Statistical estimation of a shot-noise process

Paul Ilhe

2016-ENST-0052

EDITE - ED 130

Doctorat ParisTech

T H È S E

pour obtenir le grade de docteur délivré par

TELECOM ParisTech

Spécialité « Signal et Images »

présentée et soutenue publiquement le 29/09/2016 par

Paul ILHE

Estimation statistique des éléments d'un processus shot-noise

Directeurs de thèse : **François Roueff**
Éric Moulines
Encadrant : **Antoine Souloumiac**

M. Jérome Idier, Directeur de recherche, CNRS & IRCCyN — Rapporteur
M. Mathieu Lerasle, Chargé de recherche, CNRS & Université Paris-Sud — Rapporteur
Mme. Catherine Matias, Directrice de recherche, CNRS & LPMA — Présidente
M. Eric Moulines, Professeur, CNRS & École Polytechnique — Directeur de thèse
M. François Roueff, Professeur, Télécom ParisTech — Directeur de thèse
M. Antoine Souloumiac, CEA List — Encadrant

TELECOM ParisTech
école de l'Institut Mines-Télécom - membre de ParisTech
46 rue Barrault 75013 Paris - (+33) 1 45 81 77 77 - www.telecom-paristech.fr

T
H
È
S
E

Remerciements

> *Rubens pense : "le rire est, de toutes les expressions du visage, la plus démocratique : l'immobilité des visages rend clairement discernable chacun des traits qui nous distinguent les uns des autres ; mais dans la convulsion, nous sommes tous pareils."*
>
> – Milan Kundera, *L'immortalité*

Au terme de trois années de travail dont cette thèse marque l'accomplissement, je tiens à remercier tous celles et ceux sans qui ce doctorat n'aurait pas eu la même saveur.

En premier lieu, je souhaite exprimer toute ma gratitude à mes directeurs de thèse, François Roueff et Éric Moulines. François, merci de m'avoir tant appris d'un point vue scientifique. Ta constante disponibilité et inépuisable patience m'ont permis d'en connaître plus sur les processus β-mixing et les processus ponctuels que je n'aurais espéré.

Éric, je ne saurais dater notre premier échange ni même le nombre de conseils que tu m'as donné tant ils ont jallonné mes études. Cette thèse n'aurait probablement jamais eu lieu sans ton concours et toute l'énergie que tu as su me transmettre. En revanche, cette mauvaise foi concernant le niveau réel de joueurs girondins (tel Cheick Diabaté) t'es entièrement pardonnée. Cela va me manquer.

Antoine, je te remercie d'avoir rendu cette thèse, en partenariat avec le CEA, possible, et tiens à adresser à l'ensemble de mes collègues du LADIS de chaleureuses salutations. Grâce à toi, j'ai hérité du goût de l'expérimentation numérique (délaissée il est vrai en début de thèse), mais suis aussi devenu un fin connaisseur de l'ensemble des techniques de diagonalisation conjointe de matrices. Prochaine étape, conquérir les différents secteurs de Fontainebleau !

Il serait toute fois réducteur de ne remercier seulement les derniers formateurs d'un long parcours académique sans n'exprimer une pensée envers l'ensemble des professeurs qui ont su communiquer leur passion si singulière pour les sciences, la mathématique et les applications qu'elle entrouvre. René Adad, merci pour ton soutien, tant durant la prépa que lors de ma première année d'école.

Merci également à Catherine Mattias de m'avoir fait l'honneur de présider ledit jury et aux rapporteurs Jérome Idier and Mathieu Lerasle d'avoir accepté de relire ce manuscrit.

Je tiens ègalement à remercier l'école Télécom ParisTech, d'une part pour les six années d'étude passées en son sein, mais surtout car elle a vu naître ce magnifique groupe d'amis

surnommé le “5ème & co”. Même éparpillés aux quatre coins du globe, je vous remercie de votre bonne humeur et présence continuelle. Une pensée également à chacun de mes colocataires, Tarik, Nicolas, Ivan, Gillian et bien sûr Florian, pour tous ces moments passés à vos côtés, ces sessions de grimpe, de five, ces soirées, ces belotes ou encore ces parties d’échecs. Encore merci à l’ensemble de mes collègues de thèse du CEA et de Télécom, notamment Maxime F., Maxime S., Benoit B., Eugène N. et Alain D. pour toutes nos discussions variées et repas animés ainsi que mes amis carcassonnais, Jean, Charles, Amandine et Damien.[1]

Enfin, et surtout, je remercie mes parents et mon frère pour leur soutien indéfectible durant toutes ces années d’étude, tant lorsque mes recherches se trouvaient dans un minimum local, que lorsque je traversais des étapes éprouvantes. Merci pour tout.

[1] Merci également à l’ensemble de l’équipe de Cardiologs de m’avoir permis de préparer ma soutenance de thèse dans les meilleures conditions. J’ai une pensée particulière à l’égard de Jérémy R. dont les remerciements de thèse ont grandement inspiré ceux-ci ! Maxime, ce sera toujours un plaisir de travailler à tes côtés.

Résumé

Dans cette thèse, nous nous intéressons à un problème de physique nucléaire : la spectrométrie Gamma. Il s'agit d'un problème inverse non-linéaire qui est décrit de la manière suivante. Des particules, généralement un faisceau de photons, interagissent avec un détecteur. Chacune des interactions est ensuite convertie par le détecteur en une impulsion électrique qui est proportionnelle à l'énergie déposée par le photon correspondant. La forme de l'impulsion électrique dépend du détecteur ainsi que des autres éléments composant la chaîne d'instrumentation. Même si le dépôt d'énergie associé à un photon s'avère parfois partiel, par exemple à cause de l'effet Compton, la distribution des énergies déposées reflète très étroitement celle des énergies transportées par les photons gamma. L'objectif de la spectrométrie gamma consiste alors à mesurer précisément cette dernière quantité ainsi que le nombre moyen de photons qui interagissent avec le détecteur dans un intervalle d'une seconde. Les réactions nucléaires générant les photons étant indépendantes, nous nodélisons les interactions faisceau/détecteur par un processus de Poisson homogène d'intensité λ sur la droite réelle. Le courant électrique mesuré par le détecteur correspond alors à un processus de Poisson filtré par la réponse impulsionnelle - supposée connue - et marqué multiplicativement par une amplitude aléatoire de distribution inconnue : en d'autres termes, un processus shot-noise.

Dans ce contexte, nous souhaitons retrouver la distribution des énergies photoniques à partir du courant électrique mesuré. Or, les techniques d'instrumentation nucléaire actuelles ne permettent pas de mesurer cette distribution lorsque l'intensité λ est suffisamment forte pour que les impulsions, de durée courte mais non nulle, s'empilent : ce phénomène est connu sous le nom d'empilement. Cette superposition de courants électriques empêche alors d'établir une correspondance bijective entre une impulsion et l'énergie déposée associée.

Nous introduisons dans cette thèse de nouveaux estimateurs non-paramétriques de l'intensité et de la densité des marques d'un processus shot-noise à partir d'un nombre fini d'observations du processus échantillonné à basse fréquence. Les méthodes proposées utilisent une relation non linéaire reliant la fonction caractéristique de la loi marginale du processus à la densité des marques. Elles sont particulièrement rapides et possèdent l'avantage d'être efficaces mêmes pour des intensités élevées. Les performances de ces méthodes sont étudiées quantitativement et illustrées à la fois pour des données simulées et réelles provenant du CEA Saclay. En particulier, les estimateurs de la densité des marques permettent de corriger les artefacts de pics multiples.

Abstract

In this thesis, we consider the nonlinear inverse problem of γ-spectroscopy which can be described as follows. A radioactive source randomly emits photons that interact with a detector. An acquisition system then converts each interaction into a pulse of short duration whose shape is proportional to the deposited energy. The electric shape depends on the detector and the others elements of the data acquisition system. Even though the energy deposited by a photon can not be exactly the same as the initial energy carried by the photon, for instance because of the Compton scattering, the distribution of the deposited energies is intricately linked to the gamma-ray spectra of the incoming radioactive source. The purpose of gamma-spectroscopy consists in estimating as precisely as possible this spectra as well as the mean number of photons that hit the detector in one second. Because the nuclear reactions that emit photons are independent, we choose to model the time and energy of the particles-detector interactions by a marked Poisson point process where the time arrivals are modeled by an homogeneous Poisson point process on the real line with intensity λ and the marks by independent and identically distributed random variables. Then, the analogical signal recorded by the detector corresponds to the convolution between the impulse response -assumed to be known- and the abovementioned marked Poisson point process.

In this context, we aim to estimate the density of the gamma photon energies from the electric current recorded by the detector. However, the actual instrumentation techniques do not allow to estimate this density when the intensity λ exceeds some threshold. Indeed, the probability that two or more photons hit the detector in a short interval increases as soon as the intensity λ goes large, so that the electric current generated by each particle overlap. Such a phenomenon is called pileup. This overlapping precludes to construct a one-to-one correspondence between a pulse and the associated energy deposited in the detector.

We introduce in this thesis new nonparametric estimators of the intensity and the mark's density of a shot-noise process based on a finite sample of low-frequency observations of this stochastic process. The methods developed exploit a nonlinear functional equation linking the characteristic function of the marginal law of the shot-noise with the mark's density function. They are particularly time-efficient and perform well even for processes with high intensity. The performance of the methods are quantitatively studied and illustrations are provided both on simulated datasets and real datasets stemming from the CEA Saclay. In particular, our methods correct the multiple peak artifacts that arises with classical techniques.

Contents

List of Figures

1 Introduction en français

En somme, je ne me suis jamais soucié des grands problèmes que dans les intervalles de mes petits débordements.

– Albert Camus, *La Chute*

Résumé

Dans ce chapitre, nous fournissons une introduction là la spectrométrie gamma et les enjeux inhérents à cette méthode d'instrumentation nucléaire. Après quelques rappels de physique nucléaire, nous détaillons les principales interactions photon-matière, décrivons la chaîne d'instrumentation associée aux mesures ainsi que leurs impacts sur la reconstruction des spectres de radionucléides.

Nous formalisons ensuite le problème de spectrométrie γ de manière probabiliste en le reliant à l'estimation statistique des processus stochastiques de type "shot-noise" et décrivons les enjeux majeurs poursuivis dans cette thèse.

Dans un premier temps, nous passons en revue les principales méthodes statistiques et déterministes actuellement utilisées par les physiciens. Après avoir souligné leurs limites lorsque l'intensité de la source radioactive devient trop importante, nous présentons un résumé exhaustif des méthodes d'estimation statistiques du problème développées aux Chapitres 4 et 5.

1.1 La spectrométrie gamma

L'objectif de cette partie consiste à décrire ce qu'est le rayonnement γ. À cette fin, nous rappelons quelques éléments de physique atomique.

1.1.1 Quelques bases de physique nucléaire

La notation ${}^{A}_{Z}X$ est utilisée pour qualifier un noyau formé d'un élément chimique X possédant Z neutrons, N protons et où $A = Z + N$ est communément appelé *nombre de masse* (de même, on désigne par *nombre atomique* Z). Les neutrons et protons sont appelés *nucléons*, deux noyaux possédant le même nombre de masse A, de neutrons Z sont respectivement dits *isobares* et *isotopes*.

Conformément à la théorie relative à la mécanique quantique, à tout noyau ${}^A_Z X$ correspond un nombre fini d'énergie quantifiés. Ces états d'énergie d'un noyau sont généralement distingués en deux catégories : les états stables et instables. Très souvent, l'état stable ou *fondamental* d'un noyau coïncide avec celui dont l'énergie est minimale. Les autres états sont alors qualifiés d'*excités*, signifiant qu'ils possèdent un supplément d'énergie par rapport à l'état fondamental, dû à la répartition des nucléons qui le composent sur des niveaux d'énergie supérieurs à ceux du noyau stable.
En général, les noyaux excités résultent d'un processus de désintégration α, β, ou de capture électronique (le lecteur souhaitant plus de précisions peut se référer à Tsoulfanidis (2013)). Pour retrouver un niveau stable, le noyau excité expulse l'excédant d'énergie sous forme d'un photon. Ainsi, on appelle *rayonnement* γ la réaction qui, à un noyau ${}^A_Z X^\star$ excité, associe le même noyau dans un état stable ${}^A_Z X$ ainsi qu'un photon γ dont l'énergie est exactement égale à la différence entre le niveau d'énergie du noyau instable et du noyau stable. Cette réaction est notée

1

$${}^A_Z X^\star \longrightarrow {}^A_Z X + {}^0_0 \gamma \ .$$

Le rayonnement γ fut pour la première fois observé en 1900 par Paul Villard. Du fait de la nature quantique des niveaux d'énergie dans les noyaux, il en va de même des énergies des photons gamma émis. Ainsi, **il existe une correspondance unique entre un noyau et la suite des raies énergiques caractéristiques émises par rayonnement γ**.

1.1.2 La spectrométrie γ : définition et objectifs

Définition 1 (Spectrométrie γ)**.** *La spectrométrie γ est une technique de mesure nucléaire qui, à partir d'un système d'instrumentation composé d'un détecteur (à semi-conducteur, scintillateur, etc) sensible aux rayonnements des photons, a pour objectif d'établir la densité d'énergie des particules émises par une source radioactive et d'en compter le nombre.*

La correspondance entre le spectre de raies gamma émises par un ensemble de noyaux et les radio-nucléides sous-jacents qui le composent font de la spectrométrie γ l'une des techniques privilégiées de mesure utilisées en physique nucléaire pour identifier des élements radioactifs. En pratique, la détection de rayonnements ionisants tels que le gamma s'effectue à l'aide de matériaux sensibles aux dépots d'énergie des photons, par exemple un détecteur à semi-conducteur. À la rencontre d'un photon, le détecteur va alors générer un courant électrique (voir la Figure 1.9 pour un exemple et la Section 1.1.3 pour les détails) proportionnel à l'énergie déposée. Parmi le grand nombre de mécanismes d'interaction photon-matière, nous détaillons dans la suite les trois principales interactions.

Effet photoélectrique

L'interaction photoélectrique correspond à la réaction durant laquelle un photon interagit avec un électron appartenant à la couche extérieure d'un atome composant le semi-conducteur. Lorsque l'énergie E_γ du photon incident est suffisamment élevée (c'est à dire supérieure à l'énergie de liaison d'un électron), ce dernier est alors totalement absorbé tandis que l'électron éjecté, appelé *photoélectron*, possède une énergie cinétique donnée par

$$E = E_\gamma - B_e$$

où B_e représente l'*énergie de liaison* de l'électron. L'effet photoélectrique représente l'interaction prédominante pour des photons de faibles énergies. En outre, il se produit avec une probabilité d'autant plus grande que l'énergie incidente E_γ est faible et dépend de multiples facteurs: il existe une formule dite *coefficient photoélectrique*.

$$p = aN\frac{Z^n}{E_\gamma^m}(1 - O(Z)) \ .$$

où m, n sont des constantes dépendant de E_γ comprises entre 3 et 5, N et Z respectivement le nombre de neutrons et protons de l'atome du noyau susmentionné, a une constante et $O(Z)$ est un terme correctif de premier ordre en Z. La figure 1.1 illustre comment la probabilité p évolue en fonction de l'énergie incidence E_γ et du nombre atomique Z de l'atome du semi-conducteur.

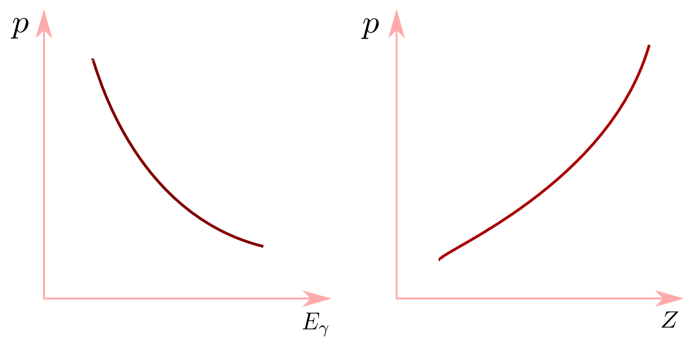


Figure 1.1 : Esquisse de l'évolution de la probabilité d'apparition de l'effet photoélectrique en fonction de l'énergie photonique E_γ et du nombre de protons Z.

Effet Compton

L'effet Compton caractérise la collision entre un photon et un électron libre (ou faiblement lié) du matériel absorbant. Contrairement à l'effet photoélectrique, le photon incident ne disparaît pas mais est dévié d'un angle polaire θ et transporte une énergie E'_γ inférieure. L'électron libre est quant à lui éjecté comme illustré par la Figure 1.2. Le photon dévié peut à son tour interagir avec le semi-conducteur avec une probabilité dépendant de paramètres comme la taille ou la position du détecteur et dépasse le cadre de ce manuscrit.

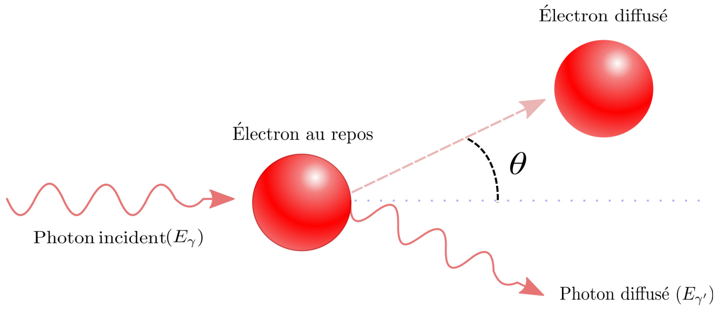


Figure 1.2 : Effet Compton

1

En appliquant les lois de conservation des moments, la formule reliant l'énergie diffusée E'_γ avec l'angle de déviation θ et l'énergie incidente est donnée par

$$E'_\gamma = \frac{E_\gamma}{1 + (1 - \cos(\theta))\, E_\gamma / m_e c^2} \, .$$

Lors d'une interaction de type Compton, le détecteur enregistre l'énergie $E_\gamma - E'_\gamma$ de l'électron diffusé. Dans le contexte de la spectrométrie γ, il est important de connaître l'intervalle dans lequel se trouve l'énergie transportée par le photon résultant. Lorsqu'il n'y a pas de déviation, correspondant au cas $\theta = 0$, il n'y a aucune perte d'énergie et $E'_\gamma = E_\gamma$. En revanche, lorsque l'angle $\theta = \pi$, c'est à dire que le photon diffusé part dans la direction opposée au premier, son énergie est donnée par

$$E'_{\gamma,\min} = \frac{E_\gamma}{1 + 2E_\gamma / m_e c^2} \, .$$

Toutefois, la diffusion Compton n'est pas isotrope, la distribution des angles θ de déviation n'étant pas uniforme. De manière plus précise, la formule de Klein-Nishina assure que la probabilité d'un photon d'énergie E_0 d'être diffusé vers un angle solide $\mathrm{d}\Omega$ est donnée par

$$\frac{\mathrm{d}\sigma_{\mathrm{KN}}}{\mathrm{d}\Omega} := \frac{r_e^2}{2} \frac{1}{(1 + \alpha(1 - \cos\theta))^2} \left(1 + \cos^2\theta + \frac{\alpha^2 (1 - \cos\theta)^2}{1 + \alpha(1 - \cos\theta)}\right) ,$$

où r_e est le rayon d'un électron et $\alpha = E_0 / m_e c^2$. La figure 1.3 montre les différentes distributions des angles de déviation θ pour diverses valeurs de α.

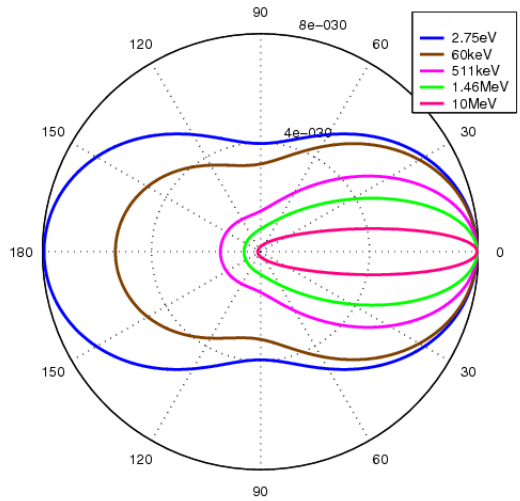
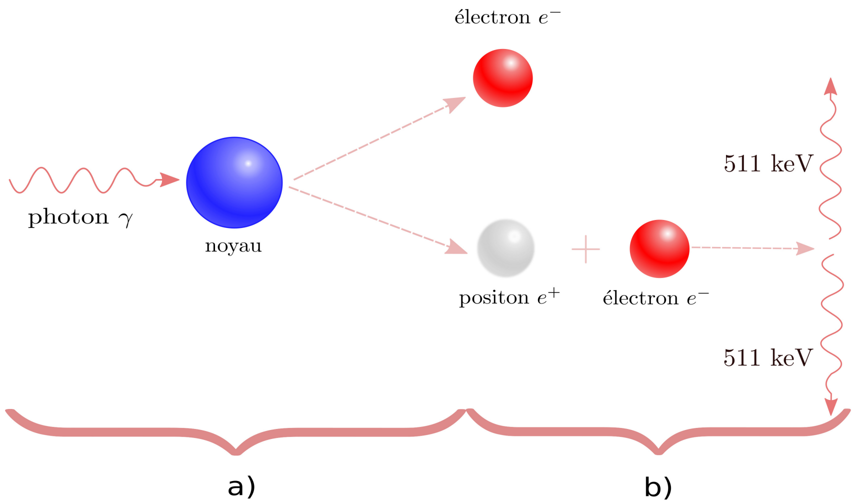


Figure 1.3 : Distribution des angles de Klein-Nishina pour 5 valeurs différentes de α

Dans le cas où $E_\gamma \ll m_e c^2$, on remarque que $E'_{\gamma,\min} \simeq E_\gamma$. Ce régime particulier caractéristique des énergies faibles porte le nom de *diffusion Thompson* : tant que les photons transportent des énergies inférieures à 45 keV, l'énergie transmise par le photon à l'électron est négligée et seule la direction du photon est modifiée. En d'autres termes, tout se passe alors comme si l'effet Compton n'avait pas d'influence sur la déformation des spectres en énergie des rayonnements γ. Cela est visible notamment en Figure 1.6.

Création de paires électron-trous

La troisième interaction photon-matière concerne les particules qui transportent des énergies supérieures à deux fois l'énergie de repos d'un électron (1.022 MeV). Lorsqu'un photon γ vérifie cette condition, le phénomène de création de paires, schématiquement représenté en Figure 1.4, devient possible. Cela correspond à une interaction entre le photon et un noyau donnant lieu à la création d'un paire électron-positon tandis que le noyau reste inchangé. Dans un second temps, le positon créé s'annihile avec un électron présent dans la matière pour donner naissance à deux photons d'énergie 511 keV émis en direction opposée. La création de paire électron-positon peut également avoir lieu, bien plus rarement, lorsque deux photons, chacun d'énergie supérieure à 511 keV, interagissent. Bien qu'il soit difficile d'exprimer analytiquement la probabilité qu'une interaction provoque une production de paires, on sait expérimentalement qu'elle est approximativement proportionnelle à la racine carrée du nombre atomique Z du noyau absorbant.



1

Figure 1.4 : Phénomène de création de paires : a) Création d'une paire électron-positon. b) Phénomène d'annihilation et émission de deux photons γ d'énergie 511 keV.

Le spectre d'énergie enregistré par le détecteur est plus compliqué que ceux produits par l'effet Compton ou photoéléctrique. En effet, si le positon s'annihile avec un électron, l'un ou les deux photons d'énergie 511 keV peuvent être absorbés par le détecteur ou au contraire, s'en échapper et interagir plus tardivement. Pour une énergie incidente E_γ supérieure à la masse de repos d'une paire électron-positon, quatre scénarios peuvent se produire :

- il n'y a pas création de paires et le détecteur enregistre l'énergie E_γ,
- le photon d'énergie $E_\gamma - 1,022$ MeV (dit en anglais *double peak escape*) interagit avec le détecteur,
- le photon d'énergie $E_\gamma - 1,022$ MeV et un photon d'énergie 511 keV (dit en anglais *single peak escape*) interagissent avec le détecteur,
- un photon provenant de l'annihilation du positon interagit avec le détecteur ; une énergie de 511 keV est enregistrée.

En résumé, les trois mécanismes d'interaction entre les photons et la matière décrits ci-dessus possèdent des comportements différents selon l'énergie incidente des particules ainsi que le nombre atomique du noyau absorbant. La figure 1.5 illustre les différentes contributions des trois interactions photon-matière selon l'énergie des rayons γ entrants et les nombres atomiques Z des atomes composant le détecteur.
La figure 1.6, issue de Pirard (2006), résume, pour un photon incident d'énergie E_γ, la distribution des différentes énergies enregistrées par le détecteur.

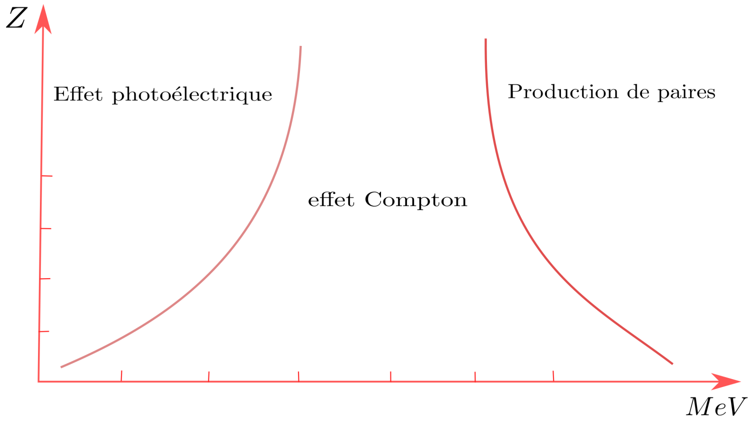


Figure 1.5 : Prépondérance des interactions photon-matière en fonction du nombre atomique du noyau et de l'énergie du photon incident.

