



HAL
open science

Decompounding with unknown noise through several independent channels

Guillaume Garnier

► **To cite this version:**

Guillaume Garnier. Decompounding with unknown noise through several independent channels. 2024. hal-04460693v2

HAL Id: hal-04460693

<https://hal.science/hal-04460693v2>

Preprint submitted on 17 May 2024

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

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

DECOMPOUNDING WITH UNKNOWN NOISE THROUGH SEVERAL INDEPENDENT CHANNELS

GUILLAUME GARNIER

Sorbonne Université, Inria, CNRS, Université de Paris, Laboratoire Jacques-Louis Lions,
4 place Jussieu, 75005 Paris, France

ABSTRACT. In this article, we consider two different statistical models. First, we focus on the estimation of the jump intensity of a compound Poisson process in the presence of unknown noise. This problem combines both the deconvolution problem and the decomposing problem. More specifically, we observe several independent compound Poisson processes but we assume that all these observations are noisy due to measurement noise. We construct an Fourier estimator of the jump density and we study its mean integrated squared error. Then, we propose an adaptive method to correctly select the estimator and we illustrate the efficiency of the method with numerical simulations. Secondly, we introduce in this article the multiplicative decomposing problem. We study this problem with an estimator based on the Mellin transform. We develop an adaptive procedure to select the optimal cutoff parameter.

Keywords: adaptive density estimation, Fourier estimator, inverse problem, nonparametric statistical inference, deconvolution, decomposing, multiplicative decomposing, Mellin transform, empirical processes.

MSC2020 subject classifications: 62C12, 62C20, 62G05, 62G07.

CONTENTS

1. Introduction	2
1.1. Motivation	2
1.2. Main results and organisation of the article	5
2. Decomposing with unknown noise	6
2.1. Notation	6
2.2. Statistical setting	7
2.3. Risk bounds	9
2.4. Speed of convergence	10
2.5. The adaptive procedure	11
3. Multiplicative decomposing	13
3.1. Preliminaries on Mellin transform	13
3.2. The Mellin estimator	14
3.3. Statistical setting: estimation procedure	15

Date: May 17, 2024.

3.4. Risk bounds	16
3.5. The adaptative procedure	16
4. Numerical experiments	17
4.1. Decompounding with noise	17
4.2. Multiplicative decompounding	20
5. Proofs	22
5.1. Proof of Theorem 1	22
5.2. Proof of Theorem 3	26
5.3. Proof of Theorem 10	30
5.4. Proof of Theorem 11	32
6. Conclusion and perspectives	34
Appendix A. The distinguished logarithm	34
Appendix B. Useful lemmas	35
B.1. Lemmas for decompounding with unknown noise	35
B.2. Lemmas for multiplicative decompounding	38
Appendix C. Acknowledgments	40
References	41

1. INTRODUCTION

1.1. **Motivation.** In the domain of non-parametric statistics, the deconvolution and the decompounding problems are classical and have been widely studied [BG03, Duv13, DK19, Joh09]. We give a brief description of the two problems:

Deconvolution. Deconvolution has applications in many fields, such as image processing [KH96], microscopy [Swe07], astronomy [SPM02, PSM17], seismology [Ulr71, Sac97] and medicine [LLMZ09, RMKB05].

Let X and ε be two independent random variables with density f and f_ε . The deconvolution problem consists in estimating f from an *i.i.d.* sample of $Y = X + \varepsilon$, *i.e.* from noisy observations

$$Y_i = X_i + \varepsilon_i, \quad i = 1, \dots, n,$$

where $(X_i)_{i=1}^n$ are *i.i.d.* random variables with density f and $(\varepsilon_i)_{i=1}^n$ are, *i.i.d.*, with density f_ε and independent of $(X_i)_{i=1}^n$.

This problem have been extensively studied in the literature. The most popular approach estimates f with Fourier estimators [CH88, Dev89, SC90, NH97]. Other approaches have also been developed, for instance by using spline-based methods [AZ09], wavelet decomposition [JKPR04, SNM03] and penalization methods [CRT06, CRT07]. In most studies, the author assume that the noise density f_ε is known, however this assumption is not necessary if an additional error sample is available [Joh09, CL11]. Many adaptive methods, which only rely on data-driven procedures, have also been developed [GST77, DK19].

Decompounding. In the literature, the term decompounding first appeared in the article of Buchmann and Grübel [BG03]. Decompounding has many applications in financial mathematics [EKM13] and queuing theory [BJD11, GCPD16].

Let $(X_k)_{k \in \mathbb{N}}$ be a family of *i.i.d.* random variable with density f and N be a homogeneous Poisson process with intensity $\lambda \in (0, \infty)$. Define $Y = (Y_t)_{t \geq 0}$ as the compound Poisson process

$$Y_t = \sum_{k=1}^{N_t} X_k, \quad t \geq 0.$$

The decompounding problem consists in estimating f from observations $(Y_{i\Delta}, i = 1, \dots, n)$ of the trajectory of Y over $[0, T]$ at a sampling rate $\Delta > 0$, *i.e.* at time $(i\Delta, i = 1, \dots, n)$. The decompounding problem has been the subject of many articles [BP10, BG04, VEU04]. Most of the time, f is estimated using its characteristic function and Fourier estimators [DK19]. The available data can be observed at high frequencies or at low frequencies [DK19, Coc18]. The decompounding problem is part of the more general framework of Levy processes that have been widely developed in recent years [Rei13, NR12, App09].

In this article, we study a model that combines both the deconvolution problem and the decompounding problem. In particular, we assume that we observe several compound Poisson processes but that all these observations are noisy due to measurement noise. More precisely, we consider a family of processes

$$Z_t^j = \left(\sum_{k=1}^{N_t^j} X_k^j \right) + \varepsilon_t^j, \quad t \geq 0, \quad j = 1, \dots, J.$$

We want to construct an estimator of the jump density of the process. Our motivation to study this type of model comes from evolutionary biology. Many studies aim to estimate the distribution of fitness effects (DFE) of cell. This probability density represents the effect of a new mutation on the fitness of a cell. An accurate determination of the DFE within a population would provide a better understanding of the evolutionary trajectory of the population [EWK07].

In 2018, Robert et al. [ROR+18] developed new experimental methods and a probabilistic model to study the DFE in a population of *Escherichia coli* (*E. coli*). The effect of each mutation on the fitness of a cell is assumed to be drawn according to a random variable X of density f where f denotes the DFE of the cell. The authors conclude that the number of mutations in a cell follows a point Poisson process and they estimate the DFE using a method of moment.

More specifically, they observe $n = 1476$ cell lines using micro-fluidic methods [WRP+10]. For each lineage $j = 1, \dots, n$, it is assume that mutations are deleterious and appear according to a Poisson point process $N^j = (N_t^j)_{t \geq 0}$ with intensity $\lambda > 0$. Each of these mutations modifies the selective value over time $(W_t^j)_{t \geq 0}$ of the j -lineage. Let us denote

t_i^j ($i \in \mathbb{N}$) the time of occurrence of the $\{i\}$ -th mutation in lineage j . The quantity

$$s_i^j = \frac{W_{t_{i-1}^j}^j - W_{t_i^j}^j}{W_{t_{i-1}^j}^j}, \quad i > 0,$$

represents the relative effect of the $\{i\}$ -th mutation on the fitness of the individual. Assuming further that these mutations have independent and identically distributed (s_i) effects, then

$$\frac{W_t^j}{W_0^j} = \prod_{i=1}^{N_t^j} (1 - s_i^j).$$

By composing by the logarithm, the evolution of each lineage is controlled by a compound Poisson process, *i.e.*

$$Y_t^j = \ln W_t^j = \sum_{i=1}^{N_t^j} \ln(1 - s_i^j).$$

Experimentally, lineages can only be observed at a sampling rate $\Delta > 0$ through discrete observations $(Y_{i\Delta}^j)_{i=1}^n$. Moreover, these observations are noisy due to measurement noise $(\varepsilon_i^j)_{i=1}^n$ on the i -th observation. In fact, it is only possible to observe the different lineages

$$(1) \quad Z_t^j = \left(\sum_{k=1}^{N_t^j} X_k^j \right) + \varepsilon_t^j, \quad t \geq 0,$$

at sampling times $(i\Delta)_{i=1, \dots, n}$, with $X_k^j = \ln(1 - s_k^j)$. By abuse of notation, we consider that the noise ε_t^j exists for all t in each lineage $j = 1, \dots, 1476$.

The main goal of this article is to provide a nonparametric estimator that is adapted to estimate the probability density f from the sample of the noisy trajectories and study its asymptotic properties. Such an approach has several advantages. First, it offers a way to study the jump density of a noisy compound Poisson process when we observe several independent processes, which has, to our knowledge, been little studied in the literature. Second this method has the advantage of studying a model very close to the biological situation.

However, using a logarithm on the model to transform it into an 'additive' model only gives us information on $\log X$. While it is true that the two models are theoretically equivalent, the additive form forces us to make assumptions about the logarithm of the density we want to reconstruct, what is not easy to justify from modelling perspective. To overcome this, we use in Section 3 a method similar to the previous one to study the multiplicative decomposing problem.

In this setting, we discretely observe one trajectory of a multiplicative compound Poisson process

$$W_t = \prod_{i=1}^{N_t} (1 - s_i),$$

where $(N_t)_{t \geq 0}$ is a Poisson process with intensity λ independent of the *i.i.d.* random variables $(s_i)_i$ with density f .

The idea of studying 'multiplicative' models was developed in [Var89] and [VZ92] where the authors studied a multiplicative deconvolution problem where the error U is uniformly distributed on $[0, 1]$. Such models can be studied using the Mellin transform, which can be seen as a multiplicative version of the Fourier transform. These estimators have been successfully used in [MCJ21, MJS23] to study the multiplicative deconvolution problem with unknown noise.

1.2. Main results and organisation of the article. In Section 2, we discretely observe J trajectory of the process $Z = (Z_t)_{t \geq 0}$ defined in Eq. (1), *i.e.*

$$Z_t^j = \sum_{k=1}^{N_t^j} X_k^j + \varepsilon_t^j, \quad t \geq 0, \quad j = 1, \dots, J,$$

where $(X_k^j)_{k \in \mathbb{N}}, j = 1, \dots, J$ is a family of *i.i.d.* random variable with density f , $(N_t^j)_{t \geq 0}, j = 1, \dots, J$ are homogeneous Poisson processes with intensity $\lambda \in (0, \infty)$ and $(\varepsilon_t^j)_{t \geq 0}, j = 1, \dots, J$ are *i.i.d.* centered random variable such that for every $t, s \geq 0$ and for every $i, j = 1, \dots, J$ the random variables ε_t^j and ε_s^i are independents. Let $\Delta > 0$. We suppose that we observe

$$(Z_{i\Delta}^j, i = 1, \dots, n, j = 1, \dots, J).$$

We aim at estimating f from these observations. In Section 2.2, we define an estimator $\widehat{f}_{m,J}$, based on Fourier methods, that rely on two different times $t_1, t_2 \in \{i\Delta, i = 1, \dots, n\}$. We establish in Theorem 1 an upper bound for its L^2 -risk

$$\begin{aligned} \mathbb{E} \left(\|\widehat{f}_{m,J} - f\|_2^2 \right) &\leq \|f_m - f\|_2^2 + \sum_{i=1}^2 \frac{4e^{4t_i}}{J(t_2 - t_1)^2} \int_{-m}^m \frac{du}{|\varphi_\varepsilon(u)|^2} \\ &\quad + \frac{4K_{J,t_1,t_2}}{(t_2 - t_1)^2} \cdot \left(\frac{\mathbb{E}[X_i^2]}{Jt_i} + \frac{\mathbb{E}[\varepsilon^2]}{Jt_i^2} + 4 \frac{m}{(Jt_i)^2} \right), \end{aligned}$$

where $\|f\|_2 = \left(\int_{\mathbb{R}} |f(x)|^2 dx \right)^{1/2}$, φ_ε denotes the characteristic function of ε and K_{J,t_1,t_2} is a constant that only depends on J, t_1 and t_2 . We discuss the rates of convergence in Section 2.4 depending on the regularity of f and f_ε . In Section 2.5, we establish an adaptive procedure to automatically select the value of the threshold m according to the data.

In Section 3, we discretely observe one trajectory of the multiplicative compound process

$$(2) \quad Y_t = \prod_{i=1}^{N_t} X_i,$$

where $(N_t)_{t \geq 0}$ is a Poisson process with constant intensity $\lambda > 0$, independent of the *i.i.d.* random variables $(X_j)_{j \in \mathbb{N}}$ with common density $f \in L^1([0, \infty)) \cap L^2([0, \infty))$. We suppose that we observe $(Y_{i\Delta}, i = 1, \dots, n)$. We aim at estimating f from these observations. In Section 3.1, we recall some properties of the Mellin transform. In Section 3.2, we define the empirical Mellin estimator

$$\widehat{\mathcal{M}}_c(t) = \frac{1}{n} \sum_{k=1}^N X_k^{c-1+it}, \quad t \in \mathbb{R}.$$

and we give some of this properties when it is well defined. In Section 3.3, we define an estimator $\widehat{f}_{m,\Delta}$ of f , based on the Mellin transform. In Section 3.4, we establish in Theorem 10 an upper bound for its L^2 -risk

$$\mathbb{E}[\|\widehat{f}_{m,\Delta} - f\|_{\omega_1}^2] \leq \|f_m - f\|_{\omega_1}^2 + \frac{1}{2\pi n \Delta^2} \int_{-m}^m \frac{1}{|\mathcal{M}_1[\Delta](s)|^2} ds + \frac{25}{\pi} \left(\frac{\mathbb{E}[\ln(X_1)^2]}{\Delta n} + 4 \frac{m}{(n\Delta)^2} \right).$$

where $\|f\|_{\omega_1} = \left(\int_0^\infty x |f(x)|^2 dx \right)^{1/2}$. In Section 3.5, we introduce an adaptive procedure m to select the optimal cutoff parameter for the Mellin estimator.

We numerically illustrate our two methods on several examples in Section 4. The proofs are postponed to the Section 5.

We rigorously define the distinguished logarithm which is a key element of our estimators in Appendix A and we describe some of its classical properties. We present some useful lemmas in Appendix B.

All of our computer code is available and documented on GitHub

<https://github.com/guimgarnier/decompounding-with-noise>

with several examples, making it easy to use on biological experimental data.

2. DECOMPOUNDING WITH UNKNOWN NOISE

2.1. Notation.

The empirical characteristic function. The characteristic function of a real random variable X is

$$\varphi_X(u) = \int_{\mathbb{R}} e^{iux} \mathbb{P}_X(dx) = \mathbb{E}[e^{iuX}], \quad u \in \mathbb{R}.$$

If X has a density $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$, then we recall the inversion formula

$$f(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \varphi_X(u) e^{-iux} du, \quad x \in \mathbb{R},$$

and Parseval's identity

$$\int_{\mathbb{R}} |f(x)|^2 dx = \frac{1}{2\pi} \int_{\mathbb{R}} |\varphi_X(u)|^2 du .$$

Let (X_1, \dots, X_N) be *i.i.d.* real-valued random variables with common characteristic function $\varphi_X(t)$. The function

$$\widehat{\varphi}_X^N(u) = \frac{1}{N} \sum_{k=1}^N e^{iuX_k} , \quad u \in \mathbb{R} ,$$

is called the empirical characteristic function associated with the sample (X_1, \dots, X_N) .

The empirical characteristic function is an unbiased estimator of φ_X *i.e.*

$$\mathbb{E}[\widehat{\varphi}_X^N(u)] = \varphi_X(u)$$

and

$$(3) \quad \mathbb{E}[|\widehat{\varphi}_X^N(u) - \varphi_X(u)|^2] = \frac{1}{N}(1 - |\varphi_X(u)|^2) .$$

For more detailed information, we refer to Ushakov's book [Ush11].

The distinguished logarithm. In the following, the notation \log denotes the distinguished logarithm that we rigorously define in Appendix A.

2.2. Statistical setting. Let $f \in L^1(\mathbb{R}) \cap L^2(\mathbb{R})$. Let $(X_i^j)_{i,j \geq 0}$ be *i.i.d.* real-valued random variables with density f and for some $J \in \mathbb{N}$, let $(N_t^j)_{t \geq 0, j = 1, \dots, J}$ be a family of *i.i.d.* Poisson point processes with intensity $\lambda \in (0, \infty)$, independent of $(X_i^j)_{i,j \geq 0}$.

We consider a family $(Y_t^j)_{t \geq 0, j = 1, \dots, J}$ of compound Poisson processes with intensity λ and jump size density f ,

$$Y_t^j = \sum_{k=1}^{N_t^j} X_k^j , \quad t \geq 0 , \quad j = 1, \dots, J .$$

It is assumed that the observations are disturbed by a random noise independent of the observed process, *i.e.* we observe a process of the form

$$Z_t^j = Y_t^j + \varepsilon_t^j = \left(\sum_{k=1}^{N_t^j} X_k^j \right) + \varepsilon_t^j , \quad t \geq 0 ,$$

for every $i, j = 1, \dots, J$. For every reals $t, s \geq 0$, the random variables ε_t^j and ε_s^i are independents.

Also, we set the following assumptions:

Assumption.

- (H₁) $\varphi_{X_1} \in L^1(\mathbb{R})$.
- (H₂) $\mathbb{E}(X_1^2) < \infty$.

- (H₃) $\forall t \in [0, \infty)$, $\mathbb{E}[\varepsilon_t] = 0$ and $\mathbb{E}[\varepsilon_t^2] < \infty$.
(H₄) $\forall m \in [0, \infty)$, $\exists c_m \in (0, \infty) : \forall u \in [0, m]$, $|\varphi_\varepsilon(u)| \geq c_m$.
(H₅) $\forall u \in \mathbb{R}$, $|\varphi_\varepsilon(u)| > 0$.

For the sake of simplicity, we write ε instead of ε_t^j in the rest of this article.

Construction of the estimator. We estimate the characteristic function φ_X of X . The main idea is to claim that if we have a good reconstruction of φ_X , by applying the inverse Fourier transform, we should have a good reconstruction of f .

For all $t \in [0, \infty)$, the characteristic function of the process on a single channel Z_t^j is given by

$$(4) \quad \varphi_{Z_t}(u) = e^{-\lambda t + \lambda t \varphi_X(u)} \cdot \varphi_\varepsilon(u), \quad \forall u \in \mathbb{R}.$$

To simplify notation, we consider in the rest of the article the particular case where $\lambda = 1$.

Consider two different times $0 < t_1 < t_2$, then

$$\frac{\varphi_{Z_{t_2}}}{\varphi_{Z_{t_1}}} = e^{-(t_2-t_1) + (t_2-t_1)\varphi_X(u)}, \quad \forall u \in \mathbb{R}.$$

Applying the distinguished logarithm, we obtain the explicit formula

$$\varphi_X(u) = 1 + \frac{1}{t_2 - t_1} \left[\log \varphi_{Z_{t_2}}(u) - \log \varphi_{Z_{t_1}}(u) \right], \quad \forall u \in \mathbb{R}.$$

This leads us to consider the estimator

$$\forall u \in \mathbb{R}, \quad \widehat{\varphi}_X^J(u) = 1 + \frac{1}{t_2 - t_1} \left[\log \widehat{\varphi}_{Z_{t_2}}^J(u) - \log \widehat{\varphi}_{Z_{t_1}}^J(u) \right],$$

with for all $u \in \mathbb{R}$, $\tau \in \{t_1, t_2\}$

$$\widehat{\varphi}_{Z_\tau}^J(u) = \frac{1}{J} \sum_{j=1}^J i Z_\tau^j e^{iu Z_\tau^j}, \quad \widehat{\varphi}_{Z_\tau}^J(u) = \frac{1}{J} \sum_{j=1}^J e^{iu Z_\tau^j}, \quad \log \widehat{\varphi}_{Z_\tau}^J(u) = \int_0^u \frac{\widehat{\varphi}_{Z_\tau}^J(z)}{\widehat{\varphi}_{Z_\tau}^J(z)} dz.$$

Since φ_X is a characteristic function, its modulus is bounded by 1. Nevertheless, this is not necessarily the case for $\widehat{\varphi}_X^J$. We avoid this explosion problem taking

$$\widetilde{\varphi}_X^J(u) = 1 + \frac{1}{t_2 - t_1} \left[\log \widehat{\varphi}_{Z_{t_2}}^J(u) \cdot \mathbf{1}_{|\log \widehat{\varphi}_{Z_{t_2}}^J(u)| \leq \ln(J)} - \log \widehat{\varphi}_{Z_{t_1}}^J(u) \cdot \mathbf{1}_{|\log \widehat{\varphi}_{Z_{t_1}}^J(u)| \leq \ln(J)} \right], \quad u \in \mathbb{R}.$$

To simplify notation, for any positive real $\tau > 0$ and for all $u \in \mathbb{R}$ we set

$$\log \widetilde{\varphi}_{Z_\tau}^J(u) = \log \widehat{\varphi}_{Z_\tau}^J(u) \cdot \mathbf{1}_{|\log \widehat{\varphi}_{Z_\tau}^J(u)| \leq \ln(J)}.$$

In particular, with this notation

$$(5) \quad \widetilde{\varphi}_X^J(u) = 1 + \frac{1}{t_2 - t_1} \left[\log \widetilde{\varphi}_{Z_{t_2}}^J(u) - \log \widetilde{\varphi}_{Z_{t_1}}^J(u) \right], \quad u \in \mathbb{R}.$$

Now, we find an estimator of f by performing an inverse Fourier transformation. However, $\tilde{\varphi}_X^J$ is not necessarily integrable. Hence, we eliminate frequencies above a threshold m before applying an inverse Fourier transform, we obtain

$$(6) \quad \hat{f}_{m,J}(x) = \frac{1}{2\pi} \int_{-m}^m e^{-iux} \tilde{\varphi}_X^J(u) du, \quad x \in \mathbb{R}, \quad m \in (0, \infty).$$

In the same way, we perform an truncated inverse Fourier transformation from the true characteristic function of X and we define

$$f_m(x) = \frac{1}{2\pi} \int_{-m}^m e^{-iux} \varphi_X(u) du, \quad x \in \mathbb{R}, \quad m \in (0, \infty).$$

2.3. Risk bounds. In this section, we study the mean integrated squared error (MISE) of the estimator $\hat{f}_{m,J}$.

For any $0 < t_1 < t_2$, we define

$$(7) \quad C_{t_1,t_2}^J = \min \left\{ m \geq 0 \mid 3t_2 - t_1 + \sup_{[-m,m]} |\log \varphi_\varepsilon(\cdot)| > \ln(J) \right\}.$$

Theorem 1. *Suppose that the assumptions (H_1) – (H_5) holds. Let $0 < t_1 < t_2$. We suppose that J is enough large such that for all $i = 1, 2$:*

- (1) $t_i < \frac{1}{4} \log(Jt_i)$,
- (2) $\sqrt{\log(Jt_i)}(Jt_i)^{2\delta_i-1/2} < 1$ with $\delta_i = t_i / \log(Jt_i)$.

Then for any $m < C_{t_1,t_2}^J$ with C_{t_1,t_2}^J defined by (7), we have

$$\begin{aligned} \mathbb{E} \left(\|\hat{f}_{m,J} - f\|_2^2 \right) &\leq \|f_m - f\|_2^2 + \sum_{i=1}^2 \frac{4e^{4t_i}}{J(t_2 - t_1)^2} \int_{-m}^m \frac{du}{|\varphi_\varepsilon(u)|^2} \\ &\quad + \frac{4K_{J,t_1,t_2}}{(t_2 - t_1)^2} \cdot \left(\frac{\mathbb{E}[X_i^2]}{Jt_i} + \frac{\mathbb{E}[\varepsilon^2]}{Jt_i^2} + 4 \frac{m}{(Jt_i)^2} \right). \end{aligned}$$

where $K_{J,t_1,t_2} = m \ln^2(J) + 8mt_2^2 + \int_{-m}^m |\log \varphi_\varepsilon(u)|^2 du$.

Remark 2. Theorem 1 is asymptotic and ensures that the variance term vanishes when $J \rightarrow \infty$. The upper bound obtained is the sum of a bias term and of a variance term.

In the variance, there is a term $V \sim \frac{4e^{4t_i}}{J(t_2 - t_1)^2}$. On one side, the presence of e^{4t_i} means that the estimator is more and more imprecise as we look the sample at a very large time t_2 . On the other side, the presence of $(t_2 - t_1)$ at the denominator means that we cannot take t_1 and t_2 too close to each other. We illustrate this trade-off in Section 4 with numerical simulations.

2.4. Speed of convergence. In this section, we study the optimal choice of m based on the regularity of the jump density and regularity of the noise density, *i.e.* that when $J \rightarrow \infty$, we look at the asymptotic behaviour of m when it minimizes the upper term in Theorem in 1. We organize the discussion according to different regularities of f and f_ε . More specifically, we consider:

- ordinary smooth densities: The characteristic function decays as $|u|^{-2a}$ (example: Gamma distribution)
- super smooth densities: The characteristic function decays as $e^{-|u|^s}$ (example: Cauchy distribution)

When f is ordinary smooth and f_ε is ordinary smooth. In this case, we assume that $f \in \mathcal{S}(\beta, L)$ for some $\beta > 0$, $L > 0$, with

$$\mathcal{S}(\beta, L) = \left\{ f \in L^2(\mathbb{R}) \mid \int_{\mathbb{R}} (1 + |u|^2)^\beta |\varphi_X(u)|^2 du \leq L \right\}.$$

We also assume that f_ε is ordinary smooth, *i.e.* there exists two reals $a > \frac{1}{2}$ and $d > 0$, such that

$$\forall u \in \mathbb{R}, \quad d \leq (1 + u^2)^a |\varphi_\varepsilon(u)|^2 \leq \frac{1}{d}.$$

It follows that $\|f - f_m\|_2^2$ is of order $m^{-2\beta}$ and that the variance term $\frac{1}{J} \int_{-m}^m \frac{du}{|\varphi_\varepsilon(u)|^2}$ is of order $\frac{1}{J} m^{2a+1}$.

Moreover, the variance term $\frac{4K_{J,t_1,t_2}}{(t_2-t_1)^2} \cdot \left(\frac{\mathbb{E}[X_i^2]}{Jt_i} + \frac{\mathbb{E}[\varepsilon^2]}{Jt_i^2} + 4\frac{m}{(Jt_i)^2} \right)$ is of order $\frac{m \ln^2(J)}{J}$.

Therefore, if the Assumptions $(H_1) - (H_5)$ hold, then

$$(8) \quad \mathbb{E} \left[\|\widehat{f}_{m,J} - f\|_2^2 \right] \lesssim m^{-2\beta} + \frac{1}{J} m^{2a+1} + m \cdot \frac{\ln^2(J)}{J},$$

up to a multiplicative constant that depends on t_1 and t_2 .

To obtain a value m^* that reaches the bias-variance compromise, we minimize the upper bound of Equation (8) by differentiating with respect to m . It follows that $m^* \sim J^{\frac{1}{2a+2\beta+1}}$ when $J \rightarrow \infty$ and that

$$\mathbb{E} \left[\|\widehat{f}_{m,J} - f\|_2^2 \right] = O \left(J^{\frac{-2\beta}{2\beta+2a+1}} \right), \quad J \rightarrow \infty.$$

When f is ordinary smooth and f_ε is super smooth. In this case, we assume that $f \in \mathcal{S}(\beta, L)$ for some reals $\beta > 0$, $L > 0$, *i.e.*

$$f \in \mathcal{S}(\beta, L) = \left\{ f \in L^2(\mathbb{R}) \mid \int_{\mathbb{R}} (1 + |u|^2)^\beta |\varphi_X(u)|^2 du \leq L \right\}.$$

We also assume that f_ε is super smooth, *i.e.* there exists $a > 0$, $s > 0$ and $d_1, d_2 > 0$ such that

$$\forall u \in \mathbb{R}, \quad d_1 \leq \exp(b \cdot |u|^s) \cdot |\varphi_\varepsilon(u)|^2 \leq d_2.$$

It follows that $\|f - f_m\|_2^2$ is of order $m^{-2\beta}$ and that the variance term $\frac{1}{J} \int_{-m}^m \frac{du}{|\varphi_\varepsilon(u)|^2}$ is of order $\frac{1}{J} m \cdot e^{b \cdot m^s}$.

$f \backslash f_\varepsilon$	ordinary smooth: $f_\varepsilon \in \mathcal{S}(a, L)$	super smooth: $f_\varepsilon \in \mathcal{A}_{b,s}(L)$
ordinary smooth: $f \in \mathcal{S}(\beta, L)$	$m^* \sim J^{\frac{1}{2a+2\beta+1}}$ $\mathbb{E}\ \widehat{f}_{m,J} - f\ _2^2 = O\left(J^{\frac{-2\beta}{2\beta+2a+1}}\right)$	$m^* \sim \frac{\ln(J)}{b}$ $\mathbb{E}\ \widehat{f}_{m,J} - f\ _2^2 = O\left(\left(\frac{\ln(J)}{b}\right)^{\frac{-2\beta}{s}}\right)$
super smooth $f \in \mathcal{A}_{c,s}(L)$	$m^* \sim \left(\frac{\ln(J)}{c}\right)^{\frac{1}{s}}$ $\mathbb{E}\ \widehat{f}_{m,J} - f\ _2^2 = O(J^{-1})$	$m^* \sim \left(\frac{\ln(J)}{b+c}\right)^{\frac{1}{s}}$ $\mathbb{E}\ \widehat{f}_{m,J} - f\ _2^2 = O\left(J^{\frac{-c}{b+c}}\right)$

TABLE 1. Speed of the optimal upper bound with respect to the regularity of f and f_ε .

Moreover, the variance term $\frac{4K_{J,t_1,t_2}}{(t_2-t_1)^2} \cdot \left(\frac{\mathbb{E}[X_i^2]}{Jt_i} + \frac{\mathbb{E}[\varepsilon^2]}{Jt_i^2} + 4\frac{m}{(Jt_i)^2} \right)$ is of order $m \cdot \frac{\ln^2(J)}{J}$.

Therefore, if the Assumptions of Theorem 1 holds, then

$$(9) \quad \mathbb{E}\|\widehat{f}_{m,J} - f\|_2^2 \lesssim m^{-2\beta} + \frac{1}{J}m \cdot e^{b \cdot m^s} + m \cdot \frac{\ln^2(J)}{J},$$

up to a multiplicative constant that depends on t_1 and t_2 .

To obtain a value m^* that reaches the bias-variance compromise, we minimize the upper bound of Equation (9) by differentiating with respect to m . It follows that $m^* \sim \left(\frac{\ln(J)}{b}\right)^{1/s}$ when $J \rightarrow \infty$ and that

$$\mathbb{E}\|\widehat{f}_{m,J} - f\|_2^2 = O\left(\left(\frac{\ln(J)}{b}\right)^{\frac{-2\beta}{s}}\right).$$

When f is super smooth. In this section, we assume that f belongs to the class of super smooth densities,

$$\mathcal{A}_{c,s}(L) = \left\{ f \in L^2(\mathbb{R}) \mid \int_{\mathbb{R}} \exp(c \cdot |u|^s) |\varphi_X(u)|^2 du \leq L \right\}.$$

It follows that $\|f - f_m\|_2^2$ is of order $\exp(-c \cdot |m|^s)$.

We apply the same strategy that in Section 2.4 and Section 2.4 to compute m^* when f_ε is ordinary smooth or super smooth.

In Table 1, we resumes the different results that we obtained with respect to the regularity of f and f_ε . For each case, we give the optimal cut-off m^* and the speed of convergence of the estimator.

2.5. The adaptive procedure. We have constructed in Section 2.3 a statistical estimator $\widehat{f}_{m,J}$ to estimate the jump size density f . However, as this estimator strongly depends on the choice of the parameter m , we would like to be able to select a value of m that depends only on the available data, without a priory knowledge on the regularity of the density f .

To do this, we aim to select the parameter m that minimizes the bound obtained in the Theorem 1.

For this, we need to make some further regularity assumptions on the measurement noise, which is to say that there is a function $g \in L^1(\mathbb{R})$ and a real $d > 0$ such that

$$\forall u \in \mathbb{R}, \quad d \cdot g(u) < |\varphi_\varepsilon(u)|^2 < \frac{g(u)}{d}.$$

In the rest of this section, we assume that the noise is ordinary smooth, *i.e.* there exists two reals $a > \frac{1}{2}$ and $d \in (0, 1)$, such that

$$\forall u \in \mathbb{R}, \quad d \leq (1 + u^2)^a \cdot |\varphi_\varepsilon(u)|^2 \leq \frac{1}{d}.$$

Nevertheless, the methods used can be adapted for any function g .

In the following, we simplify the computations by considering the particular case $a = 1$, *i.e.* we assume that

$$\forall u \in \mathbb{R}, \quad d \leq (1 + u^2) \cdot |\varphi_\varepsilon(u)|^2 \leq \frac{1}{d}.$$

How to select \widehat{m}_{J,t_2,t_1} . As we said above, we aim to select m so as to minimize the bound obtained in the theorem 1.

The dominant terms in this bound are the bias term $\int_{u \in [-m, m]} |\varphi_X(u)|^2 du$ and the variance term $\frac{4e^{4t_2}}{J(t_2 - t_1)^2} \int_{-m}^m \frac{du}{|\varphi_\varepsilon(u)|^2}$. Through differentiation, the optimal \overline{m}_J satisfies

$$|\varphi_X(\overline{m}_J)|^2 = \frac{4ae^{4t_2}}{J(t_2 - t_1)^2} (1 + \overline{m}_J^2),$$

then

$$\left| \frac{\varphi_X(\overline{m}_J)}{\sqrt{(1 + \overline{m}_J^2)}} \right|^2 = \frac{4ae^{4t_2}}{J(t_2 - t_1)^2}.$$

However we do not know φ_X , so it is impossible to calculate directly \overline{m}_J . Following the strategy developed by Duval and Kappus [DK19], we consider

$$\overline{\varphi}_X^J(u) = \widetilde{\varphi}_X^J(u) \cdot \mathbf{1}_{\left| \frac{\widetilde{\varphi}_X^J(u)}{\sqrt{1+u^2}} \right| \geq \frac{\kappa_{J,t_1,t_2}}{\sqrt{J(t_2-t_1)}}},$$

where $\kappa_{J,t_1,t_2} = 2e^{2t_2} + \kappa \sqrt{\ln(J(t_2 - t_1)^2)}$, $\kappa > 0$.

It leads us to define the empirical cutoff parameter

$$\widehat{m}_J = \max \left\{ u \geq 0 : \left| \frac{\overline{\varphi}_X(u)}{\sqrt{1+u^2}} \right| \geq \frac{\kappa_{J,t_1,t_2}}{\sqrt{J(t_2-t_1)}} \right\} \wedge (J(t_2 - t_1)^2)^\alpha, \quad \alpha \in (0, 1).$$

For simplicity, we note $m_{\max} = (J(t_2 - t_1)^2)$, keeping in mind that it depends on J, t_1 and t_2 .

We define a new estimator

$$\bar{f}_{m,J}(x) = \frac{1}{2\pi} \int_{-m}^m e^{-iux} \bar{\varphi}_X^J(u) du, \quad x \in \mathbb{R}.$$

Theorem 3. *Assume (H₁)–(H₄). For all reals $0 < t_1 < t_2$ such that $t_2 \leq \frac{1}{4} \log(Jt_2)$ and $(m_{\max})^\alpha < C_{t_1, t_2}^J$, $Jt_1 \rightarrow \infty$, $Jt_2 \rightarrow \infty$ as $J \rightarrow \infty$. Then,*

$$\begin{aligned} \mathbb{E}[\|\bar{f}_{m,J} - f\|_2^2] &\leq \inf_{m \in [0, (m_{\max})^\alpha]} \left\{ \|f_m - f\|_2^2 + C \frac{\ln(J(t_2 - t_1)^2) \cdot m \cdot (1 + m^2)}{J(t_2 - t_1)^2} + \tilde{C}A \right\} \\ &\quad + \left(2 + \frac{2 \log(J)}{(t_2 - t_1)} \right)^2 \cdot T_J \end{aligned}$$

where

$$A = \sum_{i=1}^2 \frac{4e^{4t_i}}{J(t_2 - t_1)^2} \int_{-m}^m \frac{du}{|\varphi_\varepsilon(u)|^2} + \frac{4K_{J, t_1, t_2}}{(t_2 - t_1)^2} \cdot \left(\frac{\mathbb{E}[X_i^2]}{Jt_i} + \frac{\mathbb{E}[\varepsilon^2]}{rJt_i^2} + 4 \frac{m}{(Jt_i)^2} \right)$$

and

$$(10) \quad T_J \leq C_0 (J(t_2 - t_1)^2)^{\alpha - c(\theta)^2} + \frac{C_1}{J(t_2 - t_1)^2} + \frac{C_2}{J(t_2 - t_1)^4}$$

and $c(\theta) = \kappa(t_2 - t_1) e^{2t_2} \cdot \frac{d}{\sqrt{1 + (m_{\max})^2}}$ and where C_0, C_1 and C_2 depends on $\mathbb{E}[X_1^2], \mathbb{E}[\varepsilon^2]$ and where C and \tilde{C} are two constants.

3. MULTIPLICATIVE DECOMPOUNDING

3.1. Preliminaries on Mellin transform. In this section, we first recall some classical results on Mellin transform. We define the multiplicative decomposing problem and we define a good non-parametric estimator.

Let μ be a probability measure on $(0, \infty)$. Its Mellin transform is

$$(11) \quad \mathcal{M}[\mu](s) = \int_0^\infty x^{s-1} \mu(dx), \quad s \in \mathbb{C},$$

for those values of s for which this integral is well defined.

Remark 4. If $\mathcal{M}[\mu](u)$ is well defined for some $u \in \mathbb{R}$, then $\mathcal{M}[\mu]$ converges on the vertical line $u + i\mathbb{R}$. Furthermore, if the integral exists for u and v in \mathbb{R} ($u < v$), then $\mathcal{M}[\mu](w)$ exists for $w \in (u, v)$. It follows that the Mellin transform of a probability measure is well defined on a vertical band Ξ_μ of the complex plane. The region Ξ_μ is called the fundamental strip of $\mathcal{M}[\mu]$.

Assume that μ has a density $f \in L_{\text{loc}}^1([0, \infty))$, then we denote

$$\mathcal{M}[f](s) := \mathcal{M}[f dx](s) = \int_0^\infty x^{s-1} f(x) dx, \quad s \in \mathbb{C},$$

for those values of s for which this integral is well defined. We denote Ξ_f the fundamental strip of $\mathcal{M}[f dx]$

Let $f \in L^1_{\text{loc}}([0, \infty))$ and $c \in \mathbb{R}$ that belongs to Ξ_f . Then we define

$$\mathcal{M}_c[f](t) := \mathcal{M}[f](c + it) = \int_0^\infty x^{c-1+it} f(x) dx, \quad t \in \mathbb{R}.$$

We observe that the function $x \mapsto x^{c-1} f(x)$ belongs to $L^1([0, \infty))$.

We define the norm

$$\|f\|_{\omega_c} = \left(\int_0^\infty |f(x)|^2 x^{2c-1} dx \right)^{1/2}.$$

and $L^2_{[0, \infty)}(\omega_c)$ denotes the space of function $f : [0, \infty) \rightarrow \mathbb{R}$ such that $\|f\|_{\omega_c} < \infty$

Let $f, g \in L^1_{\text{loc}}([0, \infty))$. The multiplicative convolution of f and g is the function defined by

$$f \star g(x) = \int_0^\infty f(y) g\left(\frac{x}{y}\right) dy, \quad x \in [0, \infty).$$

Proposition 5 (Multiplicative convolution and Mellin transform). *Let μ and ν be two probability measures on $[0, \infty)$. We have*

$$\mathcal{M}[\mu \star \nu](s) = \mathcal{M}[\mu](s) \mathcal{M}[\nu](s), \quad s \in \mathbb{C},$$

whenever this integral is well defined.

Proposition 6 (Parseval's theorem for Mellin transform).

$$(12) \quad \int_0^\infty \frac{|f(x)|^2}{x} dx = \frac{1}{2\pi} \int_0^\infty \mathcal{M}[f](it) dt.$$

Proposition 7. (Inversion formula) *Let $f \in L^1([0, \infty))$ and $c \in \Xi_f$. Then*

$$f(x) = \mathcal{M}^{-1} \mathcal{M}[f](x) = \frac{1}{2\pi i} \int_{\nu-i\infty}^{\nu+i\infty} x^{-s} \mathcal{M}[f](s) ds, \quad x \in [0, \infty).$$

3.2. The Mellin estimator. Let $f \in L^2_{[0, \infty)}(\omega_c)$ be a probability distribution. Let us assume that we have a sample X_1, \dots, X_N independent and identically distributed according to f .

We define the Mellin estimator of f

$$\widehat{\mathcal{M}}_c(t) = \frac{1}{n} \sum_{k=1}^N X_k^{c-1+it}, \quad t \in \mathbb{R}.$$

We use Proposition 7 to define the estimator.

$$\widehat{f}_m(x) = \frac{1}{2\pi} \int_{-m}^m x^{-c-it} \widehat{\mathcal{M}}_c(t) dt,$$

which is an estimator of

$$f_m(x) = \frac{1}{2\pi} \int_{-m}^m x^{-c-it} \mathcal{M}_1(t) dt .$$

The next proposition give a bound of the $L^2_{[0,\infty)}(\omega_c)$ -risk of \widehat{f}_k .

Proposition 8 (Miguel et al. [MCJ21] Proposition 2.1). *If $f \in L^2_{[0,\infty)}(\omega_c)$ and $\sigma_c^2 := \mathbb{E}_f(X^{2(c-1)}) < \infty$, then for all $k \in [0, \infty)$,*

$$\mathbb{E}_f(\|f - \widehat{f}_k\|_{\omega_c}^2) \leq \|f - f_k\|_{\omega_c}^2 + \frac{\sigma_c^2 k}{\pi n} .$$

By choosing $k = k_n$ such that $n^{-1}k_n \rightarrow 0$ and $k_n \rightarrow \infty$, \widehat{f}_{k_n} is a consistent estimator of f .

If X and Y are two independant random variables, then

$$\mathcal{M}[X \cdot Y](s) = \mathcal{M}[X](s) \cdot \mathcal{M}[Y](s) , s \in \mathbb{R} .$$

3.3. Statistical setting: estimation procedure. Consider the multiplicative compound process

$$(13) \quad Y_t = \prod_{i=1}^{N_t} X_i ,$$

where $(N_t)_{t \geq 0}$ is a Poisson process with constant intensity $\lambda > 0$, independent of the *i.i.d.* random variables $(X_j)_{j \in \mathbb{N}}$ with common density $f \in L^1([0, \infty)) \cap L^2([0, \infty))$.

We suppose that we observe one trajectory of (Y_t) over $[0, T]$ at equidistant time $\Delta, 2\Delta, \dots, n\Delta, T = n\Delta$. We denote $Z_{k\Delta} = \frac{Y_{k\Delta}}{Y_{(k-1)\Delta}}$ the k -th increment of Y .

Let $c \in \Xi_X \cap \Xi_\Delta$. We have $1 \in \Xi_X \cap \Xi_\Delta$. Let $\mathcal{M}_1[f]$ be the Mellin transform of X_1 and $\mathcal{M}_1[\Delta]$ be the Mellin transform of Z_Δ .

Lemma 9. *For any $s \in \mathbb{C}$, $\mathcal{M}_1[\Delta](s) = \exp(\Delta\lambda(\mathcal{M}_1[f](s) - 1))$.*

From Lemma 9, we have

$$\mathcal{M}_1[f](s) = 1 + \frac{1}{\lambda\Delta} \log \mathcal{M}_1[\Delta](s) ,$$

where $\log \mathcal{M}_1[\Delta]$ is the distinguished logarithm of $\mathcal{M}_1[\Delta]$ that we define in Appendix A.

This last equation leads us to define the estimator

$$\widehat{\mathcal{M}}_1[f](s) = 1 + \frac{1}{\lambda\Delta} \log \widehat{\mathcal{M}}_1[\Delta](s) , s \in \mathbb{R} .$$

where

$$\widehat{\mathcal{M}}_1[\Delta](t) = \frac{1}{n} \sum_{k=1}^n Z_{k\Delta}^{c-1+it} , \quad \widehat{\mathcal{M}}'_c[\Delta](t) = \frac{1}{n} \sum_{k=1}^n i \log(Z_{k\Delta}) Z_{k\Delta}^{c-1+it} ,$$

and

$$\log \widehat{\mathcal{M}}_1[\Delta](s) = \int_0^s \frac{\widehat{\mathcal{M}}_c'[\Delta](t)}{\widehat{\mathcal{M}}_1[\Delta](t)} ds .$$

However, the quantity $\log \widehat{\mathcal{M}}_1[\Delta](s)$ may explode. Then we define

$$\widetilde{\mathcal{M}}_1[f](s) = \widehat{\mathcal{M}}_1[f](s) \cdot \mathbf{1}_{|\widehat{\mathcal{M}}_1[f](s)| \leq 4}, \quad s \in \mathbb{R} .$$

Finally, we apply a Mellin inversion

$$(14) \quad \widehat{f}_{m,\Delta}(x) = \frac{1}{2\pi} \int_{-m}^m x^{-c-it} \widetilde{\mathcal{M}}_1[f](s) dt, \quad x \in (0, \infty),$$

3.4. Risk bounds.

Theorem 10. *Assume that $\mathbb{E}[X_1^2] < \infty$, $\delta \leq \frac{1}{4} \log(n\Delta)$ and $n\Delta \rightarrow \infty$ as $J \rightarrow \infty$. Then, for any $m \geq 0$, it holds*

$$(15) \quad \mathbb{E}[\|\widehat{f}_{m,\Delta} - f\|_{\omega_1}^2] \leq \|f_m - f\|_{\omega_1}^2 + \frac{1}{2\pi n\Delta^2} \int_{-m}^m \frac{1}{|\widehat{\mathcal{M}}_1[\Delta](s)|^2} ds + \frac{25}{\pi} \left(\frac{\mathbb{E}[\ln(X_1)^2]}{\Delta n} + 4 \frac{m}{(n\Delta)^2} \right) .$$

We prove this result in Appendix 5.3. In the demonstration, we need Lemma 19 and Lemma 20.

3.5. The adaptive procedure. We want to minimize the right side term in Theorem 11. The optimal cutoff \overline{m}_n is defined by

$$\overline{m}_n \in \arg \min_{m>0} \left(\|f_m - f\|_{\omega_1}^2 + \frac{1}{2\pi n\Delta^2} \int_{-m}^m \frac{1}{|\widehat{\mathcal{M}}_1[\Delta](s)|^2} ds + \frac{25}{\pi} \left(\frac{\mathbb{E}[\ln(X_1)^2]}{\Delta n} + 4 \frac{m}{(n\Delta)^2} \right) \right) .$$

In the variance term, the leading term is in $\frac{me^{4\Delta}}{n\Delta}$. Therefore, the optimal cutoff is such that

$$|\widehat{\mathcal{M}}_1[f](\overline{m}_n)|^2 = \frac{e^{4\Delta}}{n\Delta} .$$

We can define the empirical optimal cutoff by

$$|\overline{\mathcal{M}}_1[f](\overline{m}_n)|^2 = \frac{e^{4\Delta}}{n\Delta} ,$$

where

$$\overline{\mathcal{M}}_1[f](s) = \widetilde{\mathcal{M}}_1[f](s) \mathbf{1}_{|\widetilde{\mathcal{M}}_1[f](s)| \geq \kappa_{n,\Delta}/\sqrt{n\Delta}} ,$$

with

$$\kappa_{n,\Delta} = (e^{2\Delta+\kappa} \sqrt{\log(n\Delta)}), \quad \kappa > 0 .$$

We consider the empirical cutoff

$$\widehat{m}_n = \max \{ m \geq 0 : |\overline{\mathcal{M}}_1[f](m)| = \kappa_{n,\Delta}/\sqrt{n\Delta} \} \wedge (n\Delta)^\alpha .$$

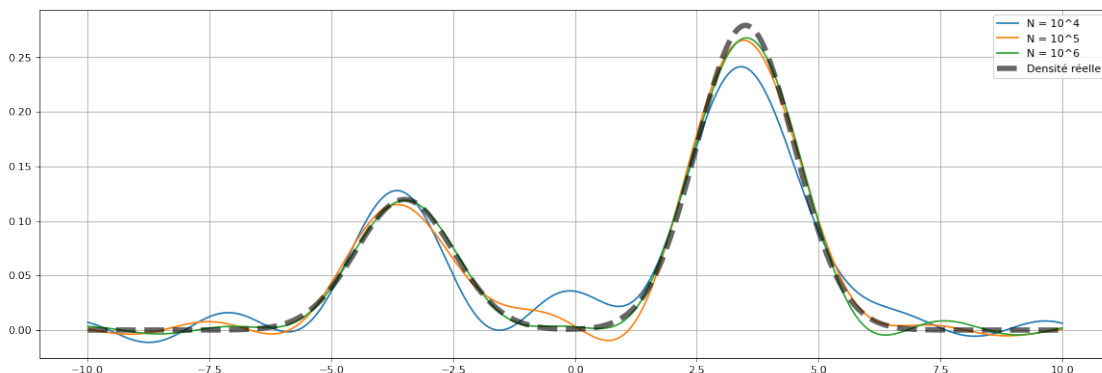


FIGURE 1. Illustration of the estimator.

Input: Jump distribution $X \sim 0.3\mathcal{N}(-3.5, 1) + 0.7\mathcal{N}(3.5, 1)$. Observations are corrupted by a Gaussian noise $\mathcal{N}(0, 1)$. $t_1 = 0.5, t_2 = 1, J = 10^5$.

Theorem 11 (Adaptative method). *Assume that $\mathbb{E}[X_1^2] < \infty$, $\delta \leq \frac{1}{4} \log(n\Delta)$ and $J\Delta \rightarrow \infty$ as $n \rightarrow \infty$. Then, there exist two positive constants A and B depending on κ and $\mathbb{E}[\ln(X_1)^2]$ such that*

$$\begin{aligned} & \mathbb{E}[\|\overline{f_{m_n, \Delta}} - f\|_{\omega_1}^2] \\ & \leq \inf_{m \in [0, (n\Delta)^\alpha]} A \left(\|f_m - f\|_{\omega_1}^2 + \frac{\log(n\Delta)m}{n\Delta} + \frac{1}{n\Delta^2} \int_{-m}^m \frac{1}{|\mathcal{M}_1[\Delta](s)|^2} ds + 4 \frac{m^2}{(n\Delta)^2} \right) \\ & \quad + B \left((n\Delta)^{\alpha - c(\Delta)^2} + \frac{1}{n\Delta} \right). \end{aligned}$$

4. NUMERICAL EXPERIMENTS

4.1. Decompounding with noise.

Implementation details. For a given estimator $\widehat{f}_{m, J}$ of f , we numerically evaluate the efficiency of $\widehat{f}_{m, J}$. More specifically, we compute the L^2 -risks $\mathbb{E}(\|\widehat{f}_{m, J} - f\|_2^2)$ by averaging the result obtained for 1000 independent simulations.

We test our method for three different jump distribution densities f :

- the mixture density $0.3\mathcal{N}(-3.5, 1) + 0.7\mathcal{N}(3.5, 1)$.
- the Gamma density $\Gamma(2)$.
- the Cauchy density $\mathcal{C}(0, 1)$.

and the noise density $\mathcal{N}(0, 1)$.

Efficiency of the estimation. The estimator is expected to be progressively more accurate as J increases. We compute the estimator for different values of J for mixed density (Figure 1) and Gamma density (Figure 2).

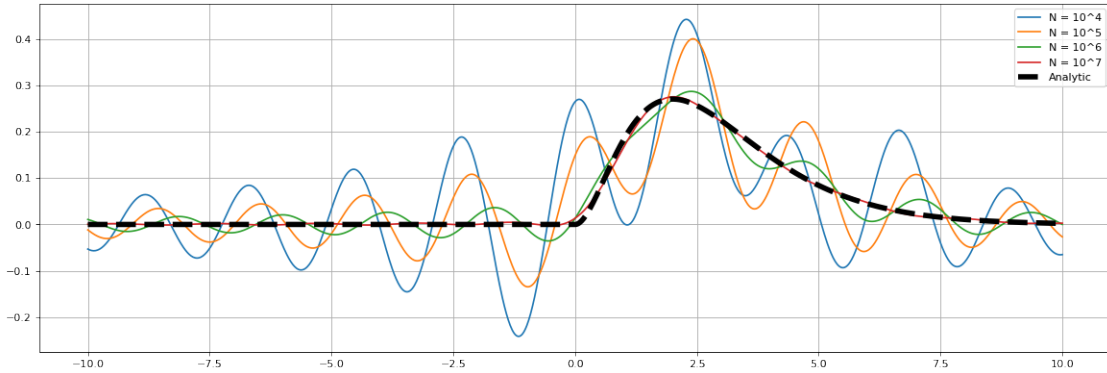


FIGURE 2. Illustration of the estimator.

Input: Jump distribution $X \sim \Gamma(3)$. Observations are corrupted by a Gaussian noise $\mathcal{N}(0, 1)$. $t_1 = 0.5, t_2 = 1, J = 10^5$.

Dependence of the L^2 -risk to the cut-off m and efficiency of the adaptive procedure. The value of m has a strong impact on the quality of the estimator $\hat{f}_{m,J}$. When m is too low or too high, we expect that the estimator should deteriorate. The adaptive procedure gives an automatic way to select a "good" m based on observations.

We illustrate the efficiency of the adaptive procedure for the mixture density (Figure 3), Gamma density (Figure 4) and the Cauchy density (Figure 5).

The results of the simulations are satisfactory, as we can see that the value of m chosen seems to minimize our error.

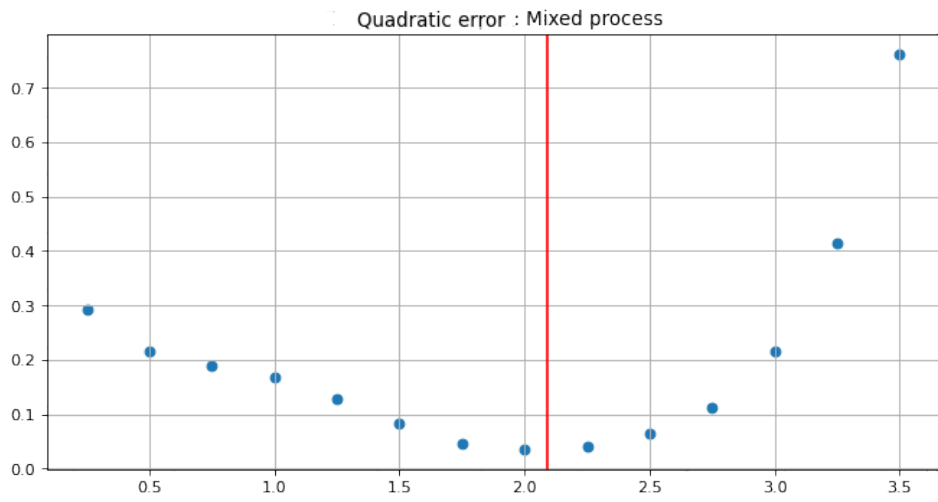


FIGURE 3. Computations of the L^2 -risks $\mathbb{E}(\|\widehat{f}_{m,J} - f\|_2^2)$ (y axis) for different values of m (x axis). The vertical line represents the value of \widehat{m} chosen automatically by the adaptive procedure.

Input: Jump distribution $X \sim 0.3\mathcal{N}(-2, 1) + 0.7\mathcal{N}(2, 1)$. Observations are corrupted by a Gaussian noise $\mathcal{N}(0, 1)$. $t_1 = 0.5, t_2 = 1, J = 10^5$.

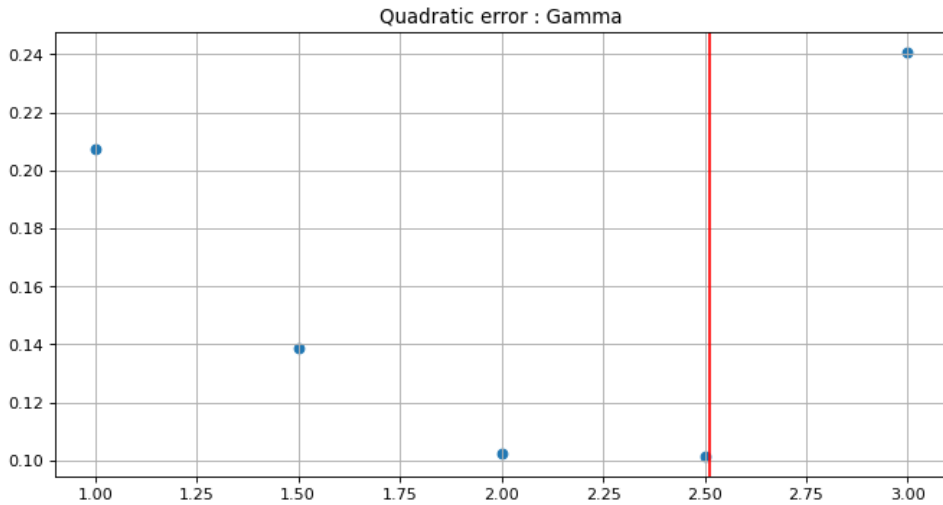


FIGURE 4. Computations of the L^2 -risks $\mathbb{E}(\|\widehat{f}_{m,J} - f\|_2^2)$ (y axis) for different values of m (x axis). The vertical line represents the value of \widehat{m} chosen automatically by the adaptive procedure.

Input: Jump distribution $X \sim \Gamma(2)$. Observations are corrupted by a Gaussian noise $\mathcal{N}(0, 1)$. $t_1 = 0.5, t_2 = 1, J = 10^5$.

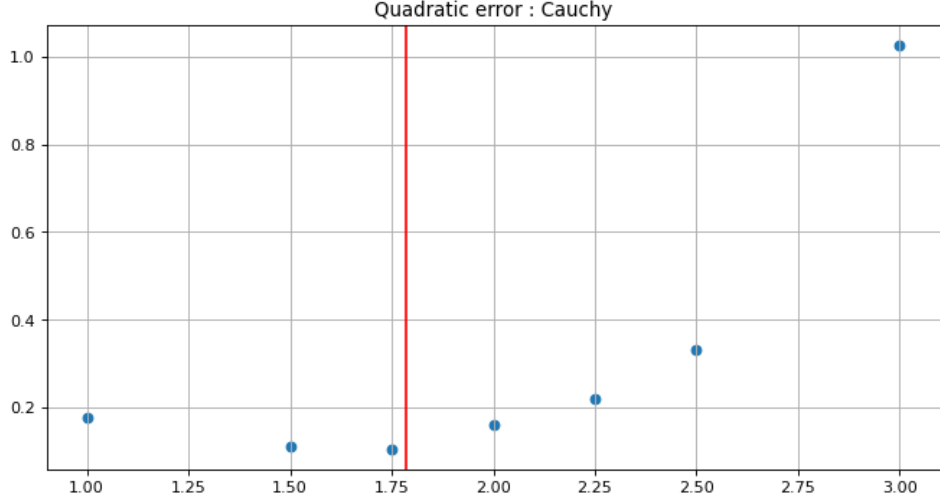


FIGURE 5. Computations of the L^2 -risks $\mathbb{E}(\|\hat{f}_{m,J} - f\|_2^2)$ (y axis) for different values of m (x axis). The vertical line represents the value of \hat{m} chosen automatically by the adaptive procedure.

Input: Jump distribution $X \sim \mathcal{C}(0, 1)$. Observations are corrupted by a Gaussian noise $\mathcal{N}(0, 1)$. $t_1 = 0.5, t_2 = 1, J = 10^5$.

How to select an optimal time t_2 . To define our estimator, we have chosen two times t_1 and t_2 in an arbitrary way. Therefore, we would like to know how to select them in an optimal way.

Theorem 1 states that t_2 must be neither too large nor too close to t_1 , which we can illustrate with numerical simulation. Figure 6 shows the value of the L^2 -risks $\mathbb{E}(\|\hat{f}_{m,J} - f\|_2^2)$ of the estimator for different values of t_2 . Each of these values is the average of the results of 1000 different simulations. Here, it seems that the estimate is better when t_2 is between 1 and 1.5, which seems to correspond roughly to the intensity of the number of jumps.

4.2. Multiplicative decomposing. In this section, we illustrate the efficiency of the Mellin estimator we built in Section 3.2 on a simulated sample of size $n = 5000$.

Here, we assume that the jump distribution densities f is the Beta law $\beta(200, 30)$. The bandwidth $m = 83.7$ was selected by the adaptive procedure described in Section 3.5. The resulting estimate is plotted in Fig. 7.

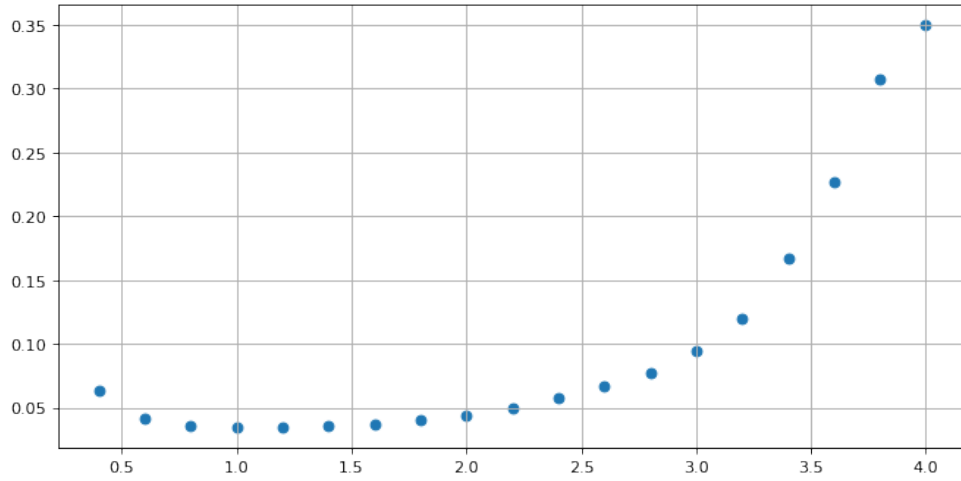


FIGURE 6. Computations of the L^2 -risks $\mathbb{E}(\|\widehat{f}_{m,J} - f\|_2^2)$ (y axis) for different values of t_2 (x axis).
 Input: Jump distribution $X \sim 0.3\mathcal{N}(-2, 1) + 0.7\mathcal{N}(2, 1)$. Observations are corrupted by a Gaussian noise $\mathcal{N}(0, 1)$. $t_1 = 0.2$, $m = 2$, $J = 10^5$.

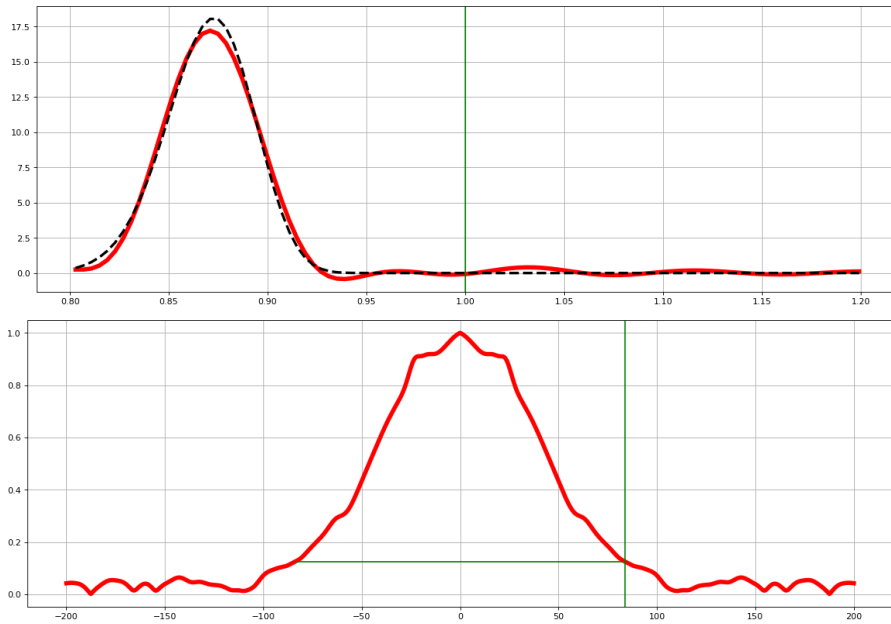


FIGURE 7. **Top:** Estimation of the Beta density $\beta(200, 30)$ for 5000 observations. The real density is the black dotted-line and the estimator is plotted in red. **Bottom:** Illustration of the adaptive procedure. The green lines show the cutoff on the Mellin transform.

5. PROOFS

5.1. Proof of Theorem 1. This proof is an adaptation of the proof of Theorem 3.1 of Duval and Kappus [DK19]. Nevertheless, we have in addition a noise term that must be taken into account in our majorations.

From the triangle inequality and Parseval's equality we have

$$\begin{aligned} \|\widehat{f}_{m,J} - f\|_2^2 &\leq \|f_m - f\|_2^2 + \|\widehat{f}_{m,J} - f_m\|_2^2 \\ &= \|f_m - f\|_2^2 + \frac{1}{2\pi} \int_{-m}^m |\widetilde{\varphi}_X^J(u) - \varphi_X(u)|^2 du . \end{aligned}$$

By recalling that for any reals $(a, b) \in \mathbb{R}^2$, $|a + b|^2 \leq 2(a^2 + b^2)$, and by using Definition 6 and Definition 5, then we have

$$\begin{aligned} &\int_{-m}^m |\widetilde{\varphi}_X^J - \varphi_X(u)|^2 du \\ &= \frac{1}{(t_2 - t_1)^2} \int_{-m}^m |\log \widetilde{\varphi}_{Z_{t_2}}^J(u) - \log \widetilde{\varphi}_{Z_{t_1}}^J(u) - \log \varphi_{Z_{t_2}}(u) + \log \varphi_{Z_{t_1}}(u)|^2 du \\ &\leq \frac{2}{(t_2 - t_1)^2} \int_{-m}^m |\log \widetilde{\varphi}_{Z_{t_1}}^J(u) - \log \varphi_{Z_{t_1}}(u)|^2 du + \frac{2}{(t_2 - t_1)^2} \int_{-m}^m |\log \widetilde{\varphi}_{Z_{t_2}}^J(u) - \log \varphi_{Z_{t_2}}(u)|^2 du . \end{aligned}$$

In the rest of the proof we establish a majorization of the two terms on the right.

Let $\tau \in \{t_1, t_2\}$ and $I_{\tau,J,m}$ denote the quantity

$$I_{\tau,J,m} = \frac{1}{(t_2 - t_1)^2} \int_{-m}^m |\log \widetilde{\varphi}_{Z_\tau}^J(u) - \log \varphi_{Z_\tau}(u)|^2 du .$$

Let $(\gamma, \zeta) \in (0, \infty)^2$ be two positive reals such that $\gamma c_m > \zeta$. In particular, we take γ such that $\gamma c_m = \frac{2\zeta}{1 \wedge (t_2 - t_1)} > \zeta$.

Consider the events

$$\Omega_{\zeta,\tau}(m) = \left\{ \forall u \in [-m, m], |\widehat{\varphi}_{Z_\tau}^J(u) - \varphi_{Z_\tau}(u)| \leq \zeta \sqrt{\frac{\log(J\tau)}{J\tau}} \right\}$$

and

$$M_{J,t}^{(\gamma c_m)} = \min \left\{ u \geq 0 : |\varphi_{Z_t}(u)| = \gamma c_m \sqrt{\log(Jt)/(Jt)} \right\} ,$$

as defined in Lemma 17 and Lemma 18.

As the events

$$\{|u| \leq m \wedge M_{J,\tau}^{(\gamma c_m)}\} \cap \Omega_{\zeta,\tau}(m) , \quad \{|u| \in [M_{J,\tau}^{(\gamma c_m)}, m]\} \cap \Omega_{\zeta,\tau}(m) \quad \text{and} \quad \Omega_{\zeta,\tau}(m)^c$$

form a partition of the sample set Ω , we only need to control $I_{\tau,J,m}$ on each of them. This leads us to define the quantities

$$\begin{aligned} I_{\tau,J,m}^{(1)} &= \frac{1}{(t_2 - t_1)^2} \int_{-m \wedge M_{J,\tau}^{(\gamma c_m)}}^{m \wedge M_{J,\tau}^{(\gamma c_m)}} \mathbb{1}_{\Omega_{\zeta,\tau}(m)} |\log \tilde{\varphi}_{Z_\tau}^J(u) - \log \varphi_{Z_\tau}(u)|^2 du, \\ I_{\tau,J,m}^{(2)} &= \frac{1}{(t_2 - t_1)^2} \mathbb{1}_{m > M_{J,\tau}^{(\gamma c_m)}} \cdot \mathbb{1}_{\Omega_{\zeta,\tau}(m)} \int_{|u| \in [M_{J,\tau}^{(\gamma c_m)}, m]} |\log \tilde{\varphi}_{Z_\tau}^J(u) - \log \varphi_{Z_\tau}(u)|^2 du, \\ I_{\tau,J,m}^{(3)} &= \frac{1}{(t_2 - t_1)^2} \mathbb{1}_{\Omega_{\zeta,\tau}(m)^c} \int_{-m}^m |\log \tilde{\varphi}_{Z_\tau}^J(u) - \log \varphi_{Z_\tau}(u)|^2 du. \end{aligned}$$

Obviously, we have the decomposition

$$I_{\tau,J,m} = I_{\tau,J,m}^{(1)} + I_{\tau,J,m}^{(2)} + I_{\tau,J,m}^{(3)}.$$

Step 1 – Control of $I_{\tau,J,m}^{(1)}$. Define the event $\mathcal{A} = \{|u| \leq m \wedge M_{J,\tau}^{(\gamma c_m)}\} \cap \Omega_{\zeta,\tau}(m)$. The triangle inequality ensures that

$$|\log \hat{\varphi}_{Z_\tau}^J(u)| \leq |\log \tilde{\varphi}_{Z_\tau}^J(u) - \log \varphi_{Z_\tau}(u)| + |\log \varphi_{Z_\tau}(u)|.$$

Then Lemma 18 applied to $\zeta < \gamma c_m$ ensures that

$$|\log \hat{\varphi}_{Z_\tau}^J(u)| \leq \frac{\gamma c_m}{\zeta} \log \left(\frac{\gamma c_m}{\gamma c_m - \zeta} \right) \frac{|\tilde{\varphi}_{Z_\tau}^J(u) - \varphi_{Z_\tau}(u)|}{|\varphi_{Z_\tau}(u)|} + |\log \varphi_{Z_\tau}(u)|.$$

Then we have by definition of $\Omega_{\zeta,\tau}(m)$,

$$|\log \hat{\varphi}_{Z_\tau}^J(u)| \leq \frac{\gamma c_m}{\zeta} \log \left(\frac{\gamma c_m}{\gamma c_m - \zeta} \right) \frac{1}{|\varphi_{Z_\tau}(u)|} \zeta \sqrt{\frac{\log(J\tau)}{J\tau}} + |\log \varphi_{Z_\tau}(u)|.$$

For $J\tau$ large enough, we have $\gamma c_m \sqrt{\frac{\log(J\tau)}{J\tau}} < 1$, hence on \mathcal{A} the definition of $M_{J,\tau}^{(\gamma c_m)}$ ensures that

$$\frac{1}{|\varphi_{Z_\tau}(u)|} \leq \frac{1}{\gamma c_m \sqrt{\frac{\log(J\tau)}{J\tau}}}.$$

It follows that

$$\begin{aligned} |\log \hat{\varphi}_{Z_\tau}^J(u)| &\leq \frac{\gamma c_m}{\zeta} \log \left(\frac{\gamma c_m}{\gamma c_m - \zeta} \right) \frac{1}{\gamma c_m \sqrt{\frac{\log(J\tau)}{J\tau}}} \zeta \sqrt{\frac{\log(J\tau)}{J\tau}} + |\log \varphi_{Z_\tau}(u)|, \\ &\leq \log \left(\frac{\gamma c_m}{\gamma c_m - \zeta} \right) + |\log \varphi_{Z_\tau}(u)|. \end{aligned}$$

By definition, this means (see also Proposition 15)

$$|\log \hat{\varphi}_{Z_\tau}^J(u)| \leq \log \left(\frac{\gamma c_m}{\gamma c_m - \zeta} \right) + |\log \varphi_{Y_\tau}(u) + \log \varphi_\varepsilon(u)|,$$

However, we know that for all $u \in \mathbb{R}$

$$\varphi_X(u) = 1 + \frac{1}{\tau} \log \varphi_{Y_\tau}(u) .$$

It follows that

$$\begin{aligned} |\log \widehat{\varphi}_{Z_\tau}^J(u)| &\leq \log \left(\frac{\gamma c_m}{\gamma c_m - \zeta} \right) + |\tau(\varphi_X(u) + 1) + \log \varphi_\varepsilon(u)| \\ &\leq \log \left(\frac{\gamma c_m}{\gamma c_m - \zeta} \right) + 2\tau + |\log \varphi_\varepsilon(u)| , \\ &\leq \log \left(\frac{\gamma c_m}{\gamma c_m - \zeta} \right) + 2\tau + \sup_{[-m, m]} |\log \varphi_\varepsilon(\cdot)| . \end{aligned}$$

Therefore,

$$\left| \frac{\log \widehat{\varphi}_{Z_\tau}^J(u)}{t_2 - t_1} \right| \leq \frac{1}{t_2 - t_1} \log \left(\frac{\gamma c_m}{\gamma c_m - \zeta} \right) + \frac{2\tau}{t_2 - t_1} + \frac{\sup_{[-m, m]} |\log \varphi_\varepsilon(\cdot)|}{t_2 - t_1} .$$

There are two different cases.

- If $t_2 - t_1 \geq 1$, then $\gamma c_m = \frac{2\zeta}{1 \wedge (t_2 - t_1)} = 2\zeta$. It follows that

$$\frac{1}{t_2 - t_1} \log \left(\frac{\gamma c_m}{\gamma c_m - \zeta} \right) = \frac{1}{t_2 - t_1} \log(2) < 1 .$$

- If $t_2 - t_1 < 1$, then $\gamma c_m = \frac{2\zeta}{1 \wedge (t_2 - t_1)} = \frac{2\zeta}{t_2 - t_1}$. It follows that

$$\begin{aligned} \frac{1}{t_2 - t_1} \log \left(\frac{\gamma c_m}{\gamma c_m - \zeta} \right) &= \frac{1}{t_2 - t_1} \log \left(\frac{\frac{2\zeta}{t_2 - t_1}}{\frac{2\zeta}{t_2 - t_1} - \frac{\zeta(t_2 - t_1)}{t_2 - t_1}} \right) \\ &= \frac{1}{t_2 - t_1} \log \left(\frac{2}{2 - (t_2 - t_1)} \right) \\ &\leq 1 , \end{aligned}$$

because for any real $x \in (0, 1)$ we have

$$0 < \frac{1}{x} \log \left(\frac{2}{2 - x} \right) < 1 .$$

It leads to

$$\begin{aligned} (23) \quad \left| \frac{\log \widehat{\varphi}_{Z_\tau}^J(u)}{t_2 - t_1} \right| &\leq 1 + \frac{2\tau}{t_2 - t_1} + \frac{\sup_{[-m, m]} |\log \varphi_\varepsilon(\cdot)|}{t_2 - t_1} \\ &\leq 1 + \frac{2t_2}{t_2 - t_1} + \frac{\sup_{[-m, m]} |\log \varphi_\varepsilon(\cdot)|}{t_2 - t_1} . \end{aligned}$$

Since we consider $0 \leq m \leq C_{t_1, t_2}^J$, we have

$$\left| \log \widehat{\varphi}_{Z_\tau}^J(u) \right| \leq 3t_2 - t_1 + \sup_{[-m, m]} |\log \varphi_\varepsilon(\cdot)| \leq \ln(J) .$$

Therefore, it comes from Equation (23) and Equation (5) that

$$\log \tilde{\varphi}_{Z_\tau}^J(u) = \log \widehat{\varphi}_{Z_\tau}^J(u) .$$

By using Lemma 18 and the definition of γ , we have

$$\begin{aligned} \mathbb{E}[I_{\tau,J,m}^{(1)}] &= \frac{1}{(t_2 - t_1)^2} \int_{-m \wedge M_{J,\tau}^{(\gamma c_m)}}^{m \wedge M_{J,\tau}^{(\gamma c_m)}} \mathbb{E} \left[\mathbb{1}_{\Omega_{\zeta,\tau}(m)} |\log \widehat{\varphi}_{Z_\tau}^J(u) - \log \varphi_{Z_\tau}(u)|^2 \right] du \\ &\leq \frac{1}{(t_2 - t_1)^2} \int_{-m}^m \frac{\mathbb{E}[|\widehat{\varphi}_{Z_\tau}^J(u) - \varphi_{Z_\tau}(u)|^2]}{|\varphi_{Z_\tau}(u)|^2} du , \end{aligned}$$

because $\frac{\gamma c_m}{\zeta} \log \left(\frac{\gamma c_m}{\gamma c_m - \zeta} \right) \leq 1$ due to the fact that $\gamma c_m > \zeta$.

Direct computations lead to

$$\begin{aligned} \mathbb{E}[|\widehat{\varphi}_{Z_\tau}^J(u) - \varphi_{Z_\tau}(u)|^2] &= \frac{1 - |\varphi_{Z_\tau}(u)|^2}{J} \\ &= \frac{1 - |\varphi_{Y_\tau}(u)|^2 |\varphi_\varepsilon(u)|^2}{J} \\ &= \frac{(1 - |\varphi_{Y_\tau}(u)|^2) |\varphi_\varepsilon(u)|^2 + 1 - |\varphi_\varepsilon(u)|^2}{J} \\ &\leq \frac{(1 - |\varphi_{Y_\tau}(u)|^2) |\varphi_\varepsilon(u)|^2 + 1 - |\varphi_\varepsilon(u)|^2}{J} \\ &\leq \frac{2}{J} . \end{aligned}$$

We deduce that

$$\mathbb{E}[I_{\tau,J,m}^{(1)}] \leq \frac{2}{J(t_2 - t_1)^2} \int_{-m}^m \frac{1}{|\varphi_{Z_\tau}(u)|^2} du .$$

Futhermore, we know that (Y_τ) is a compound Poisson process with intensity $\lambda = 1$. It leads to

$$|\varphi_{Y_\tau}(u)| \geq e^{-2\tau} , \quad u \in \mathbb{R} .$$

Then,

$$|\varphi_{Z_\tau}(u)|^2 = |\varphi_{Y_\tau}(u) \varphi_\varepsilon(u)|^2 \geq e^{-4\tau} |\varphi_\varepsilon(u)|^2 .$$

It leads to

$$\mathbb{E}[I_{\tau,J,m}^{(1)}] \leq \frac{2e^{4\tau}}{J(t_2 - t_1)^2} \int_{-m}^m \frac{1}{|\varphi_\varepsilon(u)|^2} du .$$

Step 2 – Control of $I_{\tau,J,m}^{(3)}$. on the event $\Omega_{\zeta,\tau}(m)^c$, it is more complicated to control the empirical distinguished logarithm. Nevertheless, the cut-off on the estimator allows us to write that

$$\begin{aligned}
\int_{-m}^m |\log \tilde{\varphi}_{Z_\tau}^J(u) - \log \varphi_{Z_\tau}(u)|^2 du &\leq 2 \int_{-m}^m |\log \tilde{\varphi}_{Z_\tau}^J(u)|^2 du + 2 \int_{-m}^m |\log \varphi_{Z_\tau}(u)|^2 du \\
&\leq 4m \ln^2(J) + 2 \int_{-m}^m |\log \varphi_{Y_\tau}(u) + \log \varphi_\varepsilon(u)|^2 du \\
&\leq 4m \ln^2(J) + 4 \int_{-m}^m |\log \varphi_{Y_\tau}(u)|^2 du + 4 \int_{-m}^m |\log \varphi_\varepsilon(u)|^2 du \\
&\leq 4m \ln^2(J) + 32m\tau^2 du + 4 \int_{-m}^m |\log \varphi_\varepsilon(u)|^2 du.
\end{aligned}$$

Indeed, we have that $\varphi_X = 1 + \frac{1}{\tau} \log \varphi_{Y_\tau}$. It follows that $|\log \varphi_{Y_\tau}| \leq 2\tau$. Recall that we have

$$I_{\tau, J, m}^{(3)} = \frac{1}{(t_2 - t_1)^2} \mathbf{1}_{\Omega_{\zeta, \tau}(m)^c} \int_{-m}^m |\log \tilde{\varphi}_{Z_\tau}^J(u) - \log \varphi_{Z_\tau}(u)|^2 du ,$$

which leads to write

$$\mathbb{E}[I_{\tau, J, m}^{(3)}] \leq \frac{1}{(t_2 - t_1)^2} \cdot \left(4m \ln^2(J) + 32m\tau^2 du + 4 \int_{-m}^m |\log \varphi_\varepsilon(u)|^2 du \right) \cdot \mathbb{P}(\Omega_{\zeta, \tau}(m)^c) .$$

We apply Lemma 17 with $\eta = 2$ and $\zeta > \sqrt{5\tau}$. It ensures that

$$\mathbb{E}[I_{\tau, J, m}^{(3)}] \leq \frac{1}{(t_2 - t_1)^2} \cdot \left(4m \ln^2(J) + 32m\tau^2 + 4 \int_{-m}^m |\log \varphi_\varepsilon(u)|^2 du \right) \cdot \left(\frac{\mathbb{E}[X_i^2]}{J\tau} + \frac{\mathbb{E}[\varepsilon^2]}{J\tau^2} + 4 \frac{m}{(J\tau)^2} \right) .$$

Step 3 – Control of $I_{\tau, J, m}^{(2)}$. By Assumption, $\tau \leq \delta \log(J\tau)$ and $\delta < 1/4$. It follows that for all $u \in [-m, m]$,

$$|\varphi_{Z_\tau}(u)| = |\varphi_{Y_\tau}| \cdot |\varphi_\varepsilon| > e^{-2\tau} c_m > (J\tau)^{-2\delta} c_m > \gamma c_m \sqrt{\log(J\tau)/(J\tau)} \Big\} ,$$

when $J\tau$ is enough high.

Finally, $M_{J, \tau}^{(\gamma c_m)} \geq m$, and $I_{\tau, J, m}^{(2)} = 0$ almost surely.

5.2. Proof of Theorem 3. Let $m \geq 0$. The proof of the theorem is divided in two steps: in a first one, we control $\mathbb{E} \|\widehat{f}_{\widehat{m}_n, J} - f\|_2^2$ on the event $\mathcal{E} = \{\widehat{m}_J < m\}$, then we control it on the complementary event $\mathcal{E}^c = \{\widehat{m}_J \geq m\}$.

As we know, it follows from the triangle inequality and Parseval's equality that

$$\begin{aligned} \|\bar{f}_{m,J} - f\|_2^2 &= \|f_m - f\|_2^2 + \frac{1}{2\pi} \int_{-m}^m |\tilde{\varphi}_X^J(u) - \varphi_X(u)|^2 du \\ &= \frac{1}{2\pi} \int_{|u| \in [-m, m]^c} |\varphi_X(u)|^2 du + \frac{1}{2\pi} \int_{-m}^m |\tilde{\varphi}_X^J(u) - \varphi_X(u)|^2 du. \end{aligned}$$

Therefore, by replacing the estimator $\hat{f}_{m,J}$ by $\hat{f}_{\hat{m}_J, J}$ produces a surplus of bias on \mathcal{E} and surplus of variance on \mathcal{E}^c . Then, we only need to control these surpluses.

Step 1 – Control on \mathcal{E} .

$$\begin{aligned} \mathbb{E} \left[\mathbb{1}_{\mathcal{E}} \int_{|u| \in [\hat{m}_J, m]} |\varphi_X(u)|^2 du \right] &\leq 2\mathbb{E} \left[\mathbb{1}_{\mathcal{E}} \int_{|u| \in [\hat{m}_J, m]} |\tilde{\varphi}_X^J(u)|^2 du \right] + 2\mathbb{E} \left[\mathbb{1}_{\mathcal{E}} \int_{|u| \in [\hat{m}_J, m]} |\tilde{\varphi}_X^J(u) - \varphi_X(u)|^2 du \right] \\ &\leq 2\mathbb{E} \left[\mathbb{1}_{\mathcal{E}} \int_{|u| \in [\hat{m}_J, m]} \frac{\kappa_{J, t_1, t_2}^2 \cdot (1 + u^2)}{J(t_2 - t_1)^2} du \right] + 2\mathbb{E} \left[\mathbb{1}_{\mathcal{E}} \int_{|u| \in [0, m]} |\tilde{\varphi}_X^J(u) - \varphi_X(u)|^2 du \right] \\ &\leq 4m \cdot (1 + m^2) \cdot \frac{\kappa_{J, t_1, t_2}^2}{J(t_2 - t_1)^2} + 2A, \end{aligned}$$

where $A = \sum_{i=1}^2 \frac{4e^{4t_i}}{J(t_2 - t_1)^2} \int_{-m}^m \frac{du}{|\varphi_\varepsilon(u)|^2} + \frac{4K_{J, t_1, t_2}}{(t_2 - t_1)^2} \cdot \left(\frac{\mathbb{E}[X_i^2]}{Jt_i} + \frac{\mathbb{E}[\varepsilon^2]}{Jt_i^2} + 4\frac{m}{(Jt_i)^2} \right)$ is given by Theorem 1.

Given that $\kappa_{J, t_1, t_2} = 2\sqrt{2}e^{2t_2} + \kappa\sqrt{\ln(J(t_2 - t_1)^2)}$, we have that

$$\mathbb{E} [\|\bar{f}_{\hat{m}_J} - f\|_2^2 \mathbb{1}_{\mathcal{E}}] \leq \|f_m - f\|_2^2 + C \frac{m \cdot (1 + m^2) \cdot \ln(J(t_2 - t_1)^2)}{J(t_2 - t_1)^2} + \tilde{C}A,$$

where C and \tilde{C} are two constantes.

Step 2 – Control on \mathcal{E}^c . Now, we look at the case where where $\hat{m}_J > m$. By construction, \hat{m}_J is bounded by $(m_{\max})^\alpha$. It follows that

$$\mathbb{E} \left[\int_{m < |u| < \hat{m}_J} |\tilde{\varphi}_X^J(u) - \varphi(u)|^2 du \mathbb{1}_{\mathcal{E}^c} \right] \leq \mathbb{E} \left[\int_{m < |u| < (m_{\max})^\alpha} |\tilde{\varphi}_X^J(u) - \varphi(u)|^2 du \right].$$

Let $\eta > 2$, such that $\alpha - \eta < -1$. As for the proof of the theorem 1, Lemma 17 and Lemma 18 leads to

$$\begin{aligned} \mathbb{E} [|\tilde{\varphi}_X^J(u) - \varphi(u)|^2] &\leq |\varphi_X(u)|^2 + \mathbb{E} [|\hat{\varphi}_X^J(u) - \varphi(u)|^2] \\ &\leq |\varphi_X(u)|^2 + \sum_{i=1}^2 \frac{4e^{4t_i}}{J(t_2 - t_1)^2} \frac{du}{|\varphi_\varepsilon(u)|^2} + \frac{\mathbb{E}[X_1^2]}{Jt_1} + \frac{\mathbb{E}[\varepsilon^2]}{Jt_1^2} + 4\frac{m}{(Jt_1)^\eta}. \end{aligned}$$

On the event

$$\left| \frac{\varphi_X(u)}{\sqrt{1 + u^2}} \right| \geq \frac{2e^{2t_2}}{\sqrt{J}(t_2 - t_1)},$$

we see that

$$\begin{aligned}
\mathbb{E}[|\bar{\varphi}_X^J(u) - \varphi(u)|^2] &\leq |\varphi_X(u)|^2 + \frac{8}{J(t_2 - t_1)^2} \frac{e^{4t_2}}{|\varphi_\varepsilon(u)|^2} + \frac{\mathbb{E}[X_1^2]}{Jt_1} + \frac{\mathbb{E}[\varepsilon^2]}{Jt_1^2} + 4 \frac{m}{(Jt_1)^\eta} \\
&\leq |\varphi_X(u)|^2 + \frac{2|\varphi_X(u)|^2}{1+u^2} \frac{1}{|\varphi_\varepsilon(u)|^2} + \frac{\mathbb{E}[X_1^2]}{Jt_1} + \frac{\mathbb{E}[\varepsilon^2]}{Jt_1^2} + 4 \frac{m_{\max}}{(Jt_1)^\eta} \\
&\leq |\varphi_X(u)|^2 + \frac{2|\varphi_X(u)|^2}{(1+u^2)} \frac{(1+u^2)}{d} + \frac{\mathbb{E}[X_1^2]}{Jt_1} + \frac{\mathbb{E}[\varepsilon^2]}{Jt_1^2} + 4 \frac{(t_2 - t_1)^{2\alpha}}{J \cdot t_1^\eta} \\
&\leq C_0 |\varphi_X(u)|^2,
\end{aligned}$$

as $J \rightarrow \infty$. From now on, all that remains to be done is to look at the event

$$\mathcal{A} = \left\{ \left| \frac{\varphi_X(u)}{\sqrt{1+u^2}} \right| \leq \frac{2e^{2t_2}}{\sqrt{J}(t_2 - t_1)} \right\}.$$

By definition, we know that

$$\bar{\varphi}_X^J(u) \leq 1 + \frac{2 \log(J)}{(t_2 - t_1)}.$$

Then, on the event \mathcal{A} we see that

$$\begin{aligned}
&\mathbb{E} \left[\int_{[m, \hat{m}_m]} |\bar{\varphi}_X^J(u) - \varphi(u)|^2 \cdot \mathbb{1}_{\mathcal{A}} \right] \\
&\leq \int_{|u| \in [m, (m_{\max})^\alpha]} |\varphi(u)|^2 du + \left(2 + \frac{2 \log(J)}{(t_2 - t_1)} \right)^2 \int_{|u| \in [m, (m_{\max})^\alpha]} \mathbb{P} \left(\left| \frac{\hat{\varphi}_X(u)}{\sqrt{1+u^2}} \right| \geq \frac{\kappa_{J, t_1, t_2}}{\sqrt{J}(t_2 - t_1)} \right) \mathbb{1}_{\mathcal{A}} du \\
&\leq \int_{|u| \in [m, (m_{\max})^\alpha]} |\varphi(u)|^2 du + \left(2 + \frac{2 \log(J)}{(t_2 - t_1)} \right)^2 \int_{|u| \in [m, (m_{\max})^\alpha]} \mathbb{P} \left(\left| \frac{\hat{\varphi}_X(u)}{\sqrt{1+u^2}} \right| \geq \frac{\kappa_{J, t_1, t_2}}{\sqrt{J}(t_2 - t_1)} \right) \mathbb{1}_{\mathcal{A}} du \\
&\leq \int_{|u| \in [m, (m_{\max})^\alpha]} |\varphi(u)|^2 du + \left(2 + \frac{2 \log(J)}{(t_2 - t_1)} \right)^2 \int_{|u| \in [m, (m_{\max})^\alpha]} \mathbb{P} \left(\left| \frac{\hat{\varphi}_X(u) - \varphi(u)}{\sqrt{1+u^2}} \right| \geq \frac{\kappa \sqrt{\ln(J(t_2 - t_1)^2)}}{\sqrt{J}(t_2 - t_1)} \right) du \\
&\leq \int_{|u| \in [m, (m_{\max})^\alpha]} |\varphi(u)|^2 du + \left(2 + \frac{2 \log(J)}{(t_2 - t_1)} \right)^2 \int_{|u| \in [m, (m_{\max})^\alpha]} \mathbb{P} \left(|\hat{\varphi}_X(u) - \varphi(u)| \geq \frac{\kappa \sqrt{\ln(J(t_2 - t_1)^2)}}{\sqrt{J}(t_2 - t_1)} \right) du \\
&\leq \int_{|u| \in [m, (m_{\max})^\alpha]} |\varphi(u)|^2 du + T_J \left(2 + \frac{2 \log(J)}{(t_2 - t_1)} \right)^2.
\end{aligned}$$

Finally, it only remains to bound T_J where

$$T_J = \int_{|u| \in [m, (m_{\max})^\alpha]} \mathbb{P} \left(|\hat{\varphi}_X(u) - \varphi(u)| \geq \frac{\kappa \sqrt{\ln(J(t_2 - t_1)^2)}}{\sqrt{J}(t_2 - t_1)} \right) du.$$

By definition, we have

$$\begin{aligned}
 & \mathbb{P}\left(\left|\widehat{\varphi}_X(u) - \varphi(u)\right| \geq \frac{\kappa\sqrt{\ln(J(t_2 - t_1)^2)}}{\sqrt{J}(t_2 - t_1)}\right) \\
 &= \mathbb{P}\left(\left|\frac{\log \widehat{\varphi}_{Z_{t_2}}^J(u) - \log \widehat{\varphi}_{Z_{t_1}}^J(u) - \log \varphi_{Z_{t_2}}(u) + \log \varphi_{Z_{t_1}}(u)}{t_2 - t_1}\right| \geq \frac{\kappa\sqrt{\ln(J(t_2 - t_1)^2)}}{\sqrt{J}(t_2 - t_1)}\right) \\
 &\leq \mathbb{P}\left(\left|\log \widehat{\varphi}_{Z_{t_1}}^J(u) - \log \varphi_{Z_{t_1}}(u)\right| + \left|\log \widehat{\varphi}_{Z_{t_2}}^J(u) - \log \varphi_{Z_{t_2}}(u)\right| \geq (t_2 - t_1) \frac{\kappa\sqrt{\ln(J(t_2 - t_1)^2)}}{\sqrt{J}(t_2 - t_1)}\right) \\
 &\leq \mathbb{P}\left(\left|\log \widehat{\varphi}_{Z_{t_1}}^J(u) - \log \varphi_{Z_{t_1}}(u)\right| \geq (t_2 - t_1) \frac{1}{2} \cdot \frac{\kappa\sqrt{\ln(J(t_2 - t_1)^2)}}{\sqrt{J}(t_2 - t_1)}\right) \\
 &\quad + \mathbb{P}\left(\left|\log \widehat{\varphi}_{Z_{t_2}}^J(u) - \log \varphi_{Z_{t_2}}(u)\right| \geq (t_2 - t_1) \frac{1}{2} \cdot \frac{\kappa\sqrt{\ln(J(t_2 - t_1)^2)}}{\sqrt{J}(t_2 - t_1)}\right).
 \end{aligned}$$

Without loss of generality, we only consider the term

$$\mathbb{P}\left(\left|\log \widehat{\varphi}_{Z_{t_2}}^J(u) - \log \varphi_{Z_{t_2}}(u)\right| \geq (t_2 - t_1) \frac{\kappa\sqrt{\ln(J(t_2 - t_1)^2)}}{2\sqrt{J}(t_2 - t_1)}\right).$$

By taking γ_Δ and $\zeta > 0$, like in the proof of Theorem 1, we have that

$$\begin{aligned}
 & \mathbb{P}\left(\left|\log \widehat{\varphi}_{Z_{t_2}}^J(u) - \log \varphi_{Z_{t_2}}(u)\right| \geq (t_2 - t_1) \frac{\kappa\sqrt{\ln(J(t_2 - t_1)^2)}}{2\sqrt{J}(t_2 - t_1)}\right) \\
 &\leq \mathbb{P}\left(\left|\widehat{\varphi}_{Z_{t_2}}^J(u) - \varphi_{Z_{t_2}}(u)\right| \geq (t_2 - t_1) |\varphi_{Z_{t_2}}(u)| \frac{\kappa\sqrt{\ln(J(t_2 - t_1)^2)}}{2\sqrt{J}(t_2 - t_1)}\right) + \mathbb{P}(\Omega_{\zeta, (t_2 - t_1)^2}^c).
 \end{aligned}$$

Using hypothesis on the regularity of $|\varphi_\varepsilon|$, we have that

$$|\varphi_{Z_{t_2}}| = |\varphi_{Y_{t_2}}| \cdot |\varphi_\varepsilon| \geq |\varphi_{Y_{t_2}}| \cdot \frac{d}{\sqrt{1 + u^2}} \geq e^{2t_2} \cdot \frac{d}{\sqrt{1 + (m_{\max})^2}}.$$

$$\text{Let } c(\theta) = \kappa(t_2 - t_1)e^{2t_2} \cdot \frac{d}{\sqrt{1 + (m_{\max})^2}}.$$

It follows from Hoeffding inequality and Lemma 17 that

$$\begin{aligned}
 & \mathbb{P}\left(\left|\log \widehat{\varphi}_{Z_{t_2}}^J(u) - \log \varphi_{Z_{t_2}}(u)\right| \geq \frac{\kappa\sqrt{\ln(J(t_2 - t_1)^2)}}{2\sqrt{J}(t_2 - t_1)}\right) \\
 &\leq 2(J(t_2 - t_1)^2)^{-c(\theta)^2} + \frac{\mathbb{E}[X_1^2]}{J(t_2 - t_1)^2} + \frac{\mathbb{E}[\varepsilon^2]}{J(t_2 - t_1)^4} + 4 \frac{m_{\max}}{(J(t_2 - t_1)^2)^\eta}.
 \end{aligned}$$

If we take $\eta > 3$ such that $2\alpha - \eta < -1$ and $\zeta > \sqrt{(t_2 - t_1)^2(1 + 2\eta)}$, then we have

$$\begin{aligned}
 & \int_{|u| \in [m, (m_{\max})^\alpha]} \mathbb{P}\left(\left|\log \widehat{\varphi}_{Z_{t_2}}^J(u) - \log \varphi_{Z_{t_2}}(u)\right| \geq \frac{\kappa\sqrt{\ln(J(t_2 - t_1)^2)}}{2\sqrt{J}(t_2 - t_1)}\right) \\
 &\leq 4(J(t_2 - t_1)^2)^{\alpha - c(\theta)^2} + \frac{C'}{J(t_2 - t_1)^2} + \frac{C'}{J(t_2 - t_1)^4}
 \end{aligned}$$

where C' depends on $\mathbb{E}[X_1^2], \mathbb{E}[\varepsilon^2]$. A similar reasoning on

$$\int_{|u| \in [m, (m_{\max})^\alpha]} \mathbb{P} \left(\left| \log \widehat{\varphi}_{Z_{t_1}}^J(u) - \log \varphi_{Z_{t_1}}(u) \right| \geq \frac{\kappa \sqrt{\ln(J(t_2 - t_1)^2)}}{2\sqrt{J}(t_2 - t_1)} \right),$$

leads to

$$T_J \leq C_0 (J(t_2 - t_1)^2)^{\alpha - c(\theta)^2} + \frac{C_1}{J(t_2 - t_1)^2} + \frac{C_2}{J(t_2 - t_1)^4}$$

where C_0, C_1 and C_2 depends on $\mathbb{E}[X_1^2], \mathbb{E}[\varepsilon^2]$.

5.3. Proof of Theorem 10. This proof is an adaptation of the proof of Theorem 3.1 of Duval and Kappus [DK19] for the settings of the multiplicative decomposing.

We have the decomposition

$$\|\widehat{f}_m - f\|_{\omega_1}^2 \leq \|f_m - f\|_{\omega_1}^2 + \|\widehat{f}_m - f_m\|_{\omega_1}^2.$$

To prove the theorem, we have to bound the variance term $\|\widehat{f}_m - f_m\|_{\omega_1}^2$. Using the isometry equation, we get

$$\|\widehat{f}_m - f_m\|_{\omega_1}^2 = \frac{1}{2\pi} \int_{-m}^m |\widetilde{\mathcal{M}}_1[f](s) - \mathcal{M}_1[f](s)|^2 ds.$$

To get a majorization of the right side term, we consider different events on which it can be control.

Fix $\gamma > \zeta$ and set $\gamma_\Delta = \frac{2\zeta}{1 \wedge \Delta}$. We consider the events

$$\Omega_{\zeta, \Delta}(m) = \left\{ \forall u \in [-m, m], |\widetilde{\mathcal{M}}_1[\Delta](u) - \mathcal{M}_1[\Delta](u)| \leq \zeta \sqrt{\frac{\log(n\Delta)}{n\Delta}} \right\}$$

and

$$\{m \leq M_{n, \Delta}^{(\gamma_\Delta)}\} \quad \text{where} \quad M_{n, \Delta}^{(\gamma_\Delta)} = \min \left\{ u \geq 0 : |\mathcal{M}_1[\Delta](u)| = \gamma c_m \sqrt{\log(Jt)/(Jt)} \right\}.$$

We have

$$\begin{aligned} \frac{1}{2\pi} \int_{-m}^m |\widetilde{\mathcal{M}}_1[f](s) - \mathcal{M}_1[f](s)|^2 ds &= \frac{1}{2\pi} \int_{-m \wedge M_{n, \Delta}^{(\gamma_\Delta)}}^{m \wedge M_{n, \Delta}^{(\gamma_\Delta)}} \mathbb{1}_{\Omega_{\zeta, \Delta}(m)} |\widetilde{\mathcal{M}}_1[f](s) - \mathcal{M}_1[f](s)|^2 ds \\ &+ \frac{1}{2\pi} \mathbb{1}_{m > M_{n, \Delta}^{(\gamma_\Delta)}} \cdot \mathbb{1}_{\Omega_{\zeta, \Delta}(m)} \int_{|s| \in [M_{n, \Delta}^{(\gamma_\Delta)}, m]} |\widetilde{\mathcal{M}}_1[f](s) - \mathcal{M}_1[f](s)|^2 ds \\ &+ \mathbb{1}_{\Omega_{\zeta, \Delta}^c(m)} \frac{1}{2\pi} \int_{-m}^m |\widetilde{\mathcal{M}}_1[f](s) - \mathcal{M}_1[f](s)|^2 ds \end{aligned}$$

Focus on the event $\mathcal{A} = \{|u| \leq m \wedge M_{n, \Delta}^{(\gamma_\Delta)}\} \cap \Omega_{\zeta, \Delta}(m)$. First, we prove that

$$\widehat{\mathcal{M}}_1[f](s) = \widetilde{\mathcal{M}}_1[f](s).$$

The triangle inequality ensures that

$$\widehat{\mathcal{M}}_1[f](s) \leq 1 + \frac{|\log \widehat{\mathcal{M}}_1[\Delta](s) - \log \mathcal{M}_1[\Delta](s)| + |\log \mathcal{M}_1[\Delta](s)|}{\Delta}.$$

and Lemma 18 ensures that

$$\widehat{\mathcal{M}}_1[f](s) \leq 1 + \frac{1}{\Delta} \log \left(\frac{\gamma}{\gamma - \zeta} \right) \frac{|\widehat{\mathcal{M}}_1[\Delta](s) - \mathcal{M}_1[\Delta](s)|}{|\mathcal{M}_1[\Delta](s)|} + \frac{|\log \mathcal{M}_1[\Delta](s)|}{\Delta}.$$

Furthermore, since

$$\mathcal{M}_1[f](s) = 1 + \frac{1}{\Delta} \log \mathcal{M}_1[\Delta](s),$$

it follows that

$$\widehat{\mathcal{M}}_1[f](s) \leq 3 + \frac{1}{\Delta} \log \left(\frac{\gamma}{\gamma - \zeta} \right) \leq 4.$$

Then $\widehat{\mathcal{M}}_1[f](s) = \widetilde{\mathcal{M}}_1[f](s)$.

Therefore,

$$\begin{aligned} \mathbb{E} \left[\frac{1}{2\pi} \int_{-m \wedge M_{n,\Delta}^{(\gamma_\Delta)}}^{m \wedge M_{n,\Delta}^{(\gamma_\Delta)}} \mathbf{1}_{\Omega_{\zeta,\Delta}(m)} |\widetilde{\mathcal{M}}_1[f](s) - \mathcal{M}_1[f](s)|^2 ds \right] \\ = \frac{1}{2\pi} \int_{-m \wedge M_{n,\Delta}^{(\gamma_\Delta)}}^{m \wedge M_{n,\Delta}^{(\gamma_\Delta)}} \mathbb{E} \left[\mathbf{1}_{\Omega_{\zeta,\Delta}(m)} |\widehat{\mathcal{M}}_1[f](s) - \mathcal{M}_1[f](s)|^2 \right] ds \\ = \frac{1}{2\pi \Delta^2} \int_{-m \wedge M_{n,\Delta}^{(\gamma_\Delta)}}^{m \wedge M_{n,\Delta}^{(\gamma_\Delta)}} \mathbb{E} \left[\mathbf{1}_{\Omega_{\zeta,\Delta}(m)} |\log \widehat{\mathcal{M}}_1[\Delta](s) - \log \mathcal{M}_1[\Delta](s)|^2 \right] ds. \end{aligned}$$

Lemma 18 ensures that

$$\begin{aligned} \mathbb{E} \left[\frac{1}{2\pi} \int_{-m \wedge M_{n,\Delta}^{(\gamma_\Delta)}}^{m \wedge M_{n,\Delta}^{(\gamma_\Delta)}} \mathbf{1}_{\Omega_{\zeta,\Delta}(m)} |\widetilde{\mathcal{M}}_1[f](s) - \mathcal{M}_1[f](s)|^2 ds \right] \\ \leq \frac{1}{2\pi \Delta^2} \int_{-m \wedge M_{n,\Delta}^{(\gamma_\Delta)}}^{m \wedge M_{n,\Delta}^{(\gamma_\Delta)}} \frac{\mathbb{E} [|\widehat{\mathcal{M}}_1[\Delta](s) - \mathcal{M}_1[\Delta](s)|^2]}{|\mathcal{M}_1[\Delta](s)|^2} ds. \end{aligned}$$

We have

$$\mathbb{E} [|\widehat{\mathcal{M}}_1[\Delta](s) - \mathcal{M}_1[\Delta](s)|^2] = \text{Var}(\widehat{\mathcal{M}}_1[\Delta](s)) \leq \frac{1}{n} \mathbb{E} (|Z_\Delta^{c-1+it}|^2).$$

Then we have

$$\mathbb{E} \left[\frac{1}{2\pi} \int_{-m \wedge M_{n,\Delta}^{(\gamma_\Delta)}}^{m \wedge M_{n,\Delta}^{(\gamma_\Delta)}} \mathbf{1}_{\Omega_{\zeta,\Delta}(m)} |\widetilde{\mathcal{M}}_1[f](s) - \mathcal{M}_1[f](s)|^2 ds \right] \leq \frac{1}{2\pi n \Delta^2} \int_{-m}^m \frac{1}{|\mathcal{M}_1[\Delta](s)|^2} ds.$$

Then, set $\zeta > \sqrt{5\Delta}$. Using Lemma 17 with $\eta = 2$ to get

$$\mathbb{E} \left[\mathbf{1}_{\Omega_{\zeta,\Delta}^c(m)} \frac{1}{2\pi} \int_{-m}^m |\widetilde{\mathcal{M}}_1[f](s) - \mathcal{M}_1[f](s)|^2 ds \right] \leq \frac{50m}{2\pi} \left(\frac{\mathbb{E}[\ln(X_1)^2]}{\Delta n} + 4 \frac{m}{(n\Delta)^2} \right).$$

Finally, we observe that $|\mathcal{M}_1(\Delta)(s)| > e^{-2\Delta}$. Following the strategy describes in Duval Kappus [DK19], we obtain that $M_{n,\Delta}^{\gamma\Delta} = \infty$ and that

$$\mathbb{E} \left[\frac{1}{2\pi} \mathbb{1}_{m > M_{n,\Delta}^{(\gamma\Delta)}} \cdot \mathbb{1}_{\Omega_{\zeta,\Delta}(m)} \int_{|s| \in [M_{n,\Delta}^{(\gamma\Delta)}, m]} |\widetilde{\mathcal{M}}_1[f](s) - \mathcal{M}_1[f](s)|^2 ds \right] = 0.$$

5.4. Proof of Theorem 11. Let $0 < m < (n\Delta)^\alpha$. As for the Theorem 3, the proof is divided in two steps: in the first one, we control $\mathbb{E}[\|\overline{f_{\widehat{m}_n,\Delta}} - f\|_{\omega_1}^2]$ on the event $\mathcal{E} = \{\widehat{m}_n < m\}$, then we control it on the complementary event $\mathcal{E}^c = \{\widehat{m}_n \geq m\}$.

Step 1 – The first event \mathcal{E} . The triangle inequality and the isometry equality leads to

$$\begin{aligned} \mathbb{E} \left[\mathbb{1}_{\mathcal{E}} \int_{|u| \in [\widehat{m}_n, m]} |\mathcal{M}_1[f](u)|^2 du \right] &\leq 2\mathbb{E} \left[\mathbb{1}_{\mathcal{E}} \int_{|u| \in [\widehat{m}_n, m]} |\widetilde{\mathcal{M}}_1[f](u)|^2 du \right] + 2\mathbb{E} \left[\mathbb{1}_{\mathcal{E}} \int_{|u| \in [\widehat{m}_n, m]} |\widetilde{\mathcal{M}}_1[f](u) - \mathcal{M}_1[f](u)|^2 du \right] \\ &\leq 2\mathbb{E} \left[\mathbb{1}_{\mathcal{E}} \int_{|u| \in [\widehat{m}_n, m]} \frac{\kappa_{n,\Delta}^2}{n\Delta} du \right] + 2\mathbb{E} \left[\mathbb{1}_{\mathcal{E}} \int_{|u| \in [0, m]} |\widetilde{\mathcal{M}}_1[f](u) - \mathcal{M}_1[f](u)|^2 du \right] \\ &\leq 4 \frac{\kappa_{n,\Delta}^2 m}{n\Delta} + \frac{1}{\pi n\Delta^2} \int_{-m}^m \frac{1}{|\mathcal{M}_1[\Delta](s)|^2} ds + \frac{50}{\pi} m \left(\frac{\mathbb{E}[\ln(X_1)^2]}{\Delta n} + 4 \frac{m}{(n\Delta)^2} \right). \end{aligned}$$

By definition, $\kappa_{n,\Delta} = e^{2\Delta} + \kappa \sqrt{\log(n\Delta)}$. It follows from Theorem 10 that

$$\begin{aligned} \mathbb{E} \left[\mathbb{1}_{\mathcal{E}} \|\overline{f_{\widehat{m}_n,\Delta}} - f\|_{\omega_1}^2 \right] &\leq \|f_m - f\|_{\omega_1}^2 + C \frac{\log(n\Delta)m}{n\Delta} \\ &\quad + \frac{1}{\pi n\Delta^2} \int_{-m}^m \frac{1}{|\mathcal{M}_1[\Delta](s)|^2} ds + \frac{50}{\pi} m \left(\frac{\mathbb{E}[\ln(X_1)^2]}{\Delta n} + 4 \frac{m}{(n\Delta)^2} \right). \end{aligned}$$

Step 2 – The second event \mathcal{E}^c . Henceforth, we considere the case where $\widehat{m}_n \geq m$. It remains to control the surplus in the variance of $\widetilde{f_{\widehat{m}_n}}$.

Since $m \leq \widehat{m}_n \leq (n\Delta)^\alpha$, it follows that

$$\mathbb{E} \left[\mathbb{1}_{\mathcal{E}^c} \int_{|u| \in [m, \widehat{m}_n]} |\overline{\mathcal{M}}_1[f](u) - \mathcal{M}[f](u)|^2 \right] \leq \left[\int_{|u| \in [m, (n\Delta)^\alpha]} \mathbb{E} [|\overline{\mathcal{M}}_1[f](u) - \mathcal{M}[f](u)|^2] \right].$$

Let $\eta > 2$ such that $\alpha - \eta < -1$. Lemma 19 and Lemma 20 ensure that

$$\begin{aligned} \mathbb{E} [|\overline{\mathcal{M}}_1[f](u) - \mathcal{M}[f](u)|^2] &\leq |\mathcal{M}_1[f](s)|^2 + \mathbb{E} [|\widehat{\mathcal{M}}_1[f](u) - \mathcal{M}[f](u)|^2] \\ &\leq |\mathcal{M}_1[f](s)|^2 + \frac{2}{n\Delta^2 |\mathcal{M}_1[\Delta](s)|^2} + \frac{\mathbb{E}[\ln(X_1)^2]}{\Delta n} + \frac{4}{n\Delta}. \end{aligned}$$

First, we assume that $\{|\mathcal{M}_1[f](s)|^2 > \frac{e^{2\Delta}}{\sqrt{n\Delta}}\}$. Given that $\mathcal{M}_1[\Delta](u) > e^{-2\Delta}$, it follows in this case that

$$\mathbb{E} [|\overline{\mathcal{M}}_1[f](u) - \mathcal{M}_1[f](u)|^2] \leq |\mathcal{M}_1[f](u)|^2 (6 + \mathbb{E}[\ln(X_1)^2]).$$

Therefore,

$$\begin{aligned} \mathbb{E} \left[\int_{|u| \in [m, \widehat{m}_n]} |\overline{\mathcal{M}}_1[f](u) - \mathcal{M}[f](u)|^2 \cdot \mathbf{1}_{\{|\mathcal{M}_1[f](s)|^2 > \frac{e^{2\Delta}}{\sqrt{n\Delta}}\}} \cdot \mathbf{1}_{\mathcal{E}^c} \right] du &\leq A \int_{|u| \in [m, \widehat{m}_n]} |\mathcal{M}_1[f](u)|^2 du \\ &\leq A \int_{[-m, m]^c} |\mathcal{M}_1[f](u)|^2 du, \end{aligned}$$

where $A = 6 + \mathbb{E}[\ln(X_1)^2]$.

Secondly, we assume that $\{|\mathcal{M}_1[f](s)|^2 \leq \frac{e^{2\Delta}}{\sqrt{n\Delta}}\}$. By construction, $|\overline{\mathcal{M}}_1[f](s)| \leq 4$. It follows that

$$\begin{aligned} \mathbb{E} \left[\int_{|u| \in [m, \widehat{m}_n]} |\overline{\mathcal{M}}_1[f](u) - \mathcal{M}[f](u)|^2 \cdot \mathbf{1}_{\{|\mathcal{M}_1[f](s)|^2 \leq \frac{e^{2\Delta}}{\sqrt{n\Delta}}\}} \cdot \mathbf{1}_{\mathcal{E}^c} \right] du \\ \leq \int_{|u| \in [m, (n\Delta)^\alpha]} |\mathcal{M}_1[f](u)|^2 du + 25 \int_{|u| \in [m, (n\Delta)^\alpha]} \mathbb{P} \left(|\widehat{\mathcal{M}}_1[f]| \geq \frac{\kappa_{n, \Delta}}{\sqrt{n\Delta}} \right) \cdot \mathbf{1}_{\{|\mathcal{M}_1[f](s)|^2 \leq \frac{e^{2\Delta}}{\sqrt{n\Delta}}\}} du \\ \leq \int_{|u| \in [m, (n\Delta)^\alpha]} |\mathcal{M}_1[f](u)|^2 du + 25 \int_{|u| \in [m, (n\Delta)^\alpha]} \mathbb{P} \left(|\widehat{\mathcal{M}}_1[f] - \mathcal{M}_1[f]| \geq \kappa \sqrt{\log(n\Delta/(n\Delta))} \right) du. \end{aligned}$$

It remains to bound the last term. For that, we use Lemma 19 and Lemma 20 by taking γ_Δ and $\zeta > 0$, like in the proof of Theorem 10. By definition, we have that

$$\begin{aligned} \mathbb{P} \left(|\widehat{\mathcal{M}}_1[f] - \mathcal{M}_1[f]| \geq \kappa \sqrt{\log(n\Delta/(n\Delta))} \right) \\ = \mathbb{P} \left(|\log \widehat{\mathcal{M}}_1[\Delta] - \log \mathcal{M}_1[\Delta]| \geq \kappa \Delta \sqrt{\log(n\Delta/(n\Delta))} \right) \\ \leq \mathbb{P} \left(|\widehat{\mathcal{M}}_1[\Delta] - \mathcal{M}_1[\Delta]| \geq |\mathcal{M}_1[\Delta]| \kappa \Delta \sqrt{\log(n\Delta/(n\Delta))} \right) + \mathbb{P}(\Omega_{\zeta, \Delta}^c((n\Delta)^\alpha)). \end{aligned}$$

Let $c(\Delta) = k\Delta e^{-2\Delta}$. The Hoeffding inequality and Lemma 19 ensure that

$$\mathbb{P} \left(|\widehat{\mathcal{M}}_1[f] - \mathcal{M}_1[f]| \geq \kappa \sqrt{\log(n\Delta/(n\Delta))} \right) \leq 2(n\Delta)^{\alpha - c(\Delta)^2} + \frac{\mathbb{E}[\ln(X_1)^2]}{n\Delta} + 4(n\Delta)^{\alpha - \eta}.$$

If we take $\eta > 3$ such that $2\alpha - \eta < -1$ and $\zeta > \sqrt{\Delta(1 + 2\eta)}$, then

$$\int_{|u| \in [m, (n\Delta)^\alpha]} \mathbb{P} \left(|\widehat{\mathcal{M}}_1[f] - \mathcal{M}_1[f]| \geq \kappa \sqrt{\log(n\Delta/(n\Delta))} \right) du \leq 4(n\Delta)^{\alpha - c(\Delta)^2} + \frac{B}{n\Delta}$$

where B depends on $\mathbb{E}[\ln(X_1)^2]$.

It follows that there exist two constant A, B such that

$$\begin{aligned} \mathbb{E}[\|\widehat{f_{m_n, \Delta}} - f\|_{\omega_1}^2] \\ \leq A \left(\|f_m - f\|_{\omega_1}^2 + \frac{\log(n\Delta)m}{n\Delta} + \frac{1}{n\Delta^2} \int_{-m}^m \frac{1}{|\mathcal{M}_1[\Delta](s)|^2} ds + 4 \frac{m^2}{(n\Delta)^2} \right) \\ + B \left((n\Delta)^{\alpha - c(\Delta)^2} + \frac{1}{n\Delta} \right). \end{aligned}$$

We conclude by taking the infimum in m .

6. CONCLUSION AND PERSPECTIVES

In this article, we develop an adaptive procedure to estimate the jump density f in a noisy compound Poisson process. In our case, we observe several noisy compound Poisson processes. We have shown that by looking at what happens at two different times, it is possible to reconstruct the density f .

In practice, the experimental data allow us to observe the process at different times. A discussion could be conducted later to understand if it is possible to combine estimators built with different t_1 and t_2 to improve our estimator.

In a future work, we plan to apply our statistical method to biological data from the article by Robert et al. [ROR⁺18].

APPENDIX A. THE DISTINGUISHED LOGARITHM

In the section, we recall the definition and some properties of the distinguished logarithm. The reader who would like to have more information can refer to the articles of Duval and Kappus [DK17, DK19] or to the article of Finkelstein et al. [FTV99].

Theorem 12. *Let $d \in \mathbb{N} - \{0\}$ and let $\varphi \in \mathcal{C}^0(\mathbb{R}^d; \mathbb{C}^*)$ be a continuous application which never vanishes and such that $\varphi(0) = 1$.*

Then there exists a unique continuous application $\psi \in \mathcal{C}^0(\mathbb{R}^d; \mathbb{C})$ such that

- (1) $\psi(0) = 0$,
- (2) for all $x \in \mathbb{R}^d$, $\varphi(x) = e^{\psi(x)}$.

The application ψ is called the Distinguished Logarithm of φ and is denoted by $\log \varphi$.

Remark 13. In general, the distinguished logarithm does not reduce to a composition by the principal determination of the logarithm. In his book [Sat99], Sato remarks to his readers that one can have $\varphi(z_1) = \varphi(z_2)$ and $\log \varphi(z_1) \neq \log \varphi(z_2)$.

Indeed, consider the application $\varphi(t) = e^{it}$, ($t \in \mathbb{R}$). It verifies all the assumptions of Theorem 12. which ensures that

$$\log \varphi(t) = it, \quad (t \in \mathbb{R}),$$

It follows that

$$\varphi(0) = \varphi(2) = 1 \quad \text{and} \quad \log \varphi(0) = 0, \quad \log \varphi(2) = 2i\pi.$$

Proposition 14. (Theorem 2. of [FTV99])

Let $d > 0$ be a positive integer and $f, (f_n)_{n \in \mathbb{N}} \in \mathcal{C}^0(\mathbb{R}^d; \mathbb{C}^*)$ be continuous functions which never vanishes and such that $f(0) = f_n(0) = 1$ for all $n \in \mathbb{N}$. If (f_n) converges uniformly to f on compact subsets of \mathbb{R}^d , then the sequence $(\log(f_n))_{n \in \mathbb{N}}$ converges uniformly to $\log(f)$ on compact subsets of \mathbb{R}^d .

Proposition 15. (Lemma 3. of [DK17])

Let φ be a characteristic function without zeroes and assume that φ is differentiable. Then, it follows that

$$\log \varphi(u) = \int_0^u \frac{\varphi'(z)}{\varphi(z)} dz.$$

Corollary 16. Let $d \in \mathbb{N} - \{0\}$ and $f_1, f_2 \in \mathcal{C}^0(\mathbb{R}^d; \mathbb{C}^*)$ be two continuous functions which never vanishes and such that $f_1(0) = f_2(0) = 1$. Assume that f_1, f_2 are both differentiable and let $h : \mathbb{R}^d \rightarrow \mathbb{C}^*$ denotes the quotient $h = \frac{f_2}{f_1}$. Then,

$$\forall u \in \mathbb{R}^d, \quad \log h(u) = \log f_2(u) - \log f_1(u).$$

Proof.

$$\begin{aligned} \log h(u) &= \int_0^u \frac{h'(z)}{h(z)} dz = \int_0^u \frac{\left(\frac{f_2}{f_1}\right)'(z)}{\left(\frac{f_2}{f_1}\right)(z)} dz = \int_0^u \frac{\left(\frac{f_2' f_1 - f_2 f_1'}{f_1^2}\right)}{\left(\frac{f_2}{f_1}\right)} dz = \int_0^u \left(\frac{f_2' f_1 - f_2 f_1'}{f_1 f_2}\right) dz \\ &= \int_0^u \left(\frac{f_1'}{f_1}\right) dz - \int_0^u \left(\frac{f_2'}{f_2}\right) dz = \log f_2(u) - \log f_1(u). \end{aligned}$$

■

APPENDIX B. USEFUL LEMMAS

B.1. Lemmas for decomposing with unknown noise.

Lemma 17. Let $m, \zeta, t \in [0, \infty)$ be positive reals. We consider the event

$$\Omega_{\zeta, t}(m) = \left\{ \forall u \in [-m, m], |\widehat{\varphi}_{Z_t}^J(u) - \varphi_{Z_t}(u)| \leq \zeta \sqrt{\frac{\log(Jt)}{Jt}} \right\}.$$

If $\mathbb{E}[X_1^2] < \infty$, then

$$\forall \eta > 0, \quad \forall \zeta > \sqrt{t(1+2\eta)}, \quad \mathbb{P}(\Omega_{\zeta, t}(m)^c) \leq \frac{\mathbb{E}[X_1^2]}{Jt} + \frac{\mathbb{E}[\varepsilon^2]}{Jt^2} + 4 \frac{m}{(Jt)^\eta}.$$

Proof. Let $c, h, \tau \in [0, \infty)$ be some positive reals. We define the events

$$A(c) = \left\{ \left| \frac{1}{J} \sum_{j=1}^J |Z_t^j| - \mathbb{E}[|Z_t|] \right| \leq c \right\},$$

$$B_{h,\tau(m)} = \left\{ \forall |k| \leq \left\lceil \frac{m}{h} \right\rceil, |\widehat{\varphi}_{Z_t}^J(kh) - \varphi_{Z_t}(kh)| \leq \tau \sqrt{\frac{\log(Jt)}{Jt}} \right\}.$$

As we know that the function $x \mapsto e^{iux}$ is 1-Lipschitz, then for all $u \in \mathbb{R}$ and for all $h \in [0, \infty)$ we have

$$\begin{aligned} |\widehat{\varphi}_{Z_t}^J(u) - \widehat{\varphi}_{Z_t}^J(u+h)| &= \left| \frac{1}{J} \sum_{j=1}^J e^{iuZ_t^j} - \frac{1}{J} \sum_{j=1}^J e^{i(u+h)Z_t^j} \right| \\ &\leq \frac{1}{J} \sum_{j=1}^J \left| e^{iuZ_t^j} - e^{i(u+h)Z_t^j} \right| \\ &\leq \frac{h}{J} \sum_{j=1}^J |Z_t^j| \\ &= h \left(\frac{1}{J} \sum_{j=1}^J [|Z_t^j| - \mathbb{E}[|Z_t|]] + \mathbb{E}[|Z_t|] \right). \end{aligned}$$

The definition of $A(c)$ ensures that

$$(27) \quad \forall u \in \mathbb{R}, h > 0, |\widehat{\varphi}_{Z_t}^J(u) - \widehat{\varphi}_{Z_t}^J(u+h)| \mathbf{1}_{A(c)} \leq h(c + \mathbb{E}[|Z_t|]).$$

Moreover we have

$$(28) \quad \mathbb{E}[|Z_t|] \leq t\mathbb{E}[|X_1|] + \mathbb{E}[|\varepsilon|].$$

If $\mathbb{E}[X_1^2] < \infty$, by applying the Markov inequalities

$$\text{Var}[|Z_t|] \leq \text{Var}[Z_t] = t\mathbb{E}[X_1^2] + \mathbb{E}[\varepsilon^2],$$

we claim that

$$(29) \quad \mathbb{P}(A(c)^c) \leq \frac{t\mathbb{E}[X_1^2]}{c^2 J} + \frac{\mathbb{E}[\varepsilon^2]}{c^2 J}.$$

Moreover we have

$$(30a) \quad \mathbb{P}(B_{h,\tau(m)}^c) = \mathbb{P}\left(\exists |k| \leq \left\lceil \frac{m}{h} \right\rceil, |\widehat{\varphi}_{Z_t}^J(kh) - \varphi_{Z_t}(kh)| > \tau \sqrt{\frac{\log(Jt)}{Jt}}\right)$$

$$(30b) \quad \leq \sum_{k=-\left\lceil \frac{m}{h} \right\rceil}^{\left\lceil \frac{m}{h} \right\rceil} \mathbb{P}\left(|\widehat{\varphi}_{Z_t}^J(kh) - \varphi_{Z_t}(kh)| > \tau \sqrt{\frac{\log(Jt)}{Jt}}\right)$$

$$(30c) \quad \leq \sum_{k=-\left\lceil \frac{m}{h} \right\rceil}^{\left\lceil \frac{m}{h} \right\rceil} \mathbb{P}\left(\left|\sum_{j=1}^J e^{ikhZ_t^j} - \mathbb{E}[e^{ikhZ_t^j}]\right| > \tau \sqrt{\frac{J \log(Jt)}{t}}\right).$$

As $|e^{ikhZ_t^j}| \leq 1$ almost surely, Hoeffding's inequality ensures that

$$(30d) \quad \leq \sum_{k=-\left\lceil \frac{m}{h} \right\rceil}^{\left\lceil \frac{m}{h} \right\rceil} 2 \exp\left(-\frac{\tau^2 \log(Jt)}{2t}\right)$$

$$(30e) \quad = 4 \left\lceil \frac{m}{h} \right\rceil (Jt)^{-\tau^2/(2t)}.$$

Let $|u| \leq m$. Then there exists k a positive integer such that $u \in [kh - \frac{h}{2}, kh + \frac{h}{2}]$. Then

$$(31) \quad \mathbb{1}_{A(c) \cap B_{h,\tau(m)}} |\widehat{\varphi}_{Z_t}^J(u) - \varphi_{Z_t}(u)| \leq \mathbb{1}_{A(c) \cap B_{h,\tau(m)}} \left(|\widehat{\varphi}_{Z_t}^J(u) - \widehat{\varphi}_{Z_t}^J(kh)| \right. \\ \left. + |\widehat{\varphi}_{Z_t}^J(kh) - \varphi_{Z_t}(kh)| + |\varphi_{Z_t}(kh) - \varphi_{Z_t}(u)| \right).$$

Since the function $x \mapsto e^{ixt}$ is 1-Lipschitz, we have

$$|\varphi_{Z_t}(kh) - \varphi_{Z_t}(u)| = |\mathbb{E}[e^{ikhZ_t} - e^{iuZ_t}]| \\ \leq h \mathbb{E}[|Z_t|].$$

By applying this last result to Equation (31)

$$\mathbb{1}_{A(c) \cap B_{h,\tau(m)}} \sup_{|u| \leq m} |\widehat{\varphi}_{Z_t}^J(u) - \varphi_{Z_t}(u)| \leq \mathbb{1}_{A(c) \cap B_{h,\tau(m)}} \left(|\widehat{\varphi}_{Z_t}^J(u) - \widehat{\varphi}_{Z_t}^J(kh)| \right. \\ \left. + |\widehat{\varphi}_{Z_t}^J(kh) - \varphi_{Z_t}(kh)| + h \mathbb{E}[|Z_t|] \right).$$

It follows from Equation (27) that

$$\mathbb{1}_{A(c) \cap B_{h,\tau(m)}} \sup_{|u| \leq m} |\widehat{\varphi}_{Z_t}^J(u) - \varphi_{Z_t}(u)| \leq \\ h(c + \mathbb{E}[|Z_t|]) + \mathbb{1}_{A(c) \cap B_{h,\tau(m)}} \left(|\widehat{\varphi}_{Z_t}^J(kh) - \varphi_{Z_t}(kh)| \right) + h \mathbb{E}[|Z_t|].$$

By using the definition of $B_{h,\tau}(m)$, we have

$$\begin{aligned}
(33) \quad \mathbb{1}_{A(c) \cap B_{h,\tau}(m)} \sup_{|u| \leq m} |\widehat{\varphi}_{Z_t}^J(u) - \varphi_{Z_t}(u)| &\leq h(c + \mathbb{E}[|Z_t|]) + \tau \sqrt{\frac{\log(Jt)}{Jt}} + h\mathbb{E}[|Z_t|] \\
&\leq hc + 2h\mathbb{E}[|Z_t|] + \tau \sqrt{\frac{\log(Jt)}{Jt}} \\
&\leq hc + 2h(t\mathbb{E}[|X_1|] + \mathbb{E}[|\varepsilon|]) + \tau \sqrt{\frac{\log(Jt)}{Jt}}.
\end{aligned}$$

In particular, if $c = t$, $h = o(\sqrt{\frac{\log(Jt)}{Jt}})$ such that $h > \frac{1}{\sqrt{Jt}}$ and $\zeta > \tau$, Equation (33), allows us to say that

$$A(c) \cap B_{h,\tau}(m) \subset \Omega_{\zeta,t}(m).$$

In addition, $h > \frac{1}{\sqrt{Jt}}$, (27) et (29)), we prove that for all $\eta > 0$,

$$\begin{aligned}
\mathbb{P}(\Omega_{\zeta,t}^c(m)) &\leq \mathbb{P}(A^c(t)) + \mathbb{P}(B_{h,\tau}^c(m)) \\
&\leq \frac{\mathbb{E}[X_1^2]}{Jt} + \frac{\mathbb{E}[\varepsilon^2]}{Jt^2} + 4 \left\lceil \frac{m}{h} \right\rceil (Jt)^{-\frac{\tau^2}{2t}} \\
&\leq \frac{\mathbb{E}[X_1^2]}{Jt} + \frac{\mathbb{E}[\varepsilon^2]}{Jt^2} + 4m(Jt)^{-\frac{t-\tau^2}{2t}}.
\end{aligned}$$

We obtain the result by taking $\tau^2 = t(1 + 2\eta)$. ■

B.2. Lemmas for multiplicative decomposing.

Lemma 18. (Lemma 5.2. of [DK19]) Let $\gamma \in [0, \infty)$, $t \in [0, \infty)$ and consider

$$M_{J,t}^{(\gamma)} = \min \left\{ u \geq 0 : |\varphi_{Z_t}(u)| = \gamma \sqrt{\log(Jt)/(Jt)} \right\},$$

with the convention $\inf\{\emptyset\} = \infty$.

Let $\zeta \in [0, \infty)$ be a positive real s.t. $0 < \zeta < \gamma$. Then

$$\mathbb{1}_{|u| \leq M_{J,t}^{(\gamma)} \wedge m, \Omega_{\zeta,t}(m)} \cdot \left| \log(\widehat{\varphi}_{Z_t}^J(u)) - \log \varphi_{Z_t}(u) \right| \leq \frac{\gamma}{\zeta} \log \left(\frac{\gamma}{\gamma - \zeta} \right) \frac{|\widehat{\varphi}_{Z_t}^J(u) - \varphi_{Z_t}(u)|}{|\varphi_{Z_t}(u)|}.$$

Lemma 19. Let $m > 0$ and $\zeta > 0$. We consider the event

$$\Omega_{\zeta,\Delta}(m) = \left\{ \forall u \in [-m, m], |\widehat{\mathcal{M}}_1[\Delta](u) - \mathcal{M}_1[\Delta](u)| \leq \zeta \sqrt{\frac{\log(n\Delta)}{n\Delta}} \right\}.$$

If $\mathbb{E}[\ln(X_1)^2] < \infty$, then

$$\mathbb{P}(\Omega_{\zeta,t}^c(m)) \leq \frac{\mathbb{E}[\ln(X_1)^2]}{\Delta n} + 4 \frac{m}{(n\Delta)^\eta}.$$

Proof. Let $c, h, \tau \in [0, \infty)$ be some positive reals. We define the events

$$A(c) = \left\{ \left| \frac{1}{n} \sum_{k=1}^n |\ln(Z_{k\Delta})| - \mathbb{E}|\ln(Z_\Delta)| \right| \leq c \right\},$$

$$B_{h,\tau(m)} = \left\{ \forall |k| \leq \left\lceil \frac{m}{h} \right\rceil, \left| \widehat{\mathcal{M}}_1[\Delta](kh) - \mathcal{M}_1[\Delta](kh) \right| \leq \tau \sqrt{\frac{\log(n\Delta)}{n\Delta}} \right\}.$$

As we know that the function $u \mapsto e^{iu}$ is 1-Lipschitz, then for all $u \in \mathbb{R}$ and for all $h \in [0, \infty)$ we have

$$\begin{aligned} \left| \widehat{\mathcal{M}}_1[\Delta](u+h) - \widehat{\mathcal{M}}_1[\Delta](u) \right| &= \left| \frac{1}{n} \sum_{k=1}^n Z_{k\Delta}^{i(u+h)} - \frac{1}{n} \sum_{k=1}^n Z_{k\Delta}^{iu} \right| \\ &= \frac{1}{n} \sum_{k=1}^n \left| Z_{k\Delta}^{i(u+h)} - Z_{k\Delta}^{iu} \right| \\ &= \frac{1}{n} \sum_{k=1}^n \left| e^{i(u+h)\ln(Z_{k\Delta})} - e^{iu\ln(Z_{k\Delta})} \right| \\ &\leq \frac{h}{n} \sum_{k=1}^n \left| \ln(Z_{k\Delta}) \right| \\ &= h \left(\frac{1}{n} \sum_{k=1}^n \left[|\ln(Z_{k\Delta})| - \mathbb{E}[|\ln(Z_\Delta)|] \right] + \mathbb{E}[|\ln(Z_\Delta)|] \right). \end{aligned}$$

The definition of $A(c)$ ensures that for all $u \in \mathbb{R}$ and $h > 0$

$$(36) \quad \left| \widehat{\mathcal{M}}_1[\Delta](u+h) - \widehat{\mathcal{M}}_1[\Delta](u) \right| \mathbf{1}_{A(c)} \leq h(c + \mathbb{E}[|\ln(Z_\Delta)|]) \leq h(c + \Delta \mathbb{E}[|\ln(X_1)|]).$$

If $\mathbb{E}[\ln(X_1)^2]$ is finite, the Markov inequality leads to

$$(37) \quad \mathbb{P}(A(c)^c) \leq \frac{\Delta \mathbb{E}[\ln(X_1)^2]}{c^2 n}.$$

Moreover we have

$$\begin{aligned} \mathbb{P}(B_{h,\tau(m)}^c) &= \mathbb{P} \left(\exists |k| \leq \left\lceil \frac{m}{h} \right\rceil, \left| \widehat{\mathcal{M}}_1[\Delta](kh) - \mathcal{M}_1[\Delta](kh) \right| > \tau \sqrt{\frac{\log(n\Delta)}{n\Delta}} \right) \\ &\leq \sum_{k=-\left\lceil \frac{m}{h} \right\rceil}^{\left\lceil \frac{m}{h} \right\rceil} \mathbb{P} \left(\left| \widehat{\mathcal{M}}_1[\Delta](kh) - \mathcal{M}_1[\Delta](kh) \right| > \tau \sqrt{\frac{\log(n\Delta)}{n\Delta}} \right) \\ &\leq \sum_{k=-\left\lceil \frac{m}{h} \right\rceil}^{\left\lceil \frac{m}{h} \right\rceil} \mathbb{P} \left(\left| \sum_{k=1}^n Z_{k\Delta}^{i(kh)} - \mathbb{E}[Z_\Delta^{i(kh)}] \right| > \tau \sqrt{\frac{n \log(n\Delta)}{\Delta}} \right). \end{aligned}$$

As $|Z_\Delta^{i(kh)}| \leq 1$ almost surely, Hoeffding's inequality ensures that

$$\mathbb{P}(B_{h,\tau(m)}^c) \leq \sum_{k=-\lceil \frac{m}{h} \rceil}^{\lceil \frac{m}{h} \rceil} 2 \exp\left(-\frac{\tau^2 \log(n\Delta)}{2t}\right) = 4 \lceil \frac{m}{h} \rceil (n\Delta)^{-\tau^2/(2\Delta)}.$$

Let $|u| \leq m$. There exists k a positive integer such that $u \in [kh - \frac{h}{2}, kh + \frac{h}{2}]$. It follows that

$$\begin{aligned} \mathbb{1}_{A(c) \cap B_{h,\tau(m)}} |\widehat{\mathcal{M}}_1[\Delta](u) - \mathcal{M}_1[\Delta](u)| &\leq \mathbb{1}_{A(c) \cap B_{h,\tau(m)}} \left(|\widehat{\mathcal{M}}_1[\Delta](u) - |\widehat{\mathcal{M}}_1[\Delta](kh)| \right. \\ &\quad \left. + |\widehat{\mathcal{M}}_1[\Delta](kh) - \mathcal{M}_1[\Delta](kh)| + |\mathcal{M}_1[\Delta](kh) - \mathcal{M}_1[\Delta](u)| \right). \end{aligned}$$

We bound the three right terms by using respectively the Equation (36), the definition of $B_{h,\tau}$ and the fact that $u \rightarrow e^{iux}$ is 1-Lipschitz. It follows that

$$\mathbb{1}_{A(c) \cap B_{h,\tau(m)}} |\widehat{\mathcal{M}}_1[\Delta](u) - \mathcal{M}_1[\Delta](u)| = 2h\Delta \mathbb{E}[|\ln(X_1)|] + hc + \tau \sqrt{\frac{\log(n\Delta)}{n\Delta}}.$$

By fixing $c = \Delta$, $h = o\left(\sqrt{\frac{\log(n\Delta)}{n\Delta}}\right)$ such that $h > \frac{1}{\sqrt{n\Delta}}$ and $\zeta > \tau$, it follows that

$$A(c) \cap B_{h,\tau(m)} \subset \Omega_{\zeta,\Delta}(m).$$

In addition, $h > \frac{1}{\sqrt{n\Delta}}$, (27), we prove that for all $\eta > 0$,

$$\begin{aligned} \mathbb{P}(\Omega_{\zeta,t}^c(m)) &= \mathbb{P}(A(c)^c) + \mathbb{P}(B_{h,\tau(m)}^c) \leq \frac{\mathbb{E}[\ln(X_1)^2]}{\Delta n} + 4 \lceil \frac{m}{h} \rceil (n\Delta)^{-\tau^2/(2\Delta)} \\ &\leq \frac{\mathbb{E}[\ln(X_1)^2]}{\Delta n} + 4m(n\Delta)^{-\frac{t-\tau^2}{2t}}. \end{aligned}$$

We obtain the result by taking $\tau^2 = t(1 + 2\eta)$. ■

Lemma 20. (Duval, Kappus [DK19] Lemma 5.2) *Let $\gamma \in [0, \infty)$ and consider*

$$M_{n,\Delta}^{(\gamma)} = \min \left\{ u \geq 0 : |\mathcal{M}_1[\Delta](u)| = \gamma c_m \sqrt{\log(Jt)/(Jt)} \right\},$$

with the convention $\inf\{\emptyset\} = \infty$.

Let $\zeta \in [0, \infty)$ be a positive real s.t. $0 < \zeta < \gamma$. Then

$$\mathbb{1}_{|u| \leq M_{J,t}^\gamma \wedge m, \Omega_{\zeta,t}(m)} \left| \log \widehat{\mathcal{M}}_1[\Delta](s) - \log \mathcal{M}_1[\Delta](s) \right| \leq \frac{\gamma}{\zeta} \log \left(\frac{\gamma}{\gamma - \zeta} \right) \frac{|\widehat{\mathcal{M}}_1[\Delta](s) - \mathcal{M}_1[\Delta](s)|}{|\mathcal{M}_1[\Delta](s)|}.$$

APPENDIX C. ACKNOWLEDGMENTS

This work is a part of the author's Ph.D thesis under the supervision of Marie Doumic, Marc Hoffmann and Lydia Robert. I would like to thank for their valuable remarks on this article. The author's research is supported by a Ph.D Inria Grant.

REFERENCES

- [App09] David Applebaum. Lévy processes and stochastic calculus. Cambridge university press, 2009.
- [AZ09] Amir Averbuch and Valery Zheludev. Spline-based deconvolution. Signal Processing, 89(9):1782–1797, 2009.
- [BG03] Boris Buchmann and Rudolf Grübel. Decompounding: an estimation problem for Poisson random sums. The annals of statistics, 31(4):1054–1074, 2003.
- [BG04] Boris Buchmann and Rudolf Grübel. Decompounding Poisson random sums: recursively truncated estimates in the discrete case. Annals of the Institute of Statistical Mathematics, 56(4):743, 2004.
- [BJD11] Mohamed Z. Babai, Zied Jemai, and Yves Dallery. Analysis of order-up-to-level inventory systems with compound Poisson demand. European Journal of Operational Research, 210(3):552–558, 2011.
- [BP10] Martin Bøgsted and Susan M Pitts. Decompounding random sums: a nonparametric approach. Annals of the Institute of Statistical Mathematics, 62(5):855–872, 2010.
- [CH88] Raymond J Carroll and Peter Hall. Optimal rates of convergence for deconvolving a density. Journal of the American Statistical Association, 83(404):1184–1186, 1988.
- [CL11] Fabienne Comte and Claire Lacour. Data-driven density estimation in the presence of additive noise with unknown distribution. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 73(4):601–627, 2011.
- [Coc18] Alberto J Coca. Efficient nonparametric inference for discretely observed compound Poisson processes. Probability Theory and Related Fields, 170(1):475–523, 2018.
- [CRT06] Fabienne Comte, Yves Rozenholc, and Marie-Luce Taupin. Penalized contrast estimator for adaptive density deconvolution. Canadian Journal of Statistics, 34(3):431–452, 2006.
- [CRT07] Fabienne Comte, Yves Rozenholc, and M-L Taupin. Finite sample penalization in adaptive density deconvolution. Journal of Statistical Computation and Simulation, 77(11):977–1000, 2007.
- [Dev89] Luc Devroye. Consistent deconvolution in density estimation. The Canadian Journal of Statistics/La Revue Canadienne de Statistique, pages 235–239, 1989.
- [DK17] Céline Duval and Johanna Kappus. Nonparametric adaptive estimation for grouped data. Journal of Statistical Planning and Inference, 182:12–28, 2017.
- [DK19] Céline Duval and Johanna Kappus. Adaptive procedure for fourier estimators : application to deconvolution and decompounding. Electronic Journal of Statistics, 13(2):3424–3452, 2019.
- [Duv13] Céline Duval. Density estimation for compound Poisson processes from discrete data. Stochastic Processes and their Applications, 123(11):3963–3986, 2013.
- [EKM13] Paul Embrechts, Claudia Klüppelberg, and Thomas Mikosch. Modelling extremal events: for insurance and finance. 33, 2013.

- [EWK07] Adam Eyre-Walker and Peter D Keightley. The distribution of fitness effects of new mutations. Nature Reviews Genetics, 8(8):610–618, 2007.
- [FTV99] Mark Finkelstein, Howard G Tucker, and Jerry Alan Veeh. Extinguishing the distinguished logarithm problems. Proceedings of the American Mathematical Society, pages 2773–2777, 1999.
- [GCPD16] Antonio Gómez-Corral and Tuan Phung-Duc. Retrial queues and related models. Annals of Operations Research, 247(1):1–2, 2016.
- [GST77] Llyod Griffiths, FR Smolka, and LD Trembly. Adaptive deconvolution: A new technique for processing time-varying seismic data. Geophysics, 42(4):742–759, 1977.
- [JKPR04] Iain M Johnstone, Gérard Kerkycharian, Dominique Picard, and Marc Raimondo. Wavelet deconvolution in a periodic setting. Journal of the Royal Statistical Society: Series B (Statistical Methodology), 66(3):547–573, 2004.
- [Joh09] Jan Johannes. Deconvolution with unknown error distribution. The Annals of Statistics, 37(5A):2301–2323, 2009.
- [KH96] Deepa Kundur and Dimitrios Hatzinakos. Blind image deconvolution. IEEE signal processing magazine, 13(3):43–64, 1996.
- [LLMZ09] Yiheng Liu, Yanmei Liang, Guoguang Mu, and Xiaonong Zhu. Deconvolution methods for image deblurring in optical coherence tomography. JOSA A, 26(1):72–77, 2009.
- [MCJ21] Sergio Brenner Miguel, Fabienne Comte, and Jan Johannes. Spectral cut-off regularisation for density estimation under multiplicative measurement errors. Electronic Journal of Statistics, 15(1):3551–3573, 2021.
- [MJS23] Sergio Brenner Miguel, Jan Johannes, and Maximilian Siebel. Multiplicative deconvolution under unknown error distribution. arXiv preprint arXiv:2308.08423, 2023.
- [NH97] Michael H Neumann and Ola Hössjer. On the effect of estimating the error density in nonparametric deconvolution. Journal of Nonparametric Statistics, 7(4):307–330, 1997.
- [NR12] Richard Nickl and Markus Reiß. A Donsker theorem for lévy measures. Journal of functional analysis, 263(10):3306–3332, 2012.
- [PSM17] Eric Pantin, Jean-Luc Starck, and Fionn Murtagh. Deconvolution and blind deconvolution in astronomy. Blind Image Deconvolution, pages 301–340, 2017.
- [Rei13] Markus Reiß. Testing the characteristics of a lévy process. Stochastic Processes and their Applications, 123(7):2808–2828, 2013.
- [RMKB05] Tyler S Ralston, Daniel L Marks, Farzad Kamalabadi, and Stephen A Boppart. Deconvolution methods for mitigation of transverse blurring in optical coherence tomography. IEEE Transactions on Image Processing, 14(9):1254–1264, 2005.
- [ROR⁺18] Lydia Robert, Jean Ollion, Jérôme Robert, Xiaohu Song, Ivan Matic, and Marina Elez. Mutation dynamics and fitness effects followed in single cells. Science, 359(6381):1283–1286, 2018.
- [Sac97] Mauricio D Sacchi. Reweighting strategies in seismic deconvolution. Geophysical Journal International, 129(3):651–656, 1997.

- [Sat99] Ken-Iti Sato. Lévy processes and infinitely divisible distributions. 1999.
- [SC90] Leonard A Stefanski and Raymond J Carroll. Deconvolving kernel density estimators. *Statistics*, 21(2):169–184, 1990.
- [SNM03] Jean-Luc Starck, Mai K Nguyen, and Fionn Murtagh. Wavelets and curvelets for image deconvolution: a combined approach. *Signal processing*, 83(10):2279–2283, 2003.
- [SPM02] Jean-Luc Starck, Eric Pantin, and Fionn Murtagh. Deconvolution in astronomy: A review. *Publications of the Astronomical Society of the Pacific*, 114(800):1051, 2002.
- [Swe07] Jason R Swedlow. Quantitative fluorescence microscopy and image deconvolution. *Methods in cell biology*, 81:447–465, 2007.
- [Ulr71] Tadeusz J. Ulrych. Application of homomorphic deconvolution to seismology. *Geophysics*, 36(4):650–660, 1971.
- [Ush11] Nikolai G Ushakov. Selected topics in characteristic functions. 2011.
- [Var89] Yehuda Vardi. Multiplicative censoring, renewal processes, deconvolution and decreasing density: nonparametric estimation. *Biometrika*, 76(4):751–761, 1989.
- [VEU04] Bert Van Es and Hae Won Uh. Asymptotic normality of nonparametric kernel type deconvolution density estimators: crossing the cauchy boundary. *Nonparametric Statistics*, 16(1-2):261–277, 2004.
- [VZ92] Yoav Vardi and Cun-Hui Zhang. Large sample study of empirical distributions in a random-multiplicative censoring model. *The Annals of Statistics*, pages 1022–1039, 1992.
- [WRP⁺10] Ping Wang, Lydia Robert, James Pelletier, Wei Lien Dang, Francois Taddei, Andrew Wright, and Suckjoon Jun. Robust growth of escherichia coli. *Current biology*, 20(12):1099–1103, 2010.