NON PARAMETRIC INFERENCE FOR PILE-UP CORRECTION IN NUCLEAR SPECTROMETRY

T. Trigano, A. Souloumiac

CEA Saclay
DRT/LIST/DETECS/SSTM/LETS
91191 Gif-sur-Yvette, France

E. Moulines, F. Roueff

Ecole Nationale Supérieure des Télécommunications
TSI
46 rue Barrault 75634 Paris, France

ABSTRACT

We consider the problem of estimating the density f of independent identically distributed couples $\{(X_k, Y_k)\}_k$ which are the marks of an homogenous Poisson process, from samples $\{(X'_k, Y'_k)\}_k$ deduced from $\{(X_k, Y_k)\}_k$ by a non-linear relation coming from the so-called Type II Counter Problem. From a relation between the probability density functions of the sequences, we propose an estimator of the pdf of Y, since common plug-in estimates are suboptimal.

1. INTRODUCTION

We consider a problem occurring in nuclear science, which will be referred in next sections as " Type II Counter Problem ". Suppose that we have a source of γ photons, which is typically a mixture of radionucleides in unknown proportion. This source emits photons at random times which may be assumed to be the points of an homogenous Poisson process. The photons energies have an unknown distribution, depending upon the type and the proportion of the radionucleides present in the mixture. The photoelectric effect taking place in the lead detector transforms each photon in a pulse of current, which is measured. In this work, since the shape of the pulses depends on many physical parameters (for example the shape of the detector, the location where the photon-electron interaction takes place, etc.) and is therefore very difficult to exploit, we are only considering the *pulse duration* (the duration of the charge collection), and the *total pulse energy*, which will be respectively denoted by X and Y. Our approach is motivated by the physical observation that lengths and energies of the detected pulses are not independent.

The sequence of random variables $\{(X_k, Y_k)\}_k$ is however not directly observable, because several photons can impinge the detector during the charge collection of any given photon and the only available data are on the durations and on the energy of the *clusters* of photons (denoted respectively by X' and Y') and the inter-arrival of these clusters. From this rudimentary information, it is desired to estimate the distribution, and more particularly the probability density function m of the photon energy Y. Figure 1 illustrates the problem encountered : assume that T_n is the arrival time of the n-th photon, X_n the length of its associated electrical pulse and Y_n its energy. When the n-th photon arrives, its energy is recorded, and thus we observe (X_n, Y_n). On the other hand, the $n+2$-th photon is detected during the $n+1$-th busy period ; we observe in that case neither (X_{n+1}, Y_{n+1}) nor (X_{n+2}, Y_{n+2}), but $(T_{n+2} + X_{n+2} - T_{n+1}, Y_{n+1} + Y_{n+2})$. We follow the terminology used in physics and call this phenomenon a *pile-up*.

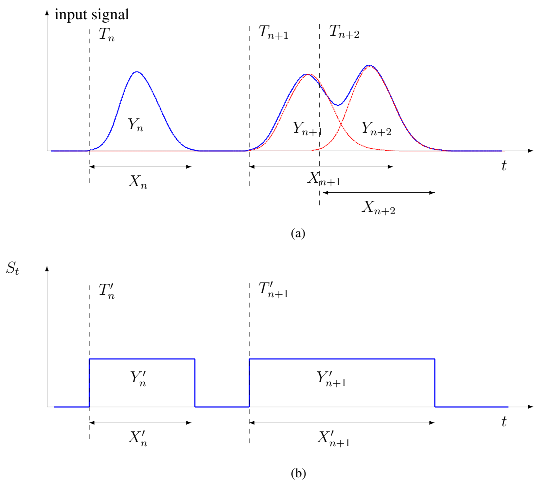

Fig. 1. Illustration of Type II Counter Problem. (a) : input signal with arrival times T_k, lengths X_k and energies Y_k, $k = n, \dots, n+2$; (b) : associated on-off observed process S_t.

Early papers dealing specifically with Type II Counter Problem, as [1] and [2] only consider the energy Y_n as the n-th mark of the point T_n, and obtain results concerning the distribution of the $\{Y'_n\}_n$, but not the distribution of interest. However this problem shares similarity with the inference of the service time in M/G/∞ queue from the duration of the busy periods, and we can find in [3] a relation between the cdfs of X and X' used in [4] to derived an estimator of the cdf of X for the study of biological signals. Remark that in the M/G/∞ and more generally in queuing systems (see e.g. in [5]) we are dealing only with the burst duration whereas in the photon problem, we are dealing with simultaneously with the duration and the energy. This relation was extended in [6] recently and gives a relation between the pdf of (X', Y') and the pdf of (X, Y). This approach yields to an estimate of the pdf of (X, Y) using the methodology described in [4].