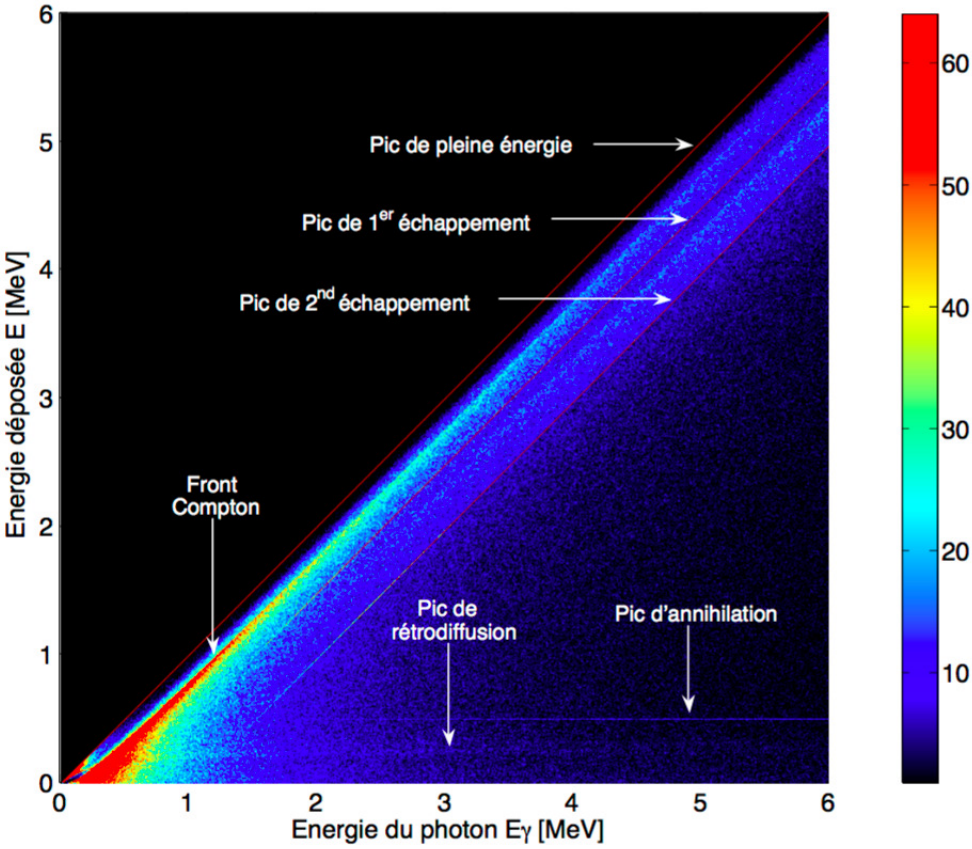


Figure 1.6 : Prépondérance des interactions photon-matière en fonction de l'énergie du photon incident. Plus la couleur est chaude, plus l'effet est prépondérant

Dans la majorité des situations réelles, ces interactions induisent une déformation du spectre initial plus ou moins conséquente. Il ne faudra donc pas s'attendre à observer ou estimer des raies d'énergie à des positions particulières déterminées par le caractère discret des énergies des noyaux radioactifs composant la source. Il paraît alors hasardeux de recourir à des modèles paramétriques afin de modéliser les spectres en énergies. Deux types de gabarits sont présentés en Figure 1.7 pour des photons monoénergétiques transportant une énergie E_γ selon qu'elle est supérieure ou inférieure à $2m_ec^2$. En outre, le type d'instrumentation associé à la détection de photons γ joue un rôle prépondérant dans la résolution des spectres que l'on peut espérer retrouver. Nous le détaillons succinctement dans la partie suivante.

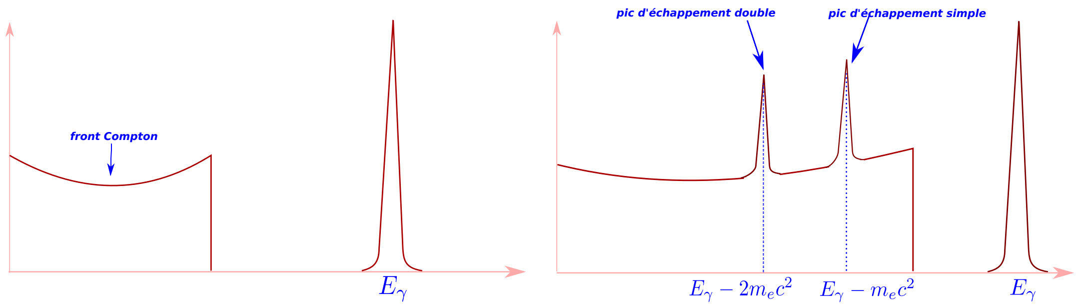


Figure 1.7 : Exemple de spectres résultant des interactions photons-matière selon que: a) $E_\gamma \leq 2m_ec^2$, b) $E_\gamma \geq 2m_ec^2$

1.1.3 Description de l'instrumentation associée

Jusqu'à présent, nous avons défini le détecteur comme l'élément de la chaîne d'acquisition qui interagit avec les photons. Plus précisément, il existe deux principaux types de détecteur, les *semi-conducteurs* et les *scintillateurs* très bien expliqués dans Pirard (2006). Nous nous contenterons d'admettre que les trois types d'interactions décrits ci-dessus provoquent une création de paires électrons-trous qui est proportionnelle à l'énergie déposée par le photon incident. En migrant vers les bornes du détecteur, elles génèrent alors une impulsion électrique mesurée à l'aide d'une chaîne d'instrumentation, schématiquement représentée par la Figure 1.8, que nous allons brièvement expliciter dans cette partie. Un exemple de signal temporel observé est fourni en Figure 1.9.
La qualité d'une chaîne d'acquisition est souvent définie en terme de *résolution*. La résolution d'un détecteur correspond à la dispersion observée en fin de chaîne par l'oscilloscope lorsqu'on mesure des dépots d'énergie identiques. Elle permet de mesurer la capacité d'un système d'acquisition à séparer les raies présentes dans le spectre. D'un point de vue physique, la résolution d'un système de détection est évaluée selon l'expérience suivante.
À partir d'une source émettant des particules transportant exactement la même énergie E_γ (modèle idéalisé), on effectue un histogramme des maxima enregistrés en sortie de la

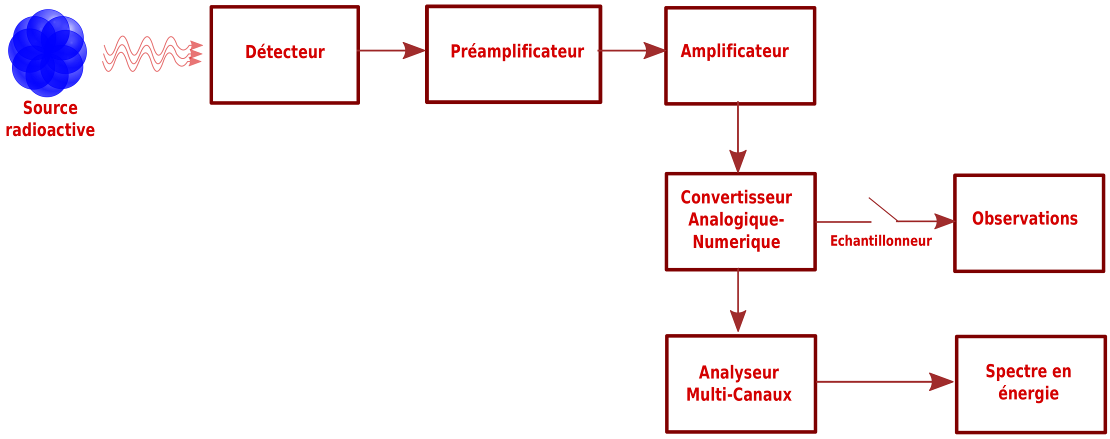


Figure 1.8 : Schéma d'une chaîne d'acquisition utilisée en spectrométrie

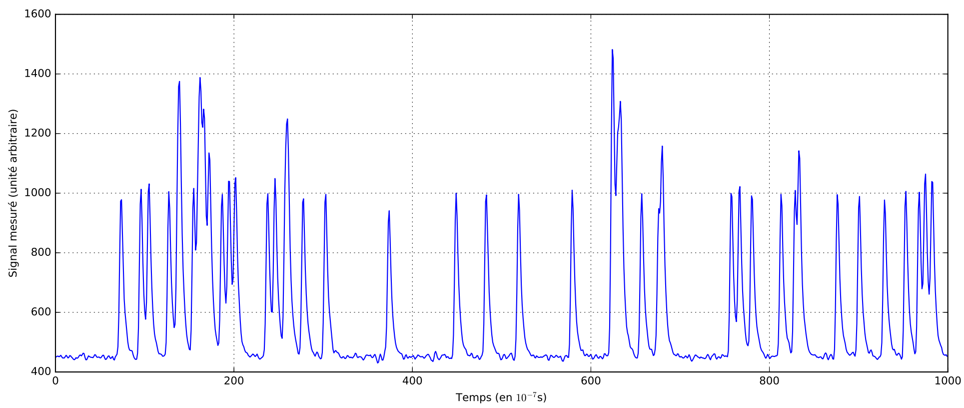


Figure 1.9 : Exemple d'un signal observé en fin de chaîne d'acquisition

chaîne d'acquisition par l'oscilloscope. Pour diverses raisons comme la fluctuation statistique du nombre de paires électrons-trous créées ou le bruit électronique des différents blocs de la chaîne, l'histogramme obtenu possède la forme d'une gaussienne centrée en E_γ. On appelle *largeur à mi-hauteur* notée Γ en eV l'intervalle qui sépare les deux valeurs pour lesquelles le signal sortant est égal à la moitié du signal maximal mesuré, comme illustré par la Figure 1.10. La *résolution en énergie* est alors définie comme le rapport entre Γ et E_γ.

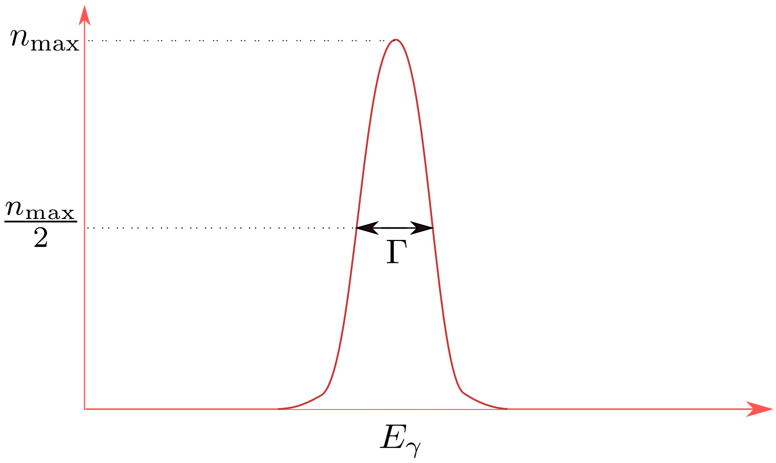


Figure 1.10 : Schéma de la résolution d'un pic relatif à une raie d'énergie E_γ

Cette résolution dépend aussi bien du type de détecteur choisi que de l'amplificateur utilisé. Outre le caractère aléatoire du nombre de paires électrons-trous créés lors d'une interaction, certains effets dûs à la chaîne d'acquisition et à la collection des charges entraînent un étalement des amplitudes enregistrées par le détecteur. De plus, la marge de manœuvre disponible au niveau du détecteur est limitée. Les détecteurs à scintillation possèdent l'attrait de générer des réponses impulsionnelles plus rapides et l'inconvénient d'avoir une mauvaise résolution tandis que les détecteurs à semi-conducteur présentent un comportement antagoniste. Comme nous le verrons dans la suite, les méthodes développées dans ce manuscrit permettent de traiter tout type de réponse impulsionnelle, tandis que celles actuellement utilisées dans l'état de l'art (voir Section 1.3) sont efficaces seulement pour des impulsions de courte durée. Il est donc préférable d'utiliser des détecteurs possédant des résolutions en énergie optimales.
Le rôle du préamplificateur consiste à fournir un couplage optimal entre le courant généré en sortie du détecteur et le reste de la chaîne d'instrumentration. D'une part, il doit être placé près du détecteur afin que le signal en sortie de ce dernier ne soit ni atténué, ni noyé dans le bruit électronique environnant. Le signal est ensuite amplifié et mis en forme. Enfin, le préamplificateur réduit l'atténuation du signal en calibrant l'impédance du détecteur avec celle de l'amplificateur.
L'amplificateur est ensuite chargé de transformer le signal en tension suivant une étape

de mise en forme suivie d'une amplification de manière à minimiser le rapport signal sur bruit. Habituellement, les amplificateurs utilisés à des fins de spectrométrie génèrent deux types de courant électrique illustrés par la Figure 1.11 lorsqu'un photon est détecté.

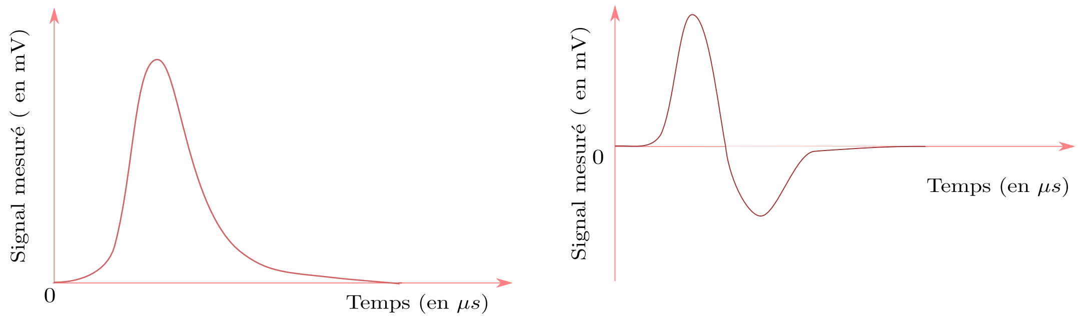


Figure 1.11 : Illustration des deux types de réponses impulsionnelles rencontrées en spectrométrie γ

L'analyseur multicanaux effectue une tâche particulièrement efficace lorsque le nombre de photons incident par seconde est faible. Pour chaque impulsion électrique associée à la détection d'un photon, il enregistre la valeur maximale de l'impulsion, qui est proportionnelle à l'énergie déposée, et stocke cette valeur dans le canal correspondant à l'intervalle contenant la valeur du maximum local mesuré. Par exemple, pour un détecteur de Germanium, une haute résolution en énergie nécessite un codage sur 13 ou 14 bits (soit 8192 et 16384 canaux respectivement) pour couvrir une plage de 0 à 8 MeV.
En pratique, un étalonnage en énergie est nécessaire lorsqu'on souhaite identifier divers éléments radioactifs par spectrométrie γ. Pour cela, une opération courante répandue consiste à utiliser le signal de sortie observé en fin de chaîne d'acquisition lorsque celle-ci est soumise à l'effet d'une source mono ou bi-énergétique à taux de comptage relativement faible. L'histogramme des maxima locaux du signal enregistré possède ainsi une allure unimodale ou bimodale associée à la distribution de la source (cf. le premier histogramme présenté en Figure 1.14). Cette procédure permet alors d'estimer précisement le facteur multiplicatif lié à la réponse impulsionnelle.

1.2 Enjeux de la thèse

Dans la cadre de cette thèse, notre objectif consiste à caractériser la source par spectrométrie gamma, c'est-à-dire à estimer à la fois la densité d'énergie des particules émises par la source radioactive ainsi que le nombre moyen de particules émises par seconde à partir d'un nombre fini d'observations d'un signal temporel similaire à celui présenté en Figure 1.9.
De nombreuses méthodes dévelopées par des physiciens et statisticiens existent déjà ; les plus utilisées ou importantes sont détaillées en Section 1.3. Cependant, la plupart d'entrs elles deviennent inefficaces lorsque l'*intensité* de la source, c'est-à-dire le nombre de

photons émis par la source radioactive au cours d'une seconde est supérieur à un certain seuil, variable selon le type d'instrumentation utilisé.

1.2.1 Le problème d'empilement

Les sous-performances de ces algorithmes s'expliquent par la présence d'*empilement*, phénomène qui correspond à la superposition d'au moins deux courants électrique générés par au moins deux interactions de photons avec le détecteur du système d'instrumentation, comme illustré en Figure 1.12. Ce phénomène apparaît généralement lorsque les taux de comptage deviennent très élevés et/ou lorsque le support (s'il est fini) de la réponse impulsionnelle de la chaîne d'acquisition n'est pas suffisamment court.

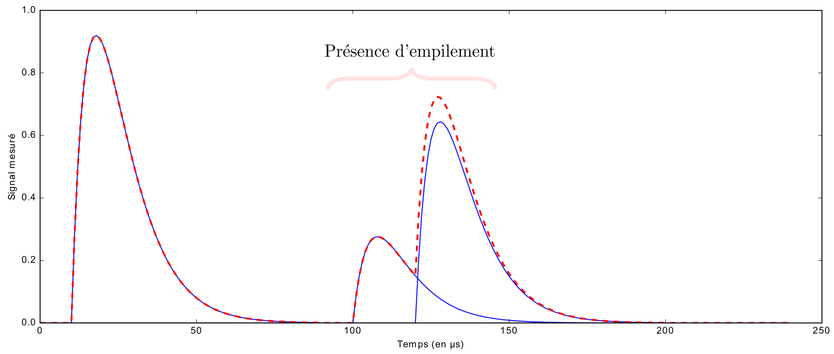


Figure 1.12 : Illustration du phénomène d'empilement des processus shot-noise

Cette thèse propose de développer des méthodes d'estimation statistique efficaces des élements d'un processus shot-noise (introduit dans la section suivante) pour des niveaux d'empilement variés[1] et vient compléter les méthodes d'estimation déjà existantes lorsque les processus shot-noise étudiés présentent peu ou pas d'empilement. En particulier, nos méthodes diffèrent de celles dites de "déconvolution" présentées par Andrieu et al. (2001) et Trigano et al. (2015b) qui proposent dans un premier temps de reconstruire le processus ponctuel marqué sous-jacent puis d'estimer la densité des marques et l'intensité du processus shot-noise. Les limitations de ces méthodes dans les cas de fort empilement sont détaillées dans la Section 1.3.

1.2.2 Formulation mathématique

Dans le contexte de cette thèse, les deux méthodes statistiques présentées aux Chapitres 4 et 5 pour résoudre le problème de la spectrométrie γ se basent sur l'hypothèse que nous

[1] On verra dans le Chapitre 3 que le processus shot-noise vérifie un théorème central limite lorsque l'intensité tend vers l'infini.

connaissons la réponse impulsionnelle générée par le détecteur et disposons d'un nombre grand mais fini d'observations du signal de sortie de chaîne.
D'un point de vue probabiliste, une manière de représenter les temps d'interaction des particules photoniques avec le détecteur, ainsi que les énergies qu'elles déposent, consiste à introduire un processus ponctuel marqué N[2]. Nous définissons alors

(**H**-1) $N = \sum_{k\in\mathbb{Z}} \delta_{T_k,Y_k}$ est un processus de Poisson ponctuel où les temps d'arrivée $(T_k)_{k\in\mathbb{Z}}$ suivent une loi de Poisson d'intensité homogène λ et dont les marques $(Y_k)_{k\in\mathbb{Z}}$ sont indépendantes des temps d'arrivée, i.i.d. et possèdent une densité de probabilité (par rapport à la mesure de Lebesgue) notée f.

Cette hypothèse se justifie par le fait qu'une source radioactive est composée d'un mélange de nombreux noyaux excités qui émettent des photons de manière quasi indépendante les uns des autres. La modélisation des temps d'arrivée $(T_k)_k$ par un processus de Poisson ponctuel d'intensité homogène λ traduit les deux propriétés suivantes adaptées à ce constat :

1

- absence de mémoire : le nombre de photons interagissant avec le détecteur avant un certain temps t est indépendant du nombre de photons interagissant avec le détecteur après ce temps t, et ce quel que soit le temps t considéré.

- homogénéité dans le temps : la loi du nombre d'arrivée dans un intervalle de durée τ fixé est indépendante de l'intervalle de temps considéré.

La densité f correspond alors au spectre de raies des noyaux composant la source déformé par les interactions et la chaîne d'acquisition tandis que l'intensité λ, exprimée usuellement en s^{-1} (également en cps signifiant *coups par seconde*), correspond au nombre moyen de photons détectés dans un intervalle de temps d'une seconde.
Lors de chaque interaction photon-détecteur, nous avons vu dans la partie 1.1.3 qu'un courant électrique proportionnel à l'énergie déposée est mesuré. Nous appelons alors *réponse impulsionnelle* la fonction h correspondant au signal mesuré et supposons qu'elle vérifie l'hypothèse

(**H**-2) h est une fonction mesurable, causale, à valeurs réelles et telle que

$$|h(t)| \leq C_0 \mathrm{e}^{-\alpha t} \ , \quad t \in \mathbb{R}_+ \ .$$

pour C_0 et α deux constantes positives.

Une telle hypothèse sur la fonction h est relativement faible et permet de considérer des fonctions mesurables non nécessairement à support compact. Dans les applications en spectrométrie γ, les réponses impulsionnelles ressemblent aux fonctions esquissées en Figure 1.11.

[2]Une définition précise est donnée dans le Chapitre 3.

Le courant électrique peut alors être modélisé par un processus shot-noise stationnaire défini par

$$X_t := \sum_{k \in \mathbb{Z}: T_k \leq t} Y_k \, h(t - T_k), \quad t \geq 0 \,. \tag{1.2.1}$$

Étant donné qu'un processus shot-noise est entièrement caractérisé par la donnée du triplet intensité-densité des marques-réponse impulsionnelle (λ, f, h), nous désignerons souvent dans cet ouvrage par "processus shot-noise de triplet (λ, f, h)" tout processus shot-noise vérifiant les hypothèses (**H**-1) et (**H**-2) d'intensité λ, de densité des marques f et de réponse impulsionnelle h. Deux exemples de trajectoires de processus shot-noise différant seulement par leur intensité sont présentés en Figure 1.13.

Toutefois, les méthodes de mesure existantes ne permettent pas d'enregistrer une version continue du courant électrique. Le signal mesuré en sortie de la chaîne d'instrumentation correspond à un échantillonnage basse fréquence[3] du processus shot-noise défini par (1.2.1). Ainsi, nous disposons seulement d'un nombre fini d'observations $X_\delta, \ldots, X_{n\delta}$, où n et δ correspondent respectivement à taille de l'échantillon disponible et à l'inverse de la fréquence d'échantillonnage.

1

1.2.3 Problématiques

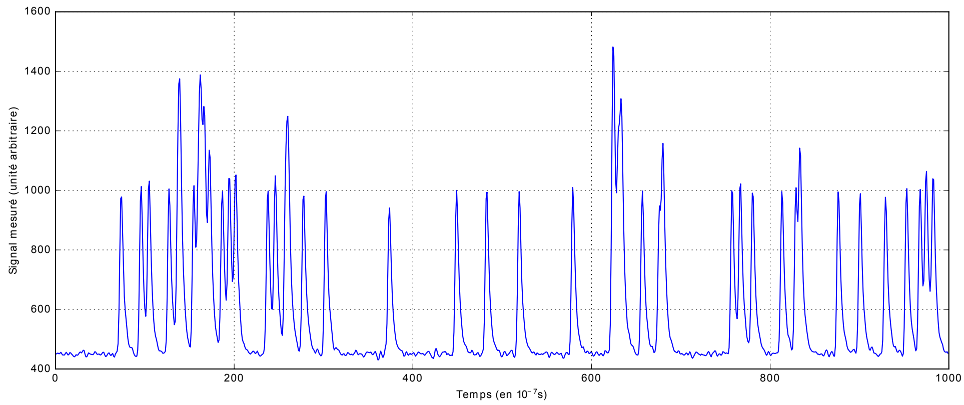
L'objectif de cette thèse consiste, dans un premier temps, à construire, à partir des observations discrètes $X_\delta, \ldots, X_{n\delta}$ du processus shot-noise $(X_t)_{t \in \mathbb{R}}$ défini par (1.2.1), des estimateurs de la densité des marques f ainsi que de l'intensité λ sous l'hypothèse additionnelle que la réponse impulsionnelle h est connue. De plus, l'échantillonnage à basse fréquence du signal rend les lois jointes de $(X_t)_{t \in \mathbb{R}}$ difficilement exploitables, de telle sorte que nous considérons dans cette thèse uniquement des méthodes reposant sur la marginale du processus.

Dans un second temps, et lorsque cela est possible, nous étudions les performances asymptotiques de ces estimateurs. Des résumés détaillés des méthodes introduites dans ce manuscrit sont disponibles en Section 1.4.1 de ce chapitre.

1.3 État de l'art

Nous présentons dans cette section une sélection des méthodes d'estimation des éléments d'un shot-noise existantes. Dans un premier temps, nous détaillons plusieurs méthodes utilisées par les physiciens.

[3] Les fréquences d'échantillonnage peuvent atteindre 100 MHz. Nous utilisons ici le terme basse fréquence par opposition à la définition usuelle en statistique du terme haute fréquence, signifiant que la fréquence d'échantillonnage peut être rendue arbitrairement grande.



