Electronic Journal of Statistics
Vol. 9 (2015) 3098–3123
ISSN: 1935-7524
DOI: 10.1214/15-EJS1103

Nonparametric estimation of mark's distribution of an exponential shot-noise process*

Paul Ilhe

LTCI, Telecom ParisTech, 46 rue Barrault, 75634 Paris Cedex 13, France, e-mail: paul.ilhe@telecom-paristech.fr

CEA, LIST, 91191 Gif-sur-Yvette Cedex, France, e-mail: paul.ilhe@cea.fr

Éric Moulines

Ecole Polytechnique, Centre de Mathématiques Appliquées- UMR 7641, Route de Saclay 91128 Palaiseau, France, e-mail: eric.moulines@polytechnique.edu

Francois Roueff

LTCI, Telecom ParisTech and CNRS, 46 rue Barrault, 75634 Paris Cedex 13, France, e-mail: francois.roueff@telecom-paristech.fr

and

Antoine Souloumiac

CEA, LIST, 91191 Gif-sur-Yvette Cedex, France, e-mail: antoine.souloumiac@cea.fr

Abstract: In this paper, we consider a nonlinear inverse problem occuring in nuclear science. Gamma rays randomly hit a semiconductor detector which produces an impulse response of electric current. Because the sampling period of the measured current is larger than the mean interarrival time of photons, the impulse responses associated to different gamma rays can overlap: this phenomenon is known as *pileup*. In this work, it is assumed that the impulse response is an exponentially decaying function. We propose a novel method to infer the distribution of gamma photon energies from the indirect measurements obtained from the detector. This technique is based on a formula linking the characteristic function of the photon density to a function involving the characteristic function and its derivative of the observations. We establish that our estimator converges to the mark density in uniform norm at a polynomial rate. A limited Monte-Carlo experiment is provided to support our findings.

MSC 2010 subject classifications: 45Q05, 60G10, 62G07, 62M05.
Keywords and phrases: Shot-noise, Lévy-driven Ornstein-Uhlenbeck process, nonparametric estimation, Markov process, empirical processes, β-mixing.

Received June 2015.

*This research was partially supported by Labex DigiCosme (project ANR-11-LABEX-0045-DIGICOSME) operated by ANR as part of the program "Investissement d'Avenir" Idex Paris-Saclay (ANR-11-IDEX-0003-02).

1. Introduction

In this paper, we consider a nonlinear inverse problem arising in nuclear science: neutron transport or gamma spectroscopy. For the latter, a radioactive source, for instance an excited nucleus, randomly emits gamma photons according to a homogeneous Poisson point process. These high frequency radiations can be associated to high energy photons which interact with matter via three phenomena: the photoelectric absorption, the Compton scattering and the pair production (further details can be found in [14]). When photons interact with the semiconductor detector (usually High-Purity Germanium (HPGe) detectors) arranged between two electrodes, a number of electron-holes pairs proportional to the photon transferred energy is created. Accordingly, the electrodes generate an electric current called impulse response whenever the detector is hit by a particle, with an amplitude corresponding to the transferred energy. In this context, a feature of interest is the distribution of this energy. Indeed, it can be compared to known spectra in order to identify the composition of the nuclear source. In practice, the electric current is not continuously observed but the sampling rate is typically smaller than the mean inter-arrival time of two photons. Therefore, there is a high probability that several photons are emitted between two measurements so that the energy deposited is superimposed in the detector, a phenomenon called pile-up. Because of the pile-up, it is impossible to establish a one-to-one correspondence between a gamma ray and the associated deposited energy.

This inverse problem can be modeled as follows. The electric current generated in the detector is given by a stationary shot-noise process $\mathbf{X} = (X_t)_{t\in\mathbb{R}}$ defined by:

$$X_t = \sum_{k:T_k\le t} Y_k\, h(t - T_k)\ , \tag{1}$$

where h is the (causal) impulse response of the detector and

(SN-1) $\sum_k \delta_{T_k,Y_k}$ is a Poisson point process with times $T_k \in \mathbb{R}$ arriving homogeneously with intensity $\lambda > 0$ and independent i.i.d. marks $Y_k \in \mathbb{R}$ having a probability density function (p.d.f.) θ and cumulative distribution function (c.d.f) F.

We wish to estimate the density θ from a regular observation sample $X_1, \dots, X_n$ of the shot noise (1). Note that the sampling rate is set to 1 without meaningful loss of generality. If a different sampling rate is used, e.g. we observe $X_\delta, \dots, X_{n\delta}$ for some $\delta \neq 1$, it amounts to change λ and to scale h accordingly.

The process (1) is well defined whenever the following condition holds on the impulse response h and the density θ

$$\int \min(1, |y\, h(s)|)\theta(y)\mathrm{d}y\mathrm{d}s < \infty\ . \tag{2}$$

As shown in [12], this condition is also necessary. Moreover, the marginal distribution of $\mathbf{X}$ belongs to the class of infinitely divisible (ID) distributions and has Lévy measure ν satisfying, for all Borel sets B in $\mathbb{R}\setminus\{0\}$,

$$\nu(B) \triangleq \lambda \int_0^\infty \mathbb{P}\left(h(s)Y_1 \in B\right) \, \mathrm{d}s \, . \tag{3}$$

The ID property of the marginal distribution shows that this estimation problem is closely related to the estimation of the Lévy measure ν. This property strongly suggests to use estimators of the Lévy triplet, see for instance [16] and [8]. However, up to our best knowledge, these estimators use the increments of the corresponding Lévy process which are i.i.d. and they assume a finite Lévy Khintchine measure. In contrast, the observations are not independent and the Lévy measure of the process is infinite since from (3), we have that

$$\nu(\mathbb{R}) = \lambda \int_0^\infty \mathbb{P}\left(h(s)Y_1 \in \mathbb{R}\right) \mathrm{d}s = \infty. \tag{4}$$

In order to tackle this estimation problem, we then propose to bypass the estimation of ν and directly retrieve the density θ of the marks distribution F from the empirical characteristic function of the measurements. Coarsely speaking, using (3), the Lévy-Khintchine representation provides an expression of the characteristic function φ_X of the marginal distribution as a functional of θ. The estimator is built upon replacing φ_X by its empirical version and inverting the mapping $\theta \mapsto \varphi_X$. A more standard marginal-based approach would be to rely on the p.d.f. of $\mathbf{X}$. However, the density of X_0 is intractable, which precludes the use of a likelihood inference method. Consequently, although shot-noise models are widespread in applications (for example, such models were used to model internet traffic [1], river streamflows [4], spikes in neuroscience ([11], [10]) and in signal processing ([19], [20])), theoretical results on the statistical inference of shot-noise appear to be limited. Recently, Xiao and al. ([24]) provide consistent and asymptotically normal estimators for parametric shot-noise processes with specific impulse responses.

In this contribution, we consider the particular case given by the following assumption.

(SN-2) The impulse response h is an exponential function with decreasing rate $\alpha > 0$:

$$h(t) \triangleq \mathrm{e}^{-\alpha t} \, \mathbb{1}_{\mathbb{R}_+}(t) \, .$$

Under (SN-2), the process $(X_t)_{t\in\mathbb{R}}$ is usually called an *exponential shot-noise*. In this case, Condition (2) becomes

$$\mathbb{E}\left[\log_+(|Y_1|)\right] < \infty \, . \tag{5}$$

Under (SN-2), the process $(X_t)_{t\geq 0}$ can alternatively be introduced by considering the following stochastic differential equation (SDE) :

$$\mathrm{d}X_t = -\alpha X_t \mathrm{d}t + \mathrm{d}L_t \, , \quad X_0 = x \in \mathbb{R} \tag{6}$$

where $\mathbf{L} = (L_t)_{t\geq 0}$ is a Lévy process defined as the compound Poisson process

$$L_t \triangleq \sum_{k=1}^{N_\lambda(t)} Y_k \quad \text{with} \quad N_\lambda(t) \triangleq \sum_k \mathbb{1}_{T_k \leq t} \, , \tag{7}$$

where $(T_k, Y_k)_{k\geq 0}$ satisfies (SN-1). The solution to the equation (6) is called a Ornstein-Uhlenbeck(O-U) process ([18, Chapter 17]) driven by $\mathbf{L}$ with initial condition $X_0 = x$ and rate α. Note that $\mathbf{L}$ defined by (7) has Lévy measure λF. Thus, by [18, Theorem 17.5], this Markov process admits a unique stationary version if (5) holds, and this stationary solution corresponds to the shot-noise process (1).

In recent works, [3, Brockwell, Schlemm] exploit the integrated version of (6) to recover the Lévy process $\mathbf{L}$ and show that the increments of $\mathbf{L}$ can be represented as:

$$L_{nh} - L_{(n-1)h} = X_{nh} - X_{(n-1)h} + \alpha \int_{(n-1)h}^{nh} X_s \mathrm{d}s.$$

These quantities are only well estimated for high frequency observations so that we cannot rely on this method in our regular sampling scheme.

To the best of our knowledge, the paper that best fits our setting is [13]. The authors propose a nonparametric estimation procedure from a low frequency sample of a stationary O-U process which exploits the self decomposability property of the marginal distribution. The authors construct an estimator of the so called canonical function k defined by:

$$\nu(\mathrm{d}x) = \frac{k(x)}{x}\mathrm{d}x.$$

The two main additional assumptions are that k is decreasing on $(0, \infty)$ and ν satisfies the integrability condition $\int_0^\infty (1 \wedge x)\nu(\mathrm{d}x) < \infty$. In our setting (i.e. when specifying the Lévy process to be the compound Poisson process defined in (7)), it is easily shown that these conditions hold and the canonical function and the cumulative distribution of the marks are related by the equation:

$$k(x) = \lambda \mathbb{P}\left(Y_0 > x\right) = \lambda\left(1 - F(x)\right).$$

In this article, we introduce an estimator of θ based on the empirical characteristic function and a Fourier inversion formula. This algorithm is numerically efficient, being able to handle large datasets typically used in high-energy physics. Secondly, we establish an upper bound of the rate of convergence of our estimator which is uniform over a smoothness class of functions for the density θ.

The paper is organized as follows. In Section 2, we introduce some preliminaries on the characteristic function of an exponential shot-noise process and provide both the inversion formula and the estimator of the density θ. In particular, we derive an upper bound of convergence for our estimator over a broad class of densities under the assumption that λ/α is known. In Section 3, we present in details the algorithm used to perform the density estimation and illustrate our findings with a limited Monte-Carlo experiment. Section 4 provides error bounds for the empirical characteristic function based on discrete-time observations and exploit the β-mixing structure of the process. Finally, Section 5 is devoted to the proofs of the various theorems.