The problem can be directly solved by plugging-in this expression, but density estimators that arise from direct inversion do not achieve optimal rates and can be subject to numerical instability, since they involve a numerical differentiation step. In this paper, we suggest an alternative approach in which the Laplace transform of the joint distribution of the cluster duration / energy is expanded first, along the lines drawn in [7] for the related problem of inference of the service time distribution from busy periods in the M/G∞ problem.

2. NOTATIONS AND MAIN THEOREM

Let $(N_t)_{t\geq 0}$ be an homogenous Poisson process, with intensity λ and associated points $\{T_n\}_{n\geq 1}$. At each point T_n is associated a mark (X_n, Y_n) where X_n represents the pulse duration and Y_n the energy of the n-th photon. Suppose that $\{(X_n, Y_n)\}_{n\geq 1}$ is *i.i.d.*, with common distribution f and that $\{\{(X_n, Y_n)\}_{n\geq 1}, \{T_m\}_{m\geq 1}\}$ are mutually independent. As mentioned in Section 1, the marks of $(N_t)_{t\geq 0}$ are not directly observable.

By analogy with queueing theory, a maximal restriction of the signal to a segment where it is positive (resp. 0) is referred to as a *busy* (resp. *idle*) *period*. It is assumed that the only available data are the durations of the busy and idle periods and the total energy collected on a busy period. We denote by $\{T'_n\}_{n\geq 1}$ the starting point of each busy period, by X'_n the duration of the n-th busy period and by Y'_n its energy. We denote by P the probability measure associated to the observed samples $\{(X'_k, Y'_k)\}_{1\leq k\leq n}$. We also naturally define the length of the n-th idle period Z_n as $Z_n \stackrel{\text{def}}{=} T'_{n+1} - (T'_n + X'_n)$. By the lack of memory property of the exponential law, the idle periods have the same distribution (exponential with parameter λ) as the inter-arrival time law. Denote, at last, by m the marginal probability density function following the second dimension associated to f, that is for all nonnegative y :

$$m(y) \stackrel{\text{def}}{=} \int f(x, y)\, dx \ .$$

Our objective is therefore to estimate m, given a sample of density P. The following result is stated in [6], and its demonstration is discussed in [8] :

Theorem 2.1 *We have for all complex couple* (s, p) *in the set* $\{(z_1, z_2) \in \mathbb{C}^2 \ ; \ \mathrm{Re}(z_1) > 0, \, \mathrm{Re}(z_2) \geq 0\}$ *:*

$$\int_0^{+\infty} \mathrm{e}^{-(s+\lambda)\tau} \left[\exp\left(\lambda \int_0^{\infty} \mathrm{e}^{-p\varepsilon} k(\tau, \varepsilon)\, d\varepsilon \right) - 1 \right] d\tau \\ = \frac{\lambda \mathcal{L}P(s,p)}{s+\lambda} \frac{1}{s + \lambda - \lambda \mathcal{L}P(s,p)} \ ,$$

where

$$k(x, \cdot) \stackrel{\text{def}}{=} \int_0^x (x - \tau) f(\tau, \cdot)\, d\tau$$

and

$$\mathcal{L}P(s,p) = \iint_{\mathbb{R}_+ \times \mathbb{R}_+} \mathrm{e}^{-s\tau} \mathrm{e}^{-p\varepsilon} P(d\tau, d\varepsilon) \ .$$

Consequently, we have from Theorem 2.1 a relation between a functional of P, which can be easily estimated from the data, and a functional of the density of interest f. Given this relation, we detail in the next section the methodology used to derive an estimator of m.

3. METHODOLOGY

As mentioned in the introduction, we focus in this contribution on the probability density function of m rather than on the joint density f. Let $c > 0$ and $T > 0$ be arbitrary constants. Let $y \rightarrow K(y)$ be a kernel function and denote by $\Phi_K : \nu \mapsto \Phi_K(\nu) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} K(y) \mathrm{e}^{-\mathrm{i}\nu y} dy$ its Fourier transform. Let h be a bandwidth parameter. Define the function $a : \mathbb{R}_+ \times \{z \in \mathbb{C}; \mathrm{Re}(z) \geq 0\} \mapsto \mathbb{C}$ by :

$$a(x, p) \stackrel{\text{def}}{=} \exp\left(\lambda \int_0^{\infty} \mathrm{e}^{-p\varepsilon} k(x, \varepsilon)\, d\varepsilon \right) \ . \qquad (1)$$

