsilon(u/\Delta_j)|^2}\,du\right) &\leq \int_{-\pi m}^{\pi m}\frac{1}{|f^*_\varepsilon(u/\Delta_j)|^4}\,du\int|f^*_\varepsilon(w/\Delta_j)|^2dw\\ &\leq \|f_\varepsilon\|^2\Delta_j\int_{-\pi m}^{\pi m}\frac{1}{|f^*_\varepsilon(u/\Delta_j)|^4}\,du\end{aligned}$$

For the second part of the variance decomposition, we have

$$\begin{aligned}&|\mathrm{cov}\left(\int_{-\pi m}^{\pi m} e^{-ixu}\frac{e^{iuZ_{k,j}}}{|f^*_\varepsilon(u/\Delta_j)|^2}\,du,\int_{-\pi m}^{\pi m} e^{-ixz}\frac{e^{izZ_{k,j'}}}{\left|f^*_\varepsilon(z/\Delta_{j'})\right|^2}\,dz\right)|\\ =\;& |\iint_{[-\pi m,\pi m]^2} e^{-ix(u-z)}(f^*_\beta(u-z)-f^*_\beta(u)f^*_\beta(-z))dudz|\\ \leq\;& 4\pi^2 m\int|f^*_\beta(w)|dw\end{aligned}$$

By gathering the terms, we obtain

$$Var\left(\widehat{f_{\beta,m}}(x)\right)\leq\frac{1}{\pi^2NJ^2}\sum_{j=1}^{J/2}\|f_\varepsilon\|^2\Delta_j\int_{-\pi m}^{\pi m}\frac{1}{|f^*_\varepsilon(u/\Delta_j)|^4}\,du+C\frac{4}{\pi}\frac{m\int|f^*_\beta(w)|dw}{N}.$$

When the observation times are equally spaced with $\Delta_j=\Delta$, we obtain the following bound

$$Var\left(\widehat{f_{\beta,m}}(x)\right)\leq 2\|f_\varepsilon\|^2\Delta\frac{\int_{-\pi m}^{\pi m}\frac{1}{|f^*_\varepsilon(u/\Delta_j)|^4}\,du}{NJ}+\frac{4\|f^*_\beta\|_1}{\pi}\frac{m}{N}.$$

Thus the MSE is bounded by

Proposition 7. *Consider Model (1) under Assumptions [A1]–[A5] with f_β integrable and square integrable, and $f_\varepsilon\in\mathbb{L}_2(\mathbb{R})$ then,*

$$\begin{aligned}\mathbb{E}\left(\left|\widehat{f_{\beta,m}}(x)-f_\beta(x)\right|^2\right) &\leq (\frac{1}{2\pi}\int_{|u|\geq\pi m}|f^*_\beta(u)|du)^2\\ &+\frac{2\|f_\varepsilon\|^2}{\pi}\frac{1}{NJ^2}\sum_{j=1}^{J/2}\Delta_j\int_{-\pi m}^{\pi m}\frac{1}{|f^*_\varepsilon(u/\Delta_j)|^4}\,du+\frac{4\|f^*_\beta\|_1}{\pi}\frac{m}{N}.\end{aligned}\tag{35}$$

Appendix B. Proofs

B.1. **Proof of Proposition 2.** We can easily calculate the expectation of the estimator $\widehat{f_{\alpha,m}}(x)$ defined by (10). We have

$$\begin{aligned}\mathbb{E}\left(\widehat{f_{\alpha,m}}(x)\right) &= \frac{1}{N}\sum_{k=1}^{N}\frac{1}{2\pi}\int_{-\pi m}^{\pi m} e^{-iux}\frac{\mathbb{E}\left(e^{iV_{k,1}u/p_1}\right)}{f_\varepsilon^*\left(\frac{u}{p_1t_2}\right)\, f_\varepsilon^*\left(\frac{-u}{p_1t_1}\right)}\,du \\ &= \frac{1}{N}\sum_{k=1}^{N}\frac{1}{2\pi}\int_{-\pi m}^{\pi m} e^{-iux}\frac{f^*_{V_{k,1}}(u/p_1)}{f_\varepsilon^*\left(\frac{u}{p_1t_2}\right)\, f_\varepsilon^*\left(\frac{-u}{p_1t_1}\right)}\,du \\ &= \frac{1}{2\pi}\int_{-\pi m}^{\pi m} e^{-iux} f^*_\alpha(u)\,du = f_{\alpha,m}(x)\end{aligned}$$

The integrated bias is therefore given by:

$$\left\|f_\alpha - \mathbb{E}\left(\widehat{f_{\alpha,m}}\right)\right\|^2 \leq \frac{1}{2\pi}\int_{|u|\geq \pi m}|f^*_\alpha(u)|^2\,du$$

The integrated variance is

$$\begin{aligned}\left\|\widehat{f_{\alpha,m}} - f_{\alpha,m}\right\|^2 &= \left\|\frac{1}{2\pi}\int_{-\pi m}^{\pi m} e^{-iux}\frac{\widehat{f^*_{V_{k,1}}}(u/p_1) - f^*_{V_{k,1}}(u/p_1)}{f_\varepsilon^*\left(\frac{u}{p_1t_2}\right)\, f_\varepsilon^*\left(\frac{-u}{p_1t_1}\right)}\,du\right\|^2 \\ &= \frac{1}{2\pi}\int_{-\pi m}^{\pi m}\left|\frac{\widehat{f^*_{V_{k,1}}}(u/p_1) - f^*_{V_{k,1}}(u/p_1)}{f_\varepsilon^*\left(\frac{u}{p_1t_2}\right)\, f_\varepsilon^*\left(\frac{u}{p_1t_1}\right)}\right|^2 du\end{aligned}$$

Therefore, as for (12),

$$\mathbb{E}\left(\left\|\widehat{f_{\alpha,m}} - f_{\alpha,m}\right\|^2\right) \leq \frac{1}{2\pi N}\int_{-\pi m}^{\pi m}\frac{1}{\left|f_\varepsilon^*\left(\frac{u}{p_1t_2}\right)\, f_\varepsilon^*\left(\frac{u}{p_1t_1}\right)\right|^2}\,du + \frac{m}{N}.$$