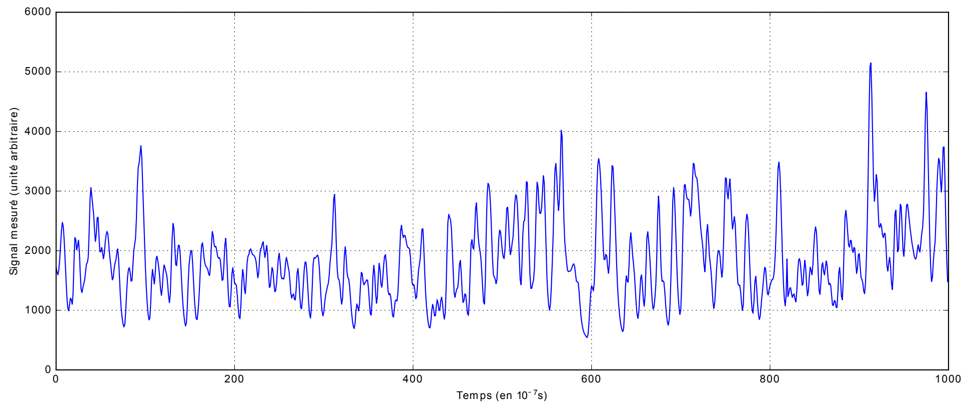


Figure 1.13 : Trajectoires de deux processus shot-noise enregistrés en spectrométrie γ dont la source est composée d'Americium 241 et de Plutonium 238. Les intensités varient et sont respectivement du haut vers le bas de 5.10^5 et 4.10^6 coups par seconde

1.3.1 Méthodes utilisées en physique

Encore aujourd'hui, les méthodes classiques d'estimation de la densité des marques sont largement utilisées en spectrométrie γ et en neutronique. La première, et la plus répandue, consiste à mesurer les maxima locaux du signal échantillonné puis d'en établir un histogramme. En effet, cette méthode permet d'estimer une version "dilatée", de la densité f associée au processus shot-noise $(X_t)_{t\geq 0}$. Néanmoins, cette technique présente une limite car elle n'est efficace que dans les cas où le support de la réponse impulsionnelle est plus court que le temps moyen entre deux interactions photon-matière. Cette configuration correspond au cas où le spectromètre est capable de discriminer de manière presque systématique chaque interaction des photons γ avec le détecteur. Une illustration de cette technique est présentée en Figure 1.14 pour des processus shot-noise de triplet (λ, f, h) dont l'intensité λ prend successivement les valeurs $\{0.2, 1, 5, 10\}$, de même réponse impulsionnelle h définie par

$$h : t \to 10t\, \mathrm{e}^{1-10t} \mathbb{1}_{t\geq 0} ,$$

et dont les marques de densité f suivent une loi Gaussienne de moyenne 10 et de variance 1.
L'histogramme obtenu fournit un estimateur de la fonction

$$\mathbb{R} \to \mathbb{R} , x \to \frac{1}{h_{\max}} f\left(\frac{x}{h_{\max}}\right) \quad \text{où} \quad h_{\max} := \sup_{t\in\mathbb{R}} h(t) .$$

Dans le cas où les données sont fortement empilées, l'histogramme résultant par cette méthode présente des "plateaux" et des pics multiples, appelés *artefacts*, qui correspondent généralement aux multiples ou combinaisons des modes de la densité f. Ceci est visible sur les deux derniers histogrammes de la Figure 1.14 dans lesquels on distingue nettement un maxima local au point d'abscisse 20. Nous voyons donc que cet estimateur est généralement trop grossier pour permettre de discriminer les raies γ caractéristiques d'un noyau donné dès lors que l'intensité augmente.
Afin de pallier partiellement aux limites de cette technique, les praticiens recourent à la méthode des cumulants détaillée dans le Chapitre 3. Il s'agit de la formule de Campbell pour les processus ponctuels, permettant d'exprimer les cumulants de la loi marginale du shot-noise $(X_t)_{t\geq 0}$ définis via la formule

$$\kappa_X^n = \lambda\, \mathbb{E}_f\left[Y^n\right] \int_0^\infty h^n(s)\mathrm{d}s , \quad n \in \mathbb{N}^\star . \tag{1.3.1}$$

Pour une réponse impulsionnelle h donnée, et lorsque la densité de probabilité des marques f ou l'intensité λ régissant les temps d'arrivée du processus shot-noise est connue, on remplace dans les équations (1.3.1) les cumulants de la loi marginale X_0 par leur version empirique $(\hat{\kappa}_X^n)_{n\geq 1}$ afin d'obtenir des estimateurs de λ ou f. Dans le cas où f

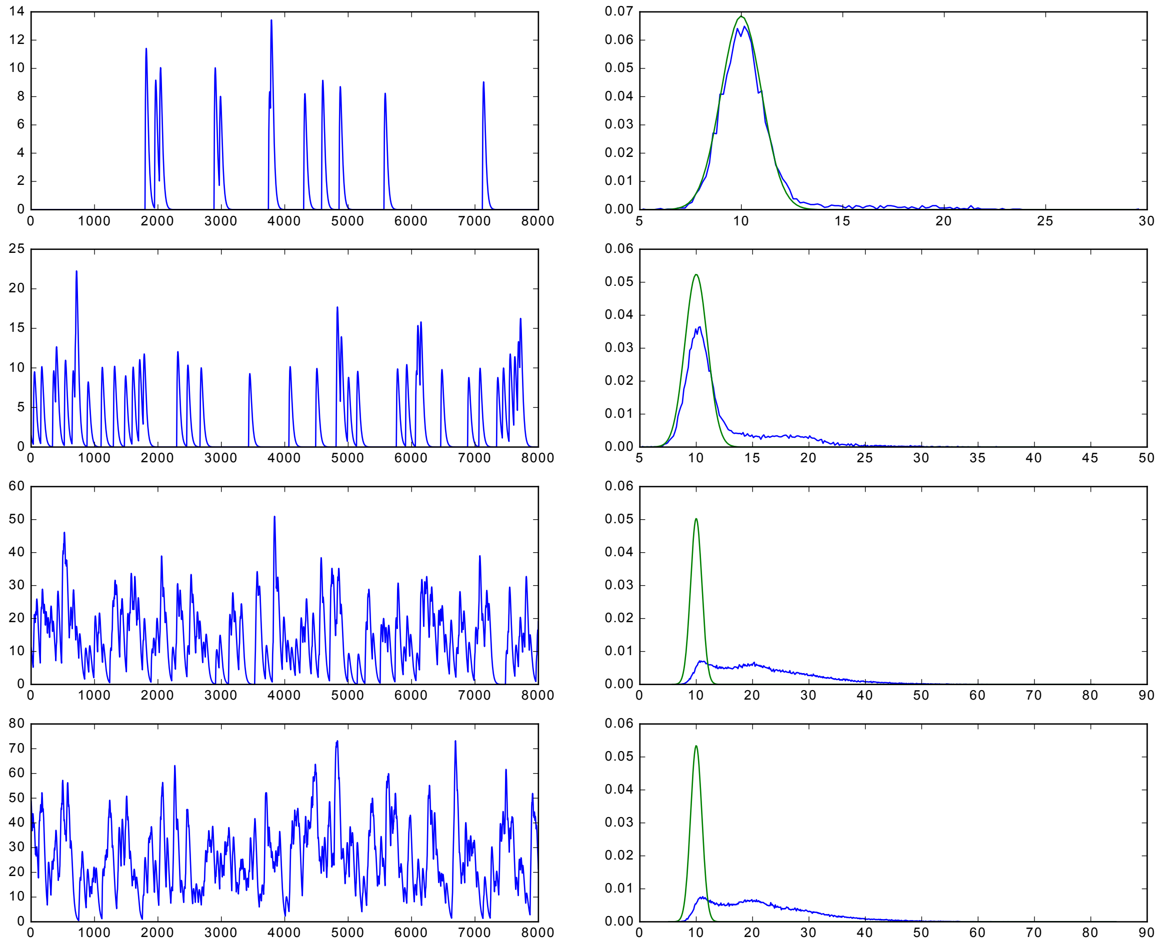


Figure 1.14 : Trajectoires de processus shot-noise (λ, f, h) simulées pour $\lambda \in \{0.2, 1, 5, 10\}$ de même densité des marques f et réponse impulsionnelle h (à gauche, de haut en bas). A droite, histogramme des maxima locaux correspondants pour des échantillons de taille $n = 5.10^6$

est connue, on dispose d'un nombre considérable d'estimateurs de λ donnés par

$$\hat{\lambda}_m = \frac{\hat{\kappa}_X^m}{\mathbb{E}_f\left[Y^m\right] \int\limits_0^\infty h^m(s)\mathrm{d}s}\ , m \in \mathbb{N}^\star\ .$$

Cette approche est considérée dans Elter et al. (2016) mais il n'est pas précisé comment les estimateurs ci-dessus sont combinés – en particulier, quelle combinaison linéaire des estimateurs $(\hat{\lambda}_n)_n$ fournit un estimateur de variance minimale – afin d'estimer λ. Toutefois, lorsque le cumulant d'ordre k est estimé à l'aide des k-statistiques, la variance de cet estimateur est proportionnelle au ratio λ^k/n où n est le nombre de points échantillonnés. Il s'ensuit que, pour une intensité λ suffisamment grande, les cumulants d'ordre élevés possèdent une variance trop importante pour permettre une estimation efficace de l'intensité à moins que l'on soit capable d'observer un grand nombre d'échantillons.

1

Dans le cas où λ est connu, un estimateur du moment d'ordre n de Y est donné par

$$\widehat{\mathbb{E}_f\left[Y^m\right]} = \frac{\hat{\kappa}_X^m}{\lambda \int\limits_0^\infty h^m(s)\mathrm{d}s}\ , \quad m \in \mathbb{N}^\star\ .$$

Il est alors possible de construire un estimateur de la fonction caractéristique de Y utilisant les M premiers moments $\widehat{\mathbb{E}_f\left[Y^1\right]}, \ldots, \widehat{\mathbb{E}_f\left[Y^M\right]}$ défini par

$$\begin{aligned}\hat{\varphi}_M(u) &:= 1 + \sum_{k=1}^{M} \frac{\widehat{\mathbb{E}_f\left[Y^k\right]}}{k!}(iu)^k \\ &= 1 + \sum_{k=1}^{M} \frac{\hat{\kappa}_X^k}{\lambda\ k! \int\limits_0^\infty h^k(s)\mathrm{d}s}(iu)^k\ , \quad u \in \mathbb{R}\ .\end{aligned} \tag{1.3.2}$$

Par suite, en appliquant la transformée de Fourier à la fonction $\hat{\varphi}_M$, on obtient un estimateur $\hat{f}_M$ de la densité f. Les résultats de cette méthode sont présentés dans le chapitre 3.

1.3.2 Avancées mathématiques récentes

La contribution des statisticiens au problème d'estimation des éléments caractéristiques (λ, f, h) du processus shot-noise défini par (1.2.1) est plus récente. Dans la thèse Trigano (2005), l'auteur étudie également les problèmes d'empilements en spectrométrie gamma et présente des méthodes nouvelles permettant de caractériser les éléments radioactifs à l'origine de la source émettrice de photons qui restent efficaces à des forts taux de comptage. Cependant, ces techniques reposent sur des estimateurs construits à partir d'échantillons différents ainsi que sur des hypothèses plus restrictives. En effet, le modèle considère des paquets d'impulsions représentés par le processus ponctuel:

- $N = \sum \delta_{T_k,(X_k,Y_k)}$ est un processus de Poisson marqué d'intensité λ sur $\mathbb{R}_+$, où la suite $(T_k)_k$ correspond aux temps d'arrivées des photons, $(X_k)_k$ à la durée des impulsions électriques respectivement associées à ces particules et $(Y_k)_k$ aux énergies émises par chacun des photons. En outre, les marques $(X_k, Y_k)_k$ sont supposées i.i.d. et indépendantes des temps d'émission $(T_k)_k$.

L'objectif poursuivi par Trigano (2005) consiste, à partir d'un échantillon $(X'_k, Y'_k, Z'_k)_{1\leq k\leq N}$ de taille N où Z'_k est un estimateur de la k-ième séquence *idle* définie par

$$Z_k := T_{k+1} - (X_k + T_k) ,$$

à construire un estimateur non-paramétrique de la loi des marques Y_0. Cependant, la caractérisation d'une impulsion par une paire de variables aléatoires (X, Y) où X et Y représentent respectivement la durée de l'impulsion et l'énergie déposée par le photon associé est difficilement généralisable à des réponses impulsionnelles à support infini. Comment définir une durée d'impulsion lorsque la réponse impulsionnelle h est de la forme

$$h : t \to \mathrm{e}^{-\alpha t}\mathbb{1}_{t\geq 0}, \quad \alpha > 0 \ ? \tag{1.3.3}$$

Une troncation "brutale" de cette fonction- par exemple en considérant que h est nulle en dehors de l'intervalle $[0, 3/\alpha]$- induirait des erreurs non négligeables à mesure que l'intensité λ croît. Nous verrons dans le Chapitre 4 comment construire un estimateur de la densité f des marques lorsque la réponse impulsionnelle h est donnée par (1.3.3).
Récemment, les travaux Lopatin et al. (2012) et Trigano et al. (2015b) se rapprochent de ceux entrepris dans cette thèse. Dans chacun de ces articles, on observe un échantillon $\mathbf{x} = (x_0, \ldots, x_{N-1})$ de taille finie N provenant d'un processus shot-noise échantillonné à basse fréquence δ^{-1} donné par

$$x_k := \sum_{k=0}^{\infty} Y_k \Phi_k(n\delta - T_k) + \epsilon_n, \quad k = 0, \ldots, N-1 , \tag{1.3.4}$$

où le vecteur $\epsilon = (\epsilon_1, \ldots, \epsilon_N)$ est tel que ses composantes sont des variables aléatoires i.i.d. suivant une loi gaussienne centrée de variance σ^2 et la fonction Φ_k est définie par

$$\Phi_k(t) \equiv \Phi_{(\theta_1^{(k)},\theta_2^{(k)})}(t) := t^{\theta_1^{(k)}} \mathrm{e}^{\theta_2^{(k)} t} \mathbb{1}_{\{0\leq t\leq (N-1)\delta\}} , \tag{1.3.5}$$

où on suppose que les paramètres $\theta_1^{(k)}$ et $\theta_2^{(k)}$ appartiennent respectivement aux intervalles $(0, \Theta_1)$ et $(0, \Theta_2)$ pour Θ_1, Θ_2 connus.
A la différence du modèle étudié dans cette thèse et présenté au Chapitre 3, celui-ci admet deux différences notables : d'une part, la version stationnaire du signal est considérée. D'autre part, la réponse impulsionnelle n'est pas invariante aux photons incidents : chaque interaction peut générer une réponse électrique propre, bien que toujours proportionnelle à l'énergie déposée dans le détecteur. Cette variation est rendue possible par une paramétrisation (θ_1, θ_2) bidimensionnelle de la fonction $\Phi_{(\theta_1,\theta_2)}$ définie par (1.3.5). L'algorithme proposé dans ces travaux suppose également que le signal

1

échantillonné est parcimonieux. Dans le cadre des processus shot-noise, cela est possible si l'on suppose que l'intensité λ est faible ou que la réponse impulsionnelle possède un support extrêmement court. L'approche abordée par le papier consiste alors à estimer les temps d'arrivée $(T_k)_k$ et les énergies $(Y_k)_k$ du processus ponctuel marqué N en utilisant la méthode itérative suivante :

- Pour p, q deux entiers, on note U_1^p et U_2^q les discrétisations régulières de $(0, \Theta_1)$ et $(0, \Theta_2)$ en p et q points respectivement.

- Pour tout $n \in \{0, \ldots, N-1\}$, on définit le vecteur A_n de toutes les contributions électriques individuelles pouvant apparaître dans l'observation x_n

$$A_n = \left[\Phi_{(j,l)}(k-n)\right]^{\mathrm{T}} \ , \quad 0 \leq k \leq N-1 \ , \ j \in U_1^p \ , \ l \in U_2^q \ ,$$

  et on appelle $\mathbf{A} = [A_0 \, A_1 \, \cdots \, A_{N-1}]$ le dictionnaire formé par concaténation des $(A_j)_{0 \leq j \leq N-1}$.

- Le vecteur des observations $\mathbf{x}$ peut alors s'écrire

$$\mathbf{x} = \mathbf{A}\beta + \epsilon + \mathbf{\Delta}$$

  où $\mathbf{x}$, ϵ sont définis par (1.3.4), $\mathbf{\Delta}$ correspond au vecteur quantifiant l'erreur d'approximation des T_k par leur temps d'échantillonnage le plus proche et β est un vecteur de taille Npq qui contient les informations relatives aux énergies $(Y_k)_{k \geq 0}$ ainsi qu'aux formes des réponses impulsionnelles associées. Sous l'hypothèse de faible empilement, ce vecteur peut raisonnablement être supposé parcimonieux. De plus, les indices non nuls correspondent aux temps d'émission des photons interagissant avec le détecteur. Une manière d'estimer ce vecteur consiste alors à utiliser la méthode LASSO (Least Absolute Shrinkage and Selection Operator, voir Tibshirani (1996) pour un exposé détaillé) qui, pour tout réel positif μ, définit le vecteur $\hat{\beta}(\mu)$ comme la solution du problème de minimisation

$$\hat{\beta}(\mu) = \underset{\beta \geq 0}{\operatorname{argmin}} \ \|\mathbf{y} - \mathbf{A}\beta\|_2^2 + \mu \|\beta\|_1 \ .$$

- Le vecteur $\hat{\beta}(\mu)$ est alors utilisé pour estimer récursivement les temps d'arrivée $(T_k)_k$ via

$$\begin{cases} \hat{T}_0 & = & \min\{0 \leq m \leq N-1 : \|\hat{\beta}(\mu)_m\|_1 \geq s\} \\ \hat{T}_k & = & \min\{\hat{T}_{k-1}/\delta < m \leq N-1 : \|\hat{\beta}(\mu)_m\|_1 \geq s, \|\hat{\beta}(\mu)_{m-1}\|_1 \leq s\} \ , \quad k \geq 1 \ . \end{cases}$$

- En définissant

$$M := \left|\{\hat{T}_k, \hat{T}_k \leq N-1\}\right| \qquad \text{et} \quad \mathbf{x}^{(0)} := \mathbf{x} \ ,$$

les énergies $\hat{Y}_1, \ldots, \hat{Y}_M$ correspondant aux temps d'arrivée $\hat{T}_1, \ldots, \hat{T}_M$ sont successivement estimées via les formules

$$\begin{cases} \hat{\beta}(\mu_k) &= \underset{\beta \geq 0}{\operatorname{argmin}} \|\mathbf{x}^{(k)} - \mathbf{A}\mathbf{W}^{(k)}\beta\|_2^2 + \mu_k \|\beta\|_1 \\ \hat{Y}_k &= \|\mathbf{A}\mathbf{W}^{(k)}\hat{\beta}(\mu_k)\|_1 \\ \mathbf{x}^{(k+1)} &= \mathbf{x}^{(k)} - \mathbf{A}\mathbf{W}^{(k)}\hat{\beta}(\mu_k) \, , \end{cases}$$

où $(\mu_k)_{k \leq N-1}$ est une suite de valeurs positives et $(\mathbf{W}^{(k)})_k$ une suite de matrices diagonales de taille N qui retranscrivent la contribution des énergies (via le vecteur $\hat{\beta}(\mu_k)$) au point d'échantillonnage $\hat{T}_k$. En général, pour un k donné, la diagonale de la matrice $\mathbf{W}^{(k)}$ correspond à l'évaluation d'un noyau gaussien de moyenne $\hat{T}_k$. La densité des marques f est ensuite evaluée en utilisant un estimateur à noyau K ainsi qu'une bande passante $h > 0$ par

$$\hat{f}_Y(x) := \frac{1}{Mh} \sum_{k=1}^{M} K\left(\frac{\hat{Y}_k - x}{h}\right), \quad x \in \mathbb{R} \, .$$

Néanmoins, bien que les hypothèses de ces papiers soient plus faibles que les nôtres et généralisent le modèle du shot-noise, l'idée selon laquelle le signal échantillonné est parcimonieux est difficilement vérifiée dès lors que le taux de comptage λ croît, pour une réponse impulsionnelle donnée, comme illustré par la Figure 1.9. Ces méthodes ou celles de Andrieu et al. (2001) ne sont alors plus efficaces en présence de fort empilement.
Dans la prochaine partie, nous détaillons les estimateurs de l'intensité λ et de la densité f obtenus dans le cadre cette thèse. En particulier, nous verrons que les estimateurs de λ sont beaucoup moins sensibles aux variations de l'intensité λ et comment les bornes de convergence établies dans le Chapitre 4 dépendent de cette quantité.

1.4 Contributions

1.4.1 Méthodes proposées et résultats obtenus

Dans cette thèse, nous traitons le problème d'estimation de la densité des marques f et de l'intensité λ comme une problème inverse linéaire. Ceci est rendu possible grâce à l'expression de la fonction caractéristique φ_{X_0} de la loi marginale X_0 du processus shot-noise $(X_t)_{t \in \mathbb{R}}$ défini par (1.2.1). En effet, on a pour tout réel u

$$\varphi_{X_0}(u) = \exp\left(\lambda \int_0^\infty [\varphi_Y(uh(s)) - 1] \, \mathrm{d}s\right) . \tag{1.4.1}$$

Lorsque $\mathbb{E}[|Y_0|] < \infty$, la fonction φ_{X_0} est dérivable et on a

$$\frac{\varphi'_{X_0}(u)}{\varphi_{X_0}(u)} = K_h[f_\lambda](u) \, , \quad u \in \mathbb{R} \, , \tag{1.4.2}$$

où $f_\lambda := \lambda f$ et K_h est un opérateur détaillé au Chapitre 5 qui est défini, pour toute fonction ξ à valeurs réelles satisfaisant $\int_0^\infty x|\xi|(x)\mathrm{d}x < \infty$ et tout réel u, par

$$K_h[\xi](u) = \int_0^\infty \left[\int_0^\infty \mathrm{i}xh(s)\mathrm{e}^{\mathrm{i}uxh(s)}\mathrm{d}s\right]\xi(x)\mathrm{d}x \,. \tag{1.4.3}$$

On dispose également d'un estimateur de la fonctionnelle $\varphi'_{X_0}/\varphi_{X_0}$ qui est obtenu par une méthode de type "plug-in" à partir des observations $X_\delta, \ldots, X_{n\delta}$. En notant respectivement $\hat{\varphi}_n$ et $\hat{\varphi}'_n$ la fonction caractéristique empirique associée à $X_\delta, \ldots, X_{n\delta}$ et sa dérivée, un estimateur de $\varphi'_{X_0}/\varphi_{X_0}$ est défini par

$$\hat{g}_n(u) := \frac{\hat{\varphi}'_n(u)}{\hat{\varphi}_n(u)} = \frac{\sum_{k=1}^n \mathrm{i}X_{k\delta}\mathrm{e}^{\mathrm{i}uX_{k\delta}}}{\sum_{k=1}^n \mathrm{e}^{\mathrm{i}uX_{k\delta}}} \qquad u \in \mathbb{R} \,.$$

1

Le Chapitre 4 traite du cas particulier où la réponse impulsionnelle h est exponentielle, c'est à dire de la forme

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

pour un certain α positif. Ce cas est considérablement différent du général car, d'une part le processus shot-noise associé est un processus de Markov stationnaire et, d'autre part, la fonctionnelle $\varphi'_{X_0}/\varphi_{X_0}$ definie par (1.4.2) se reécrit

$$\frac{\varphi'_{X_0}(u)}{\varphi_{X_0}(u)} = \frac{\lambda}{\alpha}\cdot\frac{\varphi_Y(u)-1}{u} \,, \quad u \in \mathbb{R} \,.$$

Lorsque le ratio λ/α est connu, il est alors possible d'obtenir un estimateur de la fonction f par une inversion de Fourier comme défini par (4.2.6) : pour deux suites positives $(h_n)_n$ et $(\kappa_n)_n$ bien choisies qui tendent vers 0, on définit l'estimateur $\hat{f}_n$ de f par

$$\hat{f}_n(x) := \frac{1}{2\pi}\int_{-h_n^{-1}}^{h_n^{-1}} \mathrm{e}^{-\mathrm{i}ux}\left(1 + u\frac{\alpha}{\lambda}\frac{\hat{\varphi}'_n(u)}{\hat{\varphi}_n(u)}\mathbb{1}_{\{|\hat{\varphi}_n(u)|\geq\kappa_n\}}\right)\mathrm{d}u \,, \quad x \in \mathbb{R} \,.$$

Une étude théorique permet alors de déterminer des vitesses de convergence, en norme uniforme, uniformément sur les classes de régularités Sobolev $\Theta(K, L, m, s)$ définies, pour des réels positifs K, L, m et $s > 1/2$, par

$$\Theta(K, L, m, s) := \{f : \mathbb{R}_+ \to \mathbb{R} \text{ t.q. } \int_0^\infty |x|^{4+m} f(x)\mathrm{d}x \leq K \text{ and } \int_{\mathbb{R}} |\mathcal{F}[f](u)|^2 \, |u|^{2s}\mathrm{d}u \leq L\}$$

Le Théorème 4.2.1 assure que, si les suites $(h_n)_n$ et $(\kappa_n)_n$ sont choisies sous la forme

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

pour K, L, m et $s > 1/2$ donnés et C défini par

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

alors il existe un réel M dépendant seulement des constantes K, L, m, s et λ/α tel que, pour tout $n \geq 3$,

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

L'obtention des vitesses de convergence présentées précedemment nécessite qu'on suppose que le ratio λ/α soit connu. Cependant, il est indiqué dans le Chapitre 4 comment estimer cette quantité à l'aide de l'estimateur de Hill construit à partir de l'echantillon $X_\delta^{-1}, \ldots, X_{n\delta}^{-1}$.

Dans le Chapitre 5, nous verrons qu'il est difficile d'étudier l'opérateur K_h défini par (1.4.3), et en particulier de déterminer pour quelles fonctions h celui-ci est compact. Nous privilégions alors l'usage de méthodes paramétriques afin d'estimer la fonction f_λ. Sous l'hypothèse supplémentaire que f appartient à un espace de Sobolev d'ordre 2 défini par

$$\mathcal{W}^2 := \{ f : [0,1] \to \mathbb{R} : \quad f, f' \text{ absolument continues,} \\ f(1) = f'(1) = 0 \text{ and } \int_0^1 |f''(t)|^2 \, \mathrm{d}t < \infty \} ,$$

muni de la norme

$$\|f\|_{\mathcal{W}^2} := \sqrt{\int_0^1 |f''(t)|^2 \, \mathrm{d}t} , \quad f \in \mathcal{W}^2 ,$$

nous proposons de construire un estimateur de f_λ à partir d'une grille finie de N points d'évaluation $\{u_i, i \in 1, \ldots, N\}$ pour lesquels on considère les équations

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

où $\epsilon(u_1), \ldots, \epsilon(u_N)$ sont les erreurs d'estimation de $g(u_i)$ par $\hat{\varphi}_n'(u_i)/\hat{\varphi}_n(u_i)$ pour $i = 1, \ldots, N$. Comme on suppose que f_λ appartient à l'espace fonctionnel Hilbertien $\mathcal{W}_+^2$, défini comme le sous-ensemble des fonctions positives appartenant à $\mathcal{W}^2$, nous proposons de l'estimer en considérant le problème de minimisation défini, pour toute matrice W de taille N définie positive et tout réel positif μ, par

$$\hat{f}_{n,\lambda}(\mu) := \underset{\xi \in \mathcal{W}_+^2}{\operatorname{argmin}} \|\hat{g}_n(\underline{u}) - K_{\underline{u}}[\xi]\|_{W^{-1/2}}^2 + \mu \|\xi\|_{\mathcal{W}^2}^2 , \tag{1.4.4}$$

où

$$\|\hat{g}_n(\underline{u}) - K_{\underline{u}}[\xi]\|_{W^{-1/2}}^2 := \left\| W^{-1/2} \cdot \{\hat{g}_n(\underline{u}) - K_{\underline{u}}[\xi]\} \right\|_2^2 ,$$

et

$$\hat{g}_n(\underline{u}) := [\hat{g}_n(u_1) \ \cdots \ \hat{g}_n(u_N)]^{\mathrm{T}} \quad \text{et} \quad K_{\underline{u}}[\xi] := [K_h[\xi](u_1) \ \cdots \ K_h[\xi](u_N)]^{\mathrm{T}} .$$

Une première paramétrisation est alors possible en utilisant la propriété d'espace reproduisant de $\mathcal{W}^2$ qui fournit une représentation explicite de la solution du problème (1.4.4). La seconde méthode utilise la décomposition de la fonction f_λ dans la base des B-splines cubiques. Les algorithmes sous-jacents sont ensuite utilisés à la fois sur des données simulées ainsi que sur des données réelles provenant du CEA Saclay.