2. Main result

2.1. Inversion formula

As mentioned in the introduction, it is difficult to derive the probability density function of the stationary shot-noise unless the marks are distributed according to an exponential random variable and the impulse response is an exponential function. In this case, it turns out that the marginal distribution of the shot-noise is Gamma-distributed (the reader can refer to [2] for details). In all other cases, we can only compute the characteristic function of the marginal distribution of the stationary version of the shot-noise when treating it as a filtered point process (see for example [17] for details). We have for every real u:

$$\varphi(u) \stackrel{\Delta}{=} \mathbb{E}[\mathrm{e}^{\mathrm{i}uX_0}] = \exp\left(\lambda \int_{\mathbb{R}} \int_0^{\infty} (\mathrm{e}^{\mathrm{i}uyh(v)} - 1)\mathrm{d}v\, F(\mathrm{d}y)\right). \tag{8}$$

From (8), the characteristic function of X_0 can be expressed as follows:

$$\varphi_{X_0}(u) = \exp\left(\int_{\mathbb{R}} \lambda K_h(uy) F(\mathrm{d}y)\right). \tag{9}$$

where K_h, the kernel associated to h is given by:

$$K_h(x) \stackrel{\Delta}{=} \int_0^{+\infty} (\mathrm{e}^{\mathrm{i}xh(v)} - 1)\mathrm{d}v.$$

Note that if h is integrable, then K_h is well defined since, for any real x, $\int_0^{\infty} |\mathrm{e}^{\mathrm{i}xh(s)} - 1|\mathrm{d}s \leq |x| \int_0^{\infty} |h(s)|\mathrm{d}s$. Moreover, if h is integrable, then K_h is a $\mathcal{C}^1(\mathbb{R}, \mathbb{C})$ function whose derivative is bounded and equal to:

$$K_h'(x) = \int_0^{+\infty} \mathrm{i}h(s)\mathrm{e}^{\mathrm{i}xh(v)}\mathrm{d}v.$$

Furthermore, if $\mathbb{E}[|Y_0|] < \infty$, then the characteristic function of X_0 is differentiable and we have:

$$\varphi'_{X_0}(u) = \lambda\, \varphi_{X_0}(u) \int_{\mathbb{R}} y K_h'(uy) F(\mathrm{d}y). \tag{10}$$

Under (SN-2), the kernel K_h takes the form

$$K_h(u) = \int_0^{\infty} \left(\mathrm{e}^{\mathrm{i}u\mathrm{e}^{-\alpha v}} - 1\right) \mathrm{d}v = \int_0^{u} \frac{\mathrm{e}^{\mathrm{i}s} - 1}{\alpha s} \mathrm{d}s. \tag{11}$$

With (10), we obtain that

$$\varphi'_{X_0}(u) = \varphi_{X_0}(u) \frac{\lambda}{\alpha u} \left(\varphi_{Y_0}(u) - 1\right). \tag{12}$$

Since the marginal distribution of X is infinitely divisible, we have by [18, Lemma 7.5.] that $\varphi_X(u)$ does not vanish. If in addition φ_Y is integrable, (12) provides a way to recover θ knowing α/λ, namely, for all $x \in \mathbb{R}$,

$$\theta(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}xu}\varphi_{Y_0}(u)du = \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}xu}\left(1+\frac{\alpha u}{\lambda}\frac{\varphi'_{X_0}(u)}{\varphi_{X_0}(u)}\right)\mathrm{d}u\ . \tag{13}$$

This relation shows that the estimation problem of the p.d.f. θ is directly related to the estimation of the second characteristic function.

Remark 2.1. We assume in the following that the ratio α/λ appearing in the inversion formula (13) is a known constant, as it typically depends on the measurement device. Interestingly, however, an estimator of this constant can be derived from [12, Theorem 1], where it is shown that the marginal distribution G of the stationary shot-noise is regularly varying at 0 with index λ/α, i.e.:

$$G(x) \sim x^{\lambda/\alpha}L(x) \quad , \quad x \to 0$$

with L being slowly varying at 0. Hence it is possible to estimate α/λ by applying Hill's estimator [9] to the sample $X_1^{-1}, \cdots, X_n^{-1}$.

2.2. Nonparametric estimation

Let $\hat{\varphi}_n(u) \stackrel{\Delta}{=} n^{-1}\sum_{j=1}^n \mathrm{e}^{\mathrm{i}uX_j}$ denotes the empirical characteristic function (e.c.f.) obtained from the observations and $\hat{\varphi}'_n$ its derivative. From (13), we are tempted to plug the e.c.f. of the observations to estimate the p.d.f. θ. Let $(h_n)_{n\geq 0}$ and$(\kappa_n)_{n\geq 0}$ be two sequences of positive numbers such that

$$\lim_{n\to\infty} h_n = \lim_{n\to\infty} \kappa_n = 0\ ,$$

and consider the following sequence of estimators:

$$\hat{\theta}_n(x) \stackrel{\Delta}{=} \max\left(\frac{1}{2\pi}\int_{-\frac{1}{h_n}}^{\frac{1}{h_n}} \mathrm{e}^{-\mathrm{i}xu}\left(1+\frac{\alpha u}{\lambda}\frac{\hat{\varphi}'_n(u)}{\hat{\varphi}_n(u)}\mathbb{1}_{|\hat{\varphi}_n(u)|>\kappa_n}\right)\mathrm{d}u, 0\right)\ . \tag{14}$$

Remark 2.2. We estimate $1/\varphi(u)$ by $\mathbb{1}_{\{|\hat{\varphi}_n(u)|\geq\kappa_n\}}/\hat{\varphi}_n(u)$ with a suitable choice of a sequence $(\kappa_n)_{n\geq 1}$ which converges to zero. The constant κ_n is chosen such that $|\hat{\varphi}_n(u)-\varphi(u)|$ remains smaller than $|\hat{\varphi}_n(u)|$ and $|\varphi(u)|$ with high probability in order to avoid large errors when inverting $\hat{\varphi}_n(u)$. In [16], the authors deal with the empirical characteristic function of i.i.d. random variables. In this case, the deviations of $\sqrt{n}(\hat{\varphi}_n(u)-\varphi(u))$ are bounded in probability, hence, they use $1_{\{|\hat{\varphi}_n(u)|\geq\kappa n^{-1/2}\}}/\hat{\varphi}_n(u)$ as an estimator of $1/\varphi(u)$. Here we truncate the interval of integration $\mathbb{R}$ by $\left[-h_n^{-1}, h_n^{-1}\right]$, where h_n is a bandwidth parameter. This allows us to bound the estimation error $\hat{\theta}_n - \theta$ in sup norm. The deviation of $\sqrt{n}(\hat{\varphi}_n(u)-\varphi(u))$ on $\left[-h_n^{-1}, h_n^{-1}\right]$ depends on h_n, see Theorem 4.1. The resulting κ_n is then taken slightly larger than $n^{-1/2}$.

In order to evaluate the convergence rate of our estimator, we consider particular smoothness classes for the density θ. Namely we define, for any positive constants K, L, m and $s > 1/2$,

$$\Theta(K,L,s,m) = \Big\{\theta \text{ is a density s.t. } \int |y|^{4+m}\theta(y)\,\mathrm{d}y \le K\,, \quad \int_{\mathbb{R}} (1+|u|^2)^s\,|\mathcal{F}\theta(u)|^2\,\mathrm{d}u \le L^2\Big\}\,, \tag{15}$$

where $\mathcal{F}\theta$ denotes the Fourier transform of θ

$$\mathcal{F}\theta(u) = \int \theta(y)\,\mathrm{e}^{-\mathrm{i}y\,u}\,\mathrm{d}y \quad , \quad u \in \mathbb{R}\,.$$

Hence L is an upper bound of the Sobolev semi-norm of θ. Note also that under (SN-1)–(SN-2), θ belongs to $\Theta(K,L,s,m)$ is equivalent to assuming that

$$\mathbb{E}\left[|Y_0|^{4+m}\right] \le K \quad \text{and} \quad \int_{\mathbb{R}} (1+|u|^2)^s |\varphi_{Y_0}(u)|^2 du \le L^2\,.$$

In the following, under assumptions (SN-1)–(SN-2), we use the notation $\mathbb{P}_\theta$ and $\mathbb{E}_\theta$, where the subscript θ added to the expectation and probability symbols indicates explicitly the dependence on the unknown density θ. The following result provides a bound of the risk $\mathbb{P}_\theta(\|\theta-\hat{\theta}_n\|_\infty > M_n)$ for well chosen sequences (h_n), (κ_n) and (M_n), which is uniform over the densities $\theta \in \Theta(K,L,s,m)$.

Theorem 2.1. *Assume that the process $\boldsymbol{X} = (X_t)_{t\ge 0}$ given by (1) satisfies the assumptions (SN-1)–(SN-2) for some positive constants λ and α. Let K, L, m be positive constants and $s > 1/2$. Let $C_{K,L,m,\lambda/\alpha}$ be the constant defined by*

$$C_{K,L,m,\lambda/\alpha} \triangleq \exp\left(-\frac{\lambda}{\alpha}(K^{1/(4+m)} + L)\right)\,,$$

and C be a positive constant such that $0 < C < C_{K,L,m,\lambda/\alpha}$. Set

$$h_n = n^{-1/(2s+1+2\lambda/\alpha)} \quad \text{and} \quad \kappa_n = C\left(1+h_n^{-1}\right)^{-2\lambda/\alpha}\,,$$

and define $\hat{\theta}_n$ by (14). Then, for $n \ge 3$, the density estimator $\hat{\theta}_n$ satisfies

$$\sup_{\theta\in\Theta(K,L,s,m)} \mathbb{E}_\theta\left[\left\|\theta-\hat{\theta}_n\right\|_\infty\right] \le M n^{-(2s-1)/(4s+2+4\lambda/\alpha)} \log(n)^{1/2}\,, \tag{16}$$

where $M > 0$ is a constant only depending on C, K, L, m, s and λ/α.

Proof. See Section 5.4. □

Remark 2.3. The constant C in Theorem 2.1 might be adaptively chosen. Indeed, the well known relationship (see [5, Chapter 6] for example) between the cumulant function of a filtered Poisson process and its intensity measure implies that the mean μ_θ of X_0 is given by

$$\mu_\theta = \lambda\mathbb{E}_\theta[Y_0]\int_0^\infty \mathrm{e}^{-\alpha s}\mathrm{d}s = \frac{\lambda}{\alpha}\mathbb{E}_\theta[Y_0] \le \frac{\lambda}{\alpha}K^{1/(4+m)}\,. \tag{17}$$

Since $\mathbf{X}$ is ergodic (see Section 4), the empirical mean $\hat{\mu}_n$ of the sample $X_1, \cdots, X_n$ converges to μ_θ almost surely and thus, replacing C in Theorem 2.1 by

$\hat{C}_n \triangleq \exp(-\hat{\mu}_n)/2$, leads to the same rate of convergence since

$$\lim_{n\to\infty} \hat{C}_n = \exp\left(-\frac{\lambda}{\alpha}\mathbb{E}_\theta\left[Y_0\right]\right)/2 \in (0, C_{K,L,m,\lambda/\alpha}) \quad \text{a.s.}$$

Remark 2.4. This theorem provides that the error in uniform norm converges at least at a polynomial rate $n^{-(2s-1)/(4s+2+8\lambda/\alpha)}\log(n)^{1/2}$ that depends both on the quantity λ/α and the smoothness coefficient s. For a given ratio λ/α, the convergence rate becomes faster as s increases and tend to behave as $n^{-1/2}$ when $s \to \infty$. On the other side, for a given smoothness parameter s, the rate of convergence decreases when the ratio λ/α tends to infinity. This can be interpreted as the consequence of the *pileup* effect that occurs whenever the intensity λ is large or the impulse response coefficient α is close to zero.

Remark 2.5. Based on the previous theorem, one might wonder whether the rates of convergence are optimal. According to similar but not identical problems ([16], [8]) in which authors estimate in a nonparametric fashion a Lévy triplet (with finite activity) based on a low frequency sample of the associated process, the optimal rates of convergence are identical to ours. Our estimation procedure lies on stationary but dependent infinitely divisible random variables associated to an infinite Lévy measure so that these results do not apply here. However we believe that the rates obtained in Theorem 2.1 are also optimal in this dependent context. The proof of this conjecture is left for future work.

3. Experimental results

The estimation procedure based on the estimator $\hat{\theta}_n$ given by (14) can be made time-efficient and thus well suited to a very large dataset. In nuclear applications, it is usual to deal with several million of observations while the intensity of the time-arrival point process can reach several thousand of occurrences per second. Typically, the shot-noise process in nuclear applications corresponding to the electric current is discretely observed for three minutes at a sampling rate of 10Mhz and the mean number of arrivals between two observations lies between 10 and 100. Such large values for the intensity and the number of sampled points motivate us to present a practical way to compute the estimator (14).

Practical computation of the estimator

In Section 2, we have defined the estimator of mark's density by (14). Although it theoretically converges to the true density of shot-noise marks, the evaluation of the empirical characteristic function and its derivative based on observations $X_1, \cdots, X_n$ might be time-consuming when the sample size n is large. To circumvent this issue, we propose to compute the empirical characteristic function using the fast fourier transform of an appropriate histogram of the vector $X_1, \cdots, X_n$. More precisely, for a strictly positive fixed h, we consider the grid $G = \{hl : \lfloor \min_{k\le n}(X_k)/h \rfloor \le l \le \lceil \max_{k\le n}(X_k)/h \rceil\}$ and compute the normalized histogram H of the sample sequence $(X_l)_{1\le l\le n}$ with respect to the grid G

defined by

$$H(l) = \frac{1}{n}\sum_{k=1}^{n} 1_{[G(l);G(l+1)]}(X_k), \quad \lfloor \min_{k\leq n}(X_k)/h \rfloor \leq l \leq \lceil \max_{k\leq n}(X_k)/h \rceil - 1 \ .$$

Denoting $m_n \triangleq \lfloor \min_{k\leq n}(X_k)/h \rfloor$ and $M_n \triangleq \lceil \max_{k\leq n}(X_k)/h \rceil - 1$, remark that for every real u, we have:

$$\hat{\varphi}_n(u) = \frac{1}{n}\sum_{k=1}^{n} \mathrm{e}^{\mathrm{i}uX_k} = \frac{1}{n}\sum_{k=1}^{n}\sum_{l=m_n}^{M_n} 1_{[G(l);G(l+1)]}(X_k)\ \mathrm{e}^{\mathrm{i}uX_k} \ .$$

Replacing $1_{[G(l);G(l+1)]}(X_k)\ \mathrm{e}^{\mathrm{i}uX_k}$ by $1_{[G(l);G(l+1)]}(X_k)\ \mathrm{e}^{\mathrm{i}uh(l+1/2)}$ for any real u, we get an approximation of the empirical characteristic function by defining

$$\hat{\varphi}_{h,n}(u) \triangleq \sum_{l=m_n}^{M_n} H(l)\mathrm{e}^{\mathrm{i}uh(l+1/2)} \ . \tag{18}$$

For any real u, we have the following upper bounds

$$|\hat{\varphi}_{h,n}(u) - \hat{\varphi}_n(u)| \leq \frac{h}{2}|u|$$

and

$$\left|\hat{\varphi}'_{h,n}(u) - \hat{\varphi}'_n(u)\right| \leq \frac{h}{2}\left(1 + |u|h\sum_{l=m_n}^{M_n} H(l)\ (l+1/2)\right) \quad ,$$

showing that the approximations are close to the true functions for small values of h and u. From these empirical characteristic functions, we construct an estimator of the marks' characteristic function φ_Y setting for any positive u:

$$\hat{\varphi}_{Y,h,n}(u) \triangleq 1 + \frac{\alpha}{\lambda}u\frac{\hat{\varphi}'_{h,n}(u)}{\hat{\varphi}_{h,n}(u)}\mathbb{1}_{|\hat{\varphi}_{h,n}(u)|>\kappa_n} \ .$$

The advantage of using $\hat{\varphi}_{h,n}$ is that $\hat{\varphi}_{h,n}(u)$ and $\hat{\varphi}'_{h,n}(u)$ can be evaluated on a regular grid using the fast Fourier Transform algorithm.

The last step in the numerical computation of the estimator (14) consists in evaluating the quantity

$$\int_0^{h_n^{-1}} \mathrm{e}^{-\mathrm{i}xu}\hat{\varphi}_{Y,h,n}(u)\mathrm{d}u = \int_0^{\infty} \mathrm{e}^{-\mathrm{i}xu}\hat{\varphi}_{Y,h,n}(u)1_{\left[0,h_n^{-1}\right]}(u)\mathrm{d}u \ .$$

Using the Inverse fast Fourier Transform, we approximate the integral on a regular grid $x \in$ by a Riemann sum.

Numerical results

We now illustrate the finite sample behavior of our estimator on a simulated data set when the marks (in keV energy units) density follows a Gaussian mixture $\sum_{i=1}^{3} p_i\mathcal{N}_{\mu_i,\sigma_i^2}(x)$ with

$$p = [0.3\ 0.5\ 0.2] \quad , \quad \mu = [4\ 12\ 22] \quad , \quad \sigma = [1\ 1\ 0.5] \ .$$

Furthermore, in order to fit with nuclear science applications, where detectors have a time resolution of 10 Mhz, corresponding to a sampling time $\Delta = 10^{-7}$ seconds. The parameters of the experiment are set to $\alpha = 8.10^8$, $\lambda = 10^9$. Moreover, in nuclear spectrometry, the bandwidth h_n is directly related to the known precision of the measuring instrument. For the following numerical experiment, we set it to 2.5 which is in range with the detector resolution as described in [14], Chapter 4. Figure 1 below shows a simulated sample path of such a shot-noise with its associated marked point process. As shown in Figure 2 below, our estimator $\hat{\theta}_n$ defined by (14) well retrieves the three modes of the Gaussian mixture as well as the corresponding variance from a sample of size

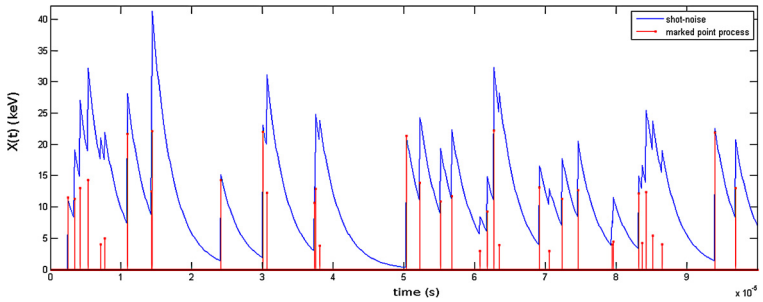

FIG 1. *Simulated Shot-Noise.*

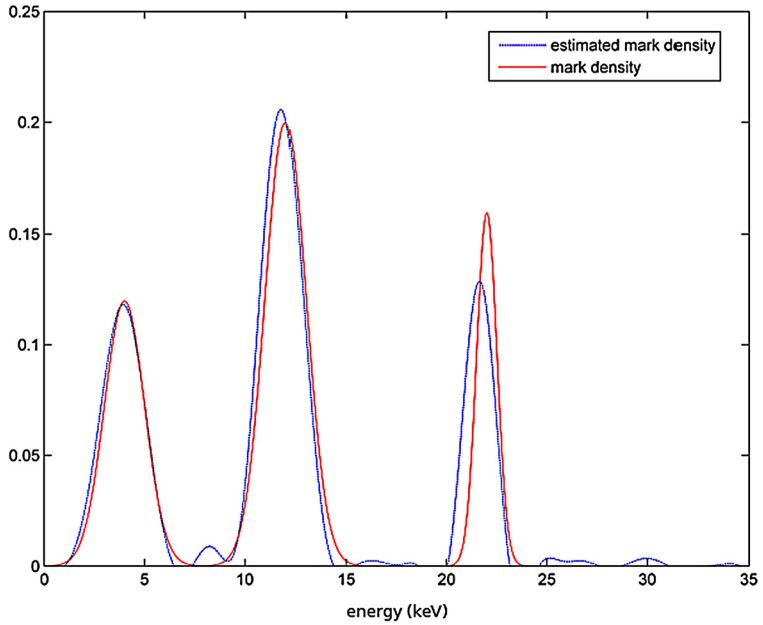

FIG 2. *Gaussian mixture case.*

TABLE 1

Mean ℓ^∞ error $\mathbb{E}_\theta[\|\theta - \hat{\theta}_n\|_\infty]$ for different sample sizes n with θ given by the Gaussian mixture defined above. The displayed values are obtained by averaging $\|\theta - \hat{\theta}_n\|_\infty$ over 100 Monte Carlo simulations. The variance of the error over the Monte Carlo runs is indicated between parentheses to assess the precision of the given estimates

Sample size (n)	Mean ℓ^∞ error (variance)
10^4	$0.1015\ (5, 10.10^{-4})$
10^5	$0.0741\ (9, 51.10^{-5})$
10^6	$0.0622\ (9, 38.10^{-6})$

10^5, which corresponds to a signal observed for one hundredth second. Current estimators used in nuclear spectrometry for similar data requires much longer measurements (up to 10 seconds). The reason is that these estimators do not consider observations where pile-up is suspected to occur, thus throwing away a large part of the available information. Moreover, an estimation of the risk $\mathbb{E}_\theta\left[\|\theta - \hat{\theta}_n\|_\infty\right]$ is provided in Table 1 for the shot-noise configuration described at the beginning of this section and for three different sample sizes.