Gathering bias and variance bounds gives the result of Proposition 2. □

B.2. **Proof of Theorem 1.** Let $S_m = \{t \in \mathbb{L}_1 \cap \mathbb{L}_2, \text{ such that } t^* = t^*\mathbb{1}_{[-\pi m,\pi m]}\}$ be the subspace of $\mathbb{L}_2$ with functions having Fourier Transforms supported by $[-\pi m, \pi m]$. Now we can notice that $\widehat{f_{\beta,m}}$ is the minimizer over S_m of the contrast

$$\gamma_{N,J}(t) = \|t\|^2 - \frac{4}{NJ}\sum_{k=1}^{N}\sum_{j=1}^{J/2}\frac{1}{2\pi}\int t^*(-u)\frac{e^{iuZ_{k,j}}}{|f^*_\varepsilon(u/\Delta_j)|^2}du$$

since $\gamma_{N,J}(t) = \|t\|^2 - 2\langle t, \widehat{f_{\beta,m}}\rangle$. Then defining

$$\nu_{N,J}(t) = \frac{2}{NJ}\sum_{k=1}^{N}\sum_{j=1}^{J/2}\frac{1}{2\pi}\int t^*(-u)\frac{(e^{iuZ_{k,j}} - \mathbb{E}(e^{iuZ_{k,j}}))}{|f^*_\varepsilon(u/\Delta_j)|^2}du$$

we have the decomposition, for all functions s, t integrable and square integrable,

$$\gamma_{N,J}(t) - \gamma_{N,J}(s) = \|t - f_\beta\|^2 - \|s - f_\beta\|^2 - 2\nu_{N,J}(t - s). \tag{36}$$

By definition of $\hat{m}_\beta$, we have that $\forall m \in \mathcal{M}_{\beta,N}$,

$$\gamma_n(\widehat{f_{\beta,\hat{m}_\beta}}) + \text{pen}_\beta(\hat{m}_\beta) \leq \gamma_n(f_{\beta,m}) + \text{pen}_\beta(m)$$

and with (36), this yields, $\forall m \in \mathcal{M}_{\beta,n}$,

$$\begin{aligned}
\|\widehat{f_{\beta,\hat{m}_\beta}} - f_\beta\|^2 &\leq \|f_\beta - f_{\beta,m}\|^2 + \text{pen}_\beta(m) + 2\nu_{N,J}(\widehat{f_{\beta,\hat{m}_\beta}} - f_{\beta,m}) - \text{pen}_\beta(\hat{m}_\beta) \\
&\leq \|f_\beta - f_{\beta,m}\|^2 + \text{pen}_\beta(m) + 2\|\widehat{f_{\beta,\hat{m}_\beta}} - f_{\beta,m}\| \sup_{t \in S_m + S_{\hat{m}_\beta}} \nu_{N,J}(t) - \text{pen}_\beta(\hat{m}_\beta) \\
&\leq \|f_\beta - f_{\beta,m}\|^2 + \text{pen}_\beta(m) + \frac{1}{4}\|\widehat{f_{\beta,\hat{m}_\beta}} - f_{\beta,m}\|^2 \\
&\quad +4 \sup_{t \in S_m + S_{\hat{m}_\beta}} \nu^2_{N,J}(t) - \text{pen}_\beta(\hat{m}_\beta) \\
&\leq \|f_\beta - f_{\beta,m}\|^2 + 2\text{pen}_\beta(m) + \frac{1}{2}\|\widehat{f_{\beta,\hat{m}_\beta}} - f_\beta\|^2 + \frac{1}{2}\|f_\beta - f_{\beta,m}\|^2 \\
&\quad +4(\sup_{t \in S_m + S_{\hat{m}_\beta}} \nu^2_{N,J}(t) - p(\hat{m}_\beta, m))
\end{aligned}$$

where $4p(m, m') \leq \text{pen}_\beta(m) + \text{pen}_\beta(m')$. We can prove

Lemma 2. *Under the assumptions of Theorem 1,*

$$\mathbb{E}\left(\sup_{t \in S_m + S_{\hat{m}_\beta}} \nu^2_{N,J}(t) - p(\hat{m}_\beta, m) \right) \leq \frac{C'}{N}$$

Applying Lemma 2, we get $\forall m \in \mathcal{M}_{\beta,N}$,

$$\mathbb{E}(\|\widehat{f_{\beta,\hat{m}_\beta}} - f_\beta\|^2) \leq 3\|f_\beta - f_{\beta,m}\|^2 + 4\text{pen}_\beta(m) + \frac{C'}{N}.$$

This is the result of Theorem 1. □

Proof of Lemma 2. Let us define $\varphi(x) = \sin(\pi x)/(\pi x)$ and $\varphi_{m,\ell}(x) = \sqrt{m}\varphi(mx - \ell)$ for $\ell \in \mathbb{Z}$. We recall that $(\varphi_{m,\ell})_{\ell \in \mathbb{Z}}$ is an orthonormal basis of S_m and it fulfills $\sum_{\ell \in \mathbb{Z}} \varphi^2_\ell \leq m$.

Let us denote by

$$\eta_{k,j} = \frac{\varepsilon_{k,2j} - \varepsilon_{k,2j-1}}{\Delta_j}.$$

Then the $(\eta_{k,j})_{k,j}$ are independent. We split $\nu_{N,J}$ in two parts and write $\nu_{N,J}(t) = \nu^{(1)}_{N,J}(t) + \nu^{(2)}_{N,J}(t)$:

$$\nu^{(1)}_{N,J}(t) = \frac{2}{NJ} \sum_{k=1}^{N} \sum_{j=1}^{J/2} \frac{1}{2\pi} \int t^*(-u) \frac{e^{iu\beta_k}(e^{iu\eta_{k,j}} - \mathbb{E}(e^{iu\eta_{k,j}}))}{|f^*_\varepsilon(u/\Delta_j)|^2} du$$

and

$$\nu^{(2)}_{N,J}(t) = \frac{1}{N} \sum_{k=1}^{N} \frac{1}{2\pi} \int t^*(-u)(e^{iu\beta_k} - f^*_\beta(u))du = \frac{1}{N} \sum_{k=1}^{N} (t(\beta_k) - \langle t, f_\beta \rangle).$$