1.4.2 Plan de la thèse

Cette thèse s'articule de la manière suivante. Le chapitre 3 fournit un complément de l'état de l'art présenté dans ce chapitre en détaillant l'émergence des processus shot-noise dans la littérature scientifique depuis le début du 20ème siècle. Dans un second temps, nous rappelons les définitions et constructions des processus de Poisson sur la droite réelle afin de définir le processus stochastique "shot-noise" donné par (1.2.1) et d'établir les propriétés particulières de ce processus. L'accent est ensuite mis sur le lien étroit existant entre l'estimation statistique des processus shot-noise et des processus de Lévy. Enfin, nous répertorions les processus stochastiques similaires aux processus shot-noise ainsi que les méthodes d'estimation existantes associées.

Le chapitre 4 étudie un estimateur de la densité des marques d'un processus shot-noise exponentiel établi à partir d'une équation reliant la fonction caractéristique des marques à une fonctionnelle de la fonction caractéristique de la loi marginale du processus shot-noise et de sa dérivée. Sous l'hypothèse que la densité appartienne à un espace de Sobolev, nous établissons que l'estimateur converge uniformément vers la densité cible à une vitesse polynomiale qui dépend de l'intensité du processus shot-noise.

Le chapitre 5 s'intéresse à l'estimation de la loi des marques d'un processus shot-noise ainsi que de l'intensité du processus de Poisson sous-jacent régissant les évènements sous l'hypothèse que la réponse impulsionnelle du système est connue. A partir d'un échantillon basse-fréquence du processus et sous l'hypothèse que la densité appartienne à l'espace de Sobolev d'ordre 2 noté $\mathcal{W}^2$, deux méthodes sont utilisées pour construire des estimateurs : l'une utilise la propriété d'espace reproduisant de l'espace $\mathcal{W}^2$ et le théorème du représentant tandis que la seconde repose sur une paramétrisation utilisant la base des B-splines. Les résultats de ces méthodes sont ensuite illustrés sur des données simulées ainsi que des données réelles provenant du CEA Saclay.

1.5 PUBLICATIONS

Le contenu de cette thèse a, en partie, fait l'objet de plusieurs publications dans des journaux ou conférences dont les détails pour chacun d'entre eux sont donnés ci-dessous :

1.5.1 Articles de journaux

- **Nonparametric estimation of mark's distribution of an exponential shot-noise process** - P. Ilhe, E. Moulines, F. Roueff et A. Souloumiac – Publié dans *Electronic Journal of Statistics*, 2016.

- **Pileup correction of a shot-noise process** - P. Ilhe, E. Moulines, F. Roueff et A. Souloumiac – En fin d'écriture.

1.5.2 Conférences avec comité de lecture

- **Nonparametric estimation of mark's distribution of an exponential Shot-noise process** - P. Ilhe, E. Moulines, F. Roueff et A. Souloumiac – Présenté au *Colloque GRETSI*, 2015.
- **Désempilement non-paramétrique de la densité d'un processus shot-noise** - P. Ilhe, E. Moulines, F. Roueff et A. Souloumiac – Présenté aux *Journées des Statistiques*, 2016.
- **Nonparametric estimation of a shot-noise process** - P. Ilhe, E. Moulines, F. Roueff et A. Souloumiac – Présenté au *IEEE Workshop on Statistical Signal Processing*, 2016.

2 Introduction

Ce sont précisément les questions auxquelles il n'est pas de réponse qui marquent les limites des possibilités humaines et qui tracent les frontières de notre existence.

– Milan Kundera, *L'insoutenable légèreté de l'être*

Summary

In this chapter, we aim to introduce the reader to the field of gamma spectroscopy and to the challenges related to this instrumentation method. After some recalls of nuclear physics, we provide a detailed presentation of the various kinds of photon-matter interactions, of the data acquisition system associated to the measured signal and of their impact on the nuclear spectra we can at best hope to recover.
Then, we interpret the gamma spectroscopy problem from a probabilistic perspective and assume that the measured electric current can be represented by a shot-noise process. The identification of radionuclide is then related to the statistical estimation of shot-noise processes and we describe the main challenges tackled in this manuscript.
Firstly, we provide a detailed account of the main statistical and deterministic methods used up to now by physicists and point out their limit when the intensity of the radioactive source becomes too large. Secondly, we present an exhaustive summary of the statistiscal estimation methods developed in Chapters 4 and 5 of this thesis.

2.1 Gamma spectroscopy

The goal of this part consists in describing what is γ-spectroscopy. To this end, we recall some background of atomic physics.

2.1.1 Reminders of nuclear physics.

The notation ${}^{A}_{Z}X$ is used to denote a nucleus made of a chemical element X with Z neutrons, N protons and where $A = Z + N$ is commonly called *mass number* (in a similar way, Z is known as the *atomic number*). The neutrons and protons are called *nucleons*

and two nucleus with the same mass number A or the same number of neutrons Z are respectively called *isobar* and *isotopes*.
According to the quantum mechanics theory, each nucleus ${}^{A}_{Z}X$ admits a finite number of quantified energies. Those energy states are usually classified into two categories: the stable and unstable states. In most cases, the stable state (also called *ground* state) coincides with the state of lower energy. The other states are called *excited* because they exhibit an higher energy than the ground state.
In general, an excited nucleus is a consequence of α or β nuclear decay (the interested reader can refer to Tsoulfanidis (2013)). In order to attain its stable state, the excited nucleus ejects a photon carrying the remaining amount of energy. Therefore, we call *γ radiation* the reaction that, from an excited nucleus ${}^{A}_{Z}X^{\star}$, results into the same nucleus in a stable state plus a gamma photon whose energy is exactly equal to the difference between the energy level of the excited nucleus minus the energy level associated to the stable nucleus. We write this reaction as

$$ {}^{A}_{Z}X^{\star} \to {}^{A}_{Z}X + {}^{0}_{0}\gamma \ . $$

2

Gamma radiation was firstly observed in 1900 by Paul Villard. Because the energy of a given nucleus can only take a finite number of values, the energy carried by the ejected photons after a gamma radiation reaction is also discrete. Thus, **there exists a unique mapping between a nucleus and the sequence of the characteristic energetic rays that are emitted by gamma radiation.**

2.1.2 Gamma spectroscopy : definition and purposes

Definition 1 (Gamma spectroscopy)**.** *Gamma spectroscopy is a nuclear instrumentation technique which, given a data acquisition system, aims to both determine the distributions of the energies emitted by gamma photons and the mean number of emitted photons in a finite interval of time.*

The mapping between the energy rays present in the gamma spectrum and the radionuclide that constitute the radioactive source implies that γ spectroscopy is one of the most used instrumentation technique in nuclear physics when one needs to identify radioactive components. In practice, the detection of ionizing radiations such as γ rays is realized using physical components that can generate time local fluctuations of electrons when hit by photons, for instance a semiconductor detector. When hit by a photon, the detector generates an electric current (see Figure 2.9 for an example and Section 2.1.3 for details) which is proportional to the energy deposited by the particle. Among the several photon-matter interaction mechanisms, we detail in the following the three majors interactions.

The photoelectric effect

The photoelectric effect corresponds to the reaction during which a photon interacts with an electron that belongs to the external shell of an atom of the semiconductor. When the gamma photon carrying an energy E_γ large enough hits the semiconductor, it is totally absorbed while the ejected electron called *photoelectron* admits a kinetic energy given by

$$E = E_\gamma - B_e$$

where B_e represents the *binding energy* of the electron.
The photoelectric effect is the major photon-matter interaction that occurs for photons with low energies. Moreover, the probability p of such an interaction becomes more important as soon as the incident energy E_γ is low and it depends on numerous factors: in fact, there exists a formula called *photoelectric coefficient* given by

$$p = aN\frac{Z^n}{E_\gamma^m}(1 - \mathcal{O}(Z)) \ .$$

where m, n are constants depending on E_γ ranging between 3 and 5, N and Z respectively are the neutrons and photons number of the aforementioned atom, a is a constant and $\mathcal{O}(Z)$ corresponds to a corrective term of first order in Z. Figure 2.1 illustrates how this probability p fluctuates in function of the incident energy E_γ and the atomic number Z of the atom of the semiconductor.

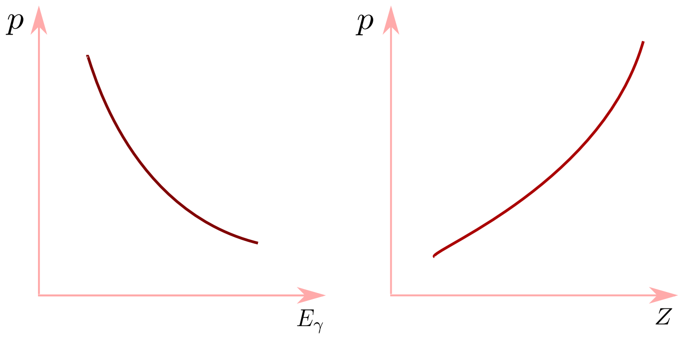


Figure 2.1 : Evolution of the probability that a photoelectric photon-matter interaction happens as a function of the photon energy E_γ and the protons' number Z

Compton scattering

The Compton scattering corresponds to the interaction between a photon and a free electron of the semiconductor. In contrast to the photoelectric effect, the incident photon does not disappear but is deviated from a polar angle θ and carries an inferior energy E'_γ while the freed electron is ejected as illustrated by Figure 2.2. The deviated photon can in turn interact with the semiconductor according to a probability that depends on parameters such as the size or the position of the semiconductor and, thus, is beyond the scope of this manuscript.

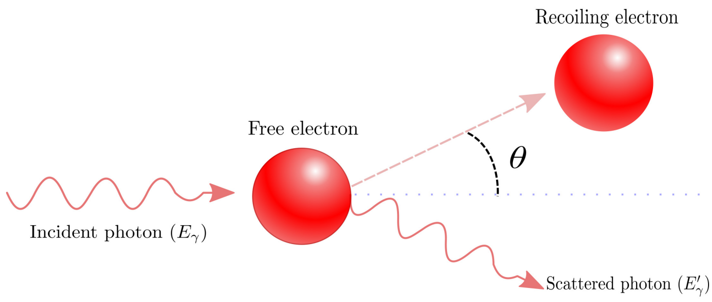


Figure 2.2 : Compton scattering

2

Applying the conservation and momentum equations, the formula that links the resulting energy E'_γ to the deviation angle θ and the incident energy E_γ is given by

$$E'_\gamma = \frac{E_\gamma}{1 + (1 - \cos(\theta))\, E_\gamma / m_e c^2} \ .$$

When a Compton scattering interaction occurs, the detector records the energy $E_\gamma - E'_\gamma$ of the freed electron. In the context of γ-spectroscopy, it is important to determine the range of energies of the scattered photon. When there is no deviation, corresponding to the case $\theta = 0$, there is no fluctuation of energy and $E'_\gamma = E_\gamma$. However, when the angle $\theta = \pi$, meaning that the scattered photon goes in the opposite direction of the incident photon, its energy is given by

$$E'_{\gamma,\mathrm{min}} = \frac{E_\gamma}{1 + 2E_\gamma / m_e c^2} \ .$$

However, the Compton diffusion is not isotropic, i.e. that the distribution of the deviation angles θ is not uniform over $[0, \pi]$. More precisely, the Klein-Nishina formula gives that the probability for a photon carrying an energy E_0 to be scattered to a solid angle $\mathrm{d}\Omega$ (defined as the 3 dimensional analogue of the planar angle) is given by

$$\frac{\mathrm{d}\sigma_{\mathrm{KN}}}{\mathrm{d}\Omega} := \frac{r_e^2}{2} \frac{1}{(1 + \alpha(1 - \cos\theta))^2} \left(1 + \cos^2\theta + \frac{\alpha^2 (1 - \cos\theta)^2}{1 + \alpha(1 - \cos\theta)}\right) \ ,$$

where r_e denotes the radius of an electron an $\alpha = E_0 / m_e c^2$. Figure 2.3 shows the different distributions of the deviation angles θ for several values of α.

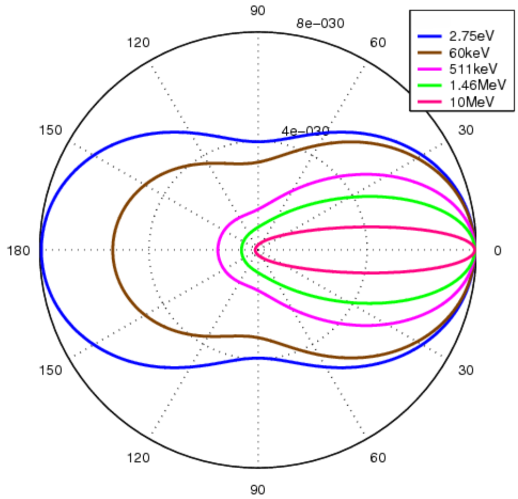


Figure 2.3 : Distribution of the deviation angles θ using the Klein-Nishina formula for 5 distinct values of α

In particular, when $E_\gamma \ll m_e c^2$, we notice that $E'_{\gamma,\min} \simeq E_\gamma$. This particular regime specifically happens at low energies and is called *Thompson diffusion*: whenever the photon energy is inferior to 45keV, the transferred energy by Compton scattering is negligible and only the photon direction is modified. In other terms, everything occurs as if the Compton scattering has no influence, meaning that one observes no distortion of the gamma ray spectrum, which can be seen in Figure 2.6.

Pair production

The third photon-matter interaction can only happen for high energy photons, namely for energies that are bigger than 1.022 MeV (i.e. twice the energy corresponding to the invariant mass of an electron). When a γ photon satisfies this condition, the phenomenon of pair production, schematically represented in Figure 2.4, becomes possible. This corresponds to an interaction between the incident photon and a nucleus that produces a creation of an electron-positron pair. The nucleus remains unchanged while the created positron annihilates with another electron and gives birth to two photons of energy 511keV emitted in opposite directions. More rarely, the pair production can also occur when two photons, carrying an energy bigger than 511 keV, interact. Even though it is difficult to analytically derive the probability that a photon-matter interaction causes a pair production, it is experimentally known that this probability is proportional to the square root of the atomic number Z of the absorbing nucleus.

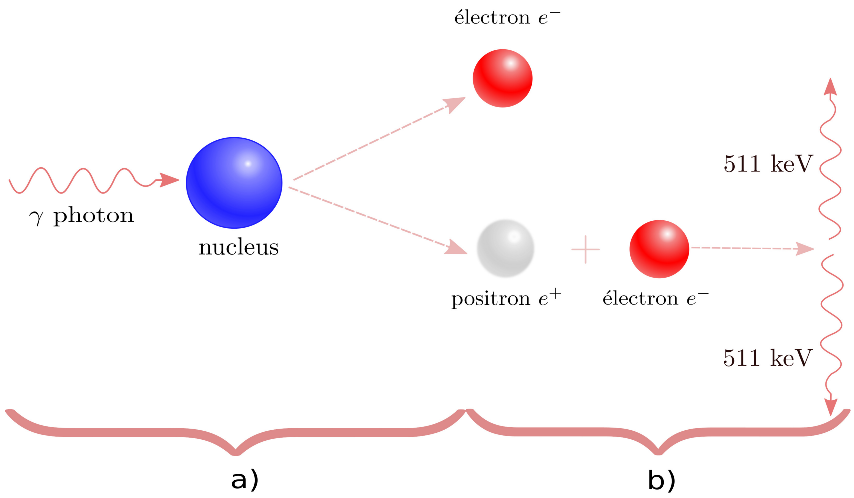


Figure 2.4 : Pair production interaction: a) Creation of an electron-positron pair. b) Annihilation phenomenon and emission of two 511 keV γ photons in opposite direction

The energy spectrum recorded by the detector is more involved than those measured in presence of Compton scattering or photoelectric effect. Indeed, if the positron annihilates with an electron, one or two gamma photons of 511 keV can be absorbed by the detector. On the contrary, these photons may also interact with matter at a later time. For an incident energy E_γ>1.022 MeV, the following four scenario can happen

- No pair production occurs and the detector records the energy E_γ,
- Only the photon carrying the energy $E_\gamma - 1,022$ MeV called *double peak escape* interacts with the detector,
- The photon with energy $E_\gamma - 1,022$ MeV and a photon resulting from the positron annihilation interact with the detector: an energy of $E_\gamma - 0,511$ MeV called *single peak espace* is detected,
- A gamma photon produced by the positron-electron annihilation interacts with the detector which records a 511 keV energy.

In summary, the three photon-matter interaction mechanisms described above follow very different behaviors according to the particles' incident energy and the atomic number of the absorbing nucleus. Figure 2.5 illustrates the different contributions of those three interactions in function of the incident energy of gamma photons and the atomic numbers Z of the atoms of the detector.
The figure 2.6, from Pirard (2006), summarizes, for an incident gamma photon with energy E_γ, the distribution of the energy spectrum recorded by the detector. The warmer the color, the more important the effect is.

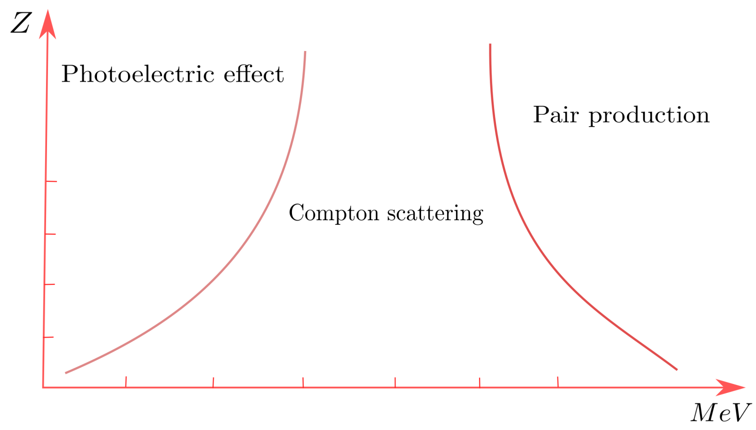


Figure 2.5 : Predominance of photon-matter interactions in function of the incident γ photon energy and the atomic number of the detector's atoms

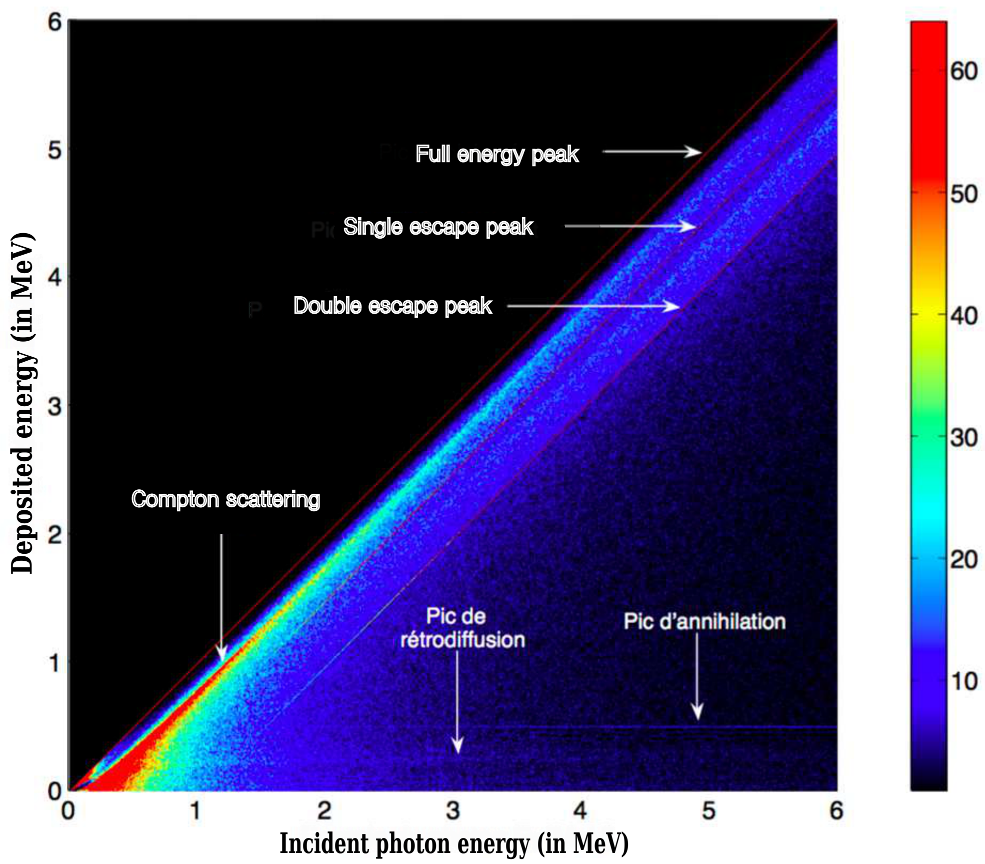


Figure 2.6 : Predominance of photon-matter interaction for different values of gamma photon incident energy E_γ

In practice, and in most cases, those interactions result in a more or loss important deformation of the gamma ray initial spectrum. It would then be unrealistic to hope to observe or estimate a discrete set of energy rays at the particular positions determined by the quantum energy levels of the radionuclides that form the radioactive source. As a consequence, we would not rely on parametric methods in order to model the gamma ray spectra of radioactive sources. Two types of spectra are displayed in Figure 2.7 for gamma photons stemming from a monoenergetic source: the two figures illustrates the different spectra one can observe whether the monoenergy E_γ is larger to $2m_ec^2 = 1.022MeV$ or not. Moreover, the data acquisition system used to perform γ-spectroscopy plays a dominant role in the resolution of the measured spectra. We briefly present the different components of this system within the next section.

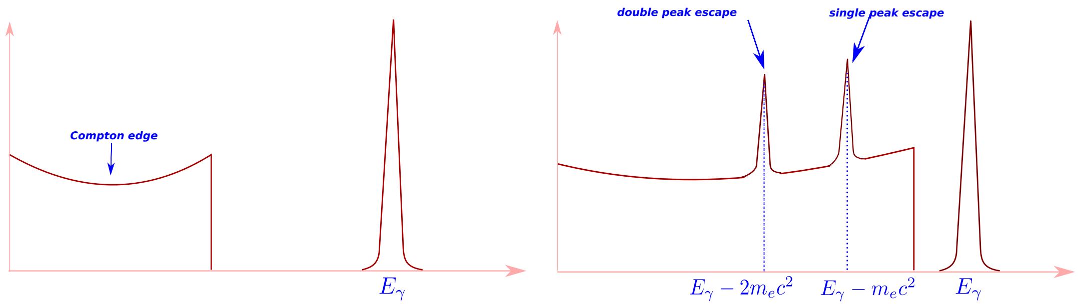


Figure 2.7 : Example of two realistic gamma ray spectra whether: a) $E_\gamma \leq 2m_ec^2$, b) $E_\gamma \geq 2m_ec^2$

2.1.3 Description of the instrumentation system

Up to now, we have said that the detector is the element of the data acquisition system that interacts with photons. More precisely, there exists two principal kinds of detectors: the *semiconductors* and the *scintillators* which are well detailed in Pirard (2006). In this manuscript, we settle for admitting that the three kinds of photon-matter interaction described above result in a creation of electron-hole pairs that is proportional to the deposited energy by the incident gamma photon. By migrating between the electrodes of the detector, they generate an electric impulsion that can be recorded with an instrumentation system, schematically illustrated in Figure 2.8. An example of the measured signal is available in Figure 2.9. We now detail the components of the instrumentation system.
The quality of a nuclear instrumentation system is often defined in terms of *resolution*. The resolution of a detector is defined as the observed variance around the width around the energy peak recorded by the detector in presence of a monoenergetic source. It allows to measure how well the nuclear instrumentation system can differentiate the various gamma energies carried by the incident photons. In Physics, the resolution of a data acquisition system is assessed via the following experiment.

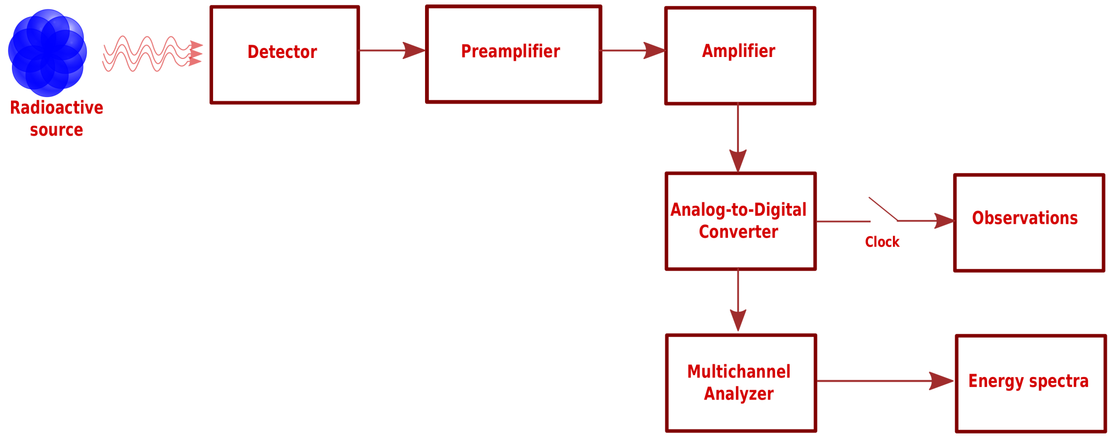


Figure 2.8 : Scheme of a data acquisition system in order to perform γ-spectroscopy

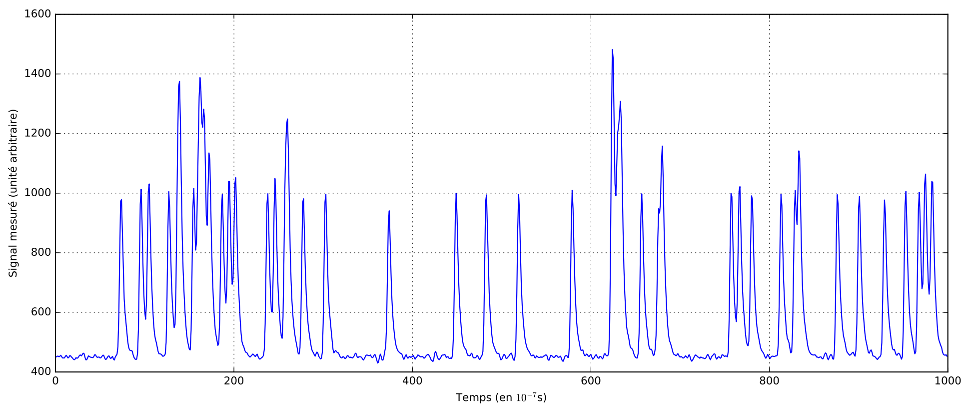


Figure 2.9 : Example of a measured signal used to perform γ-spectroscopy

From a monoenergetic source emitting photons with a unique energy E_γ, we compute the histogram of the local maxima of the signal recorded by the instrumentation system. For several reasons, like the statistical fluctuation of the number of created electron-hole pairs or the electronic noise of the different components of the acquisition system, the obtained histogram looks like a Gaussian function centered at E_γ. We call *FWHM (full width at half maximum)* the energy interval defined as the energy difference between points at which the histogram is equal to half of its maximum value, as illustrated by Figure 2.10.