4. Error bounds for the empirical characteristic function and its derivatives

To derive Theorem 2.1, since our constructed estimator involves the empirical characteristic function and its derivative, we rely on deviation bounds for

$$\mathbb{E}_\theta\left[\sup_{u\in[-h^{-1},h^{-1}]}\left|\hat{\varphi}_n^{(k)}(u) - \varphi^{(k)}(u)\right|\right] \quad , \quad k = 0, 1 \quad , \quad h > 0\,, \tag{19}$$

which are uniform over $\theta \in \Theta(K, L, s, m)$, where the smoothness class $\Theta(K, L, s, m)$ is defined by (15). Here, $\varphi^{(k)}$ and $\hat{\varphi}_n^{(k)}$ respectively denote the k-th derivative of the characteristic function and its empirical counterpart associated to the sample $X_1, \ldots, X_n$. These bounds are of independent interest and therefore are stated in this separate section. Upper bounds of the empirical characteristic function deviations have been derived in the case of i.i.d. samples: [8, Theorem 2.2] provides upper bounds of (19) for i.i.d. infinitely divisible random variables, based on general deviation bounds for the empirical process of i.i.d. samples found in [22]. Here we are concerned with a dependent sample $X_1, \cdots, X_n$ and we rely instead on [7]. We obtain upper bounds with the same rate of convergence as in the i.i.d. case but depending on the β-mixing coefficients, see Theorem 4.1. An additional difficulty in the non-parametric setting that we consider is to derive upper bounds that are uniform over smoothness classes for the density θ, and thus to carefully examine how the β coefficients depend on θ, see Theorem 4.2.

Let us first recall the definition of β-mixing coefficient (also called absolutely regular or completely regular coefficient) as introduced by Volkonskii and

Rozanov [23]. For $\mathcal{A}, \mathcal{B}$ two σ-algebras of Ω, the coefficient $\beta(\mathcal{A}, \mathcal{B})$ is defined by

$$\beta(\mathcal{A}, \mathcal{B}) \triangleq \frac{1}{2} \sup \sum_{(i,j) \in I \times J} |\mathbb{P}(A_i \cap B_j) - \mathbb{P}(A_i)\mathbb{P}(B_j)|$$

the supremum being taken over all finite partitions $(A_i)_{i \in I}$ and $(B_j)_{j \in J}$ of Ω respectively included in $\mathcal{A}$ and $\mathcal{B}$. When dealing with a stochastic process $(X_t)_{t \geq 0}$, the β-mixing coefficient is defined for every positive s by:

$$\beta(s) \triangleq \sup_{t \geq 0} \beta\left(\sigma\left(X_u, u \leq t\right), \left(\sigma\left(X_{s+u}, u \geq t\right)\right)\right.$$

The process $(X_t)_{t \geq 0}$ is said to be β-mixing if $\lim_{t \to \infty} \beta(t) = 0$ and exponentially β-mixing if there exists a strictly positive number a such that $\beta(t) = O(e^{-at})$ as $t \to \infty$.

We first state a result essentially following from [7] which specify how the β coefficients allows us to derive bounds on the estimation of the characteristic function and its derivatives.

Theorem 4.1. *Let k be a non-negative integer and $X_1, \cdots, X_n$ a sample of a stationary β-mixing process. Suppose that there exists $C \geq 1$ and $\rho \in (0,1)$ such that $\beta_n \leq C\rho^n$ for all $n \geq 1$. Let $r > 1$ and suppose that $\mathbb{E}\left[|X_1|^{2(k+1)r}\right] < \infty$.*

Then there exists a constant A only depending on C, ρ and r such that for all $h > 0$ and $n \geq 1$, we have

$$\mathbb{E}\left[\sup_{u \in [-h^{-1}, h^{-1}]} \left|\hat{\varphi}_n^{(k)}(u) - \varphi^{(k)}(u)\right|\right] \leq A \frac{\max\left(\mathbb{E}\left[|X_1|^{2kr}\right]^{1/2r}, \mathbb{E}\left[|X_1|^{2(k+1)r}\right]^{1/2r}\right)\left(1 + \sqrt{\log(1+h^{-1})}\right)}{n^{1/2}} . \quad (20)$$

Proof. The proof is deferred to Section 5.3 □

It turns out that the stationary exponential shot-noise process $\mathbf{X}$ defined by (1) is exponentially β-mixing if the first absolute moment of the marks is finite, see [15, Theorem 4.3] for a slightly more general condition. However, in order to obtain a uniform bound of the risk of our estimator $\hat{\theta}_n$ over a smoothness class, a more precise result is needed. In the sequel, we add a superscript θ to the β-mixing sequence to make explicit the dependence with respect to the mark's density θ. The following theorem provides a geometric bound for the β-mixing coefficients of the shot noise which is uniform over the class $\Theta(K, L, s, m)$.

Theorem 4.2. *Let $X_1, \cdots, X_n$ be a sample of the stationary shot-noise process given by (1) satisfying (SN-1)–(SN-2). Let $K, L, m > 0$ and $s > 1/2$. Then there exist two constants $C > 0$ and $\rho \in (0,1)$ only depending on λ, α, K, L, s and m such that, for all $n \geq 1$,*

$$\sup_{\theta \in \Theta(K,L,s,m)} \beta_n^\theta \leq C\rho^n < \infty . \quad (21)$$

Proof. See Section 5.2. □

As a corollary of Theorems 4.1 and 4.2, we obtain error bounds for the empirical characteristic function when dealing with observations $X_1, \cdots, X_n$ of the stationary shot-noise process given by (1).

Corollary 4.1. *Let $X_1, \cdots, X_n$ be a sample of the stationary shot-noise process given by (1) satisfying (SN-1)–(SN-2). Let $K, L, m > 0$ and $s > 1/2$ and let k be an integer such that $0 \leq k < 1 + n/2$. Then there exists a constant B only depending on k, λ, α, K, L, s and m such that for all and $n \geq 1$,*

$$\sup_{\theta \in \Theta(K,L,s,m)} \mathbb{E}_\theta \left[\sup_{u \in [-h^{-1}, h^{-1}]} \left| \hat{\varphi}_n^{(k)}(u) - \varphi^{(k)}(u) \right| \right] \leq B \, \frac{1 + \sqrt{\log(1 + h^{-1})}}{n^{1/2}} \, .$$

This result can be compared to [8, Theorem 2.2]. Note however that although our sample has infinitely divisible marginal distributions, it is not independent and the Lévy measure is not integrable.

5. Proofs

5.1. Preliminary results on the exponential shot noise

We establish some geometric ergodicity results on the exponential shot noise that will be needed in other proofs.

Definition 5.1 (Geometric drift condition)**.** *A Markov Kernel P satisfies a geometric drift condition (called $D(V, \mu, b)$) if there exists a measurable function $V : \mathbb{R} \to [1, \infty[$ and constants $(\mu, b) \in (0, 1) \times \mathbb{R}_+$ such that*

$$PV \leq \mu V + b \, .$$

Definition 5.2 (Doeblin set)**.** *A set C is called a (m, ϵ)-Doeblin set if there exists a positive integer m, a positive number ϵ and a probability measure ν on $\mathbb{R}$ such that, for any x in C and A in $\mathcal{B}(\mathbb{R})$*

$$P^m(x, A) \geq \epsilon \nu(A) \, .$$

The following proposition is borrowed from [21] and relates explicitly the geometrical drift condition to the convergence in V-norm (denoted by $|\cdot|_V$) to the stationary distribution.

Proposition 5.1. *Let P be a Markov kernel satisfying the drift condition $D(V, \mu, b)$. Assume moreover that for some $d > 2b(1 - \mu) - 1$, $m \in \mathbb{N} - \{0\}$ and $\epsilon \in (0, 1)$, the level set $\{V \leq d\}$ is an (m, ϵ)-Doeblin set. Then P admits a unique invariant measure π and P is V-geometrically ergodic, that is, for any $0 < u < \epsilon / \left(b_m + \mu^m d - 1 + \epsilon\right)_+ \vee 1$, $n \in \mathbb{N}$ and $x \in \mathbb{R}$,*

$$\|P^n(x, \cdot) - \pi\|_V \leq c(u)[\pi(V) + V(x)]\rho^{\lfloor n/m \rfloor}(u) \tag{22}$$

where

- $b_m = \frac{b}{\min V}\frac{1-\mu^m}{1-\mu}$
- $c(u) = u^{-1}(1-u) + \mu^m + b_m$
- $\rho(u) = (1 - \epsilon + u(b_m + \mu^m d + \epsilon - 1)) \vee \left(1 - u\frac{(1+d)(1-\mu^m-2b_m}{2(1-u)+u(1+d)}\right)$
- $\pi(V) = \mathbb{P}(Z \in V)$ *where* Z *is a random variable with distribution* π

In order to apply such a result to the sample $(X_1, \cdots, X_n)$ of the exponential shot-noise defined by (1), observe that it is the sample of a Markov chain which satisfies the autoregression equation $X_{i+1} = \mathrm{e}^{-\alpha}X_i + W_{i+1}$, where the sequence of innovations $(W_i)_{i\in\mathbb{Z}}$ is made up of i.i.d. random variables distributed as

$$W_0 \stackrel{\Delta}{=} \sum_{k=1}^{N_\lambda([0,1])} Y_k \mathrm{e}^{-\alpha U_k} \ . \tag{23}$$

where

- $N_\lambda([0,1])$ is a Poisson r.v. with mean λ,
- $(Y_i)_{i\geq 1}$ are i.i.d. r.v.'s with probability density function θ,
- $(U_i)_{i\geq 1}$ are i.i.d. and uniformly distributed on $[0,1]$,
- all these variables are independent.

In the following, we denote by Q_θ the Markov kernel associated to the Markov chain $(X_i)_{i\geq 0}$ under (SN-1)–(SN-2).

Proposition 5.2 (Uniform Geometric drift condition)**.** *Let* $K, L, m > 0$ *and* $s > 1/2$ *and let* $\theta \in \Theta(K, L, s, m)$*. Then the Markov kernel* Q_θ *satisfies the drift condition* $D(V, \mu, b)$*, where*

$$V : x \to 1 + |x| \quad , \quad \mu = \mathrm{e}^{-\alpha} \quad , \quad b = 1 + \lambda K^{1/(4+m)} - \mathrm{e}^{-\alpha} \ . \tag{24}$$

Proof. We have for all $\theta \in \Theta(K, L, s, m)$ and $x \in \mathbb{R}$,

$$\begin{aligned} Q_\theta V(x) &= \mathbb{E}_\theta\left[1 + \left|\mathrm{e}^{-\alpha}x + W_0\right|\right] \\ &\leq \mathrm{e}^{-\alpha}V(x) + 1 - \mathrm{e}^{-\alpha} + \lambda K^{1/(4+m)} = \mu V(x) + b \ . \end{aligned}$$

□

Remark 5.1. A similar result holds for the functions $V_i : x \to 1 + |x|^i$ where $i \in \{1, \cdots, \lfloor 4 + m \rfloor\}$.

Proposition 5.3 (Doeblin set)**.** *Let* $l > 1$*,* $K, L, m > 0$ *and* $s > 1/2$ *and define* V *as in (24). There exists* $\epsilon > 0$ *only depending on* l*,* α*,* λ*,* $K, L, m > 0$ *and* s *such that, for all* $\theta \in \Theta(K, L, s, m)$*, the Markov kernel* Q_θ *admits* $\{V \leq l\}$ *as an* $(1, \epsilon)$*-Doeblin set.*

Proof. Let $\theta \in \Theta(K, L, s, m)$. Denote by $\check{\theta}$ the density of random variable $Y_1\mathrm{e}^{-\alpha U_1}$ with U_1 and Y_1 two independent random variables respectively distributed uniformly on $[0,1]$ and with density θ. It is easy to show that, for all $v \in \mathbb{R}$,

$$\check{\theta}(v) = \frac{1}{\alpha v}\int_v^{v\mathrm{e}^\alpha} \theta(y)\mathrm{d}y \ . \tag{25}$$

The distribution of W_0 is thus given by the infinite mixture

$$\mathrm{e}^{-\lambda}\delta_0(\mathrm{d}\xi) + \sum_{k=1}^{\infty} \frac{\lambda^k}{k!}\mathrm{e}^{-\lambda}\check{\theta}^{*k}(\xi)\mathrm{d}\xi \triangleq \mathrm{e}^{-\lambda}\delta_0(\mathrm{d}\xi) + (1-\mathrm{e}^{-\lambda})\tilde{f}_\theta(\xi)\mathrm{d}\xi \;, \tag{26}$$

where δ_0 is the Dirac point mass at 0 and $\check{\theta}^{*k}$ denote the k-th self-convolution of $\check{\theta}$. It follows that, for all Borel set A,

$$Q_\theta(x,A) = \mathrm{e}^{-\lambda}\mathbf{1}_A(\mathrm{e}^{-\alpha}x) + \int_A (1-\mathrm{e}^{-\lambda})\tilde{f}_\theta(\xi+\mathrm{e}^{-\alpha}x)\,\mathrm{d}\xi \;.$$

In order to show that $\{V \le l\}$ is a $(\epsilon,1)$-Doeblin-set for the kernels Q_θ, it is sufficient to exhibit a probability measure ν such that, for all $|x| \le l-1$ and all Borel set A

$$\int_A \tilde{f}_\theta(\xi+\mathrm{e}^{-\alpha}x)\mathrm{d}\xi \ge (1-\mathrm{e}^{-\lambda})^{-1}\epsilon\,\nu(A) \;.$$

Hence if for each $\theta \in \Theta(K,L,s,m)$ we find $c(\theta) < d(\theta)$ such that

$$\epsilon' = \inf_{\theta\in\Theta(K,L,s,m)} \inf_{c(\theta)\le\xi\le d(\theta)} \inf_{|x|\le l-1} [d(\theta)-c(\theta)]\,\tilde{f}_\theta(\xi+\mathrm{e}^{-\alpha}x) > 0 \;,$$

the result follows by taking ν with density $[d(\theta)-c(\theta)]^{-1}\mathbf{1}_{[c(\theta),d(\theta)]}$ and $\epsilon = (1-\mathrm{e}^{-\lambda})\epsilon'$. By definition of $\tilde{f}_\theta$ above, it is now sufficient to show that there exist $c(\theta) < d(\theta)$ and $k(\theta) \ge 1$ such that

$$\inf_{\theta\in\Theta(K,L,s,m)} \inf_{c(\theta)\le\xi\le d(\theta)} \inf_{|x|\le l-1} [d(\theta)-c(\theta)]\,\check{\theta}^{*k(\theta)}(\xi+\mathrm{e}^{-\alpha}x) > 0 \;.$$

Observe that for $c \le \xi \le d$ and $|x| \le l-1$ we have $\xi+\mathrm{e}^{-\alpha}x \in [c-\mathrm{e}^{-\alpha}(l-1), d+\mathrm{e}^{-\alpha}(l-1)]$. So for any interval $[c',d']$ of length $d'-c' > 2\mathrm{e}^{-\alpha}(l-1)$, we may set $c = c'+\mathrm{e}^{-\alpha}(l-1) < d = d'-\mathrm{e}^{-\alpha}(l-1)$ so that $c \le \xi \le d$ and $|x| \le l-1$ imply $\xi+\mathrm{e}^{-\alpha}x \in [c',d']$. Hence the proof boils down to showing that for each $\theta \in \Theta(K,L,s,m)$, there exist $c'(\theta) < d'(\theta)$ with $d'(\theta)-c'(\theta) > 2\mathrm{e}^{-\alpha}(l-1)$ and $k(\theta) \ge 1$ such that

$$\inf_{\theta\in\Theta(K,L,s,m)} \inf_{c'(\theta)\le\xi\le d'(\theta)} [d'(\theta)-c'(\theta)-2\mathrm{e}^{-\alpha}(l-1)]\,\check{\theta}^{*k(\theta)}(\xi) > 0 \;. \tag{27}$$

By Lemma A.3, there exists $a > 0$, $\Delta > 0$, $\delta > 1$ and $\epsilon_0 > 0$ and such that ϵ_0 and $\Delta = b-a$ only depend on m, K, L and s (although a may depend on θ), and

$$\inf_{a<x<(a+\Delta)/\delta} \check{\theta}(x) > \epsilon_0 \;.$$

Finally, Lemma A.4 and the previous bound yield (27), which concludes the proof. □

5.2. Proof of Theorem 4.2

As explained in Section 5.1, $(X_i)_{i\geq 0}$ is a stationary V-geometrically ergodic Markov chain with Markov kernel denoted by Q_θ. By [6], the β-coefficient of the stationary Markov chain $(X_i)_{i\geq 0}$ can be expressed for all $n \geq 1$ and $\theta \in \Theta(K, L, s, m)$ as

$$\beta_n^\theta = \int_{\mathbb{R}} \|Q_\theta^n(x,\cdot) - \pi_\theta\|_{TV}\, \pi_\theta(dx)\ ,$$

where π_θ is the invariant marginal distribution and $|\cdot|_{TV}$ denotes the total variation norm, i.e. the V-norm with $V = 1$. Combining Propositions 5.2, 5.3 and 5.1, we can find constants $C > 0$ and $\rho \in (0,1)$ only depending on λ, α, K, L, s and m such that

$$\|Q_\theta^n(x,\cdot) - \pi_\theta\|_{TV} \leq \|Q_\theta^n(x,\cdot) - \pi_\theta\|_V \leq C\,(2 + \mathbb{E}_\theta\,[|X_1|] + |x|)\,\rho^n\ ,$$

where $V(x) = 1 + |x|$. The last two displays yield

$$\begin{aligned}\beta_n^\theta &\leq 2C\ (1 + \mathbb{E}_\theta\,[|X_1|])\rho^n \\ &\leq 2C\ (1 + \lambda K^{1/(4+m)})\rho^n\ ,\end{aligned} \tag{28}$$

which concludes the proof.

5.3. Proof of Theorem 4.1

In [7], the authors establish a Donsker invariance principle for the process $\{Z_n(f), f \in \mathcal{F}\}$ where $Z_n \stackrel{\Delta}{=} n^{-1/2}\sum_{i=1}^n (\delta_{X_i} - P)$ is the normalized centered empirical process associated to a stationary sequence of β-mixing random variables $(X_1, \cdots, X_n)$ with marginal distribution P and $\mathcal{F}$ is a class of functions satisfying an entropy condition. To be more precise, suppose that the sequence $(X_i)_{i\geq 1}$ is β-mixing with $\sum_{n\in\mathbb{N}} \beta_n < \infty$. The mixing rate function β is defined by $\beta(t) = \beta_{\lfloor t \rfloor}$ if $t \geq 1$ and $\beta(t) = 0$ otherwise while its càdlàg inverse β^{-1} is defined by:

$$\beta^{-1}(u) \stackrel{\Delta}{=} \inf_{t\geq 0}\{\beta(t) \leq u\}$$

Further, for any complex-valued function f, denote by Q_f the quantile function of the r.v. $|f(X_0)|$ and introduce the norm:

$$\|f\|_{2,\beta} \stackrel{\Delta}{=} \left(\int_0^1 \beta^{-1}(u) Q_f(u)^2 \mathrm{d}u\right)^{1/2}\ .$$

The space $\mathcal{L}_{2,\beta}$ is defined as the class of functions f such that $\|f\|_{2,\beta} < \infty$. In [7], the authors proved that $(\mathcal{L}_{2,\beta}, \|\cdot\|_{2,\beta})$ is a normed subspace of $\mathcal{L}_2$. A useful and trivial result from the definition of the norm $\mathcal{L}_{2,\beta}(P)$ provides the following relation:

$$|f| \leq |g| \Rightarrow \|f\|_{2,\beta} \leq \|g\|_{2,\beta}\ .$$

For any real $r > 1$, another useful (less trivial) result in [7] states that under the condition

$$\sum_{n\geq 0} \beta_n \, n^{r/(r-1)} < \infty,$$

we have $\mathcal{L}_{2r} \subset \mathcal{L}_{2,\beta}$ with the additional inequality

$$\|f\|_{2,\beta} \leq \|f\|_{2r} \sqrt{1 + r \sum_{n\geq 0} \beta_n \, n^{r/(r-1)}}\,, \tag{29}$$

where $\|f\|_{2r} = \mathbb{E}\left[|f(X_0)|^{2r}\right]^{1/2r}$ denote the usual L^{2r}-norm.

Now, we can state a result directly adapted from [7, Theorem 3] that will serve our goal to prove Theorem 2.1. For the sake of self-consistency, we recall that, given a metric space of real-valued funtions $(E, \|\cdot\|)$ and two functions l, u, the bracket $[l, u]$ represents the set of all functions f such that $l \leq f \leq u$. For any positive ϵ, $[l, u]$ is called an ϵ-bracket if $\|l - u\| < \epsilon$.

Theorem 5.1. *Suppose that the sequence $(X_i)_{i\geq 1}$ is exponentially β-mixing and that there exists $C \geq 1$ and $\rho \in (0, 1)$ such that $\beta_n \leq C\rho^n$ for all $n \geq 1$. Let $\sigma > 0$ and let $\mathcal{F} \subset \mathcal{L}_{2,\beta}$ be a class of functions such that for every f in $\mathcal{F}$, $\|f\|_{2,\beta} \leq \sigma$. Define*

$$\phi(\sigma) = \int_0^\sigma \sqrt{1 + \log\left(N_{[\,]}\left(u, \mathcal{F}, \|\cdot\|_{2,\beta}\right)\right)}\, \mathrm{d}u\,,$$

where $N_{[\,]}(u, \mathcal{F}, \|\cdot\|_{2,\beta})$ denotes the bracketing number, that is, the minimal number of u-brackets with respect to the norm $\|\cdot\|_{2,\beta}$ that has to be used for covering $\mathcal{F}$. Suppose that the two following assumptions hold.

(DMR1) $\mathcal{F}$ has an envelope function F such that $\|F\|_{2r} < \infty$ for some $r > 1$.
(DMR2) $\phi(1) < \infty$.

Then there exist a constant $A > 0$ only depending on C and ρ such that, for all integer n, we have

$$\mathbb{E}\left[\sup_{f\in\mathcal{F}} |Z_n(f)|\right] \leq A\phi(\sigma)\left(1 + \frac{\|F\|_{2r}}{\sigma\sqrt{1 - r^{-1}}}\right)\,. \tag{30}$$

Having this result at hand, we now remark that (19), for a fixed integer k, can be rewritten as

$$\mathbb{E}\left[\sup_{u\in[-h^{-1},h^{-1}]} \left|\hat{\varphi}_n^{(k)}(u) - \varphi^{(k)}(u)\right|\right] = n^{-1/2}\mathbb{E}\left[\sup_{f\in\mathcal{F}_h^k} |Z_n(f)|\right]\,, \tag{31}$$

where

$$\mathcal{F}_h^k \triangleq \{f_u : x \to (ix)^k e^{iux}, u \in \left[-h^{-1}, h^{-1}\right]\}. \tag{32}$$

The proof of Theorem 4.1 based on an application of the previous theorem is as follows.

Proof. We apply Theorem 5.1 for a fixed integer k, $\mathcal{F} = \mathcal{F}_h^k$, $F = F_k$ and $r = (4 + m)/4$ where $F_k : x \to |x|^k$.

Assumption (DMR1): Let k be a fixed integer. On the one hand, the function F_k is an envelope function of the class $\mathcal{F}_h^k$ and on the other hand, for any real $r > 1$, from (29), we have

$$\|F_k\|_{2,\beta} \leq \mathbb{E}\left[|X_1|^{2kr}\right]^{1/2r} \sqrt{1 + r \sum_{n\geq 0} \beta_n \, n^{r/(r-1)}} \triangleq \sigma_{r,k} < \infty \,. \tag{33}$$

Assumption (DMR2): For k a fixed integer, the class $\mathcal{F}_h^k$ is Lipschitz in the index parameter: indeed, we have for every s, t in $\left[-h^{-1}, h^{-1}\right]$ and every real x

$$\left|(\mathrm{i}x)^k \mathrm{e}^{\mathrm{i}sx} - (\mathrm{i}x)^k \mathrm{e}^{\mathrm{i}tx}\right| \leq |s-t| \, |x|^{k+1} \tag{34}$$

A direct application of [22, Theorem 2.7.11] for the classes $\mathcal{F}_h^k$ gives for any $\epsilon > 0$:

$$\begin{aligned} N_{[\,]}\left(2\epsilon ||F_{k+1}||_{2,\beta}, \mathcal{F}_h^k, ||\cdot||_{2,\beta}\right) &\leq N\left(\epsilon, \left[-h^{-1}, h^{-1}\right], |\cdot|\right) \\ &\leq 1 + \frac{2h^{-1}}{\epsilon} \end{aligned} \tag{35}$$

where N and $N_{[\,]}$ are respectively called the covering numbers and bracketing number (these numbers respectively represent the minimum number of balls and brackets of a given size necessary to cover a space with respect to a given norm). From (35), it follows that for any $\sigma > 0$, we have

$$\begin{aligned} \phi(\sigma) &= \int_0^\sigma \sqrt{1 + \log\left(N_{[\,]}\left(u, \mathcal{F}_h^k, \|\cdot\|_{2,\beta}\right)\right)} \mathrm{d}u \\ &\leq \int_0^\sigma \sqrt{1 + \log\left(1 + \frac{4||F_{k+1}||_{2,\beta} h^{-1}}{u}\right)} \mathrm{d}u \qquad (36) \\ &\leq \int_0^\sigma \left(1 + \frac{2||F_{k+1}||_{2,\beta}^{1/2} h^{-1/2}}{u^{1/2}}\right) \mathrm{d}u \\ &= \sigma + 4\sqrt{\sigma} \, \|F_{k+1}\|_{2,\beta}^{1/2} h^{-1/2} < \infty \qquad (37) \end{aligned}$$

because we supposed $F_{k+1} \in \mathcal{L}_{2r}$ and $\beta_n \leq C\rho^n$ which, from (29), implies that $\|F_{k+1}\|_{2,\beta} < \infty$.

Conclusion of the proof The application of Theorem 5.1 gives

$$\mathbb{E}_\theta\left[\sup_{u\in[-h^{-1},h^{-1}]} \left|\hat{\varphi}_n^{(k)}(u) - \varphi^{(k)}(u)\right|\right] \leq \tilde{A} \frac{\phi(\sigma_{r,k})}{n^{1/2}}$$

where $\tilde{A} = A(1 + 1/\sqrt{1 - r^{-1}})$ since, from (33), we have $\|F_k\|_{2r}/\sigma_{r,k} \leq 1$. Set $c_{r,\bar{\beta}} \triangleq \sqrt{1 + r\sum_{n\geq 0} \beta_n \, n^{r/(r-1)}}$. From (36) and (33), we can write

$$\phi(\sigma) \leq \int_0^\sigma \sqrt{1 + \log\left(1 + \frac{4||F_{k+1}||_{2r} c_{r,\bar{\beta}} h^{-1}}{u}\right)} \mathrm{d}u \,,$$

For $\sigma = \sigma_{r,k}$, we get after the change of variable $v = \frac{4||F_{k+1}||_{2,r} c_{r,\bar{\beta}} \sigma_{r,k} h^{-1}}{u}$

$$\phi(\sigma_{r,k}) \leq \max\left(\|F_k\|_{2r}, \|F_{k+1}\|_{2r}\right) c_{r,\bar{\beta}} \left(1 + h^{-1} \int_{h^{-1}}^{\infty} \sqrt{\log(1+v)} \frac{\mathrm{d}v}{v^2}\right) .$$

By Lemma A.6, we get for a universal constant $B > 0$ that

$$\phi(\sigma_{r,k}) \leq B \max\left(\|F_k\|_{2r} \|F_{k+1}\|_{2r}\right) c_{r,\bar{\beta}} \left(1 + \sqrt{\log\left(1 + h^{-1}\right)}\right) .$$

In the particular context of Corollary 4.1, we use the fact that $\|F_k\|_{2r}$ can be bounded by $\max(1, K^{4k/(4+m)})$ and $c_{r,\bar{\beta}}$ by a constant only depending on the parameters K, L, s, m □

5.4. *Proof of Theorem 2.1*

Proof. We denote by θ_n^0 the function defined by:

$$\theta_n^0(x) \stackrel{\Delta}{=} \max\left(0, \frac{1}{2\pi} \int_{-h_n^{-1}}^{h_n^{-1}} \mathrm{e}^{-\mathrm{i}xu} \varphi_{Y_0}(u) \mathrm{d}u\right) . \tag{38}$$

Since $\theta \in \Theta(K, L, s, m)$ and $s > 1/2$, we have that $\mathcal{F}[\theta]$ is integrable and, under $\theta \geq 0$, we have

$$\theta(x) = \max\left(0, \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}xu} \varphi_{Y_0}(u) \mathrm{d}u\right) . \tag{39}$$

We decompose the error in infinite norm as

$$\|\theta - \hat{\theta}_n\|_\infty \leq \|\theta - \theta_n^0\|_\infty + \|\theta_n^0 - \hat{\theta}_n\|_\infty . \tag{40}$$

From (38) and (39), we get

$$\begin{aligned}
\|\theta - \theta_n^0\|_\infty &\leq \frac{1}{\pi} \int_{h_n^{-1}}^{\infty} |\varphi_{Y_0}(u)| \, \mathrm{d}u \\
&= \frac{1}{\pi} \int_{h_n^{-1}}^{\infty} \left|u^{-s} u^s \varphi_{Y_0}(u)\right| \mathrm{d}u \\
&\leq \frac{1}{\pi} \left(\int_{h_n^{-1}}^{\infty} |u|^{-2s} \, \mathrm{d}u\right)^{1/2} \left(\int_{h_n^{-1}}^{\infty} \left|u^s \varphi_{Y_0}(u)\right|^2 \mathrm{d}u\right)^{1/2} \\
&\leq \frac{L}{\pi} \frac{h_n^{s-1/2}}{2s-1} = \frac{L}{(2s-1)\pi} n^{-(2s-1)/(4s+2+4\lambda/\alpha)} .
\end{aligned} \tag{41}$$

where we used the Cauchy-Schwartz inequality and the assumption that $\theta \in \Theta(K, L, s, m)$.

We conclude with a bound of the term involving $\|\theta_n^0 - \hat{\theta}_n\|_\infty$ in (40). To this end, the following inequality will be useful. Using (12) and the mean-value theorem, we have

$$\sup_{u \in \mathbb{R}} \left|\frac{\varphi'_{X_0}(u)}{\varphi_{X_0}(u)}\right| \leq \frac{\lambda}{\alpha} \sup_{u \in \mathbb{R}} \left|\varphi'_{Y_0}(u)\right| \leq \frac{\lambda}{\alpha} \mathbb{E}_\theta\left[|Y_0|\right] \leq \frac{\lambda}{\alpha} K^{1/(4+m)} . \tag{42}$$

By (12), we can bound the term $\|\theta_n^0 - \hat{\theta}_n\|_\infty$ by

$$\begin{aligned}
\|\theta_n^0 - \hat{\theta}_n\|_\infty &= \frac{\alpha}{\lambda\pi}\left|\int_{-h_n^{-1}}^{h_n^{-1}} \frac{\varphi'_{X_0}(u)}{\varphi_{X_0}(u)} - \frac{\hat{\varphi}'_n(u)}{\hat{\varphi}_n(u)}\mathbb{1}_{|\hat{\varphi}_n(u)|\geq\kappa_n}\mathrm{d}u\right| \\
&\leq \frac{\alpha}{\lambda\pi}\int_{-h_n^{-1}}^{h_n^{-1}}\left|\frac{\varphi'_{X_0}(u)}{\varphi_{X_0}(u)} - \frac{\hat{\varphi}'_n(u)}{\hat{\varphi}_n(u)}\mathbb{1}_{|\hat{\varphi}_n(u)|\geq\kappa_n}\right|\mathrm{d}u \\
&\leq \frac{2\alpha h_n^{-1}}{\lambda\pi}\sup_{|u|\leq h_n^{-1}}\left|\frac{\varphi'_{X_0}(u)}{\varphi_{X_0}(u)} - \frac{\hat{\varphi}'_n(u)}{\hat{\varphi}_n(u)}\mathbb{1}_{|\hat{\varphi}_n(u)|\geq\kappa_n}\right| \\
&\leq \frac{2\alpha h_n^{-1}}{\lambda\pi}\left(\sup_{|u|\leq h_n^{-1}}\left|\frac{\varphi'_{X_0}(u)}{\varphi_{X_0}(u)} - \frac{\hat{\varphi}'_n(u)}{\hat{\varphi}_n(u)}\right|\mathbb{1}_{|\hat{\varphi}_n(u)|>\kappa_n}\right. \\
&\quad \left. + \sup_{|u|\leq h_n^{-1}}\left|\frac{\varphi'_{X_0}(u)}{\varphi_{X_0}(u)}\right|\mathbb{1}_{|\hat{\varphi}_n(u)|\leq\kappa_n}\right) \\
&\triangleq A_{n,1} + A_{n,2}\,.
\end{aligned}$$

Writing $\frac{\varphi'_{X_0}}{\varphi_{X_0}} - \frac{\hat{\varphi}'_n}{\hat{\varphi}_n}$ as $\left(\frac{\varphi'_{X_0}}{\varphi_{X_0}} - \frac{\varphi'_{X_0}}{\hat{\varphi}_n}\right) + \left(\frac{\varphi'_{X_0}}{\hat{\varphi}_n} - \frac{\hat{\varphi}'_n}{\hat{\varphi}_n}\right)$, the term $A_{n,1}$ can be bounded as follows.

$$\begin{aligned}
&\sup_{|u|\leq h_n^{-1}}\left|\frac{\varphi'_{X_0}(u)}{\varphi_{X_0}(u)} - \frac{\hat{\varphi}'_n(u)}{\hat{\varphi}_n(u)}\right|1_{|\hat{\varphi}_n(u)|>\kappa_n} \\
&\leq \kappa_n^{-1}\sup_{|u|\leq h_n^{-1}}|\Psi(u)|\,|\hat{\varphi}_n(u) - \varphi_{X_0}(u)| + \kappa_n^{-1}\sup_{|u|\leq h_n^{-1}}\left|\hat{\varphi}'_n(u) - \varphi'_{X_0}(u)\right|
\end{aligned}$$

Thus, using (42), we get

$$\begin{aligned}
\mathbb{E}_\theta\left[A_{n,1}\right] \leq\ & \frac{2h_n^{-1}\kappa_n^{-1}K^{1/(4+m)}\mathbb{E}_\theta\left[\sup_{|u|\leq h_n^{-1}}\left|\hat{\varphi}'_n(u) - \varphi'_{X_0}(u)\right|\right]}{\pi} \\
&+ \frac{2\alpha h_n^{-1}\kappa_n^{-1}\mathbb{E}_\theta\left[\sup_{|u|\leq h_n^{-1}}|\hat{\varphi}_n(u) - \varphi_{X_0}(u)|\right]}{\pi\lambda}\,.
\end{aligned}$$

The two terms on the right hand side can be bounded using Corollary 4.1 with $r = (4+m)/4$. It gives

$$\mathbb{E}_\theta\left[\sup_{|u|\leq h_n^{-1}}\left|\hat{\varphi}'_n(u) - \varphi'_{X_0}(u)\right|\right] \leq \frac{B\ K^{4/(4+m)}\left(1 + \sqrt{\log(1+h_n^{-1})}\right)}{n^{1/2}}$$

and

$$\mathbb{E}_\theta\left[\sup_{|u|\leq h_n^{-1}}|\hat{\varphi}_n(u) - \varphi_{X_0}(u)|\right] \leq \frac{B\left(1 + \sqrt{\log(1+h_n^{-1})}\right)}{n^{1/2}}\,.$$

In the following, for two positive quantities P and Q, possibly depending on θ and n we use the notation

$$P \lesssim Q \iff \text{for all } n \geq 3, \sup_{\theta \in \Theta(K,L,s,m)} \frac{P}{Q} < \infty \,. \tag{43}$$

(P is less than Q up to a multiplicative constant uniform over $\theta \in \Theta(K, L, s, m)$). We thus have that

$$\begin{aligned}
\mathbb{E}_\theta \left[A_{n,1} \right] &\lesssim \frac{1 + \sqrt{\log(1 + h_n^{-1})}}{\kappa_n \, n^{1/2} h_n} \\
&\lesssim \frac{1 + \sqrt{\log(1 + h_n^{-1})}}{n^{1/2} h_n^{\lambda/\alpha + 1}} \\
&\lesssim n^{-(2s-1)/(4s+2+4\lambda/\alpha)} \log(n)^{1/2} \,,
\end{aligned} \tag{44}$$

where we used the fact that $\kappa_n^{-1} \leq 2C^{-1} h_n^{\lambda/\alpha}$ for any integer n.
We now bound $A_{n,2}$. From (42), remark that

$$\begin{aligned}
\mathbb{E}_\theta \left[A_{n,2} \right] &\leq \frac{2}{\pi h_n} K^{1/(4+m)} \mathbb{P}_\theta \left(\exists u \in \left[-h_n^{-1}, h_n^{-1} \right], |\hat{\varphi}_n(u)| \leq \kappa_n \right) \\
&\leq \frac{2}{\pi h_n} K^{1/(4+m)} \mathbb{P}_\theta \left(\inf_{|u| \leq h_n^{-1}} |\hat{\varphi}_n(u)| \leq \kappa_n \right) \,,
\end{aligned} \tag{45}$$

From Lemma (A.5), we have

$$\begin{aligned}
\inf_{|u| \leq h_n^{-1}} |\hat{\varphi}_n(u)| &\geq \inf_{|u| \leq h_n^{-1}} |\varphi_{X_0}(u)| - \sup_{|u| \leq h_n^{-1}} |\hat{\varphi}_n(u) - \varphi_{X_0}(u)| \\
&\geq C_{K,L,m,\lambda/\alpha} \left(1 + h_n^{-1} \right)^{-\lambda/\alpha} - \sup_{|u| \leq h_n^{-1}} |\hat{\varphi}_n(u) - \varphi_{X_0}(u)| \,.
\end{aligned}$$

It follows that

$$\begin{aligned}
&\mathbb{P}_\theta \left(\inf_{|u| \leq h_n^{-1}} |\hat{\varphi}_n(u)| \leq \kappa_n \right) \\
&\leq \mathbb{P}_\theta \left(\sup_{|u| \leq h_n^{-1}} |\hat{\varphi}_n(u) - \varphi_{X_0}(u)| \geq \left(C_{K,L,m,\lambda/\alpha} - C \right) \left(1 + h_n^{-1} \right)^{-\lambda/\alpha} \right) \,.
\end{aligned}$$

Since $0 < C < C_{K,L,m,\lambda/\alpha}$, applying Corollary 4.1 combined to the Markov's inequality, and using (45), we get

$$\begin{aligned}
\mathbb{E}_\theta \left[A_{n,2} \right] &\lesssim \frac{1 + \sqrt{\log\left(1 + h_n^{-1}\right)}}{n^{1/2} \kappa_n h_n} \\
&\lesssim n^{-(2s-1)/(4s+2+4\lambda/\alpha)} \log(n)^{1/2} \,,
\end{aligned} \tag{46}$$

Equations (41), (44) and (46) imply (16) and the proof is concluded. □

Appendix A: Useful lemmas

The following classical embedding will be useful.

Lemma A.1 (Sobolev embedding)**.** *Let $K, L, m > 0$ and $s > 1/2$. Let $\theta \in \Theta(K, L, s, m)$ defined in (15). Then, for any $\gamma \in (0, (s-1/2) \wedge 1)$, there is a constant $C > 0$ depending on L, s and γ such that, for every real numbers x, y,*

$$|\theta(x) - \theta(y)| \leq C\,|x-y|^{\gamma}\ , \tag{47}$$

where

$$C = \frac{3}{2\pi} L\ \left(\int_{\mathbb{R}} \frac{|\xi|^{2\gamma}}{(1+|\xi|^2)^s} \mathrm{d}\xi\right)^{1/2}\ . \tag{48}$$

The following result is used in the proof of Proposition 5.3.

Lemma A.2. *Let $K, L, m > 0$ and $s > 1/2$. Let $\gamma \in (0, (s-1/2) \wedge 1)$ and $\theta \in \Theta(K, L, s, m)$. Then, there exists $0 < a \leq T_K$ such that*

$$\inf_{a \leq x \leq a+\Delta} \theta(x) \geq \frac{1}{16}(2K)^{-1/(4+m)}\ ,$$

where $T_K = (2K)^{-1/(\gamma(4+m))}$, $\Delta = (2K)^{-1/(\gamma(4+m))}(16C)^{-1/\gamma}$ with C defined by (48).

Proof. We first show that, for every $T > 0$, we have

$$\sup_{|x| \leq T} \theta(x) \geq (2T)^{-1}\left(1 - T^{-(4+m)}K\right)\ ,$$

Denote by Y a random variable with p.d.f θ belonging to the class $\Theta(K, L, s, m)$. On the one side, we have

$$\mathbb{P}\left(|Y| \leq T\right) \leq 2T \sup_{|x| \leq T} \theta(x)$$

and on the other side

$$\mathbb{P}\left(|Y| \leq T\right) = 1 - \mathbb{P}\left(|Y| > T\right) \geq 1 - \mathbb{E}\left[|Y|^{4+m}\right] T^{-(4+m)} \geq \left(1 - T^{-(4+m)}K\right)\ ,$$

where the first inequality is obtained via an application of the Markov inequality. Setting $T_K = (2K)^{1/(4+m)}$, we thus have

$$\sup_{|x| \leq T_K} \theta(x) \geq (4T_K)^{-1}\ .$$

Moreover, since θ is continuous, we can without loss of generality suppose that there exists a positive number a in the interval $(0, T_K]$ such that

$$\theta(a) \geq (8T_K)^{-1}\ .$$

From Lemma A.1, there exists a positive number $\Delta = (16T_K C)^{-1/\gamma}$, independent of the choice of θ such that

$$\inf_{x \in [a, a+\Delta]} \theta(x) \geq (16T_K)^{-1}.$$

□

Lemma A.3. *Let $K, L, m, \alpha > 0$ and $s > 1/2$. Let $\gamma \in (0, (s-1/2) \wedge 1)$ and $\theta \in \Theta(K, L, s, m)$. Define $T_K = (2K)^{-1/(4+m)}$, $\Delta = (2K)^{-1/(\gamma(4+m))}(16C)^{-1/\gamma}$*

with C defined by (48) and let δ be a positive number satisfying

$$1 < \delta < \min\left(\mathrm{e}^{\alpha}, \frac{T_K + \Delta}{T_K}\right) .$$

For any strictly positive v, define the function $\check{\theta}$ by $\check{\theta}(v) = \frac{1}{\alpha v}\int_v^{v\mathrm{e}^{\alpha}} \theta(x)\mathrm{d}x$. Then, there exists $0 < a \leq T_K$ such that

$$\inf_{a \leq v \leq (a+\Delta)/\delta} \check{\theta}(v) \geq \frac{(2K)^{-1/(4+m)}(\delta - 1)}{16\alpha} .$$

Proof. From Lemma A.2, we have

$$\inf_{a \leq x \leq a+\Delta} \theta(x) \geq \frac{1}{16}(2K)^{-1/(4+m)} ,$$

for some $a \in (0, T_K]$. Let $\delta \in (1, \mathrm{e}^{\alpha} \wedge \frac{T_K+\Delta}{T_K})$. Since $(a + \Delta)/a$ is a decreasing function in a for a fixed Δ and $0 \leq a \leq T_K$, we have that

$$(a + \Delta)/a \geq \frac{T_K + \Delta}{T_K}$$

so that $\delta < (a + \Delta)/a$. For any $v \in [a, (a + \Delta)/\delta]$, we have

$$\begin{aligned}
\check{\theta}(v) = \frac{1}{\alpha v}\int_v^{v\mathrm{e}^{\alpha}} \theta(x)\mathrm{d}x &\geq \frac{1}{\alpha v}\int_v^{v\delta} \theta(x)\mathrm{d}x \\
&\geq \frac{v\delta - v}{\alpha v} \inf_{x \in [v, v\delta]} \theta(x) \\
&\geq \frac{\epsilon_K(\delta - 1)}{\alpha} .
\end{aligned}$$

which concludes the proof. □

The following elementary lemma generalizes the previous result for convolutions of lower bounded functions.

Lemma A.4. *Let θ, $\tilde{\theta}$ two positive functions such that there exist positive numbers a, b, c, d, ϵ and $\tilde{\epsilon}$ satisfying*

$$\theta(x) \geq \epsilon \mathbb{1}_{[a,b]}(x) \quad \text{and} \quad \tilde{\theta}(x) \geq \tilde{\epsilon}\mathbb{1}_{[c,d]}(x)$$

Then, for any δ satisfying $0 < \delta < (b - a) \wedge (d - c)$, we have

$$\left(\theta \star \tilde{\theta}\right)(x) \geq \min(1, \delta)\, \epsilon\tilde{\epsilon}\, \mathbb{1}_{[a+c+\delta, b+d-\delta]}(x) . \tag{49}$$

As a consequence, for any integer n in $\mathbb{N}^*$, we have

$$\theta^{\star n}(x) \geq \left(\min\left(1, \frac{b-a}{2n}\right)\right)^{n-1} \epsilon^n \mathbb{1}_{[na+(b-a)/2, nb-(b-a)/2]}(x) . \tag{50}$$

A lower bound of the decay of the absolute value of the shot-noise characteristic function is given by the following lemma.

Lemma A.5. *Assume that the process* $\boldsymbol{X}$ *is given by* (1) *under* (SN-1)–(SN-2) *with some positive constant* α *and* λ*. Let* K, L, m *and* s *be positive constants. Then for all* $\theta \in \Theta(K, L, s, m)$ *and* $u \in \mathbb{R}$*, we have*

$$|\varphi_{X_0}(u)| \geq C_{K,L,m,\lambda/\alpha} \left(1 + |u|\right)^{-\lambda/\alpha} \tag{51}$$

where $C_{K,L,m,\lambda/\alpha} \stackrel{\Delta}{=} \exp(-\lambda(L + K^{1/(4+m)})/\alpha)$.

Proof. From (9) and (11), we have for all $u \in \mathbb{R}$,

$$\varphi_{X_0}(u) = \exp\left(\frac{\lambda}{\alpha}\int_{\mathbb{R}}\left(\int_0^{ux}\frac{\mathrm{e}^{\mathrm{i}v}-1}{v}\mathrm{d}v\right)\theta(x)\mathrm{d}x\right) .$$

If follows that

$$\begin{aligned}|\varphi_{X_0}(u)| &= \exp\left(\frac{\lambda}{\alpha}\int_{\mathbb{R}}\left(\int_0^{ux}\frac{\cos(v)-1}{v}\mathrm{d}v\right)\theta(x)\mathrm{d}x\right)\\ &= \exp\left(-\frac{\lambda}{\alpha}\int_0^{|u|}\frac{1-\mathrm{Re}\left(\varphi_{Y_0}(z)\right)}{z}\mathrm{d}z\right) .\end{aligned} \tag{52}$$

First, we have for any real z and any function $\theta \in \Theta(K, L, s, m)$,

$$\begin{aligned}|1-\mathrm{Re}\left(\varphi_{Y_0}(z)\right)| &= \left|\int_{\mathbb{R}}\left(1-\mathrm{e}^{\mathrm{i}xz}\right)\theta(x)\mathrm{d}x\right|\\ &\leq \int_{\mathbb{R}}\left|1-\mathrm{e}^{\mathrm{i}xz}\right|\theta(x)\mathrm{d}x\\ &\leq 2\int_{\mathbb{R}}\left|\sin(xz/2)\right|\theta(x)\mathrm{d}x\\ &\leq \int_{\mathbb{R}}|xz|\,\theta(x)\mathrm{d}x \leq K^{1/(4+m)}|z| .\end{aligned}$$

We thus get that

$$\int_0^1\frac{1-\mathrm{Re}\left(\varphi_{Y_0}(z)\right)}{z}\mathrm{d}z \leq K^{1/(4+m)} . \tag{53}$$

Now, for $|u| \geq 1$, we have that

$$\begin{aligned}\int_1^{|u|}\frac{1-\mathrm{Re}\left(\varphi_{Y_0}(z)\right)}{z}\mathrm{d}z &\leq \log|u| + \int_1^{|u|} z^{-1}\left|\mathrm{Re}\left(\varphi_{Y_0}(z)\right)\right|\mathrm{d}z\\ &\leq \log|u| + L ,\end{aligned} \tag{54}$$

where we use the Cauchy-Schwartz inequality and

$$\left(\int_1^{\infty}\left|\mathrm{Re}(\varphi_{Y_0})\right|^2\right)^{1/2} \leq L .$$

Inserting (53) and (54) in (52), we get the result. □

Lemma A.6. *There exists a constant $B > 0$ such that, for all $u > 0$, we have*

$$u \int_u^\infty \sqrt{\log(1+v)} \frac{\mathrm{d}v}{v^2} \leq B \sqrt{\log(1+u)}$$

Proof. For all $u > 0$, we have

$$u \int_u^\infty \sqrt{\log(1+v)} \frac{\mathrm{d}v}{v^2} = \int_1^\infty \sqrt{\log(1+uy)} \frac{\mathrm{d}y}{y^2} \leq \sqrt{u} \int_1^\infty \frac{\mathrm{d}y}{y^{3/2}} = 2\sqrt{u}\,.$$

As $u \to 0$, $\sqrt{\log(1+u)}$ is equivalent to $\sqrt{u}$.
As $u \to \infty$, the Karamata's Theorem (see [17, Theorem 0.6]) applied to the function $u \to \sqrt{\log(1+u)} u^{-2}$, which is regularly varying with index -2, gives that

$$u \int_u^\infty \sqrt{\log(1+v)} \frac{\mathrm{d}v}{v^2} \underset{u\to\infty}{\sim} \sqrt{\log(1+u)}\,,$$

which concludes the proof. □

References

[1] Chadi Barakat, Patrick Thiran, Gianluca Iannaccone, Christophe Diot, and Philippe Owezarski. Modeling internet backbone traffic at the flow level. *Signal Processing, IEEE Transactions on*, 51(8):2111–2124, 2003.

[2] Lennart Bondesson. Shot-noise distributions. In *Generalized Gamma Convolutions and Related Classes of Distributions and Densities*, pages 152–159. Springer, 1992. MR1224674

[3] Peter J. Brockwell and Eckhard Schlemm. Parametric estimation of the driving lévy process of multivariate carma processes from discrete observations. *Journal of Multivariate Analysis*, 115:217–251, 2013. MR3004556

[4] Pierluigi Claps, A. Giordano, and F. Laio. Advances in shot noise modeling of daily streamflows. *Advances in Water Resources*, 28(9):992–1000, 2005.

[5] Daryl J. Daley and David Vere-Jones. *An Introduction to the Theory of Point Processes*, volume 2. Springer, 1988. MR0950166

[6] Yurii Aleksandrovich Davydov. Mixing conditions for markov chains. *Teoriya Veroyatnostei i ee Primeneniya*, 18(2):321–338, 1973. MR0321183

[7] Paul Doukhan, Pascal Massart, and Emmanuel Rio. Invariance principles for absolutely regular empirical processes. In *Annales de l'IHP Probabilités et Statistiques*, volume 31, pages 393–427. Elsevier, 1995. MR1324814

[8] Shota Gugushvili. Nonparametric estimation of the characteristic triplet of a discretely observed lévy process. *Journal of Nonparametric Statistics*, 21(3):321–343, 2009. MR2530929

[9] Bruce M. Hill. A simple general approach to inference about the tail of a distribution. *The Annals of Statistics*, 3(5):1163–1174, 1975. MR0378204

[10] Nicolas Hohn and Anthony N. Burkitt. Shot noise in the leaky integrate-and-fire neuron. *Physical Review E*, 63(3):031902, 2001.

[11] Arun V. Holden. *Models of the Stochastic Activity of Neurones*. Springer-Verlag, 1976. MR0449766

[12] Aleksander M. Iksanov and Zbigniew J. Jurek. Shot noise distributions and selfdecomposability. *Stochastic Analysis and Applications*, 21(3):593–609, 2003. MR1978236

[13] Geurt Jongbloed, Frank H. van der Meulen, and Aad W. van der Vaart. Nonparametric inference for lévy-driven ornstein-uhlenbeck processes. *Bernoulli*, 11(5):759–791, 2005. MR2172840

[14] Glenn F. Knoll. *Radiation Detection and Measurement*, volume 3. Wiley New York, 1989.

[15] Hiroki Masuda. On multidimensional ornstein-uhlenbeck processes driven by a general Lévy process. *Bernoulli*, 10(1):97–120, 2004. MR2044595

[16] Michael H. Neumann and Markus Reiss. Nonparametric estimation for lévy processes from low-frequency observations. *Bernoulli*, 15(1):223–248, 2009. MR2546805

[17] Sidney I. Resnick. *Extreme Values, Regular Variation, and Point Processes*. Springer, 2007. MR0900810

[18] Ken-Iti Sato. *Lévy Processes and Infinitely Divisible Distributions*. Cambridge university press, 1999. MR1739520

[19] Raul E. Sequeira and John A. Gubner. Intensity estimation from shot-noise data. *Signal Processing, IEEE Transactions on*, 43(6):1527–1531, 1995.

[20] Raúl E. Sequeira and John A. Gubner. Blind intensity estimation from shot-noise data. *Signal Processing, IEEE Transactions on*, 45(2):421–433, 1997.

[21] Pekka Tuominen and Richard L. Tweedie. Subgeometric rates of convergence of f-ergodic markov chains. *Advances in Applied Probability*, pages 775–798, 1994. MR1285459

[22] Aad W. van der Vaart and Jon A. Wellner. *Weak Convergence and Empirical Processes*. Springer, 1996. MR1385671

[23] V. A. Volkonskii and Yu A. Rozanov. Some limit theorems for random functions. i. *Theory of Probability & Its Applications*, 4(2):178–197, 1959. MR0121856

[24] Yuanhui Xiao and Robert Lund. Inference for shot noise. *Statistical Inference for Stochastic Processes*, 9(1):77–96, 2006. MR2224841