For $\nu^{(2)}_{N,J}(t)$, we find the bounds

$$\begin{aligned}\sup_{t\in S_m+S_{m'},\|t\|=1}[\nu^{(2)}_{N,J}(t)]^2 &\leq \sup_{t\in S_m+S_{m'},\|t\|=1}\sum_{\ell\in\mathbb{Z}} a^2_{m\vee m',\ell}\sum_{\ell\in\mathbb{Z}}[\nu^{(2)}_{N,J}(\varphi_{m\vee m',\ell})]^2\\ &\leq \sum_{\ell\in\mathbb{Z}}[\nu^{(2)}_{N,J}(\varphi_{m\vee m',\ell})]^2\end{aligned}$$

where $m\vee m'=\sup(m,m')$, and

$$\mathbb{E}\left(\sup_{t\in S_m+S_{m'},\|t\|=1}[\nu^{(2)}_{N,J}(t)]^2\right)\leq\sum_\ell\frac{1}{N}\mathrm{Var}(\varphi_{m\vee m',\ell})\leq\frac{m\vee m'}{N}:=H_2^2.$$

Moreover $\mathrm{Var}(t(\beta_1))\leq\mathbb{E}(t^2(\beta_1))\leq\|t\|_\infty\mathbb{E}(|t(\beta_1)|)\leq\|t\|_\infty\|t\|\|f_\beta\|$ and for any $t\in S_{m\vee m'}$, $\|t\|_\infty\leq\sqrt{m\vee m'}\|t\|$ (see Comte et al. [2006]) yield

$$\sup_{t\in S_m+S_{m'},\|t\|=1}\mathrm{Var}(t(\beta_1))\leq\sqrt{m\vee m'}\|f_\beta\|:=v_2$$

and

$$\sup_{t\in S_m+S_{m'},\|t\|=1}\|t\|_\infty\leq\sqrt{m\vee m'}:=b_2$$

Thus Talagrand Inequality implies

$$\begin{aligned}&\mathbb{E}\left(\sup_{t\in S_m+S_{\hat m_\beta}}[\nu^{(2)}_{N,j}(t)]^2-4\frac{m\vee\hat m_\beta}{N}\right)\leq\sum_{m'\in\mathcal{M}_{\beta,N}}\mathbb{E}\left(\sup_{t\in S_m+S_{m'}}[\nu^{(2)}_{N,j}(t)]^2-4\frac{m\vee m'}{N}\right)\\ \leq&\sum_{m'\in\mathcal{M}_{\beta,N}}K_1(\frac{v_2}{N}e^{-K_2\frac{NH_2^2}{v_2}}+\frac{b_2^2}{N^2}e^{-K_3\frac{NH_2}{b_2}})\\ \leq&\frac{K_1}{N}\sum_{m'\in\mathcal{M}_{\beta,N}}(\|f_\beta\|\sqrt{m\vee m'}e^{-K_2\sqrt{m\vee m'}/\|f_\beta\|}+\frac{m\vee m'}{N}e^{-K_3\sqrt{N}})\\ \leq&\frac{C'}{N}\end{aligned}$$

since the sums are convergent or bounded.

For $\nu^{(1)}_{N,J}$ we compute the same bounds conditionally to $(\beta_k)_{1\leq k\leq N}$. Then we proceed analogously to Comte et al. [2008] and Comte et al. [2006], and we get the result.□

B.3. **Proof of Proposition 3.** The bias order given in (13) under Assumption [A6], is $O(m^{-2b})$ which gives the announced order when choosing $m=m_{0,\beta}$. Moreover, the variance terms are made negligible by this choice. Indeed the integrated variance of $\widehat{f_{\beta,m}}$ is of order

$$\frac{(\pi m)^{-1}}{2NJ}\exp(2\sigma_\varepsilon^2\Delta_{min}^{-2}\pi^2m^2)+\frac{m}{N}$$

which leads to a variance of order

$$\frac{\sigma_\varepsilon\Delta_{min}^{-1}}{(NJ)^{1-\kappa'_\beta}\sqrt{\log(NJ)}}+\kappa'_\beta\frac{\sqrt{\log(NJ)}}{N2\pi\sigma_\varepsilon\Delta_{min}^{-1}}$$

and the convergence rate of order $(\log(NJ))^{-b}$. □

B.4. **Proof of Proposition 4.** The proof is based on the following decomposition

$$\mathbb{E}(\|\widehat{f_{\beta,\widehat{m_{0,\beta}}}}-f_\beta\|^2) \leq \mathbb{E}(\|\widehat{f_{\beta,\widehat{m_{0,\beta}}}}-f_\beta\|^2 \mathbb{1}_{|\widehat{\sigma_\varepsilon^2}-\sigma_\varepsilon^2|\leq\sigma_\varepsilon^2/2})+\mathbb{E}(\widehat{f_{\beta,\widehat{m_{0,\beta}}}}-f_\beta\|^2\mathbb{1}_{|\widehat{\sigma_\varepsilon^2}-\sigma_\varepsilon^2|>\sigma_\varepsilon^2/2}) \tag{37}$$