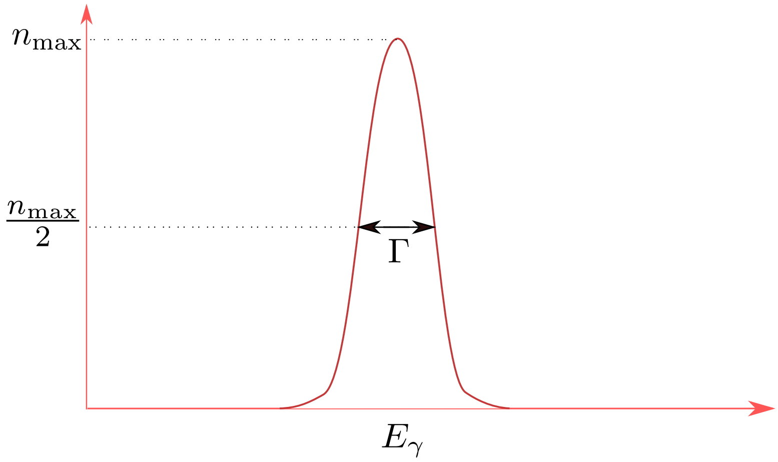


Figure 2.10 : Illustration of the detector's resolution around a peak of energy E_γ

This resolution may as well depends on the kind of the detector or on the choice of the amplifier. Moreover, in addition to the random number of electron-holes pairs created at each interaction, some effects are linked to the data acquisition system and the charge collection. Those effects then induce an sprawl that can be observed via the histogram of the recorded pulse heights. Moreover, the tuning of the parameters of the detector is quite limited. On the one hand, scintillators can generate shorter impulse responses but have a bad resolution. On the other hand, semiconductors have a resolution and longer impulse responses, which can be a hurdle when the mean number of photons emitted in one second by the source goes large. As we will see later, the methods developed in this thesis allow to handle any kind of impulse response, while the existing ones (see Section 2.3) are efficient only for short range pulses. In our case, it is thus preferable to use detectors with optimal energy resolutions.

The main purpose of the preamplifier consists in providing an optimal coupling between the output signal and the remaining of the instrumentation system. In addition, it has to be installed close to the detector in order to minimize the electric noise in the generated output signal. Then, this signal is amplified and shaped. The last step of the preamplifier consists in reducing the attenuation of the signal which is done by matching the impedance of the detector with the impedance of the amplifier.

Following the preamplifier in the data acquisition system, the amplifier converts the signal by transforming it into a impulse response in order to minimize the signal to noise ratio.

Usually, the amplifiers used to perform γ-spectroscopy generate two types of electric pulses that are illustrated by Figure 2.11 when a photon hits the detector.

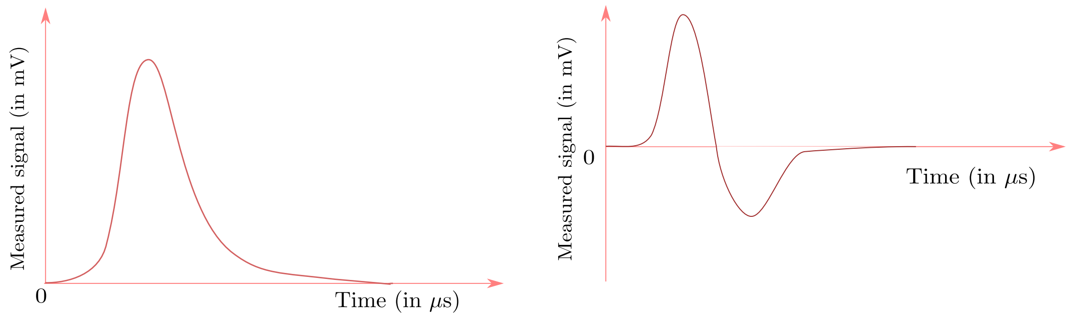


Figure 2.11 : Illustration of two impulse responses encountered in γ-spectroscopy

2

The multichannel analyzer (MCA) carries out a particular efficient procedure whenever the number of incident photons per second is low. For any electric pulse associated to the detection of a photon, the MCA records the highest value of the impulsion, which is proportional to the energy deposited at the detector. Then, this value is stored in the channel that corresponds to the interval which contains the measured value. For instance, for a Germanium detector, an high-energy resolution can be done using a 13 or 14 bits encoding (i.e. respectively 8192 and 16384 channels) in order to cover a range of energies between 0 and 8 MeV.

In practice, one needs to realize an energy calibration if we want to identify the different radioactive elements using γ-spectroscopy. To do that, a commonly manipulation consists in measuring the output signal when the instrumentation system records the photon-matter interactions that comes from a mono-energetic or bi-energetic source with a low mean number of photons by second. The histogram of the local maxima of the recorded signal has a unimodal or bimodal shape associated to the distribution of the energies of the source (as shown in the first histogram of Figure 2.14). This experiment allows to efficiently measure the multiplicative factor of the impulse response.

2.2 Challenges tackled in this thesis

In the scope of this manuscript, we aim to determine which chemical elements are in the radioactive source using γ-spectroscopy. This means that our goal consists in estimating the energy density of the emitted particles as well as the mean number of particles in an interval of one second using only a finite number of observations of the signal recorded by a data acquisition system (see Figure 2.9 for an example of such a signal).

Up until now, numerous methods have been developed by physicists and statisticians. The most used are detailed in Section 2.3. However, the vast majority of those methods becomes unusable as soon as the *intensity* of the source, i.e. the mean number of photons that the source emits in one second, becomes too important. This threshold varies in function of the nuclear instruments used to perform the measurements.

2.2.1 Pileup effect

The underperformances of those algorithms are mainly explained by the fact that *pileup* occurs: it corresponds to the phenomenon that appears when two or more photons hit the detector in such a close time interval that their associated single electric current generated by the detector overlap, as illustrated by Figure 2.12. This phenomenon generally appears when the counting rates become large or when the time support of the impulse response of the amplifier is not short enough.

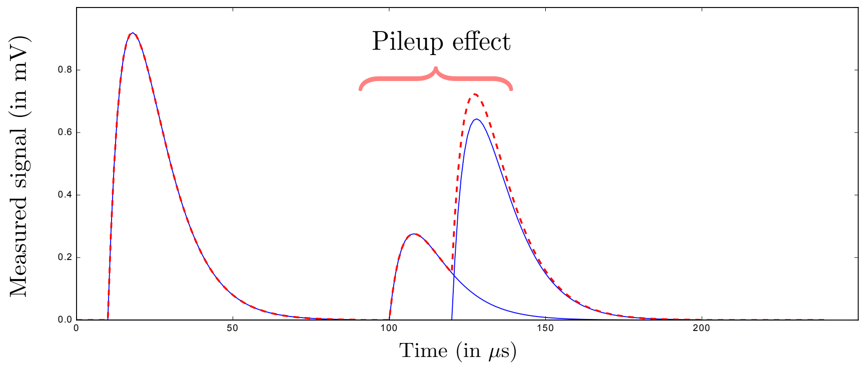


Figure 2.12 : Pileup effect

This thesis introduce new efficient statistical methods in order to estimate the different parameters of shot-noise processes (introduced in the next section) regardless of the intensity rate of the process[1]. These methods extend the already existing methods that are used to estimate the elements of shot-noise processes at low count rates. Moreover, our methods strongly differ to the "deconvolution" methods introduced in Andrieu et al. (2001) or Trigano et al. (2015b) which tackle a more involved problem: indeed, they first construct estimators of the underlying marked point process and then use it to estimate the marks' density and the intensity of the shot-noise process. The limitations of these methods in presence of pileup are detailed in Section 2.3.

2.2.2 Mathematical formulation

In this thesis, the two statistical methods introduced in Chapter 4 and 5 in order to solve the γ-spectroscopy problem rely on the assumption that we exactly know the impulse response generated by the nuclear instrumentation system and we have at hand a large but finite number of observations of the output signal.

[1]We will see in Chapter 3 that shot-noise processes satisfy a Central Limit Theorem when the intensity associated to the Poisson measure goes to infinity.

From a probabilistic perspective, one way to model the interaction times between gamma photons and the detector, as well as the associated deposited energy, consists in introducing a marked point process N[2]. We then define

(**H**-3) $N = \sum_{k\in\mathbb{Z}} \delta_{T_k,Y_k}$ is a Poisson random measure where $(T_k)_{k\in\mathbb{Z}}$ follow a Poisson point process of homogeneous intensity λ and the marks $(Y_k)_{k\in\mathbb{Z}}$ are independent from the arrival times $(T_k)_k$, i.i.d. and admit a probability density function f with respect to the Lebesgue measure.

This hypothesis can be explained by the fact that, since the number of radio-elements inside the source is colossal, these nucleus can emit photons in a quasi-independent manner from each other. The choice of modeling the arrival times $(T_k)_k$ by a Poisson point process with constant intensity λ conveys the two following properties:

- lack of memory: the number of gamma photons that interact before a certain time t is independent from the number of gamma photons that interact with the detector after this time t, for all t.
- time homogeneity: the law of the number of interactions in a given time interval of length τ does not depend on the time interval.

The density f corresponds to the gamma ray spectrum (after the deformations induced by the photon-matter interactions) of the energy levels associated to the gamma photons emitted by the radioactive source. The intensity λ is usually expressed in terms of s^{-1} (or alternatively in cps, meaning *counts per second*) and matches the mean number of detected photons in a time interval of one second.
When an photon-matter interaction occurs, we have seen in Section 2.1.3 that an electric current proportional to the deposited energy is recorded. We call *impulse response* the function h that corresponds to the measured signal and assume that it satisfies the following hypothesis

(**H**-4) h is a real-valued causal measurable function such that

$$|h(t)| \leq C_0 \mathrm{e}^{-\alpha t} , \quad t \in \mathbb{R}_+ .$$

for two positive constants C_0 and α

This assumption on h is mild and allows to consider measurable functions that are not necessarily compactly supported. In γ-spectroscopy applications, the impulse responses look like the functions drawn in Figure 2.11.
The electric current can then be modeled by a stationary shot-noise process defined by

$$X_t := \sum_{k\in\mathbb{Z}:T_k\leq t} Y_k \, h(t - T_k), \quad t \geq 0 . \tag{2.2.1}$$

[2] A detailed definition is given in Chapter 3

Since a shot-noise process is fully characterized by the knowledge of the triplet intensity-marks' density-impulse response (λ, f, h), we will often use in this manuscript the expression "shot-noise process with triplet (λ, f, h)" to denote every shot-noise process on the real line that satisfies Assumptions (**H**-1)-(**H**-2) with intensity λ, marks' density f and impulse response h. Two examples of shot-noise processes arising in γ-spectroscopy and that only differ by their respective intensity are illustrated in Figure 2.13.
However, the existing instrumentation devices do not allow to continuously record the electric current. The measured signal obtained from the data acquisition system corresponds to a low frequency sample[3] of the shot-noise process given by (2.2.1). In the following, we assume that we have at hand a sample $X_\delta, \ldots, X_{n\delta}$ where n and δ respectively denote the size of the sample and the inverse of the sampling rate.

2.2.3 Questions of the thesis

2

The purpose of this thesis is to construct, from the discrete observations $X_\delta, \ldots, X_{n\delta}$ of the shot-noise process $(X_t)_{t\in\mathbb{R}}$ given by (1.2.1), estimators of the mark's density f and of the intensity λ under the additional hypothesis that the impulse response h is known. Moreover, it is important to note the low frequency sampling of the shot-noise precludes the use of the joint laws of the shot-noise process $(X_t)_{t\in\mathbb{R}}$, so that we only rely on statistical methods that involve the marginal law of the process.
Secondly, and when it is possible, we will study the asymptotic performances of the constructed estimators. Detailed summaries of the methods introduced in this manuscript can be found in Section 2.4.1 of this Chapter.

2.3 State of the art

In this section, we give an account of the existing statistical methods used to do inference of shot-noise processes. In a first time, we detail several methods used by physicists.

2.3.1 Existing methods in Physics

Up until now, the classical methods introduced to infer the marks' density are still widely used in gamma spectroscopy and neutron transport. The first method calculates the histogram of the local maxima of the sampled signal when h is positive. Indeed, this method allows to compute an estimator proportional to a scaled version of the density f associated to the shot-noise process $(X_t)_{t\in\mathbb{R}}$. However, this method presents a main drawback because it only works for impulse response whose support is shorter than the mean average time between consecutive detected photons. This corresponds to the case where the spectrometer can discriminate, almost systematically, each

[3] The sampling frequency can reach 100 MHz. We here use the term "low frequency" in opposition to the usual definition of the term high frequency in statistics which means that the sampling rate can be chosen arbitrarily large.

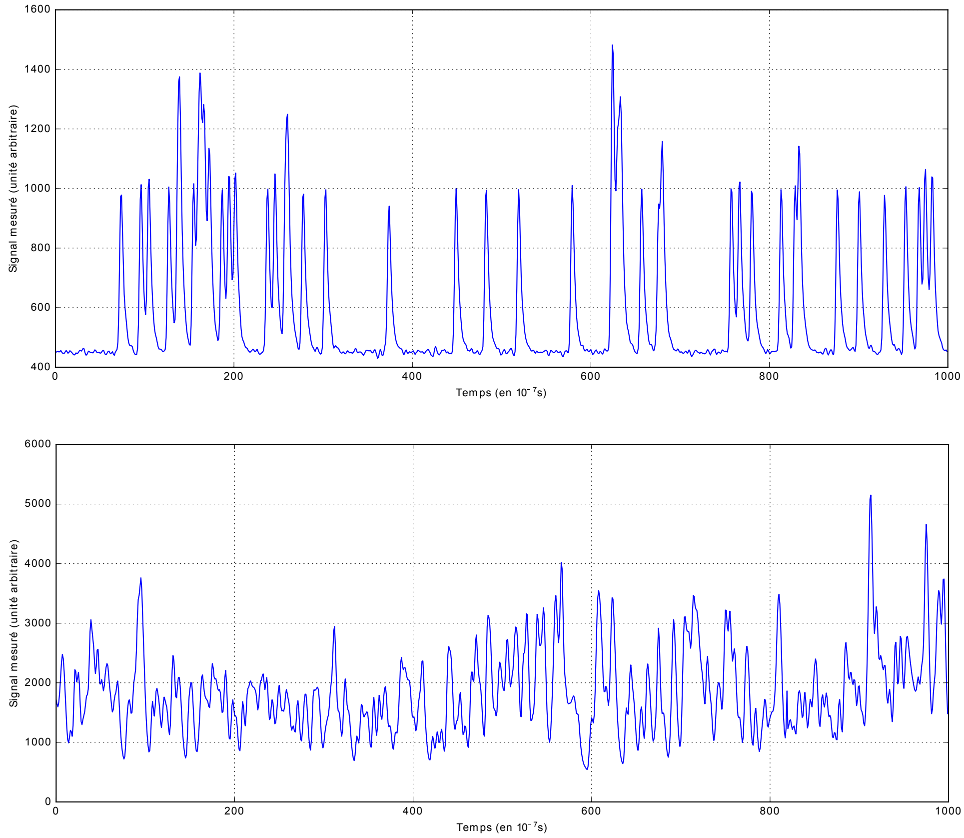


Figure 2.13 : Two shot-noise processes from γ-spectroscopy measurements: the radioactive source is a mixture of Americium 241 and Plutonium 238. The intensity rates are respectively of 5.10^5 et 4.10^6 counts by second in the upper and lower figures

photon-detector interaction. An illustration of this method is provided in Figure 2.14 for shot-noise processes with triplet (λ, f, h) where the intensity λ successively takes the values $\{0.2, 1, 5, 10\}$, the impulse response h is given by

$$h : t \to 10t\, \mathrm{e}^{1-10t} \mathbb{1}_{t \geq 0} \,,$$

and the marks' density f follow a Gaussian distribution with mean 10 and variance 1.
The computed histogram provides an estimator of the function

$$\mathbb{R} \to \mathbb{R}\,, x \to \frac{1}{h_{\max}} f\left(\frac{x}{h_{\max}}\right) \quad \text{where} \quad h_{\max} := \sup_{t \in \mathbb{R}} h(t)\,.$$

In presence of pileup, the histogram exhibits plateaus and multiple peaks called *artifacts* that appear for energies that are multiples of the modes of the marks' density. This phenomenon can be observed on the two last histograms in Figure 2.14 where we clearly see a local maxima at the abscissa point 20. In this case, we see that this estimator becomes inaccurate in order to discriminate the gamma ray spectrum associated with the radionuclide of the source as soon as the intensity increases.
In order to bypass these limitations, practitioners usually use a cumulant-based method[4] This method corresponds to the Campbell's formula for point processes which gives an expression of the cumulants of the marginal law X_0 of the shot-noise process $(X_t)_{t \geq 0}$ by

$$\kappa_X^n = \lambda\, \mathbb{E}_f\left[Y^n\right] \int_0^{\infty} h^n(s) \mathrm{d}s\,, \quad n \in \mathbb{N}^\star\,. \tag{2.3.1}$$

When either the marks' density f or the intensity λ is known, then, for a given impulse response h, the second method consists in replacing the cumulants of the marginal law X_0 in Equation (2.3.1) by their empirical counterparts $(\hat{\kappa}_X^n)_{n \geq 1}$ in order to obtain estimators of f or λ. When f is known, it gives several estimators of λ defined by

$$\hat{\lambda}_m = \frac{\hat{\kappa}_X^m}{\mathbb{E}_f\left[Y^m\right] \int_0^{\infty} h^m(s) \mathrm{d}s}\,, m \in \mathbb{N}^\star\,.$$

This approach is used in Elter et al. (2016) but the authors do not mention how the above estimators are aggregated- in particular, which linear combination provides an estimator of minimal variance- in order to estimate λ. Note that, when the cumulant of order k is estimated using the k-statistics, the variance of this estimator is proportional to the ratio λ^k/n with n the sample size. For high intensity rates λ, it follows that the high order cumulants can not be used to estimate the intensity λ unless we have at hand a large sample of data.
When the intensity λ is known, an estimator of the nth moment of Y is given by

$$\widehat{\mathbb{E}_f\left[Y^m\right]} = \frac{\hat{\kappa}_X^m}{\lambda \int_0^{\infty} h^m(s) \mathrm{d}s}\,, \quad m \in \mathbb{N}^\star\,.$$

[4]Cumulants of shot-noise processes are introduced in Chapter 3

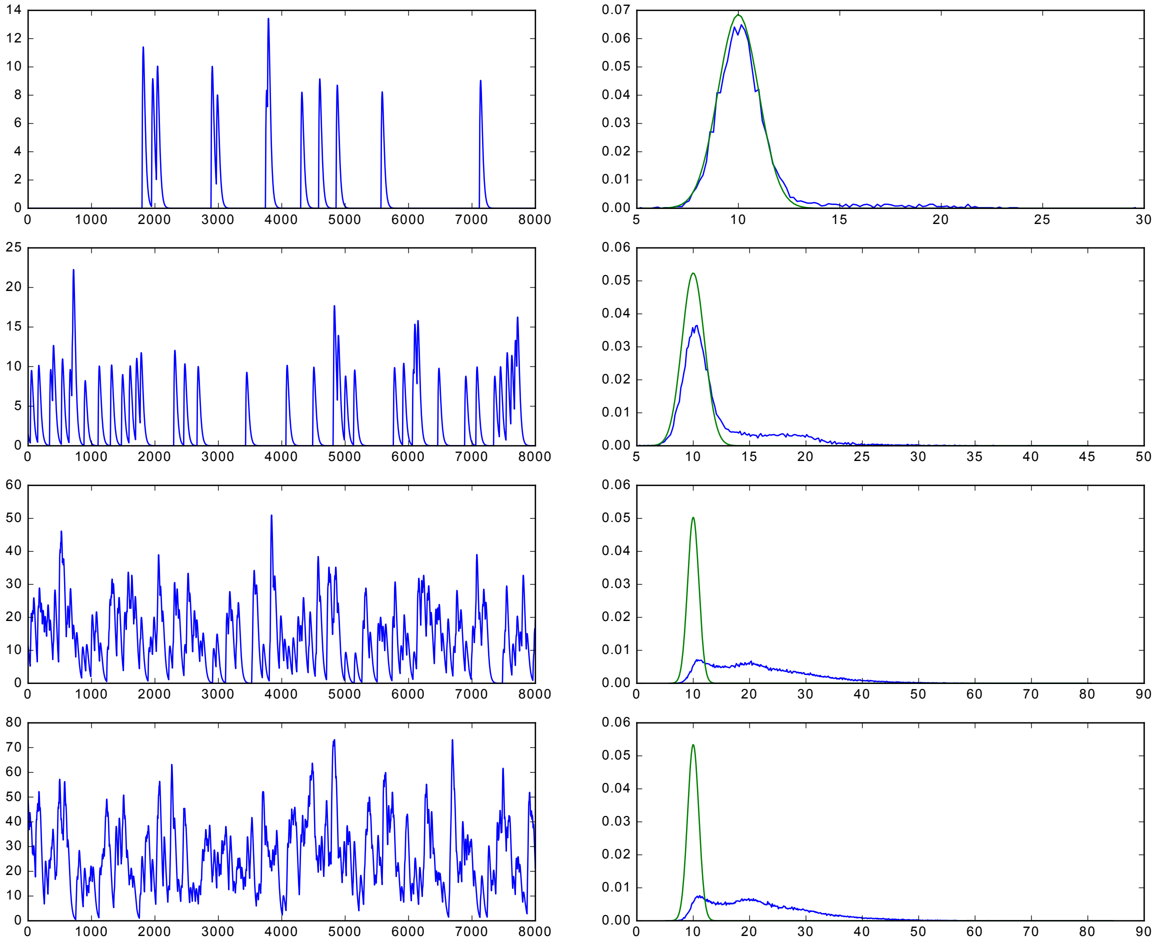


Figure 2.14 : Left: Sample paths of simulated shot-noise processes with triplet (λ, f, h) for $\lambda \in \{0.2, 1, 5, 10\}$ for a same impulse response h and the same marks' density f. Right: Histogram of the local maxima associated to a sample of the shot-noise process of size $n = 5.10^6$

2

Therefore, it is possible to construct an estimator of the characteristic function of Y by using the first M moments $\widehat{\mathbb{E}_f[Y^1]}, \dots, \widehat{\mathbb{E}_f[Y^M]}$ defined by

$$\begin{aligned} \hat{\varphi}_M(u) &:= 1 + \sum_{k=1}^{M} \frac{\widehat{\mathbb{E}_f[Y^k]}}{k!}(iu)^k \\ &= 1 + \sum_{k=1}^{M} \frac{\hat{\kappa}_X^k}{\lambda\, k! \int_0^\infty h^k(s)\mathrm{d}s}(iu)^k \ , \quad u \in \mathbb{R} \ . \end{aligned} \tag{2.3.2}$$

Then Fourier transform of $\hat{\varphi}_M$ then provides an estimator $\hat{f}_M$ of the density f. The results of this method are detailed in Chapter 3.

2

2.3.2 Recent statistical developments

Statistical inference methods for shot-noise processes are more recent. In Trigano (2005), the author also studies the issues in gamma spectroscopy caused by pileup and introduces new methods that are efficient in order to identify the radioactive elements of the source, even for shot-noise processes with high intensity rates. However, these methods rely on estimators constructed from a different sample and based on more restrictives assumptions. Indeed, the model assumes that we have a sample of pulse durations and energies that is represented by the point process:

- $N = \sum \delta_{T_k,(X_k,Y_k)}$ is a marked point process such that the time events $(T_k)_k$ follow an homogeneous Poisson point process on $\mathbb{R}_+$ with intensity λ and the marks $(X_k, Y_k)_k$ are i.i.d. and independent of $(T_k)_k$. For every time arrival T_k, $(X_k)_k$ corresponds to the duration of the electric pulse associated to the event and $(Y_k)_k$ to the energy deposited by the particle.

Then, based on a sample $(X'_k, Y'_k, Z'_k)_{1\le k\le N}$ of size N and where Z'_k is an estimator of the kth *idle* event defined by

$$Z_k := T_{k+1} - (X_k + T_k) \ ,$$

Trigano (2005) introduces a nonparametric estimator of the marks' density Y_0. However, it seems difficult to extend the notion of the duration and the idle sequence associated to a pulse when the impulse response is not compactly supported or when pileup occurs. Indeed, how would the duration be defined when the impulse response is to the form

$$h : t \to \mathrm{e}^{-\alpha t}\mathbb{1}_{t\ge 0}, \quad \alpha > 0 \ ? \tag{2.3.3}$$

A "brutal" truncation of this function (for instance, if we consider that h is zero outside the interval $[0, 3/\alpha]$) would lead to significant errors as soon as the intensity λ increases. We will see in Chapter 4 how to build an estimator of the marks' density f when the impulse response h is given by (2.3.3).