We have by differentiating (1) with respect to x :

$$\int_0^{+\infty} \mathrm{e}^{-p\varepsilon} \left(\int_0^x f(\tau, \varepsilon)\, d\tau \right) d\varepsilon = \frac{1}{\lambda} \frac{\dfrac{\partial a}{\partial x}(x, p)}{a(x, p)} \ . \qquad (2)$$

According to (2), we have

$$\frac{1}{2\pi} \int_{-\infty}^{\infty} \left[\frac{1}{\lambda} \frac{\dfrac{\partial a}{\partial x}(x, \mathrm{i}\nu)}{a(x, \mathrm{i}\nu)} \right] \Phi_K(h\nu) \mathrm{e}^{\mathrm{i}\nu y} d\nu \\ = \frac{1}{h} \int_{-\infty}^{\infty} K\left(\frac{y - \varepsilon}{h} \right) \int_0^x f(\tau, \varepsilon) d\tau d\varepsilon \ . \qquad (3)$$

Taking the limits $x \rightarrow \infty$ and $h \rightarrow 0$ in (3) leads to the following explicit inversion formula :

$$m(y) = \\ \lim_{x \to +\infty} \lim_{h \to 0} \frac{1}{2\pi} \int_{-\infty}^{+\infty} \frac{1}{\lambda} \frac{\dfrac{\partial a}{\partial x}(x, \mathrm{i}\nu)}{a(x, \mathrm{i}\nu)} \Phi_K(h\nu) \mathrm{e}^{\mathrm{i}\nu y}\, d\nu \ . \qquad (4)$$

On the other hand, a can be expressed with respect to P using Theorem 2.1. Since P is a probability measure, for any $(s,p) \in \mathbb{C} \times \mathbb{C}$, $|\mathcal{L}P(s,p)| \leq 1$, with equality if and only if $s = p = 0$. Therefore, for any s such that $\mathrm{Re}(s) > 0$, $|\mathcal{L}P(s,p)| < 1$ and $|\lambda/(s+\lambda)| < 1$. We therefore have for all $(s,p) \in \{z \in \mathbb{C}, \mathrm{Re}(z) > 0\} \times \mathbb{C}$:

$$\frac{1}{s+\lambda-\lambda\mathcal{L}P(s,p)} = \frac{1}{s+\lambda}\sum_{n\geq 0}\left(\frac{\lambda\mathcal{L}P(s,p)}{s+\lambda}\right)^n \ . \quad (5)$$

For convenience, we denote by Ψ the complex-valued function defined for all (ω,p) in $\mathbb{R} \times \{z \in \mathbb{C} \ ; \ \mathrm{Re}(z) \geq 0\}$ as

$$\Psi(\omega,p) \overset{\mathrm{def}}{=} \frac{\lambda\mathcal{L}P(c+\mathrm{i}\omega,p)}{(c+\mathrm{i}\omega+\lambda)(c+\mathrm{i}\omega+\lambda-\lambda\mathcal{L}P(c+\mathrm{i}\omega,p))} \ . \quad (6)$$

Using (5), (6) and Theorem 2.1, we get for all (T,p) in $\mathbb{R}_+^* \times \{z \in \mathbb{C} \ ; \ \mathrm{Re}(z) \geq 0\}$:

$$a(T,p) = 1 + \frac{\mathrm{e}^{(c+\lambda)T}}{2\pi}\int_{-\infty}^{+\infty}\Psi(\omega,p)\mathrm{e}^{\mathrm{i}\omega T}\,d\omega \ . \quad (7)$$

We now use (4) and (7) to construct an estimator of m. We need first to estimate the rate of the underlying Poisson process intensity. It is well-known that the idle periods are independent and identically distributed according to an exponential with intensity λ. A natural estimate of λ therefore is the sample average of the idle periods,

$$\hat{\lambda}_n = \left(\frac{1}{n}\sum_{k=1}^{n} Z_k\right)^{-1} \ . \quad (8)$$

The Laplace transform $\mathcal{L}P$ can be estimated given a sample $\{(X'_k, Y'_k)\}_{1\leq i\leq n}$ as follows :

$$\widehat{\mathcal{L}P}_n(c+\mathrm{i}\omega,\mathrm{i}\nu) = \frac{1}{n}\sum_{k=1}^{n}\mathrm{e}^{-(c+\mathrm{i}\omega)X'_k-\mathrm{i}\nu Y'_k} \ . \quad (9)$$