First, remark that when $|\widehat{\sigma_\varepsilon^2}-\sigma_\varepsilon^2| \leq \sigma_\varepsilon^2/2$, we have $\frac{1}{2}\sigma_\varepsilon^2 \leq \widehat{\sigma_\varepsilon^2} \leq \frac{3}{2}\sigma_\varepsilon^2$. Consequently $\frac{1}{\sqrt{3}}m_{0,\beta} \leq m_{0,\beta}(\widehat{\sigma_\varepsilon})/\sqrt{2} \leq m_{0,\beta}$ with $m_{0,\beta} = m_{0,\beta}(\sigma_\varepsilon)$. Thus, looking at (11), we get

$$\|\widehat{f_{\beta,\widehat{m_{0,\beta}}}} - f_{\beta,\widehat{m_{0,\beta}}}\|^2\mathbb{1}_{|\widehat{\sigma_\varepsilon^2}-\sigma_\varepsilon^2|\leq\sigma_\varepsilon^2/2} \leq \|\widehat{f_{\beta,m_{0,\beta}}} - f_{\beta,m_{0,\beta}}\|^2$$

and clearly as [A6] holds, we have

$$\|f_{\beta,\widehat{m_{0,\beta}}} - f_\beta\|^2\mathbb{1}_{|\widehat{\sigma_\varepsilon^2}-\sigma_\varepsilon^2|\leq\sigma_\varepsilon^2/2} \leq CL(\pi\widehat{m_{0,\beta}})^{-2b}\mathbb{1}_{|\widehat{\sigma_\varepsilon^2}-\sigma_\varepsilon^2|\leq\sigma_\varepsilon^2/2} \leq C'L(\pi m_{0,\beta})^{-2b}.$$

Therefore

$$\mathbb{E}(\|\widehat{f_{\beta,\widehat{m_{0,\beta}}}} - f_\beta\|^2\mathbb{1}_{|\widehat{\sigma_\varepsilon^2}-\sigma_\varepsilon^2|\leq\sigma_\varepsilon^2/2}) \leq C[\log(NJ)]^{-b}.$$

On the other hand, by using that $\widehat{m_{0,\beta}} \leq m_n$, we get

$$\|\widehat{f_{\beta,\widehat{m_{0,\beta}}}} - f_{\beta,\widehat{m_{0,\beta}}}\|^2 \leq CN$$

and $\|f_{\beta,\widehat{m_{0,\beta}}} - f_\beta\|^2 \leq \|f_\beta\|^2$, so that

$$\mathbb{E}(\|\widehat{f_{\beta,\widehat{m_{0,\beta}}}} - f_\beta\|^2\mathbb{1}_{|\widehat{\sigma_\varepsilon^2}-\sigma_\varepsilon^2|>\sigma_\varepsilon^2/2}) \leq CN\mathbb{P}(|\widehat{\sigma_\varepsilon^2}-\sigma_\varepsilon^2| > \sigma_\varepsilon^2/2).$$

The following lemma yields the result.

Lemma 3.

$$\mathbb{P}(|\widehat{\sigma_\varepsilon^2}-\sigma_\varepsilon^2| > \sigma_\varepsilon^2/2) \leq C/N^2. \tag{38}$$

Proof of Lemma 3. We take $J=3$ for simplicity. We know that $\hat\sigma_\varepsilon^2$ is a linear combination of $\hat\sigma_{Y.j}^2$, $j=1,2,3$ and (say) $\hat\sigma_{Y.2\,Y.3}^2$. Therefore as $\mathbb{P}(|X+Y|>a) \leq \mathbb{P}(|X|>a/2)+\mathbb{P}(|Y|>a/2)$, the result follows if we prove that for $j=1,2,3$,

$$\mathbb{P}(|\hat\sigma_{Y.j}^2 - \sigma_{.j}^2| > \frac{1}{2}c_j^2) \leq \frac{C_j}{N^2},\ \text{and}\ \mathbb{P}(|\hat\sigma_{Y.2\,Y.3}^2 - \sigma_{.2,3}^2| > \frac{1}{2}c_j^2) \leq \frac{C_4}{N^2}.$$

We provide few details for the first term and $j=3$. The bound relies on a Rosenthal inequality which states that for independent centered random variables $X_1,\dots,X_n$ admitting moments of order p

$$\mathbb{E}(|\sum_{i=1}^n X_i|^p) \leq c_p\left(\sum_{i=1}^n \mathbb{E}(|X_i|^p) + \left(\sum_{i=1}^n \mathbb{E}(X_i^2)\right)^{p/2}\right).$$

Consequently

$$\begin{aligned}\mathbb{P} &:= \mathbb{P}(|\hat\sigma_{Y.3}^2 - (\sigma_\alpha^2 + t_3^2\sigma_\beta^2 + 2t_3\sigma_{\alpha\beta} + \sigma_\varepsilon^2)| > \lambda)\\ &= \mathbb{P}(|\frac{1}{N}\sum_{k=1}^N (A_k^2 - \mathbb{E}(A_k^2)) - (\bar A - \mathbb{E}(\bar A))^2| > \lambda)\end{aligned}$$

where $A_k = (\alpha_k - \mathbb{E}(\alpha_k)) + t_3(\beta_k - \mathbb{E}(\beta_k)) + \varepsilon_{k,3}$ and $\bar{A} = N^{-1}\sum_{k=1}^N A_k$. Then by Markov inequality stating that $\mathbb{P}(|X| > \lambda) \le \mathbb{E}(|X|^p)/\lambda^p$, we get

$$\begin{aligned}
\mathbb{P} &\le \mathbb{P}(|\frac{1}{N}\sum_{k=1}^N (A_k^2 - \mathbb{E}(A_k^2))| > \lambda/2) + \mathbb{P}((\bar{A} - \mathbb{E}(\bar{A}))^2 > \lambda/2) \\
&\le \left(\frac{2}{\lambda}\right)^4 \mathbb{E}\left(\left|\frac{1}{N}\sum_{k=1}^N (A_k^2 - \mathbb{E}(A_k^2))\right|^4\right) + \left(\sqrt{\frac{2}{\lambda}}\right)^4 \mathbb{E}(|\bar{A} - \mathbb{E}(\bar{A})|^4) \\
&\le \left(\frac{2}{\lambda}\right)^4 \frac{\mathbb{E}(A_1^4)}{N^2} + \frac{4}{\lambda^2}(N^{-3}\mathbb{E}(A_1^4) + N^{-2}(\mathbb{E}(A_1^2))^2)
\end{aligned}$$

under a moment condition of order 4 for α, β and ε. This gives the announced result.

B.5. **Proof of Lemma 1.** (1) We first prove assertion (1) of the lemma. We denote by

$$R(u) = \frac{1}{\widetilde{(f_\varepsilon^*)^2}(u)} - \frac{1}{(f_\varepsilon^*)^2(u)}.$$

First we write a decomposition:

$$\begin{aligned}
\mathbb{E}(|R(u)|^2) &= \mathbb{E}\left(\frac{\mathbb{1}_{\widehat{(f_\varepsilon^*)^4}(u) < N^{-1/2}}}{(f_\varepsilon^*)^4(u)}\right) \\
&\quad + \frac{1}{(f_\varepsilon^*)^4(u)}\mathbb{E}\left(\mathbb{1}_{\widehat{(f_\varepsilon^*)^4}(u) \ge N^{-1/2}} \frac{\left(\widehat{(f_\varepsilon^*)^4}(u) - (f_\varepsilon^*)^4(u)\right)^2}{\widehat{(f_\varepsilon^*)^4}(u)\left[\sqrt{\widehat{(f_\varepsilon^*)^4}(u)} + (f_\varepsilon^*(u))^2\right]^2}\right)
\end{aligned}$$

Then, using that $1/[\sqrt{\widehat{(f_\varepsilon^*)^4}(u)} + (f_\varepsilon^*(u))^2]^2 \le 1/\widehat{(f_\varepsilon^*)^4}(u)$, we obtain

$$\begin{aligned}
\mathbb{E}(|R(u)|^2) &\le \frac{1}{(f_\varepsilon^*)^4(u)} + \frac{N}{(f_\varepsilon^*)^4(u)}\mathbb{E}\left[\left(\widehat{(f_\varepsilon^*)^4}(u) - (f_\varepsilon^*)^4(u)\right)^2\right] \\
&\le \frac{2}{(f_\varepsilon^*)^4(u)}
\end{aligned}$$

which is the first term of the bound.

(i) if $(f_\varepsilon^*)^4(u) \le 2N^{-1/2}$, we have $N^{-1/2}/(f_\varepsilon^*)^8(u) \le 2N^{-1}/(f_\varepsilon^*)^{12}(u)$. Moreover, starting in the same way as above, using that $1/[\sqrt{\widehat{(f_\varepsilon^*)^4}(u)} + (f_\varepsilon^*(u))^2]^2 \le 1/(f_\varepsilon^*(u))^4$, we also have

$$\begin{aligned}
\mathbb{E}(|R(u)|^2) &\le \frac{1}{(f_\varepsilon^*)^4(u)} + \frac{N^{1/2}}{(f_\varepsilon^*)^8(u)}\mathbb{E}\left[\left(\widehat{(f_\varepsilon^*)^4}(u) - (f_\varepsilon^*)^4(u)\right)^2\right] \\
&\le \frac{1}{(f_\varepsilon^*)^4(u)} + \frac{N^{-1/2}}{(f_\varepsilon^*)^8(u)} \le 3\frac{N^{-1/2}}{(f_\varepsilon^*)^8(u)}
\end{aligned}$$

(ii) If $(f_\varepsilon^*)^4 > 2N^{-1/2}$, using the Bernstein Inequality yields:

$$\begin{aligned}
\mathbb{P}\left(|\widehat{(f_\varepsilon^*)^4}(u)| < N^{-1/2}\right) &\leq \mathbb{P}\left(|\widehat{(f_\varepsilon^*)^4}(u) - (f_\varepsilon^*)^4(u)| > (f_\varepsilon^*)^4(u) - N^{-1/2}\right) \\
&\leq \mathbb{P}\left(|\widehat{(f_\varepsilon^*)^4}(u) - (f_\varepsilon^*)^4(u)| > (f_\varepsilon^*)^4(u)/2\right) \\
&\leq 2\exp\left(-N(f_\varepsilon^*)^8(u)/16\right) \\
&\leq O\left(N^{-1}(f_\varepsilon^*(u))^{-8}\right)
\end{aligned}$$

and completing the decomposition above, this yields

$$\begin{aligned}
\mathbb{E}(|R(u)|^2) &\leq \mathbb{E}\left(\frac{\mathbb{1}_{\widehat{(f_\varepsilon^*)^4}(u) < N^{-1/2}}}{(f_\varepsilon^*)^4(u)}\right) \\
&+ \frac{1}{(f_\varepsilon^*)^8(u)}\mathbb{E}\left(\mathbb{1}_{\widehat{(f_\varepsilon^*)^4}(u) \geq N^{-1/2}} \frac{\left(\widehat{(f_\varepsilon^*)^4}(u) - (f_\varepsilon^*)^4(u)\right)^2}{\left(\sqrt{\widehat{(f_\varepsilon^*)^4}(u)} + (f_\varepsilon^*)^2(u)\right)^2}\right) \\
&+ \frac{1}{(f_\varepsilon^*)^4(u)}\mathbb{E}\left(\mathbb{1}_{\widehat{(f_\varepsilon^*)^4}(u) \geq N^{-1/2}} \frac{\left(\widehat{(f_\varepsilon^*)^4}(u) - (f_\varepsilon^*)^4(u)\right)^2}{\left(\sqrt{\widehat{(f_\varepsilon^*)^4}(u)} + (f_\varepsilon^*)^2(u)\right)^2}\left(\frac{1}{\widehat{(f_\varepsilon^*)^4}(u)} - \frac{1}{(f_\varepsilon^*)^4(u)}\right)\right)
\end{aligned}$$

Then, using that $1/[\sqrt{\widehat{(f_\varepsilon^*)^4}(u)} + (f_\varepsilon^*(u))^2]^2 \leq 1/(f_\varepsilon^*(u))^4$, we get

$$\begin{aligned}
\mathbb{E}(|R(u)|^2) &\leq \frac{1}{(f_\varepsilon^*)^4(u)}\mathbb{P}\left(\widehat{(f_\varepsilon^*)^4}(u) < N^{-1/2}\right) + \frac{N^{-1}}{(f_\varepsilon^*)^{12}(u)} \\
&+ \frac{N^{1/2}}{(f_\varepsilon^*)^{12}(u)}\mathbb{E}\left[\left|\widehat{(f_\varepsilon^*)^4}(u) - (f_\varepsilon^*)^4(u)\right|^3\right] \\
&\leq \frac{CN^{-1}}{(f_\varepsilon^*)^{12}(u)} + \frac{N^{-1}}{(f_\varepsilon^*)^{12}(u)} + \frac{N^{1/2}}{(f_\varepsilon^*)^{12}(u)}N^{-3/2} \leq c\frac{N^{-1}}{(f_\varepsilon^*)^{12}(u)}
\end{aligned}$$

Thus, in that case where $N^{-1/2}/(f_\varepsilon^*)^8(u) \geq 2N^{-1}/(f_\varepsilon^*)^{12}(u)$, we get

$$\mathbb{E}(|R(u)|^2) \leq \frac{N^{-1}}{|f_\varepsilon^*(u)|^{12}}.$$

This ends the proof of assertion (1).

(2) We now prove assertion (2). Set

$$R_0(u) = \frac{1}{\widetilde{f_\varepsilon^*}(u)} - \frac{1}{f_\varepsilon^*(u)}.$$

First we write:

$$\mathbb{E}(|R_0(u)|^2) = \mathbb{E}\left(\frac{\mathbb{1}_{\widehat{(f_\varepsilon^*)^4}(u)<N^{-1/2}}}{(f_\varepsilon^*)^2(u)}\right)$$
$$+\frac{1}{(f_\varepsilon^*)^2(u)}\mathbb{E}\left(\mathbb{1}_{\widehat{(f_\varepsilon^*)^4}(u)\geq N^{-1/2}}\frac{\left(\widehat{(f_\varepsilon^*)^4}(u)-(f_\varepsilon^*)^4(u)\right)^2}{\left(\widehat{(f_\varepsilon^*)^4}(u)\right)^{1/2}\left[\sqrt{\widehat{(f_\varepsilon^*)^4}(u)}+(f_\varepsilon^*(u))^2\right]^2\left[\left(\widehat{(f_\varepsilon^*)^4}(u)\right)^{1/4}+f_\varepsilon^*(u)\right]^2}\right)$$
$$\leq \frac{1}{(f_\varepsilon^*)^2(u)}+\frac{N}{(f_\varepsilon^*)^2(u)}\mathbb{E}\left[\left(\widehat{(f_\varepsilon^*)^4}(u)-(f_\varepsilon^*)^4(u)\right)^2\right]\leq\frac{2}{(f_\varepsilon^*)^2(u)}$$

which is the first term of the bound.

(i) if $(f_\varepsilon^*)^4(u)\leq 2N^{-1/2}$, we have

$$\frac{N^{-3/4}}{(f_\varepsilon^*)^8(u)}\wedge\frac{N^{-1/2}}{(f_\varepsilon^*)^6(u)}\wedge\frac{N^{-1/4}}{(f_\varepsilon^*)^4(u)}\leq\frac{N^{-1}}{(f_\varepsilon^*)^{10}(u)}.$$

Moreover, starting in the same way as above, we also have

$$\begin{aligned}\mathbb{E}(|R_0(u)|^2) &\leq \frac{1}{(f_\varepsilon^*)^2(u)}+\left(\frac{N^{1/4}}{(f_\varepsilon^*)^8(u)}\wedge\frac{N^{1/2}}{(f_\varepsilon^*)^6(u)}\wedge\frac{N^{3/4}}{(f_\varepsilon^*)^4(u)}\right)\mathbb{E}\left[\left(\widehat{(f_\varepsilon^*)^4}(u)-(f_\varepsilon^*)^4(u)\right)^2\right]\\ &\leq \frac{1}{(f_\varepsilon^*)^2(u)}+\frac{N^{-3/4}}{(f_\varepsilon^*)^8(u)}\wedge\frac{N^{-1/2}}{(f_\varepsilon^*)^6(u)}\wedge\frac{N^{-1/4}}{(f_\varepsilon^*)^4(u)}\\ &\leq 3\frac{N^{-3/4}}{(f_\varepsilon^*)^8(u)}\wedge\frac{N^{-1/2}}{(f_\varepsilon^*)^6(u)}\wedge\frac{N^{-1/4}}{(f_\varepsilon^*)^4(u)}\end{aligned}$$

where the last line follows from the assumption $(f_\varepsilon^*)^4(u)\leq 2N^{-1/2}$.

(ii) If $(f_\varepsilon^*)^4 > 2N^{-1/2}$, we use as in the proof of (1) that $\mathbb{P}(|\widehat{(f_\varepsilon^*)^4}(u)|<N^{-1/2})\leq C_1\left(N^{-1}(f_\varepsilon^*(u))^{-8}\right)$, and completing the decomposition above, this yields

$$\mathbb{E}(|R_0(u)|^2)\leq\mathbb{E}\left(\frac{\mathbb{1}_{\widehat{(f_\varepsilon^*)^4}(u)<N^{-1/2}}}{(f_\varepsilon^*)^2(u)}\right)$$
$$+\frac{1}{(f_\varepsilon^*)^4(u)}\mathbb{E}\left(\mathbb{1}_{\widehat{(f_\varepsilon^*)^4}(u)\geq N^{-1/2}}\frac{\left(\widehat{(f_\varepsilon^*)^4}(u)-(f_\varepsilon^*)^4(u)\right)^2}{\left(\sqrt{\widehat{(f_\varepsilon^*)^4}(u)}+(f_\varepsilon^*)^2(u)\right)^2\left[\left(\widehat{(f_\varepsilon^*)^4}(u)\right)^{1/4}+f_\varepsilon^*(u)\right]^2}\right)$$

$$+\frac{1}{(f_\varepsilon^*)^4(u)}\mathbb{E}\left(\mathbb{1}_{\widehat{(f_\varepsilon^*)^4}(u)\geq N^{-1/2}}\frac{\left(\widehat{(f_\varepsilon^*)^4}(u)-(f_\varepsilon^*)^4(u)\right)^2}{\left(\sqrt{\widehat{(f_\varepsilon^*)^4}(u)}+(f_\varepsilon^*)^2(u)\right)^2\left[\left(\widehat{(f_\varepsilon^*)^4}(u)\right)^{1/4}+f_\varepsilon^*(u)\right]^2}\right.$$

$$\left.\times\left(\frac{1}{\sqrt{\widehat{(f_\varepsilon^*)^4}(u)}}-\frac{1}{(f_\varepsilon^*)^2(u)}\right)\right)$$

$$\leq \frac{1}{(f_\varepsilon^*)^2(u)}\mathbb{P}\left(\widehat{(f_\varepsilon^*)^4}(u)<N^{-1/2}\right)+\frac{N^{-1}}{(f_\varepsilon^*)^{10}(u)}+\frac{N^{1/2}}{(f_\varepsilon^*)^{10}(u)}\mathbb{E}\left[\left(\widehat{(f_\varepsilon^*)^4}(u)-(f_\varepsilon^*)^4(u)\right)^3\right]$$

$$\leq \frac{C_1N^{-1}}{(f_\varepsilon^*)^{10}(u)}+\frac{N^{-1}}{(f_\varepsilon^*)^{10}(u)}+\frac{N^{1/2}}{(f_\varepsilon^*)^{10}(u)}N^{-3/2}\leq c\frac{N^{-1}}{(f_\varepsilon^*)^{10}(u)}$$

Thus, in that case where $N^{-1}/(f_\varepsilon^*)^{10}(u)$ is smaller than the three other terms found for case (i), we get

$$\mathbb{E}(|R_0(u)|^2)\leq\frac{N^{-1}}{|f_\varepsilon^*(u)|^{10}}.$$

This ends the proof of the lemma. □

B.6. **Proof of Proposition 5.** Clearly

$$\begin{aligned}\|\widetilde{f_{\beta,m}}-f_\beta\|^2 &= \|\widetilde{f_{\beta,m}}-f_{\beta,m}\|^2+\|f_{\beta,m}-f_\beta\|^2\\ &\leq 2\|\widetilde{f_{\beta,m}}-\widehat{f_{\beta,m}}\|^2+2\|\widehat{f_{\beta,m}}-f_{\beta,m}\|^2+\|f_{\beta,m}-f_\beta\|^2\end{aligned}$$

We already know that $\mathbb{E}(\|\widehat{f_{\beta,m}}-f_{\beta,m}\|^2)\leq(4/(N(J-4))D_2(m,1)+m/N$. Moreover,

$$\begin{aligned}\|\widetilde{f_{\beta,m}}-\widehat{f_{\beta,m}}\|^2 &= \left\|\frac{1}{2\pi}\int_{-\pi m}^{\pi m}e^{-iux}\frac{2}{J-4}\sum_{j=3}^{J/2}\widehat{f_{Z_j}^*}(u)R\left(\frac{u}{\Delta}\right)du\right\|^2\\ &= \frac{1}{2\pi}\int_{-\pi m}^{\pi m}\left|\frac{2}{J-4}\sum_{j=3}^{J/2}\widehat{f_{Z_j}^*}(u)R\left(\frac{u}{\Delta}\right)\right|^2du\\ &\leq \frac{1}{\pi}\int_{-\pi m}^{\pi m}\left|\frac{2}{J-4}\sum_{j=3}^{J/2}(\widehat{f_{Z_j}^*}(u)-f_{Z_j}^*(u))\right|^2|R\left(\frac{u}{\Delta}\right)|^2du\\ &\quad+\frac{1}{\pi}\int_{-\pi m}^{\pi m}|f_\beta^*(u)|^2|f_\varepsilon^*(u/\Delta)|^4|R\left(\frac{u}{\Delta}\right)|^2du\end{aligned}$$

Then, applying Lemma 1, using the independence of $\widehat{f_{Z_j}^*}(u)$ and $R(u)$ for $j\geq 3$, and that

$$\mathbb{E}\left(\left|\frac{2}{J-4}\sum_{j=3}^{J/2}(\widehat{f_{Z_j}^*}(u)-f_{Z_j}^*(u)\right|^2\right) = \frac{1}{N}\frac{4}{(J-4)^2}\left(\sum_{j=3}^{J/2}\mathrm{Var}(e^{iuZ_j})+\sum_{j\neq j'}\mathrm{cov}(e^{iuZ_j},e^{iuZ_{j'}})\right)$$

$$\leq \frac{1}{N}\frac{4}{(J-4)^2}\left(\frac{J-4}{2}+\frac{(J-4)^2}{4}(1-|f_\beta^*(u)|^2)|f_\varepsilon^*(\frac{u}{\Delta})|^4\right)$$
$$\leq \frac{1}{N}\frac{2}{J-4}\left(1+\frac{J-4}{2}|f_\varepsilon^*(\frac{u}{\Delta})|^4\right)$$

yields

$$\begin{aligned}\mathbb{E}(\|\widetilde{f_{\beta,m}}-\widehat{f_{\beta,m}}\|^2) &\leq \frac{1}{\pi}\int_{-\pi m}^{\pi m}\frac{1}{N}\frac{2}{J-4}\left(1+\frac{J-4}{2}|f_\varepsilon^*(\frac{u}{\Delta})|^4\right)\frac{du}{|f_\varepsilon^*(\frac{u}{\Delta})|^4}\\ &+\frac{C_0}{\pi}\int_{-\pi m}^{\pi m}|f_\beta^*(u)|^2\left(\frac{N^{-1/2}}{|f_\varepsilon^*(\frac{u}{\Delta})|^4}\wedge\frac{N^{-1}}{|f_\varepsilon^*(\frac{u}{\Delta})|^8}\right)du\\ &\leq \frac{4}{N(J-4)}D_2(m,1)+\frac{2m}{N}+2C_0(\frac{D_2(m,f_\beta)}{\sqrt{N}})\wedge(\frac{D_4(m,f_\beta)}{N}).\end{aligned}$$

Now, gathering all term gives

$$\mathbb{E}(\|\widetilde{f_{\beta,m}}-f_\beta\|^2) \leq \|f_{\beta,m}-f_\beta\|^2+\frac{16}{N(J-4)}D_2(m,1)+6\frac{m}{N}+4C_0(\frac{D_2(m,f_\beta)}{\sqrt{N}})\wedge(\frac{D_4(m,f_\beta)}{N}),$$

which is the announced result. □

B.7. **Proof of Proposition 6.** The proof of Proposition 6 follows the same line as the proof of Proposition 5 in a somehow simpler setting. Clearly

$$\begin{aligned}\|\widetilde{f^0_{\alpha,m}}-f_\alpha\|^2 &= \|\widetilde{f^0_{\alpha,m}}-f_{\alpha,m}\|^2+\|f_{\alpha,m}-f_\alpha\|^2\\ &\leq 2\|\widetilde{f^0_{\alpha,m}}-\widehat{f^0_{\alpha,m}}\|^2+2\|\widehat{f^0_{\alpha,m}}-f_{\alpha,m}\|^2+\|f_{\alpha,m}-f_\alpha\|^2\end{aligned}$$

We already know that $\mathbb{E}(\|\widehat{f^0_{\alpha,m}}-f_{\alpha,m}\|^2)\leq(2\pi)^{-1}(\int_{-\pi m}^{\pi m}du/|f_\varepsilon^*(u)|^2)/N$. Moreover

$$\|\widetilde{f^0_{\alpha,m}}-\widehat{f^0_{\alpha,m}}\|^2=\frac{1}{2\pi}\int_{-\pi m}^{\pi m}|\widehat{f^*_{Y_0}}(u)R_0(u)|^2du$$

which yields the result by writing that $|\widehat{f^*_{Y_0}}(u)|^2\leq 2|\widehat{f^*_{Y_0}}(u)-f^*_{Y_0}(u)|^2+2|f^*_\alpha(u)f^*_\varepsilon(u)|^2$, using the independence of $\widehat{f^*_{Y_0}}(u)$ and R_0 and applying Lemma 1.□

References

J. Antic, C.M. Laffont, D. Chafaï, and D. Concordet. Comparison of nonparametric methods in nonlinear mixed effects models. *Comput. Stat. Data An.*, 53(3):642–656, 2009.

D. Chafaï and J.-M. Loubes. On nonparametric maximum likelihood for a class of stochastic inverse problems. *Statist. Probab. Lett.*, 76(12):1225–1237, 2006.

J. Chen, D. Zhang, and M. Davidian. A monte carlo em algorithm for generalized linear mixed models with flexible random effects distribution. *Biostatistics*, 3(3):347–360, 2002.

F. Comte and C. Lacour. Data-driven density estimation in the presence of additive noise with unknown distribution. *J. Roy. Stat. Soc. B*, 73(4):601–627, 2011.

F. Comte, Y. Rozenholc, and M.-L. Taupin. Penalized contrast estimator for adaptive density deconvolution. *Can. J. Stat.*, 34(3):431–452, 2006.

F. Comte, J. Dedecker, and M. L. Taupin. Adaptive density deconvolution with dependent inputs. *Math. Methods Statist.*, 17(2):87–112, 2008.

F. Comte, A. Samson, and J.J. Stirnemann. Deconvolution estimation of onset of pregnancy with replicate observations. *Working paper*, 2011. URL `http://hal.archives-ouvertes.fr/hal-00588235 v2/`.

A. Delaigle and I. Gijbels. Bootstrap bandwidth selection in kernel density estimation from a contaminated sample. *Ann. Inst. Statist. Math.*, 56(1):19–47, 2004.

A. Delaigle, P. Hall, and A. Meister. On deconvolution with repeated measurements. *Ann. Statist.*, 36(2):665–685, 2008.

J. Fan. On the optimal rates of convergence for nonparametric deconvolution problems. *Ann. Statist.*, 19(3):1257–1272, 1991.

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

W. Ghidey, E. Lesaffre, and G. Verbeke. A comparison of methods for estimating the random effects distribution of a linear mixed model. *Stat Methods Med Res*, 19(6): 575–600, 2010.

E Kuhn. *Maximum likelihood estimation in non linear inverse problems.* PhD thesis, Université Paris Sud - Paris XI, 2003.

N. Laird. Nonparametric maximum likelihood estimation of a mixing distribution. *J. Amer. Statistical Assoc.*, 73(364):805–811, 1978.

A. Mallet, F. Mentré, J.L. Steimer, and F. Lokiec. Nonparametric maximum likelihood estimation for population pharmacokinetics, with application to cyclosporine. *J. Pharmacokinet. Pharmacodyn.*, 16:311–327, 1988.

A. Meister and M. H. Neumann. Deconvolution from non-standard error densities under replicated measurements. *Statist. Sinica*, 20(4):1609–1636, 2010. ISSN 1017-0405.

J.S. Morris and R.J. Carroll. Wavelet-based functional mixed models. *J. Roy. Stat. Soc. B*, 68(2):179–199, 2006.

M.H. Neumann. On the effect of estimating the error density in nonparametric deconvolution. *J. Nonparametr. Statist.*, 7(4):307–330, 1997.

M.H. Neumann. Deconvolution from panel data with unknown error distribution. *J. Multivariate Anal.*, 98(10):1955–1968, 2007.

M. Pensky and B. Vidakovic. Adaptive wavelet estimator for nonparametric density deconvolution. *Ann. Statist.*, 27(6):2033–2053, 1999.

J.C. Pinheiro and D.M. Bates. *Mixed-effect models in S and Splus.* Springer-Verlag, 2000.

W. Shen and T.A. Louis. Empirical bayes estimation via the smoothing by roughening approach. *J. Comput. Graph. Stat.*, 8(4):800–823, 1999.

L. Stefanski and R.J. Carroll. Deconvoluting kernel density estimators. *Statistics*, 21(2): 169–184, 1990.

G. Verbeke and E. Lesaffre. A linear mixed-effects model with heterogeneity in the random-effects population. *J. Amer. Statistical Assoc.*, 91(433):217–221, 1996.

David M Vock, Marie Davidian, Anastasios A Tsiatis, and Andrew J Muir. Mixed model analysis of censored longitudinal data with flexible random-effects density. *Biostatistics*, 2011.

P. Wu and L.X. Zhu. An orthogonality-based estimation of moments for linear mixed models. *Scand. J. Stat.*, 37:253–263, 2010.

D. Zhang and M. Davidian. Linear mixed models with flexible distributions of random effects for longitudinal data. *Biometrics*, 57(3):795–802, 2001.