Recently, the papers Lopatin et al. (2012) and Trigano et al. (2015b) do inference of shot-noise processes with assumptions and samples close to the ones introduced in this thesis. In these articles, we have a sample $\mathbf{x} = (x_0, \ldots, x_{N-1})$ of finite size N of a low-frequency shot-noise process with a sampling frequency δ^{-1} given by

$$x_k := \sum_{k=0}^{\infty} Y_k \Phi_k(n\delta - T_k) + \epsilon_n, \quad k = 0, \ldots, N-1 \,, \tag{2.3.4}$$

where $\epsilon = (\epsilon_1, \ldots, \epsilon_N)$ is a collection of centered Gaussian i.i.d. r.v.'s with variance σ^2 and the function Φ_k is given by

$$\Phi_k(t) \equiv \Phi_{(\theta_1^{(k)}, \theta_2^{(k)})}(t) := t^{\theta_1^{(k)}} \mathrm{e}^{\theta_2^{(k)} t} \mathbb{1}_{\{0 \le t \le (N-1)\delta\}} \,, \tag{2.3.5}$$

and the parameters $\theta_1^{(k)}$ and $\theta_2^{(k)}$ are supposed to respectively belong to the intervals $(0, \Theta_1)$ and $(0, \Theta_2)$ where Θ_1, Θ_2 are assumed to be known.

2

Contrary to the model studied in this manuscript (and presented in Chapter 3), this model admits two significant difference. On the one hand, the authors do not consider the stationary version of a shot-noise process. On the other hand, the impulse response is not independent of the marks that model the incident photons energy: indeed, each interaction can generate a pulse with a varying shape that remains proportional to the deposited energy. This variation is made possible by the bi-dimensional parameterization (θ_1, θ_2) defined by (2.3.5). The proposed algorithms in this works are based on the assumptions that the sampled signal is sparse. This is true only for shot-noise processes with low intensity λ or for impulse responses h with extremely short support. The papers aim to estimate the arrival times $(T_k)_k$ and their respective energies Y_k of the underlying marked point process N using the following iterative method:

- For two positive integer p, q, let U_1^p and U_2^q respectively denote the equispaced discretizations of the intervals $(0, \Theta_1)$ and $(0, \Theta_2)$ by p and q points respectively.
- For all $n \in \{0, \ldots, N-1\}$, let A_n be the vector that represents the set of all the electric pulses that may be part of the observation point x_n

  $$A_n = \left[\Phi_{(j,l)}(k-n)\right]^{\mathrm{T}} \,, \quad 0 \le k \le N-1 \,, \; j \in U_1^p \,, \; l \in U_2^q \,,$$

  where we denote by $\mathbf{A} = [A_0 \, A_1 \, \cdots \, A_{N-1}]$ the dictionary obtained by the concatenation of $(A_j)_{0 \le j \le N-1}$.
- The observations $\mathbf{x}$ can then be written

  $$\mathbf{x} = \mathbf{A}\beta + \epsilon + \boldsymbol{\Delta}$$

  where $\mathbf{x}$, ϵ are defined by (2.3.4), $\boldsymbol{\Delta}$ corresponds to the vector measuring the approximation error that is made when replacing the arrival times (T_k) by the closest sampling. The vector β of size Npq is unknown and contains information about the marked point process $(T_k, Y_k)_{k\geq 0}$ and the corresponding impulse responses. Assuming that the intensity λ is low, the regressor β can be reasonably supposed sparse. Moreover, the nonzero indices roughly correspond to interaction times between photons and the detector. One way to estimate this vector is to use the LASSO (Least Absolute Shrinkage and Selection Operator, see Tibshirani (1996) for details) method: for all positive μ, the vector $\hat{\beta}(\mu)$ is defined as the solution of the minimization problem

$$\hat{\beta}(\mu) = \underset{\beta\geq 0}{\operatorname{argmin}} \|\mathbf{y} - \mathbf{A}\beta\|_2^2 + \mu\|\beta\|_1 \ .$$

- Then, the vector $\hat{\beta}(\mu)$ is used in order to recursively estimate the photons arrivals $(T_k)_k$ via

$$\begin{cases} \hat{T}_0 & = & \min\{0 \leq m \leq N-1 : \|\hat{\beta}(\mu)_m\|_1 \geq s\} \\ \hat{T}_k & = & \min\{\hat{T}_{k-1} < m \leq N-1 : \|\hat{\beta}(\mu)_m\|_1 \geq s, \|\hat{\beta}(\mu)_{m-1}\|_1 \leq s\} \ , \quad k \geq 1 \ . \end{cases}$$

- Setting

$$M := \left|\{\hat{T}_k, \hat{T}_k \leq N-1\}\right| \qquad \text{and} \qquad \mathbf{x}^{(\mathbf{0})} := \mathbf{x} \ ,$$

the energies $\hat{Y}_1, \ldots, \hat{Y}_M$ respectively associated to the times $\hat{T}_1, \ldots, \hat{T}_M$ are successively estimated via

$$\begin{cases} \hat{\beta}(\mu_k) & = & \underset{\beta\geq 0}{\operatorname{argmin}} \|\mathbf{x}^{(k)} - \mathbf{A}\mathbf{W}^{(k)}\beta\|_2^2 + \mu_k\|\beta\|_1 \\ \hat{Y}_k & = & \|\mathbf{A}\mathbf{W}^{(k)}\hat{\beta}(\mu_k)\|_1 \\ \mathbf{x}^{(k+1)} & = & \mathbf{x}^{(k)} - \mathbf{A}\mathbf{W}^{(k)}\hat{\beta}(\mu_k) \ , \end{cases}$$

where $(\mu_k)_{k\leq N-1}$ is a sequence of positive numbers and $(\mathbf{W}^{(k)})_k$ a sequence of square diagonal matrices of size N that reflect the different energy contributions at the sampling point $k\delta$ (via the vector $\hat{\beta}(\mu_k)$). In general, for a given k, the diagonal coefficients of the matrix $\mathbf{W}^{(k)}$ correspond to the evaluation of a Gaussian kernel with mean $\hat{T}_k$. The marks' density f is then estimated using a kernel estimator K and a bandwidth $h > 0$ given by

$$\hat{f}_Y(x) := \frac{1}{Mh}\sum_{k=1}^{M} K\left(\frac{\hat{Y}_k - x}{h}\right), \quad x \in \mathbb{R} \ .$$

Even though the assumptions of these papers are less restrictive than ours and that they consider broader shot-noise processes, the idea that the sampled signal is sparse becomes wrong when the counting rate λ increases, as illustrated by Figure 2.13. These methods or the ones used in Andrieu et al. (2001) fail to be efficient in presence of pileup. In the next section, we detail the estimators of the intensity λ and the marks' density f obtained during this thesis. In particular, we will see that the estimators of λ are still efficient for large values of λ and how the convergence rates of the density f obtained in Chapter 4 depend on λ.

2.4 Contributions

2.4.1 Methods and results

In this thesis, we tackle the estimation problem of the marks' density f and the intensity λ as a linear inverse problem. This is possible because one can analytically compute the characteristic function φ_{X_0} of the marginal law X_0 of the shot-noise process $(X_t)_{t\in\mathbb{R}}$ given by (2.2.1). Indeed, we have for all real u

$$\varphi_{X_0}(u) = \exp\left(\lambda \int_0^\infty [\varphi_Y(uh(s)) - 1]\,\mathrm{d}s\right) . \tag{2.4.1}$$

When $\mathbb{E}[|Y_0|] < \infty$, the function φ_{X_0} is differentiable and we get

$$\frac{\varphi'_{X_0}(u)}{\varphi_{X_0}(u)} = K_h[f_\lambda](u) , \quad u \in \mathbb{R} , \tag{2.4.2}$$

where $f_\lambda := \lambda f$ and K_h is an operator detailed in Chapter 5 that is defined for every real-valued function ξ which satisfies $\int_0^\infty x|\xi|(x)\mathrm{d}x < \infty$ and for all real u by

$$K_h[\xi](u) = \int_0^\infty \left[\int_0^\infty \mathrm{i}xh(s)\mathrm{e}^{\mathrm{i}uxh(s)}\mathrm{d}s\right]\xi(x)\mathrm{d}x . \tag{2.4.3}$$

From a "plug-in" method based on the observations $X_\delta, \ldots, X_{n\delta}$, we can also compute an estimator of the function $\varphi'_{X_0}/\varphi_{X_0}$. Respectively setting $\hat{\varphi}_n$ and $\hat{\varphi}'_n$ by the empirical characteristic function associated to $X_\delta, \ldots, X_{n\delta}$ and its derivative, an estimator of $\varphi'_{X_0}/\varphi_{X_0}$ is given by

$$\hat{g}_n(u) := \frac{\hat{\varphi}'_n(u)}{\hat{\varphi}_n(u)} = \frac{\sum_{k=1}^n \mathrm{i}X_{k\delta}\mathrm{e}^{\mathrm{i}uX_{k\delta}}}{\sum_{k=1}^n \mathrm{e}^{\mathrm{i}uX_{k\delta}}} \qquad u \in \mathbb{R} .$$

Chapter 4 considers the particular case of exponential impulse responses h , which corresponds to the set of functions defined by

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

for some positive α. This case is considerably different from the general one because, on the one hand, the exponential shot-noise process is indeed a Markov process and, on the other side, the function $\varphi'_{X_0}/\varphi_{X_0}$ given by (2.4.2) can be rewritten

$$\frac{\varphi'_{X_0}(u)}{\varphi_{X_0}(u)} = \frac{\lambda}{\alpha}\cdot\frac{\varphi_Y(u) - 1}{u} , \quad u \in \mathbb{R} .$$

Assuming the ratio λ/α to be known, one can estimate the function f by Fourier inversion as defined by (4.2.6): for two well chosen positive sequences $(h_n)_n$ and $(\kappa_n)_n$ that tend to zero, $\hat{f}_n$ is the estimator of f defined by

$$\hat{f}_n(x) := \frac{1}{2\pi}\int_{-h_n^{-1}}^{h_n^{-1}} \mathrm{e}^{-\mathrm{i}ux}\left(1 + u\frac{\alpha}{\lambda}\frac{\hat{\varphi}'_n(u)}{\hat{\varphi}_n(u)}\mathbb{1}_{\{|\hat{\varphi}_n(u)|\geq\kappa_n\}}\right)\mathrm{d}u , \quad x \in \mathbb{R} .$$

2

A theoretical study allows to obtain upper rates of convergence, with respect to the uniform norm, uniformly on the Sobolev regularity classes $\Theta(K, L, m, s)$ defined, for positive real numbers K, L, m and $s > 1/2$ by

$$\Theta(K, L, m, s) := \{f : \mathbb{R}_+ \to \mathbb{R} \text{ t.q. } \int_0^\infty |x|^{4+m} f(x) \mathrm{d}x \le K \text{ and } \int_{\mathbb{R}} |\mathcal{F}[f](u)|^2 \, |u|^{2s} \mathrm{d}u \le L\}$$

Theorem 4.2.1 states that, if the sequences $(h_n)_n$ and $(\kappa_n)_n$ are set to

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

for given $K, L, m,\ s > 1/2$ and where C is defined by

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

then there exists some positive number M only depending on the constants K, L, m, s and λ/α such that, for all $n \ge 3$,

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

In order to derive the above rates of converge, we assumed that the ratio λ/α is known. However, as indicated in Chapter 4, we show that this ratio can be estimated using the Hill's estimator associated to the sample $X_\delta^{-1}, \dots, X_{n\delta}^{-1}$.

In Chapter 5, we will see that the study of the functional operator K_h given by (2.4.3) is involved. In particular, it is not clear to determiner for which impulse responses h the operator K_h is compact. To bypass this issue, we introduce nonparametric techniques in order to estimate the function f_λ. If we additionally assume that f belongs to the Sobolev space of order 2 $\mathcal{W}^2$ defined by

$$\mathcal{W}^2 := \{f : [0,1] \to \mathbb{R} : \quad f, f' \text{ absolument continues,} \\ f(1) = f'(1) = 0 \text{ and } \int_0^1 |f''(t)|^2 \, \mathrm{d}t < \infty\} ,$$

equipped with the norm

$$\|f\|_{\mathcal{W}^2} := \sqrt{\int_0^1 |f''(t)|^2 \, \mathrm{d}t} , \quad f \in \mathcal{W}^2 ,$$

we propose to estimate f_λ from a finite grid of N evaluation points $\{u_i, i \in 1, \dots, N\}$ so that our problem reduces to 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 $\epsilon(u_1), \ldots, \epsilon(u_N)$ are the estimation errors of $g(u_i)$ by $\hat{\varphi}_n'(u_i)/\hat{\varphi}_n(u_i)$ for $i = 1, \ldots, N$. Since we assume that f_λ belongs to the functional Hilbert space $\mathcal{W}^2_+$, defined as the subset of positive functions in $\mathcal{W}^2$, we propose to estimate f_λ by the solution of the minimization problems defined, for any definite positive matrix W of size N and all positive μ, by

$$\hat{f}_{n,\lambda}(\mu) := \underset{\xi \in \mathcal{W}^2_+}{\operatorname{argmin}} \, \|\hat{g}_n(\underline{u}) - K_{\underline{u}}[\xi]\|^2_{W^{-1/2}} + \mu \|\xi\|^2_{\mathcal{W}^2} \,, \tag{2.4.4}$$

where

$$\|\hat{g}_n(\underline{u}) - K_{\underline{u}}[\xi]\|^2_{W^{-1/2}} := \left\| W^{-1/2} \cdot \{\hat{g}_n(\underline{u}) - K_{\underline{u}}[\xi]\} \right\|_2^2 \,,$$

and

$$\hat{g}_n(\underline{u}) := \left[\hat{g}_n(u_1) \;\cdots\; \hat{g}_n(u_N)\right]^{\mathrm{T}} \quad \text{et} \quad K_{\underline{u}}[\xi] := \left[K_h[\xi](u_1) \;\cdots\; K_h[\xi](u_N)\right]^{\mathrm{T}} \,.$$

This infinite-dimensional problem (2.4.4) can then be reduced into a finite-dimensional problem using the reproducing property of the space $\mathcal{W}^2$. The second method uses the decomposition of the function f_λ in the cubic B-spline basis. The performances of the associated algorithms are then illustrated both on simulated datasets and on real datasets from the CEA Saclay.

2.4.2 Outline of the thesis

This thesis is organized as follows. Chapter 3 gathers a brief history of shot-noise processes in science since the 20th century. Secondly, we recall how to define and to construct the Poisson point processes on the real line prior to define the shot-noise processes given by (2.2.1). Then, we lay the emphasis on the connection between the statistical inference methods of Lévy processes and our statistical estimation problem for shot-noise processes. Lastly, we provide a review of the classes of stochastic processes in connection with shot-noise processes and the related statistical estimation methods.

Chapter 4 studies a nonparametric estimator of the marks' density based on a discrete sampling of an exponential shot-noise process. The estimator is based on a formula linking the characteristic function of the photon density to a function involving the characteristic function and its derivative. Under the additional assumptions that the marks' density belongs to a Sobolev space, we establish that our estimator converges to the marks' density in uniform norm at a polynomial rate.

Chapter 5 tackles the estimation problem of the marks' density and the intensity for general impulse responses. We assume that the impulse response is known and that the marks' density belongs to the Sobolev space of order 2 denoted $\mathcal{W}^2$. Then, based on a finite low frequency sample of the shot-noise process, we introduce two statistical methods in order to constructs estimators of the quantity f_λ. The first one makes use of the reproducing property of the space $\mathcal{W}^2$ while the second one relies on a parameterization with cubic B-splines. The results of this method are then illustrated on simulated and real datasets.

2.5 Publications

Some parts of this thesis have been published and presented over the past three years. The herebelow list gathers all the contributions in journals and conferences.

2.5.1 Journal articles

- **Nonparametric estimation of mark's distribution of an exponential shot-noise process** - P. Ilhe, E. Moulines, F. Roueff et A. Souloumiac – *Electronic Journal of Statistics*, 2016.
- **Pileup correction of a shot-noise process** - P. Ilhe, E. Moulines, F. Roueff et A. Souloumiac – Nearly finished.

2.5.2 Conferences proceedings

- **Nonparametric estimation of mark's distribution of an exponential Shot-noise process** - P. Ilhe, E. Moulines, F. Roueff et A. Souloumiac – *Colloque GRETSI*, 2015.
- **Désempilement non-paramétrique de la densité d'un processus shot-noise** - P. Ilhe, E. Moulines, F. Roueff et A. Souloumiac – *Journées des Statistiques*, 2016.
- **Nonparametric estimation of a shot-noise process** - P. Ilhe, E. Moulines, F. Roueff et A. Souloumiac – *IEEE Workshop on Statistical Signal Processing*, 2016.

3 A deeper insight into shot-noise processes

Look how black the sky is, the writer said. I made it that way.
– Bret Easton Ellis, *Lunar Park*

Abstract

In this chapter, we provide an account on the history of shot-noise processes in science. Secondly, we recall the construction of a shot-noise process starting by an introduction to point processes on the real line. After recalling the mathematical definition of a shot-noise process, we detail both basic properties and results of these stochastic processes in the statistics and probability domains. Then, we introduce Lévy processes and detail the statistical inference method of the Lévy-Khintchine triplet as well as the existing connection between the marginal laws of shot-noise processes and Lévy processes. Finally, we gathers the different stochastic processes that provide extensions (ambit models, Lévy semistationary processes) or constitute subclasses (CAR, CARMA processes) of shot-noise processes and the developed statistical estimation methods.

3.1 Origins of the shot-noise

In physics, the shot-noise (also called photon noise, poisson noise, schott effect) denotes a kind of electronic noise that is generated by the inherent corpuscular nature of electrons. Before going through an historical review about shot-noise processes in litterature, let us precise what the term shot-noise means for physicists. We also recall that this manuscript is only devoted to the statistical study of the mathematical model defined by (3.2.6).

3.1.1 Shot-noise in Physics

This part is inspired from Reisch (2012)[Chapter 5] that gives an extended treatment to the various kind of noises that can occur in semiconductors. Consider the following setup: given two electrodes E_1 and E_2 and a beam of electrons as schematized in Figure 3.1, the current measured at time t is given by

$$i(t) = -\frac{e}{L}\sum_{k=1}^{N_t} v_k(t) \ .$$

where L is the distance between electrodes E_1 and E_2, v_k denotes the velocity of the electron k (the electrons are labeled in function of the time they hit electrode E_2) and N_t denotes the number of electrons between the electrodes at time t.

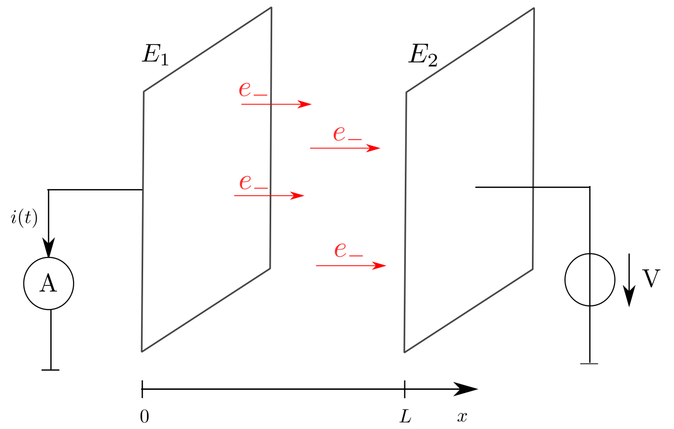


Figure 3.1 : Experimental setup in order to measure the electric current generated by electrons between two electrodes.

3

If one assumes that the velocities $v_1(t), \ldots, v_k(t)$ are mutually independent and uncorrelated to the number of electron N_t between the two electrodes at time t, then we have

$$\langle i(t)\rangle = -\frac{e}{L}\langle N_t\rangle\langle v(t)\rangle \, .$$

where the symbol $\langle x\rangle$ denotes the average value of x. Moreover, denoting by $\Delta i(t)$ the current fluctuation at time t, we have using a Taylor's expansion

$$\Delta i(t) \simeq -\frac{e}{L}\langle v(t)\rangle \Delta N_t - \frac{e}{L}\sum_{k=1}^{\langle N_t\rangle} \Delta v_k(t) \, .$$

Then, it is possible to show that

$$\langle \Delta i^2(t)\rangle = \frac{e^2}{L^2}\left(\langle v\rangle^2\langle \Delta N_t^2\rangle + \langle N_t\rangle\langle \Delta v^2(t)\rangle\right) \, . \tag{3.1.1}$$

From a mathematical standpoint, the previous equation can be coarsely rewritten

$$\mathrm{Var}(i(t)) = \mathrm{Var}\left[N_t\right]\mathbb{E}\left[v_1(t)\right]^2 + \mathbb{E}\left[N_t\right]\mathrm{Var}\left[v_1(t)\right] \, .$$

The first term in the right-hand side of (3.1.1) is referred to as *shot noise* by physicists and is associated to the random thermal motion of the electrons while the second term is called *thermal noise*.
In physics, shot noise is often described as the consequence of the discrete nature of electric charges while in this document, it rather corresponds to the convolution between a filter and a random sequence of epochs, i.e. for instance, the electric current $\{i(t) \; : t \geq 0\}$ modeled above.

3.1.2 A selected historical review

In this section, we give an overview of the developments of shot-noise processes in scientific literature since its first appearance in the early 20th century. Moreover, we rapidly draw our attention to the main statistical and probabilistic results obtained through the pioneering article of Campbell (1909b).
Indeed, the author introduces a shot-noise process to model the following experiment: charged particles originating from some source arrive at a Poisson rate on a electrometer. Namely, from an electrometer at rest (i.e. that displays a current of zero), the current θ generated by a particle with charge E is solution of the second-order differential equation

$$I\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + \mu\frac{\mathrm{d}\theta}{\mathrm{d}t} + k\theta - KE/C\mathrm{e}^{-pt} = 0\,, \quad t \geq 0\,,$$

where I is the moment of inertia of the needle, μ the coefficient of damping, k the coefficient of torsion, C the capacity of the electrode and p a constant only depending on the nature of the resistance. The solution of this equation takes the form

$$\theta(t) = A\mathrm{e}^{-\alpha t} + B\mathrm{e}^{-\beta t} + P\mathrm{e}^{-pt}\,, \quad t \geq 0\,.$$

where all the constants depend on the parameters previously introduced and α, β and p are positive constants. The measured current is then modeled as a shot-noise process, so that, at a given time t, the recorded current can be interpreted as the superposition of the individual currents generated by every particle that arrived at times $\tau_1, \ldots, \tau_m$ inferior to t via the following equation:

$$\theta(t) = \sum_{k=1}^{m} A\mathrm{e}^{-\alpha(t-\tau_k)} + B\mathrm{e}^{-\beta(t-\tau_k)} + P\mathrm{e}^{-p(t-\tau_k)}\,.$$

Note that Chapter 4 is entirely devoted to the study of shot-noise processes with causal, exponential decaying impulse response functions and literature involving exponential shot-noise processes will be developed there.
In addition, the author derives what is today known as the Campbell's theorem for Poisson point processes[1]: it states the mean and the variance of the measured number of particles are proportional to the intensity of the source. The following experiment is detailed in Campbell (1909a):

> "Let $X\tau$ be the average number of events which happen in a small time τ and let x be the deviation of that number from its mean value during the time τ: let
>
> $$\theta = f(t)$$
>
> represent the motion of the indicator of the measuring instrument at all times t subsequent to the happening of one of the events, and $\overline{\theta'^2}$ the square of the

[1] See Daley and Vere-Jones (1988) for details.

mean deviation of the indicator from its mean position for all times T from the moment of starting the observations, where T is a time long compared with the time constants of the measuring instrument. Then it is shown that

$$\overline{\theta'^2} = \overline{x^2}/\tau \int_0^T f^2(t)\mathrm{d}t \, .$$

[...] The "events" in our case are the liberation of individual electrons."

A decade later, Schottky (1918) studied the fluctuations of current in vacuum tubes and developed a particular case of the formula which is today known as the Bartlett power spectrum associated to a Point process as defined in Daley and Vere-Jones (1988): he proved that when the current $i(t)$ is modeled by a shot-noise process, its corresponding spectral density is proportional to the average current $\langle i \rangle$.

The first mathematical study was tackled by Rice (1944). A formal definition of the shot-noise process as defined in the following sections is provided under the appellation *shot effect*. Moreover, some properties related to shot-noise processes are established: one can find the first generalization of Campbell's theorem to marked shot-noise processes, the expression of the power spectra for arbitrary impulse responses and the characteristic function of every marginal of the process. All these properties are detailed in Section 3.2. However, all those calculations are derived based on the particular assumption that the Poisson process ruling the arrival of events is homogeneous.

Papoulis (1971) proved a limit theorem for the unmarked shot-noise process defined by

$$X_t = \sum_{T_k : T_k \leq t} h(t, T_k) \, , \quad t \in \mathbb{R} \tag{3.1.2}$$

where $(T_k)_{k\in\mathbb{Z}}$ are the time events associated to a Poisson random measure on $\mathbb{R}$ with inhomogeneous intensity λ and h some causal real-valued function defined on $\mathbb{R}$.

Papoulis showed that the distribution function of the standardization of the marginal X_t for every real t uniformly converges to a Gaussian distribution with same mean and variance. When the intensity λ is homogeneous and the impulse response is time invariant, i.e. such that

$$h(t, \tau) = h(t - \tau) \quad \text{for all } t, \tau \in \mathbb{R} \, , \tag{3.1.3}$$

we have the following result.