We now need to estimate the partial derivative $\frac{\partial a}{\partial x}$. Since the function $T \mapsto a(T,\mathrm{i}\nu)$ is defined as an inverse Fourier transform of an integrable function, it is tempting to estimate its derivative by multiplying the integrand by a factor $\mathrm{i}\omega$ in the Fourier domain prior to inversion. This approach however is not directly applicable, because multiplying the integrand by ω in (7) leads to a non absolutely convergent integral. As observed by [7] in a related problem, it is possible to get rid of this difficulty by finding an explicit expression of the singular part of this function, which can be computed and estimated. More precisely, we have from (5) :

$$\Psi(\omega,\mathrm{i}\nu) = A_1(\omega,\mathrm{i}\nu) + A_2(\omega,\mathrm{i}\nu) \quad (10)$$

where we have defined

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

and

$$A_2(\omega,\mathrm{i}\nu) \overset{\mathrm{def}}{=} \frac{\lambda\mathcal{L}P(c+\mathrm{i}\omega,\mathrm{i}\nu)}{c+\mathrm{i}\omega+\lambda}\Psi(\omega,\mathrm{i}\nu) \ .$$

Therefore, using (7) and (10), we obtain :

$$a(T,\mathrm{i}\nu) = 1 + a_1(T,\mathrm{i}\nu) + a_2(T,\mathrm{i}\nu) \quad (11)$$

where

$$a_1(T,\mathrm{i}\nu) = \int_0^T \lambda(T-\tau)\mathrm{e}^{\lambda\tau}\int_0^{+\infty}\mathrm{e}^{-\mathrm{i}\nu\varepsilon}P(d\tau,d\varepsilon)$$

and

$$a_2(T,\mathrm{i}\nu) = \frac{\mathrm{e}^{(c+\lambda)T}}{2\pi}\times \int_{-\infty}^{+\infty}\frac{\lambda\mathcal{L}P(c+\mathrm{i}\omega,\mathrm{i}\nu)}{c+\mathrm{i}\omega+\lambda}\Psi(\omega,p)\mathrm{e}^{\mathrm{i}\omega T}\,d\omega \ .$$

Moreover we have :

$$\frac{\partial a_2}{\partial T}(T,\mathrm{i}\nu) = \frac{\mathrm{e}^{(c+\lambda)T}}{2\pi}\times \int_{-\infty}^{+\infty}\lambda\mathcal{L}P(c+\mathrm{i}\omega,\mathrm{i}\nu)\Psi(\omega,p)\mathrm{e}^{\mathrm{i}\omega T}\,d\omega \ . \quad (12)$$

and

$$\begin{aligned}\frac{\partial a_1}{\partial T}(T,\mathrm{i}\nu) &= \lambda\int_0^T \mathrm{e}^{\lambda\tau}\int_0^{+\infty}\mathrm{e}^{-\mathrm{i}\nu\varepsilon}P(d\tau,d\varepsilon) \\ &= \iint_{\mathbb{R}_+^2}\mathbb{1}_{\{\tau\leq T\}}\mathrm{e}^{\lambda\tau-\mathrm{i}\nu\varepsilon}P(d\tau,d\varepsilon) \ . \quad (13)\end{aligned}$$

and we have using (4) and (11) :

$$\frac{1}{\lambda}\frac{\frac{\partial a}{\partial T}(T,\mathrm{i}\nu)}{a(T,\mathrm{i}\nu)} = \frac{1}{\lambda}\frac{\frac{\partial a_1}{\partial T}(T,\mathrm{i}\nu)+\frac{\partial a_2}{\partial T}(T,\mathrm{i}\nu)}{a(T,\mathrm{i}\nu)} \quad (14)$$

Estimators $\widehat{a}_n(T,\mathrm{i}\nu)$ (respectively $\left(\widehat{\frac{\partial a_2}{\partial T}}(T,\mathrm{i}\nu)\right)_n$) of a (respectively $\frac{\partial a_2}{\partial T}(T,\mathrm{i}\nu)$) can be obtained by plug-in directly estimators (8) and (9) in (7) (respectively (12)). Moreover, we can estimate $\frac{\partial a}{\partial T}(T,\mathrm{i}\nu)$ using (13) :

$$\left(\widehat{\frac{\partial a_1}{\partial T}}(T,\mathrm{i}\nu)\right)_n = \frac{\widehat{\lambda}_n}{n}\sum_{k=1}^{n}\mathbb{1}_{\{X'_k\leq T\}}\mathrm{e}^{\widehat{\lambda}_n X'_k-\mathrm{i}\nu Y'_k}$$

From (4) and and (14), we finally deduce the following estimator for the marginal energy density function,

$$\widehat{m}_{T,h,n}(y) = \frac{1}{2\pi\widehat{\lambda}_n}\times \int_{-\infty}^{+\infty}\frac{\left(\widehat{\frac{\partial a_1}{\partial T}}(T,\mathrm{i}\nu)\right)_n+\left(\widehat{\frac{\partial a_2}{\partial T}}(T,\mathrm{i}\nu)\right)_n}{\widehat{a}_n(T,\mathrm{i}\nu)}\Phi_K(h\nu)\mathrm{e}^{\mathrm{i}\nu y}\,d\nu \quad (15)$$

In practice, this inversion can be realized using a fast Fourier transform (see for instance [9] and [10] for implementation details).

4. RESULTS

In this section we present results on generated densities, and discuss on the optimal choice of the parameters used to compute our estimator. Samples were drawn following the density

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

where $\mathcal{N}_{a,b}$ denotes the gaussian distribution of mean a and standard deviation b truncated to $\mathbb{R}_+$, and piled-up to illustrate Type II Counter Problem for $\lambda = 0.04$; Figure 2 shows the ideal density and a kernel estimate of the marginal of the piled-up distribution P.

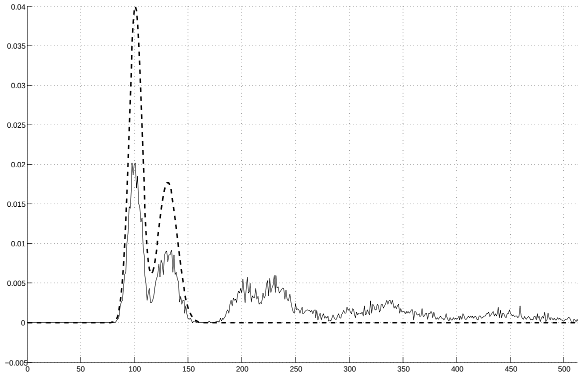

Fig. 2. Example of bimodal density (dashed curve) and piled-up observations (solid curve).

Figure 3 illustrates the differences between the ideal density and our estimator. Our approach seems valid, since the estimator $\hat{m}_{T,h,n}$ is close to m. Remark that the second mode is well recognized, though it has the same magnitude as the piled-up density.

n	ISE
1000	$4.760\,10^{-3}$
5000	$1.089\,10^{-3}$
10000	$3.852\,10^{-4}$
20000	$2.042\,10^{-4}$

Table 1. Integrated Square Error for an increasing number of samples.

Table 1 shows the Integrated Squared Error (ISE) for several values of the number of samples n. These numerical results indicates that the fluctuations in the tail of the estimator are statistical errors and decay as $n \to \infty$. Figure 4 allows us to compare the behavior of the estimate introduced in this paper and the estimate described in [6], all parameters being chosen equal for both estimators ; it shows that our estimator is less sensitive to numerical instability, which validates our approach.

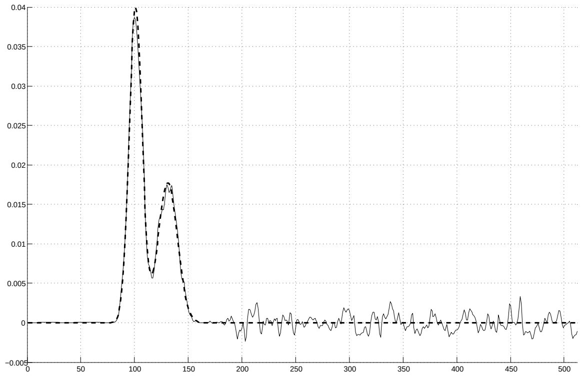

Fig. 3. Behaviour of the estimator $\hat{m}_{T,h,n}$ (solid curve) and comparison with f (dashed curve).

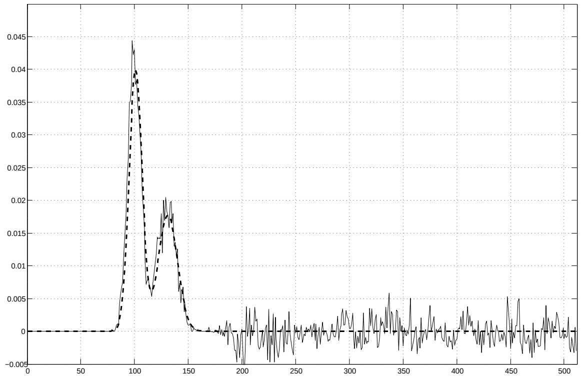

Fig. 4. Behaviour of the estimator with double differentiation introduced in [6] (solid curve) and comparison with f (dashed curve).

We now focus on the choice of the parameters c and T ; we chose T up to $\max_{i \leq n} X'_i$, and due to relations (7) and (12), c can be chosen as close as needed from 0 ; Table 2 shows the ISE for several values of c, T being chosen equal to 64, and several values of T, c being chosen equal to 0.1. The number of samples is fixed to 10^4. It appears that the parameter c has little influence, provided it is not chosen

c	ISE	T	ISE
1	$2.232\,10^{-2}$	20	$5.417\,10^{-3}$
0.1	$3.852\,10^{-4}$	40	$1.905\,10^{-4}$
0.01	$4.002\,10^{-4}$	60	$3.836\,10^{-4}$
0.001	$4.348\,10^{-4}$	80	$5.100\,10^{-4}$
0.0001	$3.852\,10^{-4}$	100	$2.229\,10^{-2}$
0.00001	$4.426\,10^{-4}$	$\max_{i\leq n} X'_i$	$2.231\,10^{-2}$

Table 2. Integrated Square Error for different values of c and T.

to big to avoid numerical problems. This is not surprising, since the Bromwich integral used to compute the inverse Laplace transform does not theorically depend on the choice of c (see e.g. [11]). The choice of the parameter T seems more crucial however : choosing T too small means that we introduce a bias when integrating from 0 to T in (15), and Table 2 indicates that T cannot be chosen as large as possible, implying the existence of an optimal sequence T_n depending on λ, c and the number of samples n.

5. CONCLUSION AND PERSPECTIVES

In this paper we presented a deconvolution problem with a nonlinear relation between both densities. We exhibit an estimator of the density of interest based on Fourier inversion that can be easily computed. Primary results showed interesting adequation on generated densities, which seems to validate our approach. A problem associated to this should be to find a good estimate of the density f in the case of real-world data and to consider the case when the input Poisson process is not homogenous anymore. These questions, as the theoritical study of the rates of convergence of our estimator, should be investigated in future papers.

6. REFERENCES

[1] A. Dvurecenskij and G.A. Ososkov, “Note on Type II Counter Problem,” *Appl. Math. (Czech.)*, vol. 29, pp. 237–249, 1984.

[2] A. Dvurecenskij and G. A. Ososkov, “On a Modified Counter with Prolonging Dead Time,” *Jour. of Appl. Prob.*, vol. 22, no. 1, pp. 678–687, 1985.

[3] L. Takacs, *Introduction to the Theory of Queues*, Oxford University Press, 1962.

[4] N. H. Bingham and S. M. Pitts, “Non-Parametric Estimation for the M/G/∞ Queue,” *Ann. Inst. Statist. Math.*, vol. 51, no. 1, pp. 71–97, 1999.

[5] F. Baccelli and P. Brémaud, *Elements of Queueing Theory*, Springer, 2002.

[6] T. Trigano, E. Barat, T. Dautremer, and A. Souloumiac, “Pile-Up Correction Algorithms for Nuclear Spectrometry,” in *International Conference on Acoustics, Speech and Signal Processing*, 2005.

[7] P. Hall and J. Park, “Nonparametric Inference about Service Time Distribution from Indirect Measurements,” *Jour. Roy. Statist. Soc.*, vol. 66, no. 4, pp. 861–875, 2004.

[8] T. Trigano, E. Barat, and T. Dautremer, “Etude probabiliste de signaux spectrométriques : méthode de désempilement de spectre en énergie pour la spectrométrie γ,” Tech. Rep. 03-018, CEA, 2003.

[9] B. Davies and B. Martin, “Numerical Inversion of the Laplace Transform : a Survey and Comparison of Methods,” *J. Comp. Phys.*, vol. 33, pp. 1–32, 1979.

[10] L. Brancik, “Programs for Fast Numerical Inversion of Laplace Transforms in Matlab Language Environment,” *Sbornik 7. Proc. Matlab 99, Prague*, pp. 27–39, 1999.

[11] G. Doetsch, *Introduction to the Theory and the Application of the Laplace Transform*, Springer-Verlag, 1974.