Nonparametric estimation of a shot-noise process

Paul Ilhe, François Roueff, Éric Moulines, Antoine Souloumiac

▶ **To cite this version:**

Paul Ilhe, François Roueff, Éric Moulines, Antoine Souloumiac. Nonparametric estimation of a shot-noise process. SSP 16 - Statistical Signal Processing Workshop, IEEE Signal Processing Society, Jun 2016, Palma de Mallorca, Spain. pp.7551709, 10.1109/SSP.2016.7551709 . hal-01418963

HAL Id: hal-01418963
https://inria.hal.science/hal-01418963

Submitted on 13 Jan 2020

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.

NONPARAMETRIC ESTIMATION OF A SHOT-NOISE PROCESS

Paul Ilhe[*]*, François Roueff*

Télécom ParisTech-Télécom ParisTech

Eric Moulines, Antoine Souloumiac

Ecole Polytechnique-CEA List

ABSTRACT

We propose an efficient method to estimate in a nonparametric fashion the marks' density of a shot-noise process in presence of *pileup* from a sample of low-frequency observations. Based on a functional equation linking the marks' density to the characteristic function of the observations and its derivative, we propose a new time-efficient method using B-splines to estimate the density of the underlying γ-ray spectrum which is able to handle large datasets used in nuclear physics. A discussion on the numerical computation of the algorithm and its performances on simulated data are provided to support our findings.

***Index Terms*—** Shot-noise, nonparametric estimation, B-splines, γ-spectroscopy, pile-up correction

1. INTRODUCTION

In this paper, we are focusing on a nonlinear inverse problem that arises in nuclear science. In the specific context of γ-spectroscopy, a photon beam stemming from a radioactive source hits a scintillator detector. Assuming that the nuclear reactions at the origin of the emitted photons are independent, the arrival times of interaction between particles and matter can be modeled by a homogeneous Poisson point process. Each interaction between a particle and the detector generates an electric current called *impulse response* whose amplitude depends on the partial or total transfer of the incident photon energy. Indeed, photons interact with matter following three different phenomena: the Compton scattering, the photoelectric absorption and the pair production (for supplementary details, see [6]). We particularly focus our attention on shot-noise processes that produce *pileup*, which is the phenomenon that occurs when two or more electric currents generated by different photons overlap. For instance, this is the case when the mean inter-arrival time between two photons is shorter than the support of the impulse response, as illustrated in Figure 1. The form of the impulse response mainly depends on the type of detector and the data acquisition setup. Since the instrumentation chain is usually set by the physicists, we assume, in this article, that we exactly know the impulse response. The measured electric current can be modeled by a shot-noise process $(X_t)_{t\geq 0}$ defined by

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

under the assumptions

(**SN**-1) $\sum_k \delta_{T_k,Y_k}$ is a real-valued homogeneous Poisson point process with times T_k arriving with intensity $\lambda > 0$ and independent real-valued i.i.d. marks Y_k with probability density function f. In addition, f is compactly supported in $[0, M]$ for some $M > 0$ and belongs to the space $\mathcal{H}^2([0, M])$ defined by

$$\mathcal{H}^2([0,M]) := \{f : [0,M] \to \mathbb{R}, f, f' \text{ are absolutely continuous and } \int_0^M |f''(t)|^2\,\mathrm{d}t < \infty\}\ .$$

(**SN**-2) the impulse response h of the detector is causal, integrable and satisfies

$$h(t) \underset{t\to\infty}{=} O\left(\mathrm{e}^{-t}\right)\ .$$

Such a process is well defined as soon as the density f and the impulse response h jointly satisfy the condition

$$\int \min\left(1, |y\ h(s)|\right) f(y)\mathrm{d}y\mathrm{d}s < \infty\ .$$

Figure 1 displays a simulated shot-noise sample path with intensity $\lambda = 3$, impulse response of the form $h : t \to 10t\mathrm{e}^{-10t}\,\mathbb{1}_{t\geq 0}$ and marks' density following a Gaussian mixture.

In γ-spectroscopy, the marks $(Y_k)_k$ represent γ-ray emitted energies that depend on the underlying nucleus. In order to identify which nucleus are present and to quantify the intensity λ, one needs to estimate the *flow*

$$f_\lambda := \lambda f$$

by comparing it to known γ-ray spectrums.

Shot-noise processes belong to the family of stochastic processes well-suited to model dynamical systems that can

[*]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).

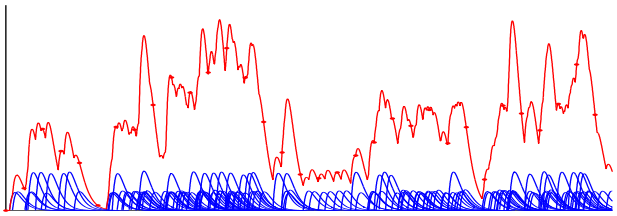

Fig. 1: Red: simulated shot-noise process and sample points Blue: corresponding singular electrical impulsions.

be interpreted as the convolution between a function (called a *filter*) and a marked point process. Among the long list of applications, one can, for instance, use it in pharmacokinetics [4], to model membrane potential dynamics over time in [10] and, more recently, to model pile-up distortion in astrophysics [2].

From a statistical standpoint, we are looking at the following problem: based on n low-frequency observations $X_\delta, \dots, X_{n\delta}$ of the shot-noise process defined by (1) and under the assumption that the impulse response h is known, we aim to estimate the function $f_\lambda := \lambda f$.

The statistical inference methods for shot-noise processes were initiated in [7], and, more recently, in [5]. Moreover, methods based on the likelihood can not be used because the marginal X_0 may not have a density (with respect to the Lebesgue measure) or,if it exists, the function can be analytically intractable. In addition, the low sampling rate of the signal precludes the use of techniques involving the joint laws of $(X_t)_{t\geq 0}$, so that we only rely on methods exploiting the marginal law of the process. Since X_0 is infinitely divisible (see [11]), one might be tempted to estimate the associated characteristic Lévy triplet (see [9] for statistical inference of Lévy processes). However, in our case, the observations $X_1, \dots, X_n$ are neither i.i.d. nor Markovian as in [5] for the function

$$h(t) = \mathrm{e}^{-\alpha t}\mathbb{1}_{t\geq 0} \, .$$

In this particular case, it is possible to establish a formula linking the characteristic function of the mark Y_0 to a function involving the characteristic function of X_0 and its derivative. For every real u, we have

$$\varphi_{Y_0}(u) = 1 + \frac{\alpha}{\lambda}\frac{u\,\varphi'_{X_0}(u)}{\varphi_{X_0}(u)} \, .$$

Under the assumption that the marks' density f belongs to some functional Sobolev space, they derive an estimator that converges at a polynomial speed. However, this paper exploits that the shot-noise process is Markovian, which is not true for general impulse responses h. This property allows to control the rates of convergence of the estimator using the β-mixing coefficients (see [3]) of the observations $(X_1, \dots, X_n)$ associated to the shot-noise process given by (1).

2. ESTIMATION PROCEDURE

In this article, we propose a method encompassing [5] that estimate the function f_λ. Without loss of generality, we assume that the sampling frequency $\delta^{-1} = 1$ that f is compactly supported in $[0, 1]$. Since the marginal law of X_0 is infinitely divisible, we can derive the analytic expression of the characteristic function φ of the marginal X_0. For every real u, we have

$$\begin{aligned}\varphi(u) :&= \mathbb{E}\left[\mathrm{e}^{\mathrm{i}uX_0}\right] \\ &= \exp\left(\int_0^\infty \int_0^1 \left[\mathrm{e}^{\mathrm{i}uxh(s)} - 1\right] f_\lambda(x)\mathrm{d}x\mathrm{d}s\right) \, . \quad (2)\end{aligned}$$

Under the mild assumption $\int |x| f_\lambda(x)\mathrm{d}x < \infty$, the differentiation of (2) with respect to u leads to the functional equation

$$g(u) = K_h\left[f_\lambda\right](u) \quad , \quad u \in \mathbb{R} \, ,$$

where $g = \varphi'/\varphi$ and K_h is a bounded linear operator only depending on h, which is defined by

$$K_h : f \to \left[u \to \int_0^\infty \int_0^1 \mathrm{i}xh(s)\mathrm{e}^{\mathrm{i}uxh(s)} f(x)\mathrm{d}x\mathrm{d}s\right] \, .$$

However, we only have access to a noisy version $\hat{g}_n$ of the function g that we compute by a "plug-in" method

$$\hat{g}_n(u) := \hat{\varphi}'_n(u)/\hat{\varphi}_n(u), \quad u \in \mathbb{R} \, ,$$

where $\hat{\varphi}_n(u) := n^{-1}\sum_{k=1}^n \mathrm{e}^{\mathrm{i}uX_k}$ corresponds to the empirical characteristic function associated to the shot-noise observations $X_1, \dots, X_n$. The general framework of those inverse problems is developed in [8] where the authors assume that one can construct unbiased estimators of the function g and explain the continuous-time singular value decomposition of the operator K_h. In general, these two hypothesis are not satisfied. Consequently, we propose to build an estimator of f_λ from a finite grid of N evaluation grid $\{u_i, i \in 1, \dots, N\}$ for which we consider the equations

$$\hat{g}_n(u_i) = K_h\left[f_\lambda\right](u_i) + \epsilon_n(u_i), \quad i \in \{1, \dots, N\} \, ,$$

where $\epsilon(u_1), \dots, \epsilon(u_N)$ correspond to the estimation errors of $g(u_i)$ by $\hat{\varphi}'_n(u_i)/\hat{\varphi}_n(u_i)$ for $i = 1, \dots, N$. In particular, the following result ensures that the normalized vector of errors can be estimated by a correlated centered Gaussian vector.

Proposition 1. *Let $\lambda > 0$ and $X_1, \cdots, X_n$ observations of the shot-noise process given by* (1) *under Assumptions (**SN**-1) and (**SN**-2). Furthermore, assume there exists some $\eta > 0$ such that $\mathbb{E}\left[|Y_1|^{4+\eta}\right] < \infty$. Let φ and φ' respectively denote*

the characteristic function of X_1 and its derivative. Then, for any integer N and any N-tuple of real numbers $\underline{u} = (u_1, \cdots, u_N)$, we have

$$\sqrt{n}\,(\epsilon_n(u_1), \ldots, \epsilon_n(u_N)) \Rightarrow \tilde{Z}(\underline{u})\ ,$$

where $\tilde{Z}(\underline{u})$ is a centered complex-valued Gaussian vector with covariance matrix denoted by $W(\underline{u})$.

For the sake of brevity, we omit here to give the explicit formula of $W(\underline{u})$. However, one can estimate it using several methods: "plug-in", generalized method of moments or bootstrap. In the sequel, we will use the bootstrap method for the numerical applications.

B-splines parameterization

Since f_λ lies in the functional space $\mathcal{H} = \mathcal{H}^2([0,1])$, the estimation problem can be translated as the following

$$\hat{f}_{n,\lambda}(\alpha) := \underset{f\in\mathcal{H}}{\operatorname{argmin}} \|\hat{g}_n(\underline{u}) - K_h[f](\underline{u})\|^2_{\hat{W}_n^{-1}} + \alpha\|f\|^2_{\mathcal{H}}\ , \quad (3)$$

where $\|f\|^2_{\mathcal{H}} := \int_0^1 |f''(t)|^2\,\mathrm{d}t$. In order to tackle this problem, we propose to estimate f_λ using the B-splines basis. This approach is suggested by [1], which studies a closely related problem. Given two positive integers q, k and a set of real numbers $\underline{\mathbf{t}} = (t_1, \ldots, t_k)$ called *knots* that satisfy $t_1 < \cdots < t_k$, the space $\mathcal{S}^q_{\underline{\mathbf{t}}}$ *of B-splines of degree q associated to the knots* $\underline{\mathbf{t}}$ is the set of functions v that can be written

$$v(t) = \sum_{j=1}^{k+q} \omega_j B^q_{j,k}(t) =: \omega^T \mathbf{B}^q_k(t)\ ,$$

for some vector $(\omega_1, \ldots, \omega_{k+q})$ of real numbers and where the basis functions $\mathbf{B}^q_k := (B^q_{1,k}, \ldots, B^q_{k+q,k})$ are defined by

$$B_{j,0}(t) := \begin{cases} 1 & \text{if} \quad t_j \leqslant t < t_{j+1} \\ 0 & \text{sinon} \end{cases}$$

for $q = 0$ and satisfy the recursive relationship

$$B_{j,l}(t) := \frac{t - t_j}{t_{j+l} - t_j} B_{j,l-1}(t) + \frac{t_{j+l+1} - t}{t_{j+l+1} - t_{j+1}} B_{j+1,l-1}(t)\ ,$$

for $l \in \{1, \ldots, k+q\}$. We consider such a parameterization for at least three reasons. On the one hand, the functions $(B_{j,q})_{1\leq j\leq k+q}$ (for given k, q and $\mathbf{t}$) are compactly supported: this allows us to model the detector's resolution in nuclear applications, which is known *à priori*. On the other hand, if a function v belongs to the B-splines vector space $\mathcal{S}$, then there exists a linear relationship between v and its derivatives. More precisely, if for some integers $q \geq 3$, $k \geq 1$, and some vector ω,

$$v = \omega^T \mathbf{B}^q_k, \quad \text{then} \quad v'' = {\omega^{(2)}}^T \mathbf{B}^{q-2}_k\ ,$$

with

$$\omega^{(2)} = A^q_k \omega \quad \text{and} \quad A^q_k \in \mathbb{R}^{(k+q-2)\times(k+q)}\ .$$

Moreover, the B-splines vector spaces are of particular interest to approximate the functions belonging to $\mathcal{H}$. Thus, it is possible to construct an estimator of f_λ able to adapt to the local regularity of the target function by

$$\hat{f}_{n,\lambda}(\alpha) := \hat{\omega}_{n,\lambda}(\alpha)^T \mathbf{B}\ .$$

where $\hat{\omega}_{n,\lambda}(\alpha)$ is the solution of the quadratic minimization problem

$$\min \|\hat{g}_n(\underline{u}) - K_h[\mathbf{B}](\underline{u})^T\omega\|^2_{\hat{W}_n^{-1}} + \alpha\omega^T D_2^{k+q}\omega\ , \quad (4)$$

for $\omega \in \mathbb{R}^{k+q}$ and where D_2^{k+q} is a matrix only depending on k and q.

Penalty cross-validation

The choice of the penalty coefficient α in the minimization problem (4) is related to the quadratic problems considered in [12]. The author proposes a practical and efficient way to choose, for a given number of penalty terms $\alpha_1, \ldots, \alpha_p$, which one is the closest to the usual leave-one out cross-validation score. Coarsely speaking, it boils down to choose the real number $\alpha \in \{\alpha_1, \ldots, \alpha_p\}$ that minimizes the function

$$\mathrm{CV}(\alpha) := \frac{\|\hat{g}_n(\underline{u}) - A_n(\alpha)\hat{g}_n(\underline{u})\|^2_{\hat{W}_n^{-1}}}{\left(1 - \mathrm{Tr}(A_n(\alpha))/N\right)^2} \quad (5)$$

with

$$A_n(\alpha) := K_h[\mathbf{B}^q_k]\left(K_h[\mathbf{B}^q_k]^\star \hat{W}_n^{-1} K_h[\mathbf{B}^q_k] + \alpha D_2^{k+q}\right) K_h[\mathbf{B}^q_k]^\star \hat{W}_n^{-1}.$$

3. ALGORITHM AND NUMERICAL RESULTS

The detailed estimation procedure described in the previous section is summed up in the following pseudo-code. In particular, we propose a practical manner to set hyperparameters and illustrate the performances of this algorithm on simulated datasets.

We illustrate the performance of our algorithm on simulated observations of a shot-noise process with intensity and impulse response given by

$$\lambda = 3 \quad \text{and} \quad h : t \to 10\ t\ \mathrm{e}^{-10t}\mathbb{1}_{t\geq 0}\ ,$$

and marks $(Y_k)_k$ following a Gaussian mixture with density

$$f = 0{,}7\ \mathcal{N}\left(0{,}3; 2{,}5.10^{-3}\right) + 0{,}3\ \mathcal{N}\left(0{,}6; 2{,}5.10^{-3}\right)\ ,$$

where $\mathcal{N}(\mu; \sigma^2)$ is the probability density function of Gaussian random variable with mean μ and variance σ^2. Figures display the densities obtained using Algorithm 1 for several number of observations and a boxplot of the integrated

Algorithm 1 B-splines estimator

Inputs : Shot-noise observations $X_1, \cdots, X_n$.

1. Choice of $\Delta_n > 0$. Computation of the histogram $(H_i)_{i\in\mathbb{Z}}$ based on the observations $X_1, \cdots, X_n$:
   $H_i := \{X_k \in [\Delta_n i; \Delta_n(i+1)]\}/n$ and $(dH_i)_{i\in\mathbb{Z}}$ defined by $dH_i = \Delta i H_i$
2. Choice of $N := 2^{\lfloor \log(n) \rfloor}$ and computation of $X\phi = \text{FFT}(H, N)$, $dX\phi = \text{FFT}(dH, N)$.
   For $k = 1, \cdots, N$, we set $\hat{g}_n\left(\frac{2\pi(k-1)}{N\Delta}\right) := \frac{\overline{dX\phi_k}}{X\phi_k}$
3. Choice of the knots' number $k = N$, the degree $q = 3$ and computation of the design matrix $(K_h[B_{j,q}](2\pi(i-1)/N\Delta))_{i,j}$ for $i \in \{1, \ldots, N\}, j \in \{1, \ldots k+q\}$
4. Estimation of $\hat{W}_n$ by bootstrap.
5. Estimation of $\hat{\omega}_n(\alpha)$ for $\alpha \in \{2^{-15}, \ldots, 2^{15}\}$ via (4) and choice by the cross-validation score using (5).

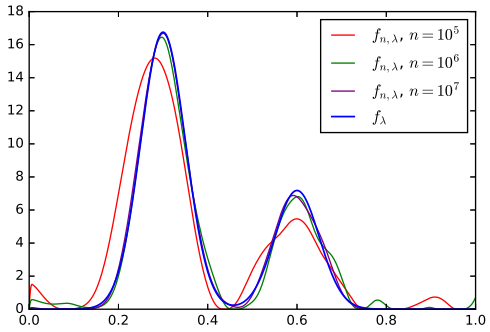

Fig. 2: Estimation of f_λ for $n \in \{10^5, 10^6, 10^7\}$

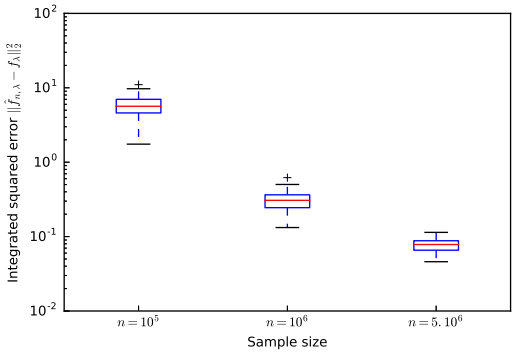

Fig. 3: Boxplot of the Integrated Squared Error $\|\hat{f}_{n,\lambda} - f_\lambda\|_2^2$ for $n \in \{10^5, 10^6, 5.10^6\}$ and 100 independent runs of Algorithm 1 over independent input data sets.

squared error $\hat{f}_{n,\lambda}$ and f_λ for 100 independent runs of Algorithm 1 and three different sample sizes.

As shown in Figure 2 and 3, the estimator $\hat{f}_{n,\lambda}$ converges to f_λ. In addition, we see that the estimator well retrieves the modes of f_λ, even with a restricted number of observations . This presents an interest for γ-spectroscopy because it permits to identify the characteristic peaks of a given radionuclide.

4. REFERENCES

[1] Cardot, Hervé, Spatially adaptive splines for statistical linear inverse problems, *Journal of Mutivariate Analysis*, (2002)

[2] Chaplin, V., Bhat, N., Briggs, M. S., Connaughton, V. Analytical modeling of pulse-pileup distortion using the true pulse shape; applications to Fermi-GBM. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 717, 21-36. (2013)

[3] Doukhan, Paul, Mixing, *Springer*, (1994)

[4] Helfenstein, U., Steiner, M. . Random drug excess. *European journal of drug metabolism and pharmacokinetics*, 15(4), 309-315. (1990)

[5] Ilhe, Paul, Moulines, Eric, Roueff, François and Souloumiac Antoine. Nonparametric estimation of mark's distribution of an exponential Shot-noise process. *Electronic Journal of Statistics*, 9(2). (2015)

[6] Knoll, Glenn F, Radiation detection and measurement, *Wiley New York*, (1989)

[7] Macchi, Odile and Picinbono, Bernard C., Estimation and detection of weak optical signals, *Information Theory, IEEE Transactions on*,18(5),562-573, (1972)

[8] Mair, B. A., Ruymgaart, F. H., Statistical inverse estimation in Hilbert scales. *SIAM Journal on Applied Mathematics*, 56(5), 1424-1444. (1996)

[9] Neumann, M. H., Reiss, M. Nonparametric estimation for Lvy processes from low-frequency observations. *Bernoulli*, 15(1), 223-248. (2009)

[10] Rudolph, M., Pospischil, M., Timofeev, I., Destexhe, A. Inhibition determines membrane potential dynamics and controls action potential generation in awake and sleeping cat cortex. *The Journal of neuroscience*, 27(20), 5280-5290, (2007)

[11] Sato, Ken-iti, Lévy processes and infinitely divisible distributions, *Cambridge University Press*, (1999)

[12] Wood, Simon N, Modelling and smoothing parameter estimation with multiple quadratic penalties, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, (2000)