Theorem 3.1.1. *Let $(X_t)_{t\in\mathbb{R}}$ be the shot-noise process given by (3.1.2) and assume that the random variables $(T_k)_{k\in\mathbb{Z}}$ follow a homogeneous Poisson point process with intensity λ and that the impulse response h satisfies (3.1.3). For every real number t, the standardization $\tilde{X}_t$ of X_t is defined as*

$$\tilde{X}_t := \frac{X_t - \mathbb{E}[X_t]}{\sqrt{Var(X_t)}} \, .$$

Let F be the cumulative distribution function of $\tilde{X}_t$ and G be the cumulative distribution of a centered Gaussian random variable with unit variance. If $\int_{\mathbb{R}} |h|^3(t)\mathrm{d}t < \infty$, then

$$\sup_{x\in\mathbb{R}} |F(x) - G(x)| \leq \frac{4}{3}\sqrt{\frac{2\pi}{\lambda}} \frac{\int_{\mathbb{R}} |h|^3(t)\mathrm{d}t}{\left(\int_{\mathbb{R}} |h|^2(t)\mathrm{d}t\right)^{3/2}} .$$

This results establishes that the standardization of a stationary shot-noise process converges in distribution to a standard Gaussian random variable when the intensity λ goes large at a rate $\lambda^{-1/2}$. Such a result is of particular importance when we aim to estimate the marks' density because we can no longer hope to estimate precisely this density when the observations become nearly distributed as Gaussian r.v.'s.
Actually, this result was already known since Lapierre and Fortet (1953) but in the more restrictive case where the impulse response is bandlimited, i.e. when the Fourier transform of h is compactly supported. These results were also extended by Heinrich and Schmidt (1985) for the multi-dimensional shot-noise processes introduced by Daley (1971).
The investigation of the distribution function of the marginal of a shot-noise process was before carried by Gilbert and Pollak (1960) and still is an active topic of research [2]: considering the shot-noise process defined by

$$X_t = \sum_{T_k : T_k \leq t} H(t - T_k)$$

where $(T_k)_{k\in\mathbb{Z}}$ is a homogeneous Poisson random measure on the real line with intensity λ and H a measurable function, he proved that the distribution function F of the marginal X_0 given by

$$F(x) := \mathbb{P}(X_0 \leq x) , \quad x \in \mathbb{R} ,$$

satisfies the following integral equation

$$xF(x) = \int_{-\infty}^{x} F(y)\mathrm{d}y + \lambda \int_{\mathbb{R}} F(x - H(t))H(t)\mathrm{d}t , \quad x \in \mathbb{R} .$$

Moreover, they solved these functional equations for particular impulse responses H such as

$$H(t) = \mathrm{e}^{-t}\mathbb{1}_{t\geq 0} , \quad t \in \mathbb{R} .$$

In the same spirit, Gubner (1996) and Lowen and Teich (1990) also developed techniques to numerically compute both the distribution and the density of the shot-noise marginal.
From a more statistical point of view, Roberts (1966) suggested to use Edgeworth's series in order to estimate the marginal distribution function as detailed in Chapter 1 , Eq. (1.3.2).
Finally, the study of the level crossings of stochastic processes is another field of research related to shot-noise process which has received a lot of attention since the second half of the 20th century. Indeed, researchers use shot-noise processes to model quantities such as water level or concentration level of a given molecule in blood (see for

[2] See for instance Taguchi et al. (2010).

instance Helfenstein and Steiner (1990)) and are interested in determining the expected number of crossings over some threshold, since, in the both cases, it is important to keep the water level or the molecule's concentration under some fixed level. Leadbetter et al. (1966) used Rice's formula to compute the number of crossings for a large variety of stochastic processes while Bar-David and Nemirovsky (1972) then Biermé et al. (2012) focused on shot-noise processes. Rice's formula, introduced in Rice (1944), states that, given a level u and an ergodic stationary, differentiable in the L^2 sense, stochastic process $(X_t)_{t\in\mathbb{R}}$, the expected number of crossings D_u is given by

$$\mathbb{E}\left[D_u\right] = \int_{\mathbb{R}} |y| p(u, y) \mathrm{d}y \ ,$$

where $p(\cdot, \cdot)$ denotes the joint probability density function of (X_0, X_0').
Other statistical methods and probability results related to shot-noise processes are presented in Chapters 4 and 5.

3

3.2 Model, properties and existing results

The shot-noise process defined by (1.2.1) in Chapter 1 is build upon an underlying point process we have not yet formally defined. For the sake of completeness, we provide a short introduction in the following section which is inspired from the excellent textbook Reiss (2012).

3.2.1 Point processes on the real line

In this section, we formally introduce the concept of point processes on the real line and their marked versions. The point processes on the real line constitute a particular class of point processes that enjoys a natural interpretation and which can be entirely characterized by the finite dimensional laws of its interarrival times. They are of particular interest and provide very effective results when applied in neuroscience or queuing theory for instance.

Definition of a point process

Loosely speaking, a point process N corresponds to a collection of real-valued random "points" $(T_k)_{k\in I}$ where I is a countable subset of $\mathbb{Z}$. In order to specify how those points are distributed on the real line, we are interested to count, for any subset I of $\mathbb{Z}$, the number of points $(T_k)_{k\in I}$ that lie in a given interval $B \subset \mathbb{R}$ which is given by

$$|\{T_k \ , k \in I\} \cap B| \ .$$

Setting $\tilde{\mathbb{M}} = \{(T_k)_{k\in I} \ ; I \subset \mathbb{Z}\}$, we define for a given interval B of $\mathbb{R}$ the projection $\tilde{\pi}_B$ as the function

$$\tilde{\pi}_B : \tilde{\mathbb{M}} \to \mathbb{N} \ , \quad \{T_k \ , k \in I\} \to |\{T_k \ , k \in I\} \cap B| \ .$$

Then, the configuration of points $(T_k)_{k\in I}$ is fully specified by the projections $\{\tilde{\pi}_B\ , B \in \mathcal{B}(\mathbb{R})\}$ where $\mathcal{B}(\mathbb{R})$ denotes the Borel σ-field of $\mathbb{R}$. Another way to look at a projection $\tilde{\pi}_B$ consists in rewriting the number of points in $\{T_k\ , k \in I\}$ that belongs to a given Borel set $B \in \mathcal{B}(\mathbb{R})$ by

$$|\{T_k\ , k \in I\} \cap B| = \sum_{k\in I} \mathbb{1}_B(T_k) = \sum_{k\in I} \delta_{T_k}(B)$$

where δ_x corresponds to the Dirac measure at the point x. This suggests to represent the set $\{T_k\ , k \in I\}$ of points by the random measure μ called *point measure* defined by

$$\mu = \sum_{k\in I} \delta_{T_k}$$

which satisfies for a given collection of points $(T_k)_{k\in I}$ all the properties of a measure on the Borel σ-algebra on real numbers. Denoting by $\mathbb{M}\,(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ the space of point measures on $\mathcal{B}(\mathbb{R})$, it is possible to endow it with the smallest σ-algebra $\mathcal{M}\,(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ such that the projections $\{\pi_B\ , B \in \mathcal{B}(\mathbb{R})\}$ defined by

$$\pi_B : \mathbb{M} \to \mathbb{N}\ ,\ \mu \to \mu(B)$$

are measurable. Note that, since there is a one-to-one mapping between the space of point measures $\mathbb{M}$ and the set of configurations $\tilde{\mathbb{M}}$, we have, for any Borelian set B

$$\pi_B(\mu) = \pi_B\left(\sum_{k\in I} \delta_{T_k}\right) = \tilde{\pi}_B\left(\{T_k\ , k \in I\}\right)\ .$$

We now have in hand all the tools to define a point process on the real line.

Definition 2 (Point process, Reiss (2012))**.** *Let* $(\Omega, \mathcal{A}, \mathbb{P})$ *be a probability space. A mapping*

$$N : \Omega \to \mathbb{M}\,(\mathbb{R}, \mathcal{B}(\mathbb{R}))$$

is a point process on the real line (also called random point measure) if it is measurable with respect to $\mathcal{A}$ *and* $\mathcal{M}\,(\mathbb{R}, \mathcal{B}(\mathbb{R}))$*. In particular, for a given Borel set B,* $N(B)$ *defined by* $N(B) = \pi_B \circ N$ *represents the random number of points in B.*

Among all the point processes that are used in the statistical field, we present two of them that are of particular importance.

Example 1 (Empirical process). Let n be an integer, $T_1, \ldots, T_n$ be i.i.d. real-valued random variables defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$ and $N_n := \sum_{k=1}^{n} \delta_{T_k}$. Following Definition 2, N_n is a point process and, for any Borel set B, the random variable

$$N_n(B) = \sum_{k=1}^{n} \mathbb{1}_{T_k}(B)$$

follows a binomial distribution with parameters n and $p = \mathbb{P}\,(T_1 \in B)$.

Example 2 (Mixed empirical process). Let $(T_k)_{k\geq 1}$ be i.i.d. real-valued random variables defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$. Let also ζ be an integer-valued random variable defined on the same probability space and independent of the family of random variables $(T_k)_{k\geq 1}$. Then,

$$N := \sum_{k=1}^{\zeta} \delta_{T_k}$$

is a point process called *mixed empirical process*.

For any point process, it is possible to define an application called *intensity* which is particularly useful to define Poisson point processes on the real line. We recall its definition herebelow.

3

Definition 3 (Intensity of a point process)**.** *Let N be a point process on the real line defined on the probability space $(\Omega, \mathcal{A}, \mathbb{P})$. The intensity measure ν associated to the point process N is the application:*

$$\nu : \mathcal{B}(\mathbb{R}) \to \mathbb{R}\ ,\ B \to \mathbb{E}\left[N(B)\right]\ .$$

In fact, ν is a measure on $\mathcal{B}(\mathbb{R})$.

Example 3 (Empirical process). Let n be a non-negative integer and N_n be the empirical point process defined in Example 1. Then, for any Borel set B, we have

$$\mathbb{E}\left[N_n(B)\right] = n\, \mathbb{P}\left(T_1 \in B\right)\ .$$

Example 4 (Mixed empirical process). Let N be the empirical point process defined in Example 2. Then, for any Borel set B, we have

$$\mathbb{E}\left[N(B)\right] = \mathbb{E}\left[\zeta\right]\ \mathbb{P}\left(T_1 \in B\right)\ .$$

We have seen that the projections $\{\pi_B\ , B \in \mathcal{B}(\mathbb{R})\}$ play a major role in the definition of a point process. The following theorem will not be proved in this manuscript but will turn essential when defining Poisson point processes. It states that two point processes follow the same law if, and only if, their finite-dimensional marginals are equal in law.

Proposition 1 (Theorem 1.1.1, Reiss (2012))**.** *Let N_0 and N_1 be two point processes on the real line. Then, the two following statements are equivalent:*

1. $N_0 \overset{\mathcal{L}}{=} N_1$.

2. *For every non-negative integer k and $B_1, \ldots, B_k$ Borel sets,*

$$(N_0(B_1), \ldots, N_0(B_k)) \stackrel{\mathcal{L}}{=} (N_1(B_1), \ldots, N_1(B_k)) \ .$$

Poisson point process on the real line

Before defining what is a Poisson point process on the real line, we need to introduce the Poisson distribution.

Definition 4. *Let λ denote a positive real number and $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space. An integer-valued random variable X follows a Poisson distribution with parameter λ denoted by $\mathcal{P}_\lambda$ when*

$$\mathbb{P}(X = k) = \frac{\mathrm{e}^{-\lambda}\lambda^k}{k!} \ , \quad k \in \mathbb{N} \ .$$

Furthermore, if X, Y are two independent random variables respectively following the Poisson distributions with parameter λ_1 and λ_2 for λ_1, λ_2 positive numbers, then $X + Y$ also follows a Poisson distribution with parameter $\lambda_1 + \lambda_2$.

Interestingly, the Poisson distribution corresponds to the limiting case of a Binomial distribution $\mathcal{B}(n, p_n)$ when the number of trials n tends to infinity and the probability of success $p_n = \lambda/n$ tends to zero for a positive number λ.
Thanks to all the previous elements, the following definition introduces the concept of *Poisson point process* on the real line.

Definition 5 (Poisson point process)**.** *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and ν be a measure on $\mathcal{B}(\mathbb{R})$. A point process N with intensity measure ν is a Poisson point process that satisfies the two following assumptions:*

1. *For a Borel set B, the random variable $N(B)$ follows a Poisson distribution with parameter $\nu(B)$.*

2. *For any non-negative integer k and any Borel distinct sets $B_1, \ldots, B_k$, the random variables $N(B_1), \ldots, N(B_k)$ are independent.*

Furthermore, the point process N is called a stationary or homogeneous Poisson point process with intensity λ for a positive number λ whenever the intensity measure ν satisfies

$$\forall B \in \mathcal{B}(\mathbb{R}) \ , \quad \nu(B) = \lambda |B| \ .$$

In order to avoid confusions, we say that a Poisson point process on the real line is stationary if, for a given positive integer k and Borel sets $B_1, \ldots, B_k$, we have for every real number t,

$$(N(B_1 + t), \ldots, N(B_k + t)) \stackrel{\mathcal{L}}{=} (N(B_1), \ldots, N(B_k)) \ .$$

The assumptions in Definition 5 allow to fully specify the finite-dimensional distribution of the point process N. In particular, one can show that if A and B are two disjoint Borel sets, then, $N(A \cup B) \stackrel{\mathcal{L}}{=} N(A) + N(B)$. In Daley and Vere-Jones (1988), it is shown that the stationary Poisson process with intensity λ can be fully defined by the equations

$$\mathbb{P}\left(N\left((a_i, b_i]\right) = n_i\ , i = 1, \ldots, k\right) = \prod_{i=1}^{k} \frac{\mathrm{e}^{-\lambda(b_i - a_i)} \lambda (b_i - a_i)^{n_i}}{n_i!}\ , n_1, \ldots, n_k \in \mathbb{N}\ ,\ k \in \mathbb{N} - \{0\}\ , \tag{3.2.1}$$

where $a_1, \ldots, a_k, b_1, \ldots, b_k$ are real numbers such that $a_i \leq b_i < a_{i+1}$ for $i = 1, \ldots, k-1$ and $N\left((a_i, b_i]\right)$ represents the random number of points falling in the set $(a_i, b_i]$ for $i = 1, \ldots, k$.

Representation of a Poisson process on the real line

Though the definition of a Poisson point process is clear, we would like both to exploit the representation of a point process as a point measure and to introduce a clearer notion of time in order to gain a better interpretation. To do so, we define the integer-valued stochastic process $(N_t)_{t \in \mathbb{R}}$ as follows

$$N_t := \begin{cases} N\left((0, t]\right) & \text{if } t > 0 \\ 0 & \text{if } t = 0 \\ -N\left((0, t]\right) & \text{if } t < 0 \end{cases} . \tag{3.2.2}$$

Since this process is increasing and right-continuous, we can define the sequence $(T_k)_{k \in \mathbb{Z}}$ by

$$T_k := \{t \in \mathbb{R}\ : N_t \geq k\}\ , \quad k \in \mathbb{Z}\ , \tag{3.2.3}$$

which can be understood as a sequence of points at which some event occurs. Then, we have the following theorem.

Theorem 3.2.1. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and N be a stationary Poisson point process with intensity λ defined on this probability space. Also, let $(N_t)_{t \in \mathbb{R}}$ be the stochastic process defined by (3.2.2) and $(T_k)_{k \in \mathbb{Z}}$ be the sequence of numbers defined by (3.2.3). Then,*

$$N = \sum_{k \in \mathbb{Z}} \delta_{T_k}\ .$$

Furthermore, the inter-arrival times $(\tau_k)_{k \in \mathbb{Z}}$ defined for any integer k by $\tau_k := T_{k+1} - T_k$ are i.i.d. and follow an exponential distribution with parameter λ.

From the previous theorem, we see that the knowledge of the sequence of inter-arrival times $(\tau_k)_{k \in \mathbb{Z}}$ and the point T_0 is sufficient to fully determine its associated Poisson point process. In the rest of this thesis, we will mainly deal with this representation and with stationary Poisson point processes in order to model the interaction times between the gamma photons and the detector.

Marked Poisson point processes

The stationary Poisson point processes presented in the previous sections are particularly useful to model several physical phenomena. For instance, assume that we want to model the electric current generated by a pulse of electrons (with arrival times given by $(T_k)_{k\in\mathbb{Z}}$) at the anode of a vacuum tube. Assuming that, for a single electron, we observe through the oscilloscope a current represented by a causal function $t \to h(t)$, then the overall current at time t is given by

$$U_t = \sum_{k\in\mathbb{Z}\,:T_k\leq t} h(t-T_k)\ . \tag{3.2.4}$$

Such a stochastic process is called a *filtered point process*.
Now, assume that we want to model a similar dynamical system which depends on an underlying point process that models the time arrivals of particles of various nature or causes, for instance river streamflow. Suppose that, at some fixed spatial point of a given river, we have in hand an instrument able to record the river level at this precise point. We suppose that the river level has a still water level corresponding to a measure equal to zero that can only increase in case of rain[3]. Modeling rainfalls using a point process, we assume that the river level jumps after a rainfall from the associated quantity of rain and then decreases until its still water level according to a function h. Then, we can not model the river level using a formula similar to Equation (3.2.4) because it would signify that the same amount of rain falls for each rainfall. Rather, we would like to write something like

$$W_t = \sum_{k\in\mathbb{Z}\,:\,T_k\leq t} Y_k h(t-T_k)\ , \quad t\in\mathbb{R}\ ,$$

where Y_k is the amount of rain that fell at the rainfall that occurred at time T_k. In order to properly define the previous equation, we need to formally introduce the concept of marked point process.

Definition 6 (Daley and Vere-Jones (1988), Marked point process)**.** *Let $\mathcal{E}$ be a countable separable metric space. We say that N defined by*

$$N = \sum_{k\in\mathbb{Z}} \delta_{(T_k,Y_k)}\ ,$$

is a marked point process with locations on $\mathbb{R}$ and marks on $\mathcal{E}$ if it is a point process such that the underlying process N_g defined for every Borel set A by

$$N_g(A) := N(A\times\mathcal{E}) = \sum_{k\in\mathbb{Z}} \delta_{T_k}(A)\ ,$$

is also a point process.
Moreover, we say that N has independent marks if, conditionally on the ground process $N_g = \sum_{k\in\mathbb{Z}}\delta_{T_k}$, the random variables $(Y_k)_{k\in\mathbb{Z}}$ are mutually independent.

[3]Of course, this is a very simplistic assumption.

In this thesis, we will consider marked point processes N whose marks $(Y_k)_{k\in\mathbb{Z}}$ are i.i.d. and independent from the arrival times $(T_k)_{k\in\mathbb{Z}}$.

Example 5 (Poisson marked point processes). Let $N = \sum_{k\in\mathbb{Z}} \delta_{(T_k,Y_k)}$ be a marked point process with locations and marks on the real line $\mathbb{R}$ such that:

1. the *ground process* N_g given by Definition 6 is a homogeneous point process on the real line with intensity λ.
2. the marks $(Y_k)_{k\in\mathbb{Z}}$ are i.i.d. and admit a probability density function f with respect to the Lebesgue measure on $\mathbb{R}$.

Then, the intensity ν defined on $\mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R})$ associated to the marked point process is of the form

$$\nu(A \times B) = \lambda|A| \int_B f(x)\mathrm{d}x = \lambda|A|\mathbb{P}\left(Y_0 \in B\right) , \quad A, B \in \mathcal{B}(\mathbb{R}) .$$

Thanks to this rigorous development, we are now able to introduce the concept of shot-noise processes.

3.2.2 Shot-noise processes

In its more general form, a real-valued shot-noise process is constructed using a marked point process

$$N = \sum_{k\in\mathbb{Z}} \delta_{(T_k,Z_k)}$$

with locations $(T_k)_{k\in\mathbb{Z}}$ on the real line and i.i.d. marks $(Z_k)_{k\in\mathbb{Z}}$ which belong to the space of applications $\mathbb{R}^{\mathbb{R}}$. The shot-noise process associated to N is then defined by the following equation

$$X_t := \sum_{k\in\mathbb{Z}} Z_k(t - T_k) , \quad t \in \mathbb{R} , \tag{3.2.5}$$

and each stochastic process $Z_k(t)$ for a given integer k is called an *impulse*.
In this manuscript, we only consider the following particular case: there exists a sequence of i.i.d. real-valued random variables $(Y_k)_{k\in\mathbb{Z}}$ independent of $(T_k)_{k\in\mathbb{Z}}$ and a causal, measurable, real-valued function h such that, for every integer k, the stochastic process Z_k can be written

$$Z_k(t) = Y_k\, h(t) , \quad t \in \mathbb{R} .$$

From now on, a shot-noise process will refer to the following definition.

Definition 7 (Shot-noise process)**.** *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, λ be a positive number, f a probability density function on $\mathbb{R}$ and $N = \sum_{k\in\mathbb{Z}} \delta_{(T_k,Y_k)}$ be the stationary Poisson marked point process with intensity function ν given by*

$$\nu(A \times B) = \lambda|A| \int_B f(y)\mathrm{d}y \ , \quad A, B \in \mathcal{B}(\mathbb{R}) \ ,$$

as defined in Example 5.
Let also h be a real-valued, measurable and causal function. Then, the stochastic process $(X_t)_{t\in\mathbb{R}}$ defined by

$$X_t := \sum_{k\in\mathbb{Z}} Y_k \ h(t - T_k) = \sum_{k:T_k\le t} Y_k \ h(t - T_k) \ , \quad t \in \mathbb{R} \tag{3.2.6}$$

is called a stationary shot-noise process with characteristic triplet (λ, f, h).
In order to be consistent with the vocabulary introduced in Chapter 2, the function h is called an impulse response.

3

A major question arises about the definition of a shot-noise process by Definition 7: is the process well defined for any triplet (λ, f, h)? In order to answer to this question, remark that for all t, we have almost surely

$$X_t \le |X_t| \le \sum_{k\in\mathbb{Z}} |Y_k| \, |h| \, (t - T_k) \ , \quad t \in \mathbb{R} \ .$$

Using the monotone convergence theorem, we have for all t in $\mathbb{R}$

$$\begin{aligned}\mathbb{E}\,[X_t] \le \sum_{T_k\le Z} \mathbb{E}\,[|Y_k| \, |h| \, (t - T_k)] &= \mathbb{E}\,[|Y_0|] \sum_{T_k\le t} \mathbb{E}\,[|h|(t - T_k)] \\ &= \lambda \int_{\mathbb{R}} |x| f(x)\mathrm{d}x \int_0^\infty |h|(s)\mathrm{d}s \ .\end{aligned}$$

where the last equality called *Campbell's formula* will be proved in the next subsection of this chapter. Then, a sufficient condition for a shot-noise process to be well defined is that its associated characteristic triplet satisfies the condition

$$\lambda \int_{\mathbb{R}} |x| f(x)\mathrm{d}x \int_0^\infty |h|(s)\mathrm{d}s < \infty \ .$$

It turns out that Iksanov and Jurek (2003) prove that a shot-noise process with a characteristic triplet (λ, f, h) is well defined if, and only if,

$$\lambda \int_0^\infty \int_{\mathbb{R}} \min\left(1, y \ h(s)\right) f(y)\mathrm{d}y\mathrm{d}s < \infty \ .$$

In Chapters 4 and 5, we will introduce different assumptions on f and h that ensure the shot-noise process to be well defined.

Laplace transform of a shot-noise process

The cornerstone of all the estimation procedures derived in this manuscript is based on the Laplace functional of a shot-noise process. Indeed, due to the context of our applications, we only have at disposal a low-frequency sample $X_\delta, \ldots, X_{n\delta}$ of the shot-noise process defined by (3.2.6) where n is an integer and δ some positive number. Such observations preclude the use of the joint laws of the samples so that we only deal with the marginal law of the shot-noise process.
In addition, it is in general impossible to analytically derive the probability density function of the random variable

$$X_0 = \sum_{T_k \leq 0} Y_k \, h(-T_k) \; .$$

Up to now, the only case where the analytical expression is known can be found in Bondesson (1992). If the sequence of the marks $(Y_k)_{k\in\mathbb{Z}}$ are i.i.d., follow an exponential distribution with parameter ρ and the function h is the form

$$h_\alpha : \mathbb{R}_+ \to \mathbb{R}_+ \; , \; t \to \mathrm{e}^{-\alpha t} \; , \quad \alpha > 0 \; ,$$

then the marginal X_0 of the shot-noise process associated to this triplet follows a Gamma distribution with shape parameter λ/α and scale parameter ρ.
We then need to find a mathematical object able to establish a link between the finite sample $X_\delta, \ldots, X_{n\delta}$ at hand of the shot-noise and its triplet using only the available knowledge of the marginal X_δ. This can be done using the *Laplace transform* of the random variable X_0. Indeed, this is the main object to handle random variables that can be viewed as filtered point processes and, more generally, as we will see in the next section, *infinitely divisible* random variables. Let us prove that the marginal X_0 of the stationary shot-noise process enjoys this property.
Let $(X_t)_{t\in\mathbb{R}}$ be the stationary shot-noise process defined by (3.2.6) with triplet (λ, f, h) and underlying point process N. Since the marked point process N can be seen as a point measure, then, for every real number t, the random variable X_t can be rewritten as follows[4].

$$X_t = \sum_{k\,:\,T_k \leq t} Y_k \, h(t - T_k) = \int_{\mathbb{R}^2} y \, h(t-s) \mathbb{1}_{\{t \geq s\}} \, N(\mathrm{d}s, \mathrm{d}y) =: N\left[F_t\right] \; ,$$

where for any real t, we set F_t as the real-valued function acting on $\mathbb{R}^2$ defined by

$$F_t(s, y) := y h(t-s) \mathbb{1}_{\{t \geq s\}} \; .$$

In addition, several random variables $N(\mathrm{f})$ can be constructed when replacing the functions $(F_t)_{t\in\mathbb{R}}$ by functions f that belong to the set

$$\mathcal{F}_N := \{\mathrm{f} : \mathbb{R}^2 \to \mathbb{R} \;:\; |N(\mathrm{f})| < \infty \text{ a.s.}\} \; .$$

[4]A good introduction to integration calculus with respect to Poisson random measures can be found in Applebaum (2009)[Section 2.3.2.]

For a given point process N, the *Laplace functional* corresponds to the mapping between $\mathcal{F}_N$ and $\mathbb{R}_+$ defined by

$$\Psi_N : \mathcal{F}_N \to \mathbb{R}_+ \ , \ \mathrm{f} \to \mathbb{E}\left[\mathrm{e}^{-N(\mathrm{f})}\right] \ . \tag{3.2.7}$$

However, probabilists often prefer to work with the characteristic function of a random variable because it fully determines the probability distribution and is defined on the real line $\mathbb{R}$. We recall that the characteristic function φ_X of a real-valued random variable X is the function defined on the real line by

$$\varphi_X(u) := \mathbb{E}\left[\mathrm{e}^{\mathrm{i}uX}\right] \ , \qquad u \in \mathbb{R} \ .$$

The following theorem provides an analytical expression of the Laplace transform and the characteristic function of a filtered point process on $\mathbb{R}^d$ for a positive integer d.

Theorem 3.2.2. *Let $(\Omega, \mathcal{A}, \mathbb{P}))$ be a probability space and N be a point process with intensity measure ν defined on $\mathbb{R}^d$ for some positive integer d.*
Then, for any function f *belonging to $\mathcal{F}_N$, we have*

$$\Psi_N(\mathrm{f}) = \mathbb{E}\left[\mathrm{e}^{-N(\mathrm{f})}\right] = \exp\left(\int_{\mathbb{R}^d} \left(\mathrm{e}^{-\mathrm{f}(x)} - 1\right) \nu(\mathrm{d}x)\right) \ . \tag{3.2.8}$$

Similarly, for every real number u, we have

$$\varphi_{N(\mathrm{f})}(u) = \mathbb{E}\left[\mathrm{e}^{\mathrm{i}uN(\mathrm{f})}\right] = \exp\left(\int_{\mathbb{R}^d} \left(\mathrm{e}^{\mathrm{i}u\mathrm{f}(x)} - 1\right) \nu(\mathrm{d}x)\right) \ . \tag{3.2.9}$$

Proof. The proof is available in Resnick (1992), p 353. □

As a consequence, the characteristic function of the marginal of the stationary shot-noise process defined by (3.2.6) can be written for all u in $\mathbb{R}$

$$\begin{aligned} \varphi_{X_0}(u) &= \exp\left(\lambda \int_0^\infty (\varphi_{Y_0}(uh(s)) - 1)\,\mathrm{d}s\right) \qquad (3.2.10) \\ &= \exp\left(\lambda \int_0^\infty \int_{\mathbb{R}} \left(\mathrm{e}^{\mathrm{i}uyh(s)} - 1\right) f(y)\mathrm{d}y\mathrm{d}s\right) \\ &=: \exp\left(\psi(u)\right) \ . \end{aligned}$$

Moments of a shot-noise process

It is well known that the moment of order n of a random variable X_0, when it exists, can be obtained via the formula

$$\mathbb{E}\left[X_0^n\right] = (-\mathrm{i})^n \varphi_{X_0}^{(n)}(0) \ .$$

Since the calculation of high order terms can be painful, we would rather prefer to use the cumulants of X_0. When it makes sense, the cumulant of order n of a random variable X_0 is given by

$$\kappa_{X_0}^n := (-\mathrm{i})^n \left.\frac{\mathrm{d}^n \mathrm{Log}\,(\varphi_{X_0})}{\mathrm{d}u^n}\right|_{u=0} \ ,$$

where for a complex-valued function z, Log(z) represents, when it exists, the continuous function that satisfies

$$z = \exp(\mathrm{Log}(z)) \ .$$

As we will see in a few lines, $\mathrm{Log}(\varphi_{X_0})$ exactly coincides with the function ψ defined by (3.2.10). Thus, the cumulant of order n of X_0 is given by

$$\begin{aligned} \kappa^n_{X_0} &= (-\mathrm{i})^n \psi^{(n)}(0) \\ &= \lambda \int_{\mathbb{R}} x^n f(x) \mathrm{d}x \int_0^\infty h^n(s) \mathrm{d}s \ . \end{aligned} \tag{3.2.11}$$

For instance, Elter et al. (2016) use Equations (3.2.11) when both the impulse response h and the mark's density f are known in order to estimate either the intensity λ.
However, we believe that only exploiting the successive derivatives of ψ at zero is not as efficient as exploiting the information contained in the whole function ψ or its derivative. In Chapters 4 and 5, we introduce statistical inference methods based on the function ψ' which are strongly suggested by those used for the statistical inference of Lévy processes. Indeed, as we will develop in the next section, the marginal X_0 of the shot-noise process given by (3.2.5) is *infinitely divisible* and infinitely divisible distributions are related to stochastic processes called *Lévy processes*. The next section provides an introduction to Lévy processes and statistical estimations techniques used by Neumann and Reiß (2009) and Gugushvili (2009) that rely on nonparametric methods that make use of the *Lévy or characteristic exponent*.

3

3.3 Lévy processes

This section is devoted to Lévy processes and the related statistical methods. We first introduce the concept of infinitely divisible distribution and derive some properties regarding the characteristic function of such a distribution. Then, we define when a stochastic process $(X_t)_{t\geq 0}$ is qualified of Lévy process and detail how, from a finite sample $X_1, \ldots, X_n$ of a Lévy process, we can exploit the characteristic function of X_1 to fully determine the distribution of the aforementioned process.

3.3.1 Definitions and properties

Prior to introducing Lévy process, we need to define what is an *infinitely divisible* random variable.

Definition 8. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and X a real-valued random variable. X is said to be infinitely divisible if, for every integer n, there exists n i.i.d. real-valued random variables $X_1^{(n)}, \ldots, X_n^{(n)}$ such that*

$$X \overset{\mathcal{L}}{=} X_1^{(n)} + \cdots + X_n^{(n)} \ .$$

This definition can be generalized to stochastic processes as well as probability measure in the following manner: a probability measure μ on $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$ is infinitely divisible if, for every positive integer n, one can find a probability measure μ_n such that

$$\mu = \mu_n^{\star n} ,$$

where the right-hand side term of the above equation denotes the n-fold convolution of μ_n with itself and the convolution product for two probability measures μ, ν is defined by

$$\mu \star \nu(A) := \int_{\mathbb{R}} \int_{\mathbb{R}} \mathbb{1}_A(x+y)\mu(\mathrm{d}x)\nu(\mathrm{d}y) , \quad A \in \mathcal{B}(\mathbb{R}) .$$

Example 6 (Gaussian r.v.). Any nondegenerate Gaussian random variable X with mean μ and variance $\sigma^2 > 0$ is infinitely divisible. Indeed, for a positive integer n, we can write

$$X \stackrel{\mathcal{L}}{=} X_1^{(n)} + \cdots + X_n^{(n)} ,$$

where $X_1^{(n)}, \ldots, X_n^{(n)}$ are Gaussian i.i.d. r.v.'s with mean μ/n and variance σ^2/n.

3

Example 7 (Compound Poisson distribution). Let $(Y_k)_{k\geq 1}$ a sequence of real-valued i.i.d. random variables and N be an integer-valued random variable following a Poisson distribution with parameter $\lambda > 0$ and independent of the sequence (Y_k). The random variable X defined by

$$X := \begin{cases} \sum_{k=1}^{N} Y_k & \text{if } N \neq 0 \\ 0 & \text{else} \end{cases} .$$

is said to follow a *Compound Poisson distribution* and is infinitely divisible.
To see that, let n be a positive integer, $(N_n^{(1)}, \ldots, N_n^{(n)})$ be i.i.d. random variables following a Poisson distribution with parameter λ/n and $\left((Y_k^{(1)})_{k\geq 1}, \ldots, (Y_k^{(n)})_{k\geq 1}\right)$ be i.i.d. sequences of i.i.d. r.v.'s distributed as Y_1 that are independent of the r.v.'s $(N_n^{(1)}, \ldots, N_n^{(n)})$. Then, setting for $k = 1, \ldots, n$

$$X_n^{(k)} := \begin{cases} \sum_{j=1}^{N_n^{(k)}} Y_j^{(k)} & \text{if } N_n^{(k)} \neq 0 \\ 0 & \text{else} \end{cases} ,$$

it leads to

$$X \stackrel{\mathcal{L}}{=} X_n^{(1)} + \cdots + X_n^{(n)} .$$

The characteristic function of an infinitely divisible random variable X is a particularly fruitful mathematical object to characterize the distribution of X. As explained in the next lines, there exists a one-to-one mapping between an infinitely divisible r.v. X with values in $\mathbb{R}$ and a triplet (γ, Σ, ν) called *Lévy-Khintchine triplet* where

- γ is some vector in $\mathbb{R}$,

- Σ is a positive semi-definite matrix of size d,
- ν is a measure on $\mathbb{R} - \{0\}$ with additional properties.

Prior to that, let us state an important result that is used throughout this manuscript.

Proposition 2. *Let d denote a positive integer, X be a infinitely divisible random variable taking values in $\mathbb{R}^d$ and φ denote its characteristic function. Then,*

$$\forall u \in \mathbb{R}^d \ , \ \varphi(u) \neq 0 \ . \tag{3.3.1}$$

Proof. For a given positive integer n, let φ_n denote the characteristic function of the i.i.d. random variables $(X_k^{(n)})_{1 \leq k \leq n}$ such that

$$X \overset{\mathcal{L}}{=} X_1^{(n)} + \cdots + X_n^{(n)} \ .$$

It thus gives for all $u \in \mathbb{R}^d$ that

$$\varphi_n(u) = \varphi(u)^{1/n} \ .$$

For a given vector $u \in \mathbb{R}^d$, let $\varphi_\infty(u)$ denote the simple limit of the sequence of complex numbers $\{\varphi_n(u) \ , \ n \in \mathbb{N} - \{0\}\}$. Since $|\varphi| \leq 1$, it is obvious that the function φ_∞ can only takes the values 0 or 1. If we prove that φ_∞ actually is a characteristic function, then by continuity, it follows that $\varphi_\infty \equiv 1$ or $\varphi_\infty \equiv 0$. Since $\varphi_\infty(0) = \lim_{n\to\infty} \varphi_n(0) = 1$, we necessarily have that $\varphi_\infty \equiv 1$ and then φ does not vanish.
The proof that φ_∞ is a characteristic function can be established by invoking the Bochner's theorem or Lévy's continuity theorem. □

This result provides a simple criterion to determine whether a random variable is infinitely divisible or not, knowing its characteristic function. Furthermore, the previous result ensures that we can define the logarithm of the characteristic function of a infinitely divisible distribution[5].
The study of the characteristic function of an infinitely divisible distribution is an active field of research, essentially because it governs the efficiency of the statistical inference methods used for Lévy processes. Recently, Trabs (2014) develops criteria based on the Lévy-Khintchine triplet in order to determine when the characteristic function decays at a polynomial speed. Before stating these results and applying it to shot-noise processes, we present the major theorem of this section and make explicit the Lévy-Khintchine triplet.

Theorem 3.3.1 (Lévy-Khintchine's theorem)**.** *Let X be a $\mathbb{R}$-valued random variable and φ denote its characteristic function. Then, X is infinitely divisible if, and only if, there exists a triplet (γ, σ, ν) called Lévy-Khintchine triplet where*

1. *γ is a real number,*
2. *σ^2 is a nonnegative number,*

[5] At least its principal branch. Theorem 3.3.1 explicitly gives the continuous branch of the logarithm.

3. ν is a Lévy measure on $\mathbb{R} - \{0\}$,

such that, for every u in $\mathbb{R}$, we have

$$\varphi(u) = \exp\left(\mathrm{i}\gamma u - \frac{1}{2}u^2\sigma^2 + \int_{\mathbb{R}-\{0\}} \left(\mathrm{e}^{\mathrm{i}ux} - 1 - \mathrm{i}ux 1_{|x|\le 1}\right)\nu(\mathrm{d}x)\right) . \qquad (3.3.2)$$

A Borel measure ν on $\mathbb{R} - \{0\}$ is said to be a Lévy measure whenever it satisfies the condition

$$\int_{\mathbb{R}-\{0\}} \min(1, |y|^2)\nu(\mathrm{d}y) < \infty .$$

Remark 1. To an infinitely divisible distribution corresponds a unique Lévy triplet (γ, σ, ν). However, some authors define another representation and replace the term

$$\mathrm{e}^{\mathrm{i}ux} - 1 - \mathrm{i}ux 1_{|x|\le 1} , \quad u, x \in \mathbb{R}$$

by

$$\mathrm{e}^{\mathrm{i}ux} - 1 - \mathrm{i}\frac{ux}{1 + |x|^2} , \quad u, x \in \mathbb{R} .$$

It does not impact the parameters σ and ν in the representation but γ has to be modified accordingly. In fact, as mentioned in Applebaum (2009), the last term can be replaced by any term of the form

$$\mathrm{e}^{\mathrm{i}ux} - 1 - \mathrm{i}uc(x)$$

whenever the function c is chosen so that the above term is ν-integrable.

The reader eager to read a proof can refer to Sato (1999). What this theorem says is extremely powerful. It states that every infinitely random variable is the sum of three independent random variables: a constant, a centered Gaussian random variable with variance σ^2 and a *pure-jump process*[6], i.e. a superposition of independent Poisson point processes with different jump sizes.
Up to this point, let us come back to the stationary shot-noise process $(X_t)_{t\in\mathbb{R}}$ introduced in (3.2.6). Since the marginal X_0 is infinitely divisible, it would be interesting to relate the triplet (λ, f, h) of the shot-noise process with the Lévy triplet (γ, σ, ν). We then have the following proposition

Proposition 3. *Let λ be a positive number, h be a causal, measurable, real-valued function and f be a probability density function such that the following condition is satisfies*

$$\lambda \int_0^\infty \int_{\mathbb{R}} \min(1, yh(s)) f(y)\mathrm{d}y\mathrm{d}s < \infty .$$

Also, let $(X_t)_{t\in\mathbb{R}}$ be the stationary shot-noise process with triplet (λ, f, h) defined by (3.2.6) and Y be a random variable with p.d.f. f. Then, the marginal X_0 admits the characteristic triplet $(\gamma, 0, \nu)$ where

$$\nu(B) = \int_0^\infty \mathbb{P}\left(h(s)Y \in B\right) \mathrm{d}s ,$$

[6]Further details are available in Applebaum (2009).

and

$$\gamma = \int_{|y|\leq 1} y\nu(\mathrm{d}y) \ .$$

TRANSITION

Definition 9 (Independent and stationary increments)**.** *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(X_t)_{t\geq 0}$ a stochastic process. We say that $(X_t)_{t\geq 0}$ has independent increments when, for every positive integer n and for every choice of nonnegative numbers $(t_k)_{0\leq k\leq n}$ such that*

$$0 \leq t_0 < t_1 < \cdots < t_n \ ,$$

the random variables

$$X_{t_1} - X_{t_0}, \ldots, X_{t_n} - X_{t_{n-1}}$$

are independent. When we have

$$X_{t_{k+1}} - X_{t_k} \overset{\mathcal{L}}{=} X_{t_{k+1}-t_k} - X_0 \quad \textit{for } k = 0, \ldots, n-1 \ ,$$

we say the stochastic process $(X_t)_{t\geq 0}$ has stationary increments.

3

Definition 10. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(X_t)_{t\geq 0}$ a stochastic process. We say that $(X_t)_{t\geq 0}$ is a Lévy process when*

- *$X_0 = 0$ almost surely,*
- *$(X_t)_{t\geq 0}$ has stationary and independent increments*
- *$(X_t)_{t\geq 0}$ is stochastically continuous: for every positive number ϵ, we have*

$$\lim_{t\searrow 0} \mathbb{P}\left(|X_t| > \epsilon\right) = 0 \ .$$

Remark 2. Some authors define a Lévy process with the additional property that its sample paths are càdlàg (i.e. almost surely right continuous with left limits). However, it is not necessary to assume it, since, from a Lévy process defined according to Definition 10, one can always construct a càdlàg version.

From these definitions, the two following properties can be easily established.

Proposition 4. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(X_t)_{t\geq 0}$ be a $\mathbb{R}$-valued Lévy process. Then, for all positive t, X_t is infinitely divisible.*

Proposition 5. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space, $(X_t)_{t\geq 0}$ be a $\mathbb{R}$-valued Lévy process and let φ_{X_t} denote the characteristic function of the marginal X_t for any positive number t. Then, for every $t \geq 0$, we have for all real u*

$$\varphi_{X_t}(u) = \exp\left(t\eta(u)\right)$$

where η is the Lévy exponent of the random variable X_1, which, according to Proposition 4, admits a Lévy characteristic triplet (γ, σ, ν), such that for every u in $\mathbb{R}$,

$$\eta(u) = \mathrm{i}\gamma u - \frac{1}{2}u^2\sigma^2 + \int_{\mathbb{R}-\{0\}} \left(\mathrm{e}^{\mathrm{i}ux} - 1 - \mathrm{i}ux\mathbb{1}_{\{|x|\leq 1\}}\nu(\mathrm{d}x)\right) \ .$$

Proposition 5 shows that a Lévy process is fully determined by the knowledge of its characteristic triplet. We are now in possession of all the tools needed to detail the recent statistical methods developed for Lévy processes.

3.3.2 Statistical inference of the triplet of a Lévy process

In Neumann and Reiß (2009) and Gugushvili (2009), the authors introduce nonparametric methods, using a finite sample $X_1, \dots, X_n$ of a real-valued Lévy process with triplet (γ, σ, ν), in order to estimate the aforementioned triplet. In these two papers, the main idea consists in exploiting the characteristic function of the Lévy process $(X_t)_{t\geq 0}$ given by Proposition 5. More precisely, the random variables $(X_k - X_{k-1})_{k=1,\dots,n}$ (called *increments*) are i.i.d. and follow the same law as X_1 with associated characteristic function φ_1. The empirical characteristic function of the increments is then defined by

$$\hat{\varphi}_n(u) := n^{-1} \sum_{k=1}^{n} \mathrm{e}^{\mathrm{i}u(X_k - X_{k-1})} , \quad u \in \mathbb{R} .$$

By Glivenko-Cantelli's theorem, we have that $\hat{\varphi}_n$ converges to φ_1 almost surely.
In Gugushvili (2009), the author additionally assumes that the Lévy measure ν has a finite total mass (i.e. $\nu(\mathbb{R}) < \infty$) and admits a density ρ with respect to the Lebesgue measure such that

$$\int_{\mathbb{R}} x^2 \rho(x) \mathrm{d}x < \infty .$$

In this case, we have for every real u that

$$\frac{\varphi_1'(u)}{\varphi_1(u)} = \mathrm{i}\gamma - \sigma^2 u + \mathrm{i} \int_{\mathbb{R}} \mathrm{e}^{\mathrm{i}ux} x \rho(x) \mathrm{d}x .$$

Differentiating the above equation with respect to u and using Fourier's inversion, it is easy to obtain that for every real x, we have

$$x^2 \rho(x) = \frac{1}{2\pi} \int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}ux} \left(\frac{\left(\varphi_1'(u)\right)^2 - \varphi_1''(u)\varphi_1(u)}{(\varphi_1(u))^2} - \sigma^2 \right) \mathrm{d}u .$$

Gugushvili (2009) then constructs an estimator $\hat{\sigma}_n^2$ of σ^2 using a clever transformation of the functional $\log |\varphi_1|$. Then, it is possible to estimate the density ρ of the Lévy measure ν using the plug-in procedure given by

$$\hat{\rho}_n(x) = \frac{1}{2\pi x} \int_{|u|\leq h_n} \mathrm{e}^{-\mathrm{i}ux} \left(\frac{(\hat{\varphi}_n'(u))^2 - \hat{\varphi}_n''(u)\hat{\varphi}_n(u)}{(\hat{\varphi}_n(u))^2} - \hat{\sigma}_n^2 \right) \mathrm{d}u , \quad x \neq 0 ,$$

where $(h_n)_{n\geq 0}$ is a sequence of positive numbers that tends to infinity that can be chosen adaptively. The performance of the estimator $\hat{\rho}_n$ are then studied using the mean square integrated error (MISE) defined by

$$\mathrm{MISE}[\hat{\rho}_n] := \mathbb{E}\left[\int_{\mathbb{R}} |\hat{\rho}_n(x) - \rho(x)|^2 \, \mathrm{d}x \right] , \tag{3.3.3}$$

3

and the author proves both upper and lower rates of convergence assuming that the density ρ satisfies the conditions given by

$$\int_{\mathbb{R}} x^2 \rho(x) \mathrm{d}x \leq K \ , \qquad \int_{\mathbb{R}} \rho(x) \mathrm{d}x \leq \Lambda \ .$$

and

$$\int_{\mathbb{R}} |u|^{2m} \left|\mathcal{F}[\rho](u)\right|^2 \mathrm{d}u \leq L$$

where Λ, K, L, m are positive numbers and $\mathcal{F}[\rho]$ denotes the Fourier transform of ρ.
In Neumann and Reiß (2009), the estimators of the Lévy triplet are constructed using a minimum distance estimator that we briefly recall before going into details in their procedure.

Definition 11 (Minimum distance estimation)**.** *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be a probability space and $(X_1, \ldots, X_n)$ random variables with joint law P_{θ_0}, where θ_0 belongs to some set Θ. Assume that we can construct an E-valued estimator*

$$\hat{g}_n \equiv \hat{g}_n(X_1, \ldots, X_n) \ ,$$

that converges in probability to a function $g(\theta_0)$ where (E, d) is some metric space.
Then, the minimum distance-fit estimator $\hat{\theta}_n$ of θ_0 associated to the distance d is defined by

$$\hat{\theta}_n := \underset{\theta \in \Theta}{argmin}\ d\left(\hat{g}_n, g(\theta)\right) \ .$$

In their article, the authors write the characteristic function of X_1 for every real u

$$\begin{aligned}
\varphi_1(u) &= \exp\left(\mathrm{i}u\gamma - \sigma^2 u^2/2 + \int_{\mathbb{R}} \left(\mathrm{e}^{\mathrm{i}ux} - 1 - \frac{\mathrm{i}ux}{1+x^2}\right)\nu(\mathrm{d}x)\right) \\
&= \exp\left(\mathrm{i}u\gamma + \int_{\mathbb{R}} \frac{\left(\mathrm{e}^{\mathrm{i}ux} - 1\right)\left(1 + x^2\right) - \mathrm{i}ux}{x^2} \nu_\sigma(\mathrm{d}x)\right) \\
&=: \varphi(u, \gamma, \nu_\sigma) \ ,
\end{aligned}$$

where ν_σ is the measure defined by

$$\nu_\sigma(\mathrm{d}x) = \sigma^2 \delta_0(\mathrm{d}x) + \frac{x^2}{1+x^2}\nu(\mathrm{d}x) \ .$$

This measure is used for estimation purpose since it is statistically impossible to estimate the volatility parameter σ in a uniformly consistent way (see Neumann and Reiß (2009)[Remark 3.2.]): instead, the authors estimate the parameters (γ, ν_σ). Note that the measure ν_σ satisfies

$$\mathcal{F}[\nu_\sigma](u) = -\frac{\left(\varphi_1'(u)\right)^2 - \varphi_1''(u)\varphi_1(u)}{\left(\varphi_1(u)\right)^2} \qquad u \in \mathbb{R} \ ,$$

which shows that the estimation of ν_σ is closely connected to the estimation of the characteristic function φ_1 in a C^2-sense. From the finite sample $X_1, \ldots, X_n$ of a Lévy process with triplet (γ, σ, ν), Neumann and Reiß (2009) then construct a sequence of estimators $(\hat{\gamma}_n, \hat{\nu}_{\sigma,n})_{n\in\mathbb{N}}$ using a minimum distance-fit between the empirical characteristic function of the increments (and its first and second derivate) and the characteristic function φ_1 of X_1 by

$$(\hat{\gamma}_n, \hat{\nu}_{\sigma,n}) := \underset{\tilde{\gamma}\in\mathbb{R}, \tilde{\nu}_\sigma\in\mathcal{M}(\mathbb{R})}{\operatorname{argmin}} \; d\left(\hat{\varphi}_n, \varphi(\cdot, \tilde{\gamma}, \tilde{\nu}_\sigma)\right) ,$$

where d denotes the metric defined by

$$d(\varphi_1, \varphi_2) := \sum_{k=0}^{2} \|\varphi_1^{(k)} - \varphi_2^{(k)}\|_{L^\infty(w)}$$

for w a positive bounded function. While it can be easily established that the sequence of estimators $(b_n)_{n\geq 0}$ converges to b in $O_{\mathbb{P}}\left(n^{-1/2}\right)$, the performance of the estimators $(\hat{\nu}_{\sigma,n})_n$ is evaluated using the metric defined for $s \geq 0$ by

$$l_s(\hat{\nu}_{\sigma,n}, \nu_\sigma) := \sup_{f\in F_s} \left| \int_{\mathbb{R}} f(x)\hat{\nu}_{\sigma,n}(\mathrm{d}x) - \int_{\mathbb{R}} f(x)\nu_\sigma(\mathrm{d}x) \right| ,$$

where F_s is the smoothness class of measurable functions given by

$$F_s := \{f : \mathbb{R} \to \mathbb{R} \;:\; f \text{ is measurable and } \int_{\mathbb{R}} (1 + |u|)^s \, |\mathcal{F}[f](u)| \, \mathrm{d}u \leq 1\} .$$

Then, Neumann and Reiß (2009) derive optimal rates of convergence depending on how fast the characteristic function φ_1 decays. For instance, if the Lévy process admits a Gaussian part, i.e. $\sigma > 0$, then it is possible to find a positive constant C such that for every real u, we have

$$|\varphi_1(u)| \geq C\mathrm{e}^{-\sigma^2 u^2/2} .$$

Given a positive s, the optimal rate of convergence with respect to the metric l_s is in $O_{\mathbb{P}}\left((\log n)^{-s/2}\right)$.

When the Lévy process has no Gaussian part and that it is possible to find[7] two positive constants C and d such that

$$|\varphi_1(u)| \geq C\,(1 + |u|)^{-d} \;, \quad u \in \mathbb{R} ,$$

the rates of convergence for l_s become polynomial in $O_{\mathbb{P}}\left(n^{-\min(s/2d, 1/2)}\right)$.

From this, we see that the rates of convergence strongly depends on how the characteristic function decays. Such results can be compared to the problem of deconvolution[8] (see

[7] The marginal of the exponential stationary shot-noise introduced in Chapter 4 is an example of such a distribution.

[8] which consists in estimating the density f of X given a finite sample of i.i.d. observations from $Y = X + \epsilon$, where ϵ is an error term.

3

Fan (1991) for details): the faster the characteristic function of the error ϵ decays, and thus the smoother the distribution of ϵ is, the harder it is to estimate the density of X. Recently, Trabs (2014) has established necessary and sufficient conditions under which the characteristic function φ_1 associated to a Lévy triplet (γ, σ, ν) decays in a polynomial fashion.
The estimator of the mark's density of the exponential shot-noise process defined in Chapter 4 uses similar arguments. Moreover, the rates of convergence obtained are polynomial because the marginal laws of the process belongs to a subclass of infinitely divisible distributions called *self-decomposable* (see Sato (1999) for a definition and subsequent properties).

3.4 Related stochastic processes

3

Shot-noise processes have a close relationship with other stochastic processes. Among them, we detail two of them that are particularly popular because they cast jump diffusion and allow to model complex dynamics over time.

3.4.1 Lévy semi-stationary processes

The *Lévy semi-stationary processes* are the unidimensional analogue of the *Ambit fields*. Usually, a Lévy semi-stationary process is a stochastic process $(X_t)_{t\in\mathbb{R}}$ defined by

$$X_t := \int_{-\infty}^{t} h(t-s)\sigma_s \mathrm{d}L_s + \int_{-\infty}^{t} q(t-s)a_s \mathrm{d}s + \gamma\,, \quad t \in \mathbb{R}\,, \tag{3.4.1}$$

where

- γ is a real number,
- σ and a are càdlàg processes,
- h and q and measurable causal functions,
- $(L_t)_{t\in\mathbb{R}}$ is a two-sided Lévy process.

When only the first in the right-hand side of (3.4.1) is nonzero, the process is called a *Lévy-driven moving average process* because there are the continuous-time analogue of the discrