Bernoulli **13**(2), 2007, 365–388
DOI: 10.3150/07-BEJ5184

Nonparametric inference of photon energy distribution from indirect measurement

E. MOULINES[1,*], F. ROUEFF[1,**], A. SOULOUMIAC[2] and T. TRIGANO[3]

[1]*ENST Paris, CNRS/LTCI, 34 rue Dareau, 75014 Paris, France. E-mails: *moulines@tsi.enst.fr, **roueff@tsi.enst.fr*
[2]*SSTM/LETS, CEA Saclay, 91191 Gif/Yvette CEDEX, France. E-mail: antoine.souloumiac@cea.fr*
[3]*Department of Statistics, Hebrew University of Jerusalem, Jerusalem 91905, Israel. E-mail: trigano@mscc.huji.ac.il*

We consider a density estimation problem arising in nuclear physics. Gamma photons are impinging on a semiconductor detector, producing pulses of current. The integral of this pulse is equal to the total amount of charge created by the photon in the detector, which is linearly related to the photon energy. Because the inter-arrival times of photons can be shorter than the charge collection time, pulses corresponding to different photons may overlap leading to a phenomenon known as pile-up. The distortions on the photon energy spectrum estimate due to pile-up become worse when the photon rate increases, making pile-up correction techniques a must for high counting rate experiments. In this paper, we present a novel technique to correct pile-up, which extends a method introduced by Hall and Park for the estimation of the service time from the busy period in $M/G/\infty$ models. It is based on a novel formula linking the joint distribution of the energy and duration of the cluster of pulses and the distribution of the energy of the photons. We then assess the performance of this estimator by providing an expression for its integrated square error. A Monte Carlo experiment is presented to illustrate, with practical examples, the benefits of the pile-up correction.

Keywords: indirect observations; marked Poisson processes; nonlinear inverse problems; nonparametric density estimation

1. Introduction

We consider a problem occurring in nuclear spectroscopy. A radioactive source (a mixture of radionuclides) emits photons which impinge on a semiconductor detector. Photons (X- and gamma rays) interact with the semiconductor crystal to produce electron–hole pairs. The migration of these pairs in the semiconductor produces a finite-duration pulse of current. Under appropriate experimental conditions (ultra-pure crystal, low temperature), the integral over time of this pulse of current corresponds to the total number of electron–hole pairs created in the detector, which is proportional to the energy deposited in the semiconductor (see Knoll 1989; Leo 1994). In most classical semiconductor radiation detectors, the pulse amplitudes are recorded and sorted to produce a histogram which is used as an estimate of the photon energy distribution (referred to in the nuclear physics literature as the *energy spectrum*).

The inter-arrival times of photons are independent of their electrical pulses, and can therefore be shorter than the typical duration of the charge collection, thus creating *clusters* (see Figure 1). In gamma ray spectrometry, this phenomenon is referred to as *pile-up*. The pile-up phenomenon induces a distortion of the acquired energy spectrum which becomes more severe as the incoming counting rate increases. This problem has been extensively studied in the field of nuclear instrumentation since the 1960s – see Bristow (1990) for a detailed review of these early contributions; classical pile-up rejection techniques are detailed in American National Standards Institute (1999).

In mathematical terms, the problem can be formalized as follows. Denote by $\{T_k,\ k \geqslant 1\}$ the sequence of arrival times of the photons, assumed to be the ordered points of a homogeneous Poisson process. The current intensity as a function of time can be modelled as a shot-noise process

$$W(t) \stackrel{\text{def}}{=} \sum_{k \geqslant 1} F_k(t - T_k), \tag{1}$$

where $\{F_k(s),\ k \geqslant 1\}$ are the contributions of each individual photon to the overall intensity. By analogy with queuing models, we call $\{W(t),\ t \geqslant 0\}$ the *workload* process. The current pulses $\{F_k(s),\ k \geqslant 1\}$ are assumed to be independent copies of a continuous time stochastic process $\{F(s),\ s \geqslant 0\}$. The pulse duration (the duration of the charge collection), defined as $X \stackrel{\text{def}}{=} \sup\{t : F(t) > 0\}$, is assumed to be finite almost surely and the support of the path of F is assumed to be of the form $[0, X]$ almost surely, so that a busy-period arrival corresponds to a pulse arrival and a pulse cannot belong to several busy periods. The integral of the pulse $Y \stackrel{\text{def}}{=} \int_0^X F(u)\mathrm{d}u$ is equal to the total amount of charge collected for a single photon. Under appropriate experimental condition, this quantity may be shown to be linearly related to the photon energy; for convenience, Y is referred to as the *energy* in the following. For the kth photon, we define the couple (X_k, Y_k) accordingly with respect to F_k. The restriction of the

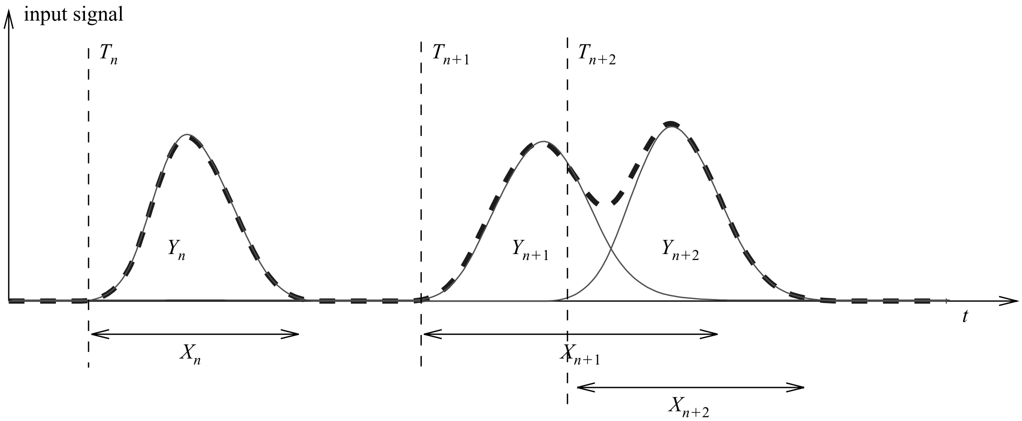


Figure 1. Illustration of the pile-up phenomenon: input signal with arrival times T_j, lengths X_j and energies Y_j, $j = n$, $n+1$, $n+2$. Here $X'_n = X_n$, $Y'_n = Y_n$, $X'_{n+1} = T_{n+2} - T_{n+1} + X_{n+2}$ and $Y'_{n+1} = Y_{n+1} + Y_{n+2}$.

workload process to a maximal segment where it is positive is referred to as a *busy period*, and where it is 0 as *idle*. An idle period followed by a busy period is called a *cycle*.

In our experimental setting, the sequence of pulse duration and energy $\{(X_k, Y_k), k \geqslant 1\}$ is not directly observed. Instead, the only available data are the durations of the busy and idle periods and the total amounts of charge collected on busy periods. Define the on–off process

$$S_t = \sum_{k \geqslant 1} \mathbf{1}_{T'_k, T'_k + X'_k}(t), \tag{2}$$

where $\{T'_k, k \geqslant 1\}$ is the ordered sequence of busy-period arrivals and $\{X'_k, k \geqslant 1\}$ the corresponding sequence of durations. We further define, for all $k \geqslant 1$,

$$Y'_k \stackrel{\text{def}}{=} \int_{T'_k}^{T'_k + X'_k} W(t)\mathrm{d}t,$$

the total amount of charge of the kth busy period. Finally, we denote by Z_k the duration of the kth idle period, $Z_1 = T_1$ and, for all $k \geqslant 2$, $Z_k = T'_k - (T'_{k-1} + X'_{k-1})$. We consider the problem of estimating the distribution of the photon energy Y with n cycles $\{(Z_k, X'_k, Y'_k), k = 1, \ldots, n\}$ observed. In the terminology introduced by Pyke (1958), this corresponds to a type II counter.

The problem shares some similarity with service-time distribution from busy and idle measurements in a $M/G/\infty$ model (see Baccelli and Brémaud 2003). Note, indeed, that the $M/G/\infty$ model is a particular instance of the above setting, as it corresponds to $F = \mathbf{1}_{[0,X)}$, so that $X = Y$. There exists a vast literature for this particular case. Takács (1962) (see also Hall 1988) has derived a closed-form relation linking the cumulative distribution functions (cdfs) of the service time X and the busy period X'. Bingham and Pitts (1999) derived from this formula an estimator of the service-time distribution X, which they apply to the study of biological signals. An alternative estimator was recently introduced in Hall and Park (2004), in which a kernel estimator of the probability density function (pdf) of X is derived in a nonparametric framework, together with a bound on the pointwise error.

Although our estimator can be applied to the $M/G/\infty$ framework – thus allowing a comparison with Hall and Park (2004) in this special case – we stress the fact that we are dealing here simultaneously with durations and energies, without assuming any particular dependence structure between them. Secondly, the main emphasis in the photon problem is on estimating the distribution of the photon energy and not the distribution of the duration, in sharp contrast to the $M/G/\infty$ problem.

The paper is organized as follows. In Section 2 we give the notation and main assumptions, and list the basic properties of the model. In Section 3 we present an inversion formula relating the Laplace transform of the cluster duration/energy to the Laplace transform of the density function of interest. We also derive an estimator of this function, which is based on an empirical version of the inversion formula and kernel smoothing. Our main result is presented in Section 4, showing that this estimator achieves standard minimax rates in the sense of the integrated squared error when the pulse duration is almost surely upper-bounded. The study of this error is detailed is Section 5. Some applications and examples are presented in Section 6. Since the present paper is directed towards establishing a theory, practical aspects are not discussed in much detail and we refer to Trigano *et al.* (2006) for a thorough discussion of the implementation and applications to real data. Proofs of the various propositions are presented in Appendices A–C.

2. Notation and main assumptions

Throughout the paper, we make the following assumptions:

(H1) $\{T_k,\ k \geqslant 1\}$ is the ordered sequence of points of a homogeneous Poisson process on the positive half-line with intensity λ.

(H2) $\{(X, Y), (X_k, Y_k),\ k \geqslant 1\}$ is a sequence of independent and identically distributed $(0, \infty)^2$-valued random variables, independent of $\{T_k,\ k \geqslant 1\}$. In addition, $\mathbb{E}[X]$ and $\mathbb{E}[Y]$ are finite.

In other words, $\{(T_k, X_k, Y_k),\ k \geqslant 1\}$ is a Poisson point process with control measure $\lambda \mathrm{Leb} \otimes P$, where Leb denotes the Lebesgue measure on the positive half-line and P denotes the probability distribution of (X, Y). Let us recall a few basic properties satisfied under this assumption by the sequence $\{(Z_k, X'_k, Y'_k),\ k \geqslant 1\}$ defined in the Introduction. By the memorylessness property of the exponential distribution, the idle periods are independent and identically distributed with common exponential distribution with parameter λ. Moreover, they are independent of the busy periods, which also are independent and identically distributed. We denote by (X', Y') a couple having the same distribution as the variables of the sequence $\{(X'_k, Y'_k),\ k \geqslant 1\}$ and by P' its probability measure. Given that $\mathbb{E}[X]$ and $\mathbb{E}[Y]$ are finite, it is easily shown that

$$\mathbb{E}[X'] = \{\exp(\lambda \mathbb{E}[X]) - 1\}/\lambda$$

$$\mathbb{E}[Y'] = \mathbb{E}[Y]\exp(\lambda \mathbb{E}[X]).$$

Our goal is the nonparametric estimation of the distribution of Y; hence we assume that:

(H3) Y admits a probability density function denoted by m, that is,

$$\int_{x>0} P(\mathrm{d}x, \mathrm{d}y) = m(y)\mathrm{Leb}(\mathrm{d}y).$$

As mentioned in Section 1, the marks $\{(X_k, Y_k),\ k \geqslant 1\}$ are not directly observed but, instead, we observe the sequence $\{(T'_k, X'_k, Y'_k),\ k = 1, \ldots, n\}$ – the arrival times, duration and integrated energy of the successive busy periods. These quantities are recursively defined as follows. Let $T'_1 = T_1$ and, for all $k \geqslant 2$,

$$T'_k = \inf\left\{ T_i : T_i > \left(T'_{k-1} \vee \max_{j \geqslant i-1} (T_j + X_j) \right) \right\}; \tag{3}$$

for all $k \geqslant 1$,

$$X'_k = \max_{T_i \in [T'_k, T'_{k+1}[} \{T_i + X_i\} - T'_k, \tag{4}$$

$$Y'_k = \sum_{i \geq 1} Y_i \mathbf{1}(T'_k \leq T_i < T'_{k+1}).$$

Equations (3) and (4) are used for simulation procedures.

Remark 2.1. In this paper, it is assumed that the experiment involves collecting a number n of cycles. Hence, the total duration of the experiment is equal to $T'_n + X'_n$ and is therefore random. Using the renewal property and the law of large numbers, as $n \to \infty$, $(T'_n + X'_n)/n$ converges almost surely to the mean duration of a cycle, $1/\lambda + \mathbb{E}[X'] = \lambda^{-1} \exp(\lambda \mathbb{E}[X])$. Another approach, which appears more sensible in certain scenarios, is to consider that the total duration of the experiment is given, say equal to T. In this case, the number of cycles is random, equal to the renewal process of the busy cycles, $N_\mathsf{T} = \sum_{k-1}^{\infty} \mathbf{1}\{T'_k + X'_k \leq \mathsf{T}\}$. Given that T lies between $T'_{N_\mathsf{T}} + X'_{N_\mathsf{T}}$ and $T'_{N_\mathsf{T}+1} + X'_{N_\mathsf{T}+1}$, as $\mathsf{T} \to \infty$, we have $N_\mathsf{T}/\mathsf{T} \to \lambda \exp(-\lambda \mathbb{E}[X])$, showing that the asymptotic theory in both cases can be easily related.

3. Inversion formula and estimation

Let $\tilde{P}$ be a probability measure on $\mathbb{R} \times \mathbb{R}$ equipped with the Borel σ-algebra; for all $(s, p) \in \mathbb{C}^+ \times \mathbb{C}^+$, where $\mathbb{C}^+ = \{z \in \mathbb{C}, \mathrm{Re}(z) \geq 0\}$, we define its Laplace transform (or moment-generating function) $\mathcal{L}\tilde{P}$ as:

$$\mathcal{L}\tilde{P}(s, p) = \int\int \mathrm{e}^{-su-pv} \tilde{P}(\mathrm{d}u, \mathrm{d}v).$$

The following theorem provides a relation between the joint distribution of the individual pulse energies and durations P and the moment-generating function of the distribution of the energies and durations of the busy periods $\mathcal{L}P'$; this key relation will be used to derive an estimator of m.

Theorem 3.1. *Under Assumptions* (H1) *and* (H2)*, for all* $(s, p) \in \mathbb{C}^+ \times \mathbb{C}^+$*,*

$$\int_{u=0}^{+\infty} \mathrm{e}^{-(s+\lambda)u} \{a(u, p) - 1\} \mathrm{d}u = \frac{\lambda \mathcal{L}P'(s, p)}{s + \lambda} \frac{1}{s + \lambda - \lambda \mathcal{L}P'(s, p)}, \tag{5}$$

where

$$a(u, p) \stackrel{\text{def}}{=} \exp(\lambda \mathbb{E}[\mathrm{e}^{-pY}(u - X)_+]). \tag{6}$$

Proof. See Appendix C for a rigorous proof. Since (5) may not seem very intuitive, we here give an outline of the demonstration. Denote by $\overline{Y}_x$ the total accumulated energy at time x. Observe, following Takács (1962), that reliable information on the energy can only be gathered when being in an idle period. Consequently, the probability $\mathbb{P}(S_x = 0;\ \overline{Y}_x \leq y)$ is of interest. This probability can be computed by considering the idle and busy periods $\{(T'_k, X'_k, Y'_k),\ k \geqslant 1\}$, using renewal properties, thus giving the right-hand side of (5) in the Laplace space. On the other hand, considering $\{(T_k, X_k, Y_k),\ k \geqslant 1\}$, the same probability can be computed using standard properties of the homogeneous Poisson process, thus giving the left-hand side of (5), hence the result. □

Remark 3.1. Observe that the integral in (5) can be replaced by $\int_{u=-\infty}^{\infty}$ since, in (6), $a(u, p) = 0$ for $u < 0$. Moreover, from (6), we trivially obtain $|a(u, p)| \leqslant \exp(\lambda u)$ for $\mathrm{Re}(p) \geqslant 0$; hence this integral is well defined for $\mathrm{Re}(s) > 0$ and $\mathrm{Re}(p) \geqslant 0$.

Since we will make extensive use of relation (5), we introduce some auxiliary notation for the sake of readability, and define

$$\Phi(s, p; \lambda, \mathcal{L}P') \stackrel{\text{def}}{=} \frac{\lambda \mathcal{L}P'(s, p)}{s+\lambda} \frac{1}{s+\lambda-\lambda\mathcal{L}P'(s, p)}.$$

It is perhaps not immediately obvious how this relation may lead to an estimator of the distribution of the energy. By logarithmic differentiation with respect to x, (6) implies

$$\frac{\partial}{\partial x} \log a(x, p) = \lambda \mathbb{E}[\mathrm{e}^{-pY} \mathbf{1}(X \leqslant x)]. \tag{7}$$

We consider a kernel function K that integrates to 1 and denote by K^* its Fourier transform, $K^*(\nu) = \int_{-\infty}^{+\infty} K(y)\mathrm{e}^{-\mathrm{i}\nu y}\,\mathrm{d}y$, so that $K^*(0) = 1$. We further assume that K^* is integrable; hence, for any $y \in \mathbb{R}$,

$$K(y) = \frac{1}{2\pi} \int_{-\infty}^{\infty} K^*(\nu)\mathrm{e}^{\mathrm{i}\nu y}\,\mathrm{d}\nu.$$

Hence, from (7) and Fubini's theorem, we have, for any bandwidth parameter $h > 0$ and all $y \in \mathbb{R}$,

$$\begin{aligned} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{\lambda} \frac{\partial}{\partial x} \log a(x, \mathrm{i}\nu) K^*(h\nu)\mathrm{e}^{\mathrm{i}\nu y}\,\mathrm{d}\nu &= \mathbb{E}\left[\frac{1}{2\pi} \int_{-\infty}^{\infty} K^*(h\nu)\mathrm{e}^{\mathrm{i}\nu(y-Y)} \mathbf{1}(X \leqslant x)\mathrm{d}\nu\right] \\ &= \mathbb{E}\left[\frac{1}{h} K\left(\frac{y-Y}{h}\right) \mathbf{1}(X \leqslant x)\right]. \end{aligned} \tag{8}$$

Taking the limits $x \to \infty$ and $h \to 0$ in the previous equation leads to the following explicit inversion formula which will be used to derive our estimator. For any continuity point y of the density m, we have

$$m(y) = \lim_{h\to 0} \lim_{x\to+\infty} \left\{\frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1}{\lambda} \frac{\partial}{\partial x} \log a(x, \mathrm{i}\nu) K^*(h\nu)\mathrm{e}^{\mathrm{i}\nu y}\,\mathrm{d}\nu\right\} \tag{9}$$

We now observe that for any $p \in \mathbb{C}^+$, the right-hand side of (5) is integrable on a line $\{c + \mathrm{i}\omega,\ \omega \in \mathbb{R}\}$ where c is an arbitrary positive number. By inverting the Laplace transform, (5) implies that, for all $p \in \mathbb{C}^+$ and $x \in \mathbb{R}_+$,

$$a(x,\ p) = 1 + \frac{1}{2\pi}\int_{-\infty}^{+\infty} \Phi(c + \mathrm{i}\omega,\ p;\ \lambda,\ \mathcal{L}P'\}\mathrm{e}^{(c+\lambda+\mathrm{i}\omega)x}\,\mathrm{d}\omega. \tag{10}$$

Our estimator of m is based on (9) and (10), but we need first to estimate λ, the intensity of the underlying Poisson process. Since the idle periods are independent and identically distributed according to an exponential distribution with intensity λ, we use the maximum-likelihood estimator based on the durations of the idle periods $\{Z_k,\ k = 1,\ \ldots,\ n\}$, namely,

$$\hat{\lambda}_n \overset{\mathrm{def}}{=} \left(\frac{1}{n}\sum_{k=1}^{n} Z_k\right)^{-1}. \tag{11}$$

The function $a(x,\ \mathrm{i}\nu)$ can be estimated from $\{(X'_k,\ Y'_k),\ k = 1,\ \ldots,\ n\}$ by plugging into (10) an estimate of the Laplace transform $\mathcal{L}P'$ of the joint distribution of the busy-period duration and energy. More precisely, let $\widehat{P'}_n$ be the associated empirical measure: for any bivariate measurable function g, we write

$$\widehat{P'}_n g \overset{\mathrm{def}}{=} \int\!\!\int g(x,\ y)\widehat{P'}_n(\mathrm{d}x,\ \mathrm{d}y) = \frac{1}{n}\sum_{k=1}^{n} g(X'_k,\ Y'_k).$$

We consider the estimator

$$\hat{a}_n(x,\ \mathrm{i}\nu) = 1 + \frac{1}{2\pi}\int_{-\infty}^{+\infty} \Phi(c + \mathrm{i}\omega,\ \mathrm{i}\nu;\ \hat{\lambda}_n,\ \mathcal{L}P'_n)\mathrm{e}^{(\hat{\lambda}_n+c+\mathrm{i}\omega)x}\,\mathrm{d}\omega, \tag{12}$$

where

$$\widehat{\mathcal{L}P'}_n(c + \mathrm{i}\omega,\ \mathrm{i}\nu) \overset{\mathrm{def}}{=} \mathcal{L}\widehat{P'}_n(c + \mathrm{i}\omega,\ \mathrm{i}\nu) = \frac{1}{n}\sum_{k=1}^{n} \mathrm{e}^{-(c+\mathrm{i}\omega)X'_k - \mathrm{i}\nu Y'_k}. \tag{13}$$

In practice, the numerical computation of this integral (and also the one in (20) below) can be done by using efficient numerical packages – see Gautschi (1996) for an overview of numerical integration methods. Since the integrand is infinitely differentiable and has a modulus decaying as $|\omega|^{-2}$ when $\omega \to \pm\infty$, the errors in computing this integral numerically can be made arbitrary small. The numerical error will thus not be taken into account here for the sake of brevity.

In order to estimate $\lambda^{-1}\partial \log a/\partial x$, we also need to estimate the partial derivative $\partial a/\partial x$. Because the function $x \mapsto a(x,\ \mathrm{i}\nu)$ (see (10)) is defined as an inverse Fourier transform of an integrable function, it is tempting to estimate its partial derivative simply by multiplying its Fourier transform by a factor $\lambda + c + \mathrm{i}\omega$ prior to inversion. However, this approach is not directly applicable, because multiplying the integrand by ω in (10) leads to an absolutely non-convergent integral. As observed by Hall and Park (2004) in a related problem, it is possible to get rid of this difficulty by finding an explicit expression for the singular part of this function, which can be computed and estimated. Note first that, for any s and p with non-negative real parts, $|\mathcal{L}P'(s,\ p)| \leqslant 1$; on the other hand, $\mathrm{Re}(s) > 0$ implies $|\lambda/(s + \lambda)| < 1$. Therefore, for all $(\omega,\ \nu) \in \mathbb{R} \times \mathbb{R}$,

$$\frac{1}{c+\mathrm{i}\omega+\lambda-\lambda\mathcal{L}P'(c+\mathrm{i}\omega,\,\mathrm{i}\nu)}=\frac{1}{c+\mathrm{i}\omega+\lambda}\sum_{n\geqslant 0}\left(\frac{\lambda\mathcal{L}P'(c+\mathrm{i}\omega,\,\mathrm{i}\nu)}{\lambda+c+\mathrm{i}\omega}\right)^{n}.$$

Using the latter equation, we obtain

$$\Phi(c+\mathrm{i}\omega,\,\mathrm{i}\nu;\,\lambda,\,\mathcal{L}P')=A_1(\omega,\,\mathrm{i}\nu)+A_2(\omega,\,\mathrm{i}\nu), \tag{14}$$

where we have defined

$$A_1(\omega,\,\mathrm{i}\nu)\stackrel{\mathrm{def}}{=}\frac{\lambda\mathcal{L}P'(c+\mathrm{i}\omega,\,\mathrm{i}\nu)}{(c+\mathrm{i}\omega+\lambda)^2},$$

$$A_2(\omega,\,\mathrm{i}\nu)\stackrel{\mathrm{def}}{=}\frac{\lambda\mathcal{L}P'(c+\mathrm{i}\omega,\,\mathrm{i}\nu)}{c+\mathrm{i}\omega+\lambda}\Phi(c+\mathrm{i}\omega,\,\mathrm{i}\nu;\,\lambda,\,\mathcal{L}P').$$

It is easily seen that the functions $\omega\mapsto A_k(\omega,\,\mathrm{i}\nu)$, $k=1,2$ are integrable. Hence, we may define, for $k=1,2$, and all real numbers x and ν,

$$a_k(x,\,\mathrm{i}\nu)\stackrel{\mathrm{def}}{=}\frac{1}{2\pi\lambda}\int_{\omega=-\infty}^{\infty}A_k(\omega,\,\mathrm{i}\nu)\mathrm{e}^{(\lambda+c+\mathrm{i}\omega)x}\,\mathrm{d}\omega, \tag{15}$$

and therefore, using (10) and (14),

$$a(x,\,\mathrm{i}\nu)=1+\lambda a_1(x,\,\mathrm{i}\nu)+\lambda a_2(x,\,\mathrm{i}\nu)$$

which finally yields

$$\frac{1}{\lambda}\frac{\partial}{\partial x}\log a(x,\,\mathrm{i}\nu)=\frac{1}{a(x,\,\mathrm{i}\nu)}\left[\frac{\partial a_1}{\partial x}+\frac{\partial a_2}{\partial x}\right](x,\,\mathrm{i}\nu). \tag{16}$$

Recall that the moment-generating function of a gamma distribution with shape parameter 2 and scale parameter λ is given by $x\mapsto\lambda^2/(\lambda-x)^2$. It follows that, for all $u\in\mathbb{R}$,

$$\frac{1}{2\pi}\int_{\omega=-\infty}^{\infty}\frac{\mathrm{e}^{(c+\mathrm{i}\omega)u}}{(\lambda+c+\mathrm{i}\omega)^2}\,\mathrm{d}\omega=u_+\mathrm{e}^{-\lambda u}.$$

Using Fubini's theorem and this equation, we obtain, for all real numbers x and ν,

$$\begin{aligned}a_1(x,\,\mathrm{i}\nu)&=\frac{1}{2\pi}\int_{\omega=-\infty}^{\infty}\frac{\mathbb{E}[\mathrm{e}^{-((c+\mathrm{i}\omega)X'+\mathrm{i}\nu Y')}]\mathrm{e}^{(\lambda+c+\mathrm{i}\omega)x}}{(\lambda+c+\mathrm{i}\omega)^2}\,\mathrm{d}\omega\\&=\mathrm{e}^{\lambda x}\mathbb{E}[\mathrm{e}^{-\mathrm{i}\nu Y'}(x-X')_+\mathrm{e}^{-\lambda(x-X')}]=\mathbb{E}[(x-X')_+\mathrm{e}^{\lambda X'-\mathrm{i}\nu Y'}],\end{aligned}$$

and, differentiating this latter expression with respect to x, we obtain

$$\frac{\partial a_1}{\partial x}(x,\,\mathrm{i}\nu)=\mathbb{E}[\mathbf{1}(X'\leqslant x)\mathrm{e}^{\lambda X'-\mathrm{i}\nu Y'}]. \tag{17}$$

On the other hand, since $|A_2(\omega,\,\mathrm{i}\nu)|=O(|\omega|^{-3})$ as $\omega\to\pm\infty$, the derivative of a_2 can (and will) be computed by multiplying the integrand in (15) by $\lambda+c+\mathrm{i}\omega$:

$$\frac{\partial a_2}{\partial x}(x,\,\mathrm{i}\nu)=\frac{1}{2\pi}\int_{-\infty}^{+\infty}\mathcal{L}P'(c+\mathrm{i}\omega,\,\mathrm{i}\nu)\Phi(c+\mathrm{i}\omega,\,\mathrm{i}\nu;\,\lambda,\,\mathcal{L}P')\mathrm{e}^{(\lambda+c+\mathrm{i}\omega)x}\,\mathrm{d}\omega. \tag{18}$$

Equations (18) and (17) then yield the following estimators for $\partial a_k/\partial x$, $k = 1, 2$:

$$\hat{I}_{1,n}(x, \mathrm{i}\nu) = \frac{1}{n}\sum_{k=1}^{n} \mathbf{1}(X'_k \leqslant x)\mathrm{e}^{\hat{\lambda}_n X'_k - \mathrm{i}\nu Y'_k} \tag{19}$$

$$\hat{I}_{2,n}(x, \mathrm{i}\nu) = \frac{\mathrm{e}^{(c+\hat{\lambda}_n)x}}{2\pi}\int_{-\infty}^{+\infty} \widehat{\mathcal{L}P'}_n(c+\mathrm{i}\omega, \mathrm{i}\nu)\Phi)c+\mathrm{i}\omega, \mathrm{i}\nu; \hat{\lambda}_n, \widehat{\mathcal{L}P'}_n)\mathrm{e}^{\mathrm{i}\omega x}\,\mathrm{d}\omega, \tag{20}$$

where $\hat{\lambda}_n$ and $\widehat{\mathcal{L}P'}_n$ are given by (11) and (13), respectively. From (9) and (16), we finally define the following estimator for the energy distribution density function:

$$\hat{m}_{x,h,n}(y) = \frac{1}{2\pi}\int_{-\infty}^{+\infty}\left[\frac{\hat{I}_{1,n}+\hat{I}_{2,n}}{\hat{a}_n}(x, \mathrm{i}\nu)\right]K^*(h\nu)\mathrm{e}^{\mathrm{i}\nu y}\,\mathrm{d}\nu, \tag{21}$$

where $\hat{a}_n$, $\hat{I}_{1,n}$ and $\hat{I}_{2,n}$ are defined in (12), (19) and (20), respectively.

4. Main result

We denote by $\|\cdot\|_\infty$ the infinite norm, by $\|\cdot\|_2$ the L^2-norm and by $\|\cdot\|_{\mathcal{W}(\beta)}$ the Sobolev norm of exponent β, that is, the norm endowing the Sobolev space

$$\mathcal{W}(\beta) \stackrel{\text{def}}{=} \left\{ g \in L^2(\mathbb{R});\ \|g\|^2_{\mathcal{W}(\beta)} \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} (1+|\nu|)^{2\beta}|g^*(\nu)|^2\,\mathrm{d}\nu < \infty \right\},$$

where g^* denotes the Fourier transform of g. In order to control the error terms, it is necessary to make a standard assumption on the kernel:

(H4) K^* has a compact support, and there exists constants $C_K > 0$ and $l \geqslant \beta$ such that, for all $\nu \in \mathbb{R}$,

$$|1 - K^*(\nu)| \leqslant C_K \frac{|\nu|^l}{(1+|\nu|)^l}.$$

We may now state the main result of this section, which establishes the rate of convergence of the integrated square error.

Theorem 4.1. *Let β, C and x be positive numbers. Assume* (H1)–(H4) *and suppose that $X \leqslant x$ almost surely and $\|m\|_{\mathcal{W}(\beta)} \leqslant C$. Then, for all $M > 0$,*

$$\limsup_{n\to\infty} \mathbb{P}(n^{\beta/(1+2\beta)}\|\hat{m}_{x,h_n,n} - m\|_2 \geqslant M) \leqslant C' M^{-2}, \tag{22}$$

where $h_n \asymp n^{-1/(1+2\beta)}$ and C' is a positive constant depending only on K, λ, c, β, x and C.

The proof is postponed to Appendix B.

Remark 4.1. In the application we have considered, the condition $X \leqslant x$ is always satisfied. Indeed, the pulse duration corresponds to the duration of the charge collection, and therefore to the lifetime of the electron–hole pairs in the semiconductor detector. This lifetime is always finite and depends primarily on the geometry of the detector. However, the condition $X \leqslant x$ almost surely can actually be circumvented if, for fixed $x > 0$, we consider $\hat{m}_{x,h_n,n}$ as an estimator of m_x, defined as the density of the measure $\int_{u=0}^{x} P(\mathrm{d}u, \mathrm{d}y)$, which is always defined under assumption (H3).

Remark 4.2. If X and Y are independent, then $m_x(y) = m(y)\mathbb{P}(X \leqslant x)$ so that, for all x such that $\mathbb{P}(X \leqslant x) > 0$, we obtain an estimator of m up to a multiplicative constant.

Remark 4.3. In the $M/G/\infty$ case, that is, if $X = Y$ almost surely, $m_x = m\mathbf{1}_{[0,x]}$. Hence, since

$$\|(\hat{m}_{x,h_n,n} - m)\mathbf{1}_{[0,x]}\|_2^2 \leqslant \|(\hat{m}_{x,h_n,n} - m\mathbf{1}_{[0,x]})\|_2^2,$$

our results apply to the locally integrated error for estimating $m\mathbf{1}_{[0,x]}$. As a comparison, the rate of our estimator is given by the smoothness of $m\mathbf{1}_{[0,x]}$, whereas the rate of the estimator proposed in Hall and Park (2004) for estimating the time service density is given by the smoothness of the pdf of X' (see Hall and Park 2004: Eq. (3.8)).

Remark 4.4. The estimators in (22) are functions of $\{(Z_k, X'_k, Y'_k), k = 1, \ldots, n\}$, where n is the number of observed cycles. For $t \in \mathbb{R}_+$, denote by $\mathcal{N}_t$ the renewal process corresponding to the arrivals of the photons, $\mathcal{N}_t \stackrel{\text{def}}{=} \sum_{k-1}^{\infty} \mathbf{1}\{T_k \leqslant t\}$. The number of arrivals in n cycles is equal to $\tilde{n} = \mathcal{N}_{T'_n+X'_n}$ and is therefore random. As n tends to infinity, $(T'_n + X'_n)/n$ converges almost surely to the mean of the cycle duration, $\exp(\lambda\mathbb{E}[X])/\lambda$ and it can easily be shown that the nth return to an idle period (that is, $T'_n + X'_n$) is a stopping time with respect to the natural history of $\mathcal{N}_t$. Therefore, by the Blackwell theorem, $\mathcal{N}_{T'_n+X'_n}/(T'_n + X'_n)$ converges to λ. Therefore, $\tilde{n}/n = \mathcal{N}_{T'_n+X'_n}/n$ converges almost surely to $\exp(\lambda\mathbb{E}[X])$. It is well known that the minimax integrated rate for estimating m from $\{Y_k, k = 1, \ldots, \tilde{n}\}$ with m in a β-Sobolev ball is $\tilde{n}^{\beta/(1+2\beta)}$ (see Schipper 1996), the only non-standard feature being that the density estimator is calculated by using a random number of data, which does not alter the density's estimator first-order property. Since $\tilde{n}/n$ converges almost surely to a constant, Theorem 4.1 shows that the rate achieved by our estimator is the minimax integrated rate.

5. Decomposition of the error

In this section we give theoretical results for the proposed estimators. We first introduce auxiliary variables which will be used in the proof of the main theorem. For any positive numbers W, x and $\tilde{\lambda}$, define

$$\hat{\Delta}_n(W) \stackrel{\text{def}}{=} \sup_{(\omega,\nu)\in[-W,W]^2} |\mathcal{L}P'(c+\mathrm{i}\omega,\, \mathrm{i}\nu) - \widehat{\mathcal{L}P'}_n(c+\mathrm{i}\omega,\, \mathrm{i}\nu)|,$$

$$\hat{E}_n(W;\, x,\, \tilde{\lambda}) \stackrel{\text{def}}{=} \sup_{\nu\in[-W,W]} \left|\int \mathbf{1}_{[0,x]}(u)\mathrm{e}^{\tilde{\lambda}(u-x)}\mathrm{e}^{-\mathrm{i}\nu y}(P' - \widehat{P'}_n)(\mathrm{d}u,\, \mathrm{d}y)\right|.$$

Proposition 5.1 provides bounds for the random variables $\hat{\Delta}_n$ and $\hat{E}_n$.

Proposition 5.1. *Assume* (H1) *and* (H2). *Then* $M_1 \stackrel{\text{def}}{=} \mathbb{E}(\max\{X',\, Y'\})$ *is finite and the following inequalities hold for all* $\varepsilon > 0$, $r > 0$ *and* $W > 1$:

$$\mathbb{P}(|\hat{\Delta}_n(W)| \geqslant \varepsilon) \leqslant \frac{4rM_1}{\varepsilon} + \left(1 + \frac{W}{r}\right)^2 \exp\left(-\frac{n\varepsilon^2}{16}\right), \tag{23}$$

$$\sup_{x,\tilde{\lambda}>0} \mathbb{P}(|\hat{E}_n(W;\, x,\, \tilde{\lambda})| \geqslant \varepsilon) \leqslant \frac{4rM_1}{\varepsilon} + \left(1 + \frac{W}{r}\right) \exp\left(-\frac{n\varepsilon^2}{16}\right). \tag{24}$$

The proof of Proposition 5.1 is omitted here for brevity, but can be found in Trigano (2005: Proposition 3.4.2). Since our estimate depends on $\hat{\lambda}_n$ and $\widehat{\mathcal{L}P'}_n$, we introduce auxiliary functions to exhibit both dependencies. Define the following functions depending on h, x, $\tilde{\lambda}$ and on any probability measure $\tilde{P}$:

$$\tilde{a}(x,\, \mathrm{i}\nu;\, \tilde{\lambda},\, \tilde{P}) \stackrel{\text{def}}{=} 1 + \frac{\mathrm{e}^{(c+\tilde{\lambda})x}}{2\pi} \int_{-\infty}^{+\infty} \Phi(c+\mathrm{i}\omega,\, \mathrm{i}\nu;\, \tilde{\lambda},\, \mathcal{L}\tilde{P})\mathrm{e}^{\mathrm{i}\omega x}\, \mathrm{d}\omega, \tag{25}$$

$$\tilde{I}_1(x,\, \mathrm{i}\nu;\, \tilde{\lambda},\, \tilde{P}) \stackrel{\text{def}}{=} \int\int_{\mathbb{R}_+{}^2} \mathbf{1}_{\{u\leqslant x\}}\mathrm{e}^{\tilde{\lambda}u - \mathrm{i}\nu v}\tilde{P}(\mathrm{d}u,\, \mathrm{d}v)$$

$$\tilde{I}_2(x,\, \mathrm{i}\nu;\, \tilde{\lambda},\, \tilde{P}) \stackrel{\text{def}}{=} \frac{\mathrm{e}^{(\tilde{\lambda}+c)x}}{2\pi} \int_{-\infty}^{+\infty} \mathcal{L}\tilde{P}(c+\mathrm{i}\omega,\, \mathrm{i}\nu)\Phi(c+\mathrm{i}\omega,\, \mathrm{i}\nu;\, \tilde{\lambda},\, \mathcal{L}\tilde{P})\mathrm{e}^{\mathrm{i}\omega x}\, \mathrm{d}\omega.$$

Define also

$$\tilde{m}(y;\, x,\, h,\, \tilde{\lambda},\, \tilde{P}) \stackrel{\text{def}}{=} \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left[\frac{\tilde{I}_1 + \tilde{I}_2}{\tilde{a}}(x,\, \mathrm{i}\nu;\, \tilde{\lambda},\, \tilde{P})\right] K^*(h\nu)\mathrm{e}^{\mathrm{i}\nu y}\, \mathrm{d}\nu \tag{26}$$

whenever the integral is well defined. Hence, by (6), (12), (18), (17), (19) and (20), for $i = 1, 2$,

$$\tilde{a}(x,\, \mathrm{i}\nu;\, \lambda,\, P') = a(x,\, \mathrm{i}\nu) \quad \text{and} \quad \tilde{I}_i(x,\, \mathrm{i}\nu;\, \lambda,\, P') = \frac{\partial a_i}{\partial x}(x,\, \mathrm{i}\nu), \tag{27}$$

$$\tilde{a}(x,\, \mathrm{i}\nu;\, \hat{\lambda}_n,\, \widehat{P'}_n) = \hat{a}_n(x,\, \mathrm{i}\nu) \quad \text{and} \quad \tilde{I}_i(x,\, \mathrm{i}\nu;\, \hat{\lambda}_n,\, \widehat{P'}_n) = \hat{I}_{i,n}(x,\, \mathrm{i}\nu), \tag{28}$$

and $\hat{m}_{x,h,n}(y) = \tilde{m}(y;\, x,\, h,\, \hat{\lambda}_n,\, \widehat{P'}_n)$. Now define

$$b_1(y) \stackrel{\text{def}}{=} m(y) - \mathbb{E}\left[\frac{1}{h}K\left(\frac{y-Y}{h}\right)\right], \tag{29}$$

$$b_2(y) \stackrel{\text{def}}{=} \mathbb{E}\left[\frac{1}{h}K\left(\frac{y-Y}{h}\right)\right] - \tilde{m}(y;\, x,\, h,\, \lambda,\, P'), \tag{30}$$

$$V_1(y) \stackrel{\text{def}}{=} \tilde{m}(y;\, x,\, h,\, \lambda,\, P') - \tilde{m}(y;\, x,\, h,\, \hat{\lambda}_n,\, P'), \tag{31}$$

$$V_2(y) \stackrel{\text{def}}{=} \tilde{m}(y;\, x,\, h,\, \hat{\lambda}_n,\, P') - \hat{m}_{x,h,n}(y), \tag{32}$$

so that, by definition,

$$\hat{m}_{x,h,n} - m = b_1 + b_2 + V_1 + V_2. \tag{33}$$

In this decomposition, b_1 and b_2 are deterministic functions and V_1, V_2 are random processes. We now provide bounds for these quantities in the L^2 sense.

Theorem 5.1. *Let β, x and h be positive numbers and n be a positive integer. Assume* (H1)–(H4). *If $m \in \mathcal{W}(\beta)$, then we have*

$$\|b_1\|_2^2 \leqslant C_K^2 h^{2\beta} \|m\|_{\mathcal{W}(\beta)}^2, \tag{34}$$

$$\|b_2\|_2^2 \leqslant \|K\|_2^2 h^{-1} \mathbb{P}[X > x]. \tag{35}$$

Moreover, there exist positive constants M and η, depending only on c and λ, such that the following two assertions hold:

(i) *We have*

$$\|V_1\|_2^2 \leqslant M^2 \|K\|_2^2 (1+x)^2 h^{-1} \mathrm{e}^{4(c+2\lambda)x} (\hat{\lambda}_n - \lambda)^2 \tag{36}$$

on the event

$$E_1 \stackrel{\text{def}}{=} \{|\hat{\lambda}_n - \lambda| \leqslant \eta (1+x)^{-1} \mathrm{e}^{-(c+2\lambda)x}\}. \tag{37}$$

(ii) *For all $W \geqslant 1$ such that $[-Wh, Wh]$ contains the support of K^*, we have*

$$\|V_2\|_2^2 \leqslant M^2 \|K\|_2^2 h^{-1} \mathrm{e}^{4(c+2\lambda)x} [\hat{\Delta}_n(W) + W^{-1} + \hat{E}_n(W;\, x,\, \hat{\lambda}_n)]^2 \tag{38}$$

on the event E_1 intersected with the event

$$E_2 \stackrel{\text{def}}{=} \{\hat{\Delta}_n(W) + W^{-1} \leqslant \eta \mathrm{e}^{-(c+2\lambda)x}\}. \tag{39}$$

The proof is given in Appendix A.

In this result, b_1 is the usual bias in kernel nonparametric estimation; b_2 is a non-standard bias term which only vanishes when X is bounded and corresponds to the fact that the limit $x \to \infty$ is not attained in (9); the fluctuation term V_1 accounts for the error in the estimation of λ by $\hat{\lambda}_n$ and is of order $h^{-1}\sqrt{n}$ for fixed x; and V_2 accounts for the error in the estimation of $\mathcal{L}P'$ by $\widehat{\mathcal{L}P'}_n$ and, by using Proposition 5.1, can be shown to be 'almost' of order $h^{-1}\sqrt{n}$ for W chosen to diverge quickly enough with respect to n. The events E_1 and E_2 have probability tending to 1 as n tends to infinity; they are induced by the fraction present in the definition (21) of the estimator as they primarily avoid the denominator approaching zero.

We now give a result on the consistency of our estimator, and also on a rate of convergence, based on Theorem 5.1 and Proposition 5.1, by imposing a superexponential tail for X.

Corollary 5.1. *Let $\beta > 0$ and $\gamma > 1$. Assume* (H1)–(H4) *and suppose that $m \in \mathcal{W}(\beta)$ and $\mathbb{P}[X > x] = O(\mathrm{e}^{-|x|^\gamma})$. Then, for all $\epsilon > 0$, as $n \to +\infty$,*

$$\|m - \hat{m}_{x_n,h_n,n}\|_2^2 = O_{\mathbb{P}}(n^{\epsilon - 2\beta/(1+2\beta)}), \tag{40}$$

where $h_n \asymp n^{-1/(1+2\beta)}$ and $x_n \asymp (\log n)^{\gamma'}$ with $\gamma' \in (\gamma^{-1}, 1)$.

Proof. We set $W_n \stackrel{\text{def}}{=} n$. By Proposition 5.1, we obtain, by choosing $\epsilon = C(\log(n)/n)^{1/2}$ and $r = n^{-1/2}$,

$$\mathbb{P}(|\hat{\Delta}_n(W_n)| \geqslant C(\log(n)/n)^{1/2}) \leqslant \frac{4}{C}\log^{-1/2}(n) + (1+\sqrt{n})^2 n^{-C/16},$$

which tends to 0 as $n \to \infty$ for $C > 32$. Hence,

$$|\hat{\Delta}_n(W_n)| = O_{\mathbb{P}}\{(\log(n)/n)^{1/2}\}.$$

Similarly, since $\hat{\lambda}_n$ is independent of $\{(X'_k, Y'_k),\ k = 1, \ldots, n\}$, we obtain $|\hat{E}_n(W_n; x_n, \hat{\lambda}_n)| = O_{\mathbb{P}}\{\log(n)/n)^{1/2}\}$. Observe that, for any $\delta_1 \geqslant 0$, $\delta_2 > 0$ and $\epsilon > 0$, $x_n^{\delta_1}\exp(\delta_2 x_n) = o(n^\epsilon)$. Since $\hat{\lambda}_n = \lambda + O_{\mathbb{P}}(n^{-1/2})$ and $|\hat{\Delta}_n(W_n) + W_n^{-1}| = O_{\mathbb{P}}((\log(n)/n)^{1/2})$, E_1 and E_2 have a probability tending to one, so that the bounds of Theorem 5.1 finally give, for any $\epsilon > 0$,

$$\|V_i\| = O_{\mathbb{P}}((\mathrm{h}_n n)^{\epsilon - 1/2}), \qquad i = 1, 2.$$

Now using the superexponential tail assumption for X, we have $\mathbb{P}(X > x_n) = O(\exp\{-\log^{\gamma\gamma'}(n)\}) = o(n^\epsilon)$ for all $\epsilon > 0$, and the result follows. □

As seen from (40), the estimator almost achieves the standard nonparametric minimax rate $n^{-\beta/(1+2\beta)}$ that one would obtain by observing $\{(X_k, Y_k),\ k = 1, \ldots, n\}$ directly. If X is bounded, then the rate can be made more precise as in Theorem 4.1: by taking x equal to an upper bound for X (so that $b_2 = 0$) and $h_n \asymp n^{-1/(1+2\beta)}$, one easily obtains from the above proof that

$$\|m - \hat{m}_{x,h_n,n}\|_2^2 = O_{\mathbb{P}}(\log(n) n^{-2\beta/(1+2\beta)}), \tag{41}$$

thus a loss of $\log(n)$ in comparison with the claimed rate. This $\log(n)$ can in fact be removed, as shown in Appendix B.

6. Applications and discussion

The present paper is directed towards the construction of an estimator and deriving elements of its asymptotic theory. We will therefore content ourselves with providing simple examples and refer the reader to Trigano *et al.* (2006) for an in-depth discussion of the selection of the setting parameters (the kernel bandwidth, the truncation bound, etc.) and the analysis of various different data sets.

We first consider a simple simulated data set. Samples are drawn according to the bimodal density

$$f(x, y) = \mathcal{N}_{20,3}(x) \times (0.6\mathcal{N}_{100,6}(y) + 0.4\mathcal{N}_{130,9}(y)), \tag{42}$$

where $\mathcal{N}_{a,b}$ denotes the Gaussian distribution with mean a and standard deviation b truncated to $\mathbb{R}_+$. The intensity of the Poisson process is set to $\lambda = 0.04$. Some plots of the results can be found in Section D.6 of Trigano (2005). Numerical values of the mean integrated squared error (MISE) are presented in Table 1 for a fixed bandwidth parameter $h = 2.0$ and different values of n, c and x. It is evident that c has little influence on the error. This is hardly surprising, since the Bromwich integral used to compute the inverse Laplace transform does not theoretically depend on the choice of c (see Doetsch 1974). Concerning the influence of x, knowing that X has distribution $\mathcal{N}_{20,3}$, 'reasonable' values (displayed in the three first rows) all give equally good results but the last row shows that the 'naive' data-driven choice $x = \max_{i \leqslant n} X_i'$ significantly worsens the estimate. Indeed, in view of the upper bounds of Theorem 5.1, on the one hand, choosing x too large does not ensure that the variance terms V_1 and V_2 are controlled, since in this case conditions (37) and (39) may not be satisfied; on the other hand, x too small introduces a bias in (21), since the control of the bias term b_2 is not guaranteed in that case.

We now present some results using a more realistic model of the energy distribution of the caesium 137 radionuclide (including the Compton effect). We draw $n = 500\,000$ samples of (X, Y) using the adaptive rejection sampling algorithm, according to the density

$$f(x, y) = m(y) \times f_{X|Y}(x|y),$$

Table 1. Mean integrated square error Monte Carlo estimates as n, c or x varies

n	MISE	c	MISE	x	MISE
1 000	4.760×10^{-3}	0.01	4.002×10^{-4}	40	1.905×10^{-4}
5 000	1.089×10^{-3}	0.001	4.348×10^{-4}	60	3.852×10^{-4}
10 000	3.852×10^{-4}	0.0001	3.853×10^{-4}	80	5.100×10^{-4}
20 000	2.042×10^{-4}	0.00001	4.426×10^{-4}	$\max_{i \leqslant n} X_i'$	2.231×10^{-2}
$(c = 10^{-4}, x = 60)$		$(n = 10^4, x = 60)$		$(n = 10^4, c = 10^{-4})$	

where m is represented by the dashed plot of Figure 2(a) and the conditional distribution $f_{X|Y}(\cdot|y)$ is a gamma distribution with unit scale parameter, shape parameter equal to $2 + y/1024$ and truncated at $T = 4 + y/2048$; the number of samples may appear to be large, but such large numbers are commonly used in nuclear spectrometry, especially when active sources are measured. Figure 2(a) also shows the pile-up distribution (solid curve), based on the observations Y'_k, $k = 1, \ldots, n$, to illustrate the difference from m; note that the Compton continuum (which is the smooth part of the density on the left of the spike) is also distorted, since electrical pulses generated by Compton photons are also susceptible to overlap. Figure 2(b) illustrates the behaviour of our estimator. We observe that the pile-up effect is well corrected.

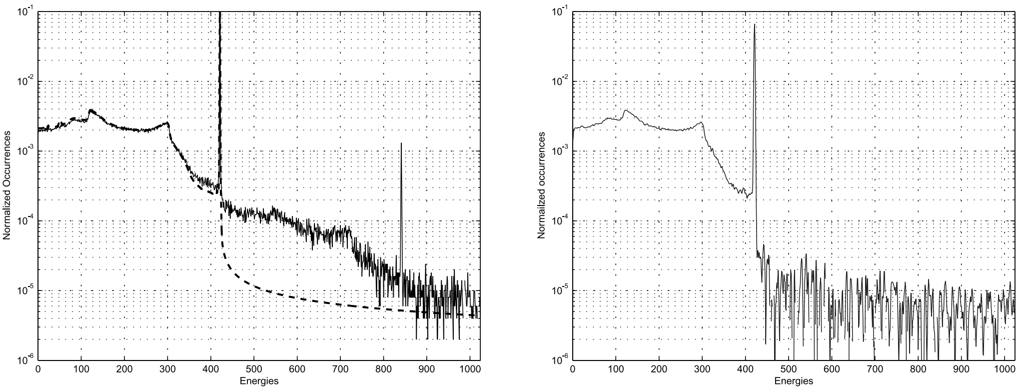
We now briefly discuss the choice of the bandwidth parameter h. In standard nonparametric estimation, there are several data-driven ways of choosing a bandwidth parameter. It is not yet clear how these methods can be adapted to this non-standard density estimation scenario, except in special cases. For instance, a possible approach would then involve using an automatic bandwidth selector (such as cross-validation) on the observations $\{Y'_k, k = 1, \ldots, n\}$, and using the optimal bandwidth obtained for the estimator $\hat{m}_{x,h,n}$. Further insights on the data-driven choices of c, x and h and discussion of the practical applications can be found in the companion paper, Trigano *et al.* (2006).

Appendix A. Proof of Theorem 5.1

The following lemma will be used repeatedly:

Lemma A.1. *Let* $c > 0$ *and* $\eta_0 > 0$. *For any complex-valued functions* z_1 *and* z_2 *satisfying*

$$\sup_{\omega \in \mathbb{R}, i=1,2} |z_i(\omega)| \leqslant 1, \tag{43}$$



(a) (b)

Figure 2. Energy spectrum of Cs 137: (a) ideal probability density function (dashed curve) and kernel estimates of pile-up distribution (solid curve); (b) estimate $\hat{m}_{x,h,n}$.

let $z = (z_1, z_2)$ *and denote by* Ψ_z *the function defined on* $\mathbb{R}_+ \times \mathbb{R}$ *by*

$$\Psi_z(\tilde{\lambda}, \omega) \stackrel{\text{def}}{=} \frac{z_1(\omega)}{(c + \mathrm{i}\omega + \tilde{\lambda})(c + \mathrm{i}\omega + \tilde{\lambda} - \tilde{\lambda} z_2(\omega))}.$$

Then the following assertions hold:

(i) *The function* $\tilde{\lambda} \mapsto \int_{-\infty}^{+\infty} \Psi_z(\tilde{\lambda}, \omega)\mathrm{d}\omega$ *is continuously differentiable on* $\mathbb{R}_+$ *and its derivative is bounded independently of* z *over* $\tilde{\lambda} \in [0, \eta_0]$.

(ii) *There exists* $K > 0$, *depending only on* c *and* η_0, *such that, for any* $W \geqslant 1$ *or* $W = \infty$ *and any function* $\tilde{z} = (\tilde{z}_1, \tilde{z}_2)$ *also satisfying* (43),

$$\sup_{\tilde{\lambda} \in [0, \eta_0]} \left| \int_{-\infty}^{+\infty} \Psi_z(\tilde{\lambda}, \omega)\mathrm{d}\omega - \int_{-\infty}^{+\infty} \Psi_{\tilde{z}}(\tilde{\lambda}, \omega)\mathrm{d}\omega \right| \leqslant K \left(\max_{i=1,2} \sup_{\omega \in [-W, W]} |z_i(\omega) - \tilde{z}_i(\omega)| + \frac{1}{W} \right).$$

The proof of Lemma A.1 can be found in Trigano (2005: Lemma 3.2.1).

Bound for b_1. Observe that b_1 is the usual bias in nonparametric kernel estimation. The bound of the integrated error is classically given, for densities in a Sobolev space, by

$$\|b_1\|_2^2 = \int_{-\infty}^{\infty} |1 - K^*(h\nu)|^2 |m^*(\nu)|^2 \,\mathrm{d}\nu \leqslant C_K^2 h^{2\beta} \|m\|_{\mathcal{W}(\beta)}^2,$$

which shows (34).

Bound for b_2. By (8), (16), (26) and (27), we find that

$$b_2(y) = \mathbb{E}\left[\frac{1}{h} K\left(\frac{y - Y}{h} \right) \mathbf{1}(X > x) \right].$$

An application of the Cauchy–Schwarz inequality yields (35).

Bound for V_1. We will show below that there exist positive constants M and η such that, on E_1 (as defined in (37)),

$$\sup_{\nu \in \mathbb{R}} \left| \frac{\partial}{\partial \tilde{\lambda}} \left[\frac{\tilde{I}_1 + \tilde{I}_2}{\tilde{a}} \right] (x, \mathrm{i}\nu; \hat{\lambda}_n, P') \right| \leqslant M(1 + x)\mathrm{e}^{(2c + 4\lambda)x}. \tag{44}$$

Using (26) and (31), the Parseval theorem and the latter relation imply

$$\|V_1\|_2^2 \leqslant \int_{-\infty}^{\infty} |M(1 + x)\mathrm{e}^{(2c+4\lambda)x}(\hat{\lambda}_n - \lambda)|^2 |K^*(h\nu)|^2 \mathrm{d}\nu$$

on the event E_1, which yields (36). Hence, it remains to show (44).

First observe that, by the definition of $\tilde{I}_1$, one trivially obtains, for all $\tilde{\lambda} > 0$,

$$|\tilde{I}_1(x, \mathrm{i}\nu; \tilde{\lambda}, P')| \leqslant \mathrm{e}^{\tilde{\lambda}x} \quad \text{and} \quad \left| \frac{\partial}{\partial \tilde{\lambda}} \tilde{I}_1(x, \mathrm{i}\nu; \tilde{\lambda}, P') \right| \leqslant x\mathrm{e}^{\tilde{\lambda}x}. \tag{45}$$

Substituting (6) into (27) gives, for all $\nu \in \mathbb{R}$,

$$|\tilde{a}(x, \mathrm{i}\nu; \lambda, P')| = \exp(\lambda \mathbb{E}[\cos(\nu Y)(x - X)_+]) \in [\mathrm{e}^{-\lambda x}, \mathrm{e}^{\lambda x}]. \tag{46}$$

Let $\eta_0 > 0$, to be chosen later. From (25), Lemma A.1 shows that there exists a constant M_1 depending only on λ, c and η_0 such that, for all $\tilde{\lambda}$ in $[0, \lambda + \eta_0]$ and $\nu \in \mathbb{R}$,

$$|(\tilde{a}(x, \mathrm{i}\nu; \tilde{\lambda}, P') - 1)\mathrm{e}^{-(\tilde{\lambda}+c)x} - (\tilde{a}(x, \mathrm{i}\nu; \lambda, P') - 1)\mathrm{e}^{-(\lambda+c)x}| \leqslant M_1|\tilde{\lambda} - \lambda|. \tag{47}$$

From (46) and (47) and since, for all real y, $|1 - \mathrm{e}^y| \leqslant |y|\mathrm{e}^{|y|}$, we obtain, for all $\tilde{\lambda}$ in $[0, \lambda + \eta_0]$ and $\nu \in \mathbb{R}$,

$$|\tilde{a}|(x, \mathrm{i}\nu; \tilde{\lambda}, P') \geqslant |\tilde{a}|(x, \mathrm{i}\nu; \lambda, P')\mathrm{e}^{(\tilde{\lambda}-\lambda)x} - |\mathrm{e}^{(\tilde{\lambda}-\lambda)x} - 1| - M_1|\tilde{\lambda} - \lambda|\mathrm{e}^{(\tilde{\lambda}+c)x},$$

hence

$$|\tilde{a}|(x, \mathrm{i}\nu; \tilde{\lambda}, P') \geqslant \mathrm{e}^{(\tilde{\lambda}-2\lambda)x} - [M_1\mathrm{e}^{(\tilde{\lambda}+c)x} + x\mathrm{e}^{|\tilde{\lambda}-\lambda|x}]|\tilde{\lambda} - \lambda|.$$

Note that, taking $\eta_0 = c$ and $M_1' = M_1 \vee 1$, the term in brackets is at most $M_1'\mathrm{e}^{(\tilde{\lambda}-2\lambda)x}(1 + x)\mathrm{e}^{(c+2\lambda)x}$ for $\tilde{\lambda} \in [\lambda - \eta_0, \lambda + \eta_0]$ so that we obtain, on the event E_1 with $\eta \leqslant \eta_1 \stackrel{\text{def}}{=} \{\eta_0 \wedge (M_1')^{-1}/2\}$,

$$|\tilde{a}|(x, \mathrm{i}\nu; \hat{\lambda}_n, P') \geqslant \frac{1}{2}\mathrm{e}^{(\hat{\lambda}_n-2\lambda)x}. \tag{48}$$

From (25) and using similar bounds to those in Lemma A.1, one can easily show that, for some constant M_2 depending only on λ, c and η_1, for all $\tilde{\lambda} \in \mathbb{R}$ such that $|\lambda - \tilde{\lambda}| \leqslant \eta_1$,

$$|\tilde{a}(x, \mathrm{i}\nu; \tilde{\lambda}, P')| \leqslant M_2\mathrm{e}^{(\tilde{\lambda}+c)x} \quad \text{and} \quad \left|\frac{\partial\tilde{a}}{\partial\tilde{\lambda}}(x, \mathrm{i}\nu; \tilde{\lambda}, P')\right| \leqslant M_2(1 + x)\mathrm{e}^{(\tilde{\lambda}+c)x}. \tag{49}$$

$$|\tilde{I}_2(x, \mathrm{i}\nu; \tilde{\lambda}, P')| \leqslant M_2\mathrm{e}^{(\tilde{\lambda}+c)x} \quad \text{and} \quad \left|\frac{\partial\tilde{I}_2}{\partial\tilde{\lambda}}(x, \mathrm{i}\nu; \tilde{\lambda}, P')\right| \leqslant M_2(1 + x)\mathrm{e}^{(\tilde{\lambda}+c)x}. \tag{50}$$

Bringing together (45) and (48)–(50) shows that (44) holds on E_1, for any $\eta \leqslant \eta_1$.

Bound for V_2. Since the support of K^* is included in $[-Wh, Wh]$, by Parseval's theorem, (26) and (32), the claimed bound is implied by

$$\sup_{|\nu|\leqslant W}\left|\frac{\tilde{I}_1 + \tilde{I}_2}{\tilde{a}}(x, \mathrm{i}\nu; \hat{\lambda}_n, P') - \frac{\tilde{I}_1 + \tilde{I}_2}{\tilde{a}}(x, \mathrm{i}\nu; \hat{\lambda}_n, \widehat{P'}_n)\right|$$

$$\leqslant M\mathrm{e}^{2(c+2\lambda)x}[\hat{\Delta}_n(W) + W^{-1} + \hat{E}_n(W; x, \hat{\lambda}_n)], \tag{51}$$

which we now show. Using (25), we may write

$$|\tilde{a}(x, \mathrm{i}\nu; \tilde{\lambda}, \tilde{P}) - \tilde{a}(x, \mathrm{i}\nu; \tilde{\lambda}, P')| = \left|\frac{\tilde{\lambda}\mathrm{e}^{(c+\tilde{\lambda})x}}{2\pi}\int_{-\infty}^{+\infty}[\Psi_{\tilde{z}}(\tilde{\lambda}, \omega) - \Psi_z(\tilde{\lambda}, \omega)]\mathrm{d}\omega\right|, \tag{52}$$

where Ψ is defined in Lemma A.1, and where the complex functions $\tilde{z}$ and z are defined as

$$z(\omega) \stackrel{\text{def}}{=} (\mathrm{e}^{\mathrm{i}\omega x}\mathcal{L}P'(c + \mathrm{i}\omega, \mathrm{i}\nu); \mathcal{L}P'(c + \mathrm{i}\omega, \mathrm{i}\nu))$$

and

$$\tilde{z}(\omega) \stackrel{\text{def}}{=} (\mathrm{e}^{\mathrm{i}\omega x}\mathcal{L}\tilde{P}(c + \mathrm{i}\omega, \mathrm{i}\nu); \mathcal{L}\tilde{P}(c + \mathrm{i}\omega, \mathrm{i}\nu)).$$

Using (52) and assertion (ii) of Lemma A.1, there exists $M_1 > 0$ such that, for all $\tilde{\lambda} \leqslant \lambda + \eta_1$,

$$\sup_{|\nu| \leqslant W} |\tilde{a}(x, \mathrm{i}\nu; \tilde{\lambda}, P') - \tilde{a}(x, \mathrm{i}\nu; \tilde{\lambda}, \widehat{P'}_n)| \leqslant M_1 \mathrm{e}^{(c+\tilde{\lambda})x} \left(\hat{\Delta}_n(W) + \frac{1}{W} \right). \tag{53}$$

It is also clear that for all $\tilde{\lambda} \leqslant \lambda + \eta_1$,

$$\sup_{|\nu| \leqslant W} |\tilde{I}_1(x, \mathrm{i}\nu; \tilde{\lambda}, \widehat{P'}_n) - \tilde{I}_1(x, \mathrm{i}\nu; \tilde{\lambda}, P')| \leqslant \mathrm{e}^{\tilde{\lambda}x} \hat{E}_n(W; \tilde{\lambda}, x) \tag{54}$$

and

$$|\tilde{I}_2(x, \mathrm{i}\nu; \tilde{\lambda}, \tilde{P}) - \tilde{I}_2(x, \mathrm{i}\nu; \tilde{\lambda}, P')| = \left| \frac{\mathrm{e}^{(c+\tilde{\lambda})x}}{2\pi} \int_{-\infty}^{+\infty} [\Psi_{\tilde{z}}(\tilde{\lambda}, \omega) - \Psi_z(\tilde{\lambda}, \omega)] \mathrm{d}\omega \right|,$$

with

$$z(\omega) \stackrel{\text{def}}{=} (\mathrm{e}^{\mathrm{i}\omega x} (\mathcal{L}P'(c + \mathrm{i}\omega, \mathrm{i}\nu))^2; \mathcal{L}P'(c + \mathrm{i}\omega, \mathrm{i}\nu))$$

and

$$\tilde{z}(\omega) \stackrel{\text{def}}{=} (\mathrm{e}^{\mathrm{i}\omega x} (\mathcal{L}\tilde{P}(c + \mathrm{i}\omega, \mathrm{i}\nu))^2; \mathcal{L}\tilde{P}(c + \mathrm{i}\omega, \mathrm{i}\nu)).$$

Consequently, using assertion (ii) of Lemma A.1, we have for all $\tilde{\lambda} \leqslant \lambda + \eta_1$,

$$\sup_{|\nu| \leqslant W} |\tilde{I}_2(x, \mathrm{i}\nu; \tilde{\lambda}, \widehat{P'}_n) - \tilde{I}_2(x, \mathrm{i}\nu; \tilde{\lambda}, P')| \leqslant M_2 \mathrm{e}^{(\tilde{\lambda}+c)x} \left(\hat{\Delta}_n(W) + \frac{1}{W} \right). \tag{55}$$

We now derive a lower bound for $\hat{a}_n(x, \mathrm{i}\nu) = a(x, \mathrm{i}\nu; \hat{\lambda}_n, \widehat{P'}_n)$. By (53), we obtain

$$\inf_{|\nu| \leqslant W} |\hat{a}|(x, \mathrm{i}\nu) \geqslant \inf_{|\nu| \leqslant W} |\tilde{a}|(x, \mathrm{i}\nu; \hat{\lambda}_n, P') - M_1 \mathrm{e}^{(c+\hat{\lambda}_n)x} \left(\hat{\Delta}_n(W) + \frac{1}{W} \right).$$

Recall that E_1 and E_2 are defined in (37) and (39), respectively. Using (48), which holds on E_1 for any $\eta \leqslant \eta_1$, we obtain, on $E_1 \cap E_2$,

$$\inf_{|\nu| \leqslant W} |\tilde{a}(x, \mathrm{i}\nu; \hat{\lambda}_n, \widehat{P'}_n) \hat{a}_n(x, \mathrm{i}\nu)| \geqslant \frac{1}{2} \left[\frac{1}{2} - M_1 \eta \right] \mathrm{e}^{(2\hat{\lambda}_n - 4\lambda)x}. \tag{56}$$

Hence we set $\eta \stackrel{\text{def}}{=} (4M_1)^{-1} \wedge \eta_1$, so that the term in brackets is at least $1/4$. Finally, using the fact that, for all complex numbers x, y, z, x', y', z',

$$\frac{x + y}{z} - \frac{x' + y'}{z'} = \frac{(z' - z)(x + y) + z(x - x') + z(y - y')}{zz'},$$

and bringing together (45), (49), (50) and (53)–(56) leads to (51).

Appendix B. Proof of Theorem 4.1

In this section we denote by η_i, M_i and C_i, $i = 0, 1, 2, \ldots$, some positive constants depending only on $\|m\|_{\mathcal{W}(\beta)}$, K, λ, c and x. We will also use the notation introduced in Appendix A. As shown in this section, we have $\|b_1\|_2^2 \leqslant C_K^2 \|m\|_{\mathcal{W}(\beta)}^2 h^{2\beta}$ and, since $\mathbb{P}(X > x) = 0$, we have $b_2 = 0$.

By (6), because $\mathbb{E}[X] < \infty$, it is easily seen that $|a(x, \mathrm{i}\nu)| \geqslant \mathrm{e}^{-\lambda x}$. This can be used in the ratio appearing in (21) to lower-bound $\hat{a}_n$ for large n by using the fact that $\hat{a}_n(x, \mathrm{i}\nu)$ converges to $a(x, \mathrm{i}\nu)$. However, this will not allow bounds of the ratio in the mean square sense. For obtaining mean square error estimates, we consider the following modified estimator which (artificially) circumvent this difficulty. Let $\eta_0 > \lambda$ and denote by A_n the set

$$A_n \stackrel{\text{def}}{=} \{\hat{\lambda}_n \leqslant \eta_0\} \cap \left\{ \inf_{h_n\nu \in \operatorname{Supp}(K^*)} |\hat{a}_n|(x, \mathrm{i}\nu) \geqslant \frac{1}{5} \exp(-\hat{\lambda}_n x) \right\},$$

where $\operatorname{Supp}(K^*)$ denotes the (compact) support of K^*. Define

$$\check{m}_{x,h,n}(y) = \mathbf{1}_{A_n} \hat{m}_{x,h,n}(y). \tag{57}$$

We will show that

$$\sup_{n \geqslant 1} n^{2\beta/(1+2\beta)} \mathbb{E}\|\check{m}_{x,h_n,n} - m\|_2^2 \leqslant C_0, \tag{58}$$

Let V_3 be the random process

$$V_3(y) \stackrel{\text{def}}{=} \tilde{m}(y;\, x,\, h_n,\, \lambda,\, P') - \hat{m}_{x,h_n,n}(y),$$

so that

$$\|\check{m}_{x,h,n} - m\|_2^2 \leqslant \|b_1\|_2^2 + \|m\|_2^2 \mathbf{1}_{A_n^c} + \|V_3\|_2^2 \mathbf{1}_{A_n}. \tag{59}$$

We will show that there exists $C_1 > 0$ such that, for n large enough,

$$\mathbb{P}(A_n^c) \leqslant C_1 n^{-1}, \tag{60}$$

$$\mathbb{E}[\|V_3\|_2^2 \mathbf{1}_{A_n}] \leqslant C_1 (h_n n)^{-1}. \tag{61}$$

Since $\|b_1\|_2^2 \leqslant C_K^2 \|m\|_{\mathcal{W}(\beta)}^2 h^{2\beta}$ and $\|m\|_2 \leqslant \|m\|_{\mathcal{W}(\beta)}$, (59)–(61) yield the bound (58). The bound (22) then follows by writing

$$\begin{aligned}
&\mathbb{P}(n^{\beta/(1+2\beta)} \|\hat{m}_{x,h_n,n} - m\|_2 \geqslant M) \\
&\quad \leqslant \mathbb{P}(n^{\beta/(1+2\beta)} \|\check{m}_{x,h_n,n} - m\|_2 \geqslant M) + \mathbb{P}(A_n^c) \leqslant C_0 M^{-2} + C_1 n^{-1},
\end{aligned}$$

where we have used the Markov inequality, (58) and (60). It now remains to show (60) and (61).

Proof of bound (60). We set $W_n \stackrel{\text{def}}{=} n$, so that for n large enough, $h_n\nu \in \operatorname{Supp}(K^*)$ implies $|\nu| \leqslant W_n$. As in (56), we have, on $E_1 \cap E_2$,

$$\inf_{|\nu|\leqslant W_n} |\hat{a}_n|(x, \mathrm{i}\nu) \geqslant \frac{1}{4}\, \mathrm{e}^{(\hat{\lambda}_n - 2\lambda)x}.$$

Hence the intersection of $\{\hat{\lambda}_n \leqslant \eta_0\}$, E_1, E_2 and $\{\exp((\hat{\lambda}_n - 2\lambda)x)/4 \geqslant \exp(-\hat{\lambda}_n x)/5\}$ is included in A_n. Since the last inequality and E_2 both contain $|\hat{\lambda}_n - \lambda| \leqslant \eta_2$ for $\eta_2 > 0$ small enough, we obtain

$$\mathbb{P}(A_n^c) \leqslant \mathbb{P}(\hat{\lambda}_n > \eta_0) + \mathbb{P}(|\hat{\lambda}_n - \lambda| > \eta_2) + \mathbb{P}(\hat{\Delta}_n(n) + n^{-1} > \eta_3).$$

Clearly the first two probabilities on the right-hand side are $O(n^{-1})$. For n large enough, the last probability is less than $\mathbb{P}(\hat{\Delta}_n(n) > \eta_3/2)$, which is $o(n^{-1})$ by applying Proposition 5.1, say with $r = n^{-2}$. We thus get (60) for n large enough.

Proof of bound (61). By (26)–(28), V_3 is defined as the inverse Fourier transform of

$$V_3^*(\nu) = K^*(h_n\nu)\left[\frac{\partial_x a_1 + \partial_x a_2}{a} - \frac{\widehat{I_{1,n}} + \widehat{I_{2,n}}}{\hat{a}_n}\right](x, \mathrm{i}\nu),$$

where $\partial_x a_i$ is shorthand for $\partial a_i/\partial x$. Using the fact that

$$\left|\frac{\partial_x a_1 + \partial_x a_2}{a} - \frac{\widehat{I_{1,n}} + \widehat{I_{1,n}}}{\hat{a}_n}\right| \leqslant \frac{1}{|\hat{a}_n|}\left[\sum_{i=1,2}\left|\partial_x a_i - \widehat{I_{i,n}}\right| + \left|\frac{\partial_x a_1 + \partial_x a_2}{a}\right||\hat{a}_n - a|\right],$$

we obtain that, on the set A_n defined above, for all $\nu \in \mathbb{R}$,

$$|V_3^*(\nu)| \leqslant 5|K^*(h_n\nu)|\left[\mathcal{E}_{1,n} + \mathcal{E}_{2,n} + \mathcal{E}_n\left|\frac{\partial_x a_1 + \partial_x a_2}{a}\right|(x, \mathrm{i}\nu)\right].$$

where, for $i = 1, 2$, we define

$$\mathcal{E}_{i,n} \stackrel{\text{def}}{=} \mathrm{e}^{\hat{\lambda}_n x}|\partial_x a_i - \widehat{I_{i,n}}|(x, \mathrm{i}\nu) \quad \text{and} \quad \mathcal{E}_n \stackrel{\text{def}}{=} \mathrm{e}^{\hat{\lambda}_n x}|\hat{a}_n - a|(x, \mathrm{i}\nu).$$

Multiplying by $\mathbf{1}_{A_n}$, taking the expectation and applying Parseval's theorem yields

$$\mathbb{E}[\|V_3\|_2^2\mathbf{1}_{A_n}] \leqslant C_2\left[h_n^{-1}\|K\|_2^2\sum_{i=1}^{2}\sup_{\nu\in\mathbb{R}}\mathbb{E}[\mathbf{1}_{A_n}\mathcal{E}_{i,n}{}^2] + M_1^2\sup_{\nu\in\mathbb{R}}\mathbb{E}[\mathbf{1}_{A_n}\mathcal{E}_n]\right], \tag{62}$$

where, by (8), (16) and Parseval's theorem,

$$M_1^2 \stackrel{\text{def}}{=} \int_{-\infty}^{\infty}|K^*(h_n\nu)|^2\left|\frac{\partial_x a_1 + \partial_x a_2}{a}\right|^2(x, \mathrm{i}\nu)\mathrm{d}\nu \leqslant \|K^*\|_\infty\|m\|_2^2.$$

By (17) and (19), we have

$$|\partial_x a_1 - \widehat{I_{1,n}}|(x, \mathrm{i}\nu) \leqslant |\partial_x a_1(x, \mathrm{i}\nu) - \tilde{I}_1(x, \mathrm{i}\nu; \hat{\lambda}_n, P')| + |\tilde{I}_1(x, \mathrm{i}\nu; \hat{\lambda}_n, P') - \tilde{I}_1(x, \mathrm{i}\nu; \hat{\lambda}_n, \widehat{P'}_n)|.$$

Using this decomposition in $\mathcal{E}_{1,n}$, the independence of $\hat{\lambda}_n$, (X'_k, Y'_k), $k = 1, \ldots, n$, the fact that $\operatorname{var}(\mathbf{1}(X' \leqslant x)\mathrm{e}^{\tilde{\lambda}X' - \mathrm{i}\nu Y'}) \leqslant 2\mathrm{e}^{2\tilde{\lambda}x}$ and the bound $\hat{\lambda}_n \leqslant \eta_0$ on A_n, we obtain

$$\mathbb{E}[\mathbf{1}_{A_n}\mathcal{E}_{1,n}{}^2] \leqslant M_2\{\mathbb{E}[\mathbf{1}_{A_n}|\partial_x a_1(x, \mathrm{i}\nu) - \tilde{I}_1(x, \mathrm{i}\nu; \hat{\lambda}_n, P')|^2] + n^{-1}\}.$$

Using (45) and the mean value theorem for bounding the first expectation shows that the first term is $O(1/n)$ and thus

$$\sup_{\nu\in\mathbb{R}} \mathbb{E}[\mathcal{E}_{1,n}^2] \leqslant C_3 n^{-1}. \tag{63}$$

By (18) and (20), we have

$$|\partial_x a_2 - \widehat{I_{2,n}}|(x, \mathrm{i}\nu) \leqslant |\partial_x a_2(x, \mathrm{i}\nu) - \tilde{I}_2(x, \mathrm{i}\nu; \hat{\lambda}_n, P')| + |\tilde{I}_2(x, \mathrm{i}\nu; \hat{\lambda}_n, P') - \tilde{I}_2(x, \mathrm{i}\nu; \hat{\lambda}_n, \widehat{P'}_n)|.$$

Using (27) and (50), by the mean value theorem, we obtain

$$\sup_{\nu\in\mathbb{R}} \mathbb{E}[\mathbf{1}_{A_n} \mathrm{e}^{2\hat{\lambda}_n x} |\partial_x a_2(x, \mathrm{i}\nu) - \tilde{I}_2(x, \mathrm{i}\nu; \hat{\lambda}_n, P')|^2] \leqslant M_3 n^{-1}.$$

Hence, for all $\nu \in \mathbb{R}$, on the set $\{\hat{\lambda}_n \leqslant \eta_0\}$,

$$|\tilde{I}_2(x, \mathrm{i}\nu; \hat{\lambda}_n, P') - \tilde{I}_2(x, \mathrm{i}\nu; \hat{\lambda}_n, \widehat{P'}_n)| \leqslant M_4 \int_{-\infty}^{\infty} g(\omega) |\mathcal{L}P' - \widehat{\mathcal{L}P'}_n|(c + \mathrm{i}\omega, \mathrm{i}\nu)\mathrm{d}\omega,$$

where g is an integrable function depending only on c and λ. Inserting the three last bounds in the definition of $\mathcal{E}_{2,n}$, we obtain

$$\sup_{\nu\in\mathbb{R}} \mathbb{E}[\mathbf{1}_{A_n} \mathcal{E}_{2,n}^2] \leqslant C_4 n^{-1}.$$

Comparing (10) with (18) and (12) with (20), one can easily see that a similar argument applies for bounding $\mathcal{E}_n$ on the set A_n, giving

$$\sup_{\nu\in\mathbb{R}} \mathbb{E}[\mathbf{1}_{A_n} \mathcal{E}_n^2] \leqslant C_5 n^{-1}.$$

Inserting (63) and the two last displays into (62) shows (61).

Appendix C. Proof of Theorem 3.1

Denote by $\overline{Y}_x$ the integrated workload at time x, that is,

$$\overline{Y}_x \stackrel{\text{def}}{=} \int_0^x W(t)\,\mathrm{d}t, \tag{64}$$

where $\{W(x), x \geqslant 0\}$ is the workload process given in (1). Recall that $\{S_x, x \geqslant 0\}$ denotes the on–off process equal to 0 in idle periods and equal to 1 in busy periods (see (2)). Define by $\rho(x, y)$ the probability

$$\rho\{x, y\} = \mathbb{P}(S_x = 0, \overline{Y}_x \leqslant y). \tag{65}$$

In a first step, we calculate the Laplace transform $\mathcal{L}\rho$ of ρ using the renewal process of the idle and busy periods. Note that this renewal process is stationary. Define by $\{R_n, n \geqslant 1\}$ the successive time instants of the end of the busy periods and by $\{A_n, n \geqslant 1\}$ the integrated workload at the end of the busy periods:

$$R_n \stackrel{\text{def}}{=} T'_n + X'_n = \sum_{k=1}^{n} (Z_k + X'_k), \quad A_n \stackrel{\text{def}}{=} \sum_{k=1}^{n} Y'_k, \qquad n \geqslant 1, \tag{66}$$

where we have set $R_0 \stackrel{\text{def}}{=} 0$ and $A_0 \stackrel{\text{def}}{=} 0$.

Proposition C.1. *Under assumptions* (H1) *and* (H2), *for any* $(s, p) \in \mathbb{C}^2$ *such that* $\mathrm{Re}(s) > 0$ *and* $\mathrm{Re}(p) > 0$,

$$\mathcal{L}\rho(s/p) = \frac{1}{s + \lambda - \lambda \mathcal{L}P'(s, p)} \times \frac{\lambda \mathcal{L}P'(s, p)}{p(s+\lambda)} + \frac{1}{p(s+\lambda)}.$$

Proof. The proof is based on classical renewal arguments and the fact that for all integers k, the idle period Z_k is distributed according to an exponential distribution with scale parameter λ, $\mathcal{E}_\lambda$. Note that the event $\{S_x = 0, \overline{Y}_x \leqslant y\}$ may be decomposed as

$$\begin{aligned}\{S_x = 0, \overline{Y}_x \leqslant y\} &= \{x < T'_1\} \cup \left(\bigcup_{n \geqslant 1} \left\{ T'_n + X'_n \leqslant x < T'_{n+1}, \sum_{k=1}^{n} Y'_k \leqslant y \right\} \right) \\ &= \{x < T'_1\} \cup \left(\bigcup_{n \geqslant 1} \{R_n \leqslant x < R_n + Z_{n+1}, A_n \leqslant y\} \right),\end{aligned} \tag{67}$$

where A_n and R_n are defined in (66). Since Z_{n+1} is independent of these variables, we obtain

$$\rho\{x, y\} - \mathrm{e}^{-\lambda x} = \sum_{n \geqslant 1} \int_0^{+\infty} \mathbb{P}(x - u < R_n \leqslant x, A_n \leqslant y) \lambda \mathrm{e}^{-\lambda u} \, \mathrm{d}u.$$

Writing

$$\begin{aligned}&\int_0^{+\infty} \mathbb{P}(x - u < R_n \leqslant x, A_n \leqslant y) \lambda \mathrm{e}^{-\lambda u} \, \mathrm{d}u \\ &\quad = \mathbb{P}(R_n \leqslant x, A_n \leqslant y) - \lambda \int_0^{+\infty} \mathbb{P}(R_n \leqslant u - x, A_n \leqslant y) \mathrm{e}^{-\lambda u} \, \mathrm{d}u,\end{aligned}$$

the proof follows from the identity

$$\int_0^{\infty} \int_0^{\infty} \mathbb{P}(R_n \leqslant x, A_n \leqslant y) \mathrm{e}^{-sx} \mathrm{e}^{-py} \, \mathrm{d}x \, \mathrm{d}y = \frac{1}{sp} \left(\frac{\lambda}{s + \lambda} \mathcal{L}P'(s, p) \right)^n.$$

□

We will now derive another expression for $\mathcal{L}\rho$, using standard properties of the Poisson process.

Proposition C.2. *Under assumptions* (H1) *and* (H2), *for any* $(s, p) \in \mathbb{C}^2$ *such that* $\mathrm{Re}(s) > 0$ *and* $\mathrm{Re}(p) > 0$,

$$\mathcal{L}\rho(s/p) = \frac{1}{p(s+\lambda)} + \frac{1}{p}\int_0^{+\infty} \mathrm{e}^{-(s+\lambda)x}\left[\exp\left(\lambda\int_0^{\infty} \mathrm{e}^{-pv}\kappa(x,\, \mathrm{d}v)\right) - 1\right]\mathrm{d}x.$$

Proof. Denote by $\{\mathcal{N}_t,\, t \geqslant 0\}$ the counting process associated with the homogeneous Poisson process $\{T_k,\, k \geqslant 0\}$ of the arrivals – more explicitly, $\mathcal{N}_t = \sum_{n=1}^{\infty} \mathbf{1}\{\mathrm{T}_n \leqslant t\}$. By conditioning the event $\{S_x = 0,\, \overline{Y}_x \leqslant y\}$ on the event $\{\mathcal{N}_x = n\}$,

$$\rho(x,\, y) = \mathrm{e}^{-\lambda x} + \sum_{n\geqslant 1} \mathbb{P}(\mathcal{N}_x = n)\mathbb{P}\left(\{T_i + X_i \leqslant x\}_{i=1}^{n},\, \sum_{k=1}^{n} Y_k \leqslant y \middle| \mathcal{N}_x = n\right). \tag{68}$$

The conditional distribution of the arrival times $(T_1,\, \ldots,\, T_n)$ given $\{\mathcal{N}_x = n\}$ is equal to the distribution of the order statistics of n independent and identically distributed uniform random variables on $[0,\, x]$; hence, for any n-tuple $(x_1,\, \ldots,\, x_n)$ of positive real numbers,

$$\mathbb{P}(T_1 \leqslant x_1,\, \ldots,\, T_n \leqslant x_n | \mathcal{N}_x = n) = \mathbb{P}(U_{(1)} \leqslant x_1,\, \ldots,\, U_{(n)} \leqslant x_n), \tag{69}$$

where $\{U_k\}_{k=1}^{n}$ are independent and identically distributed random variables uniformly distributed on $[0,\, x]$ and $U_{(1)} \leqslant \ldots \leqslant U_{(n)}$ are the order statistics. Therefore, (68) and (69) imply that

$$\begin{aligned} A &\stackrel{\text{def}}{=} \mathbb{P}\left(\{T_i + X_i \leqslant x\}_{i=1}^{n},\, \sum_{k=1}^{n} Y_k \leqslant y \middle| \mathcal{N}_x = n\right) \\ &= \frac{1}{x^n}\int \cdots \int \prod_{k=1}^{n} \mathbf{1}(u_k + x_k \leqslant x\}\mathbf{1}\left\{\sum_{k=1}^{n} y_k \leqslant y\right\} \prod_{k=1}^{n} P(\mathrm{d}x_k,\, \mathrm{d}y_k)\mathrm{d}u_k, \end{aligned}$$

since the latter integral is invariant by permuting the indices. An application of the Fubini theorem leads to

$$A = \frac{1}{x^n}\int \cdots \int \mathbf{1}\left\{\sum_{k=1}^{n} y_k \leqslant y\right\} \prod_{k=1}^{n} \kappa(x,\, \mathrm{d}y_k),$$

where $\kappa(x,\, y)$ is the probability kernel defined by

$$\kappa(x,\, y) \stackrel{\text{def}}{=} \int (x-u)\mathbf{1}\{u \leqslant x\}P(\mathrm{d}u,\, \mathrm{d}y). \tag{70}$$

We obtain, for any p such that $\mathrm{Re}(p) > 0$,

$$\begin{aligned} \int_0^{\infty} \rho(x,\, v)\mathrm{e}^{-pv}\,\mathrm{d}v &= \frac{\mathrm{e}^{-\lambda x}}{p} + \frac{1}{p}\sum_{n\geqslant 1} \frac{\lambda^n}{n!}\,\mathrm{e}^{-\lambda x}\left[\int_0^{\infty} \kappa(x,\, \mathrm{d}v)\mathrm{e}^{-pv}\right]^n \\ &= \frac{\mathrm{e}^{-\lambda x}}{p} + \frac{\mathrm{e}^{-\lambda x}}{p}\left[\exp\left(\lambda\int_0^{\infty} \kappa(x,\, \mathrm{d}v)\mathrm{e}^{-pv}\right) - 1\right], \end{aligned}$$

and hence

$$\mathcal{L}\rho(s,\ p) = \frac{1}{p(s+\lambda)} + \frac{1}{p}\int_0^{+\infty} \mathrm{e}^{-su}\mathrm{e}^{-\lambda u}\left[\exp\left(\lambda\int_0^{\infty} \mathrm{e}^{-pv}\kappa(u,\ \mathrm{d}v)\right) - 1\right]\mathrm{d}u.$$

□

The proof of Theorem 3.1 is then a direct consequence of Propositions C.1 and C.2 and the fact that

$$a(x,\ p) = \exp\left(\lambda\int_0^{\infty} \mathrm{e}^{-pv}\kappa(x,\ v)\right).$$

The result is extrapolated on the line $\mathrm{Re}(p) = 0$ by continuity in p at fixed s such that $\mathrm{Re}(s) > 0$.

References

American National Standards Institute (1999) *American National Standard for Calibration and Use of Germanium Spectrometers for the Measurement of Gamma-Ray Emission Rates of Radionuclides*, ANSI N42.14-1999. New York: American National Standards Institute.

Baccelli, F. and Brémaud, P. (2003) *Elements of Queueing Theory*. 2nd edn. Berlin: Springer-Verlag.

Bingham, N.H. and Pitts, S.M. (1999) Non-parametric estimation for the $M/G/\infty$ queue. *Ann. Inst. Statist. Math.*, **51**, 71–97.

Bristow, Q. (1990) Theoretical and experimental investigations of coincidences in Poisson distributed pulse trains and spectral distortion caused by pulse pileup. Ph.D. thesis, Carleton University, Ottawa, Canada.

Doetsch, G. (1974) *Introduction to the Theory and the Application of the Laplace Transform*. Berlin: Springer-Verlag.

Gautschi, W. (1996) *Numerical Analysis: An Introduction*. Boston: Birkhaüser.

Hall, P. (1988) *Introduction to the Theory of Coverage Processes*. New York: Wiley.

Hall, P. and Park, J. (2004) Nonparametric inference about service time distribution from indirect measurements. *J. Roy. Statist. Soc. Ser. B*, **66**, 861–875.

Knoll, G. (1989) *Radiation Detection and Measurement*, 2nd edn. New York: Wiley.

Leo, W.R. (1994) *Techniques for Nuclear and Particle Physics Experiments: A How-To Approach*, 2nd rev. edn. Berlin: Springer-Verlag.

Pyke, R. (1958) On renewal processes related to type I and type II counter models. *Ann. Math. Statist.*, **29**, 737–754.

Schipper, M. (1996) Optimal rates and constants in L_2-minimax estimation of probability density functions. *Math. Methods Statist.*, **5**, 253–274.

Takács, L. (1962) *Introduction to the Theory of Queues*. New York: Oxford University Press.

Trigano, T. (2005) *Statistical signal processing in nuclear spectrometry: An application to pileup correction in gamma spectrometry*. Ph.D. thesis, ENST.

Trigano, T., Montagu, T., Moulines, E., Roueff, F. and Souloumiac, A. (2006) Analytical pileup correction method for HPGe detectors. *IEEE Trans. Signal Process*. To appear.

Received December 2005 and revised July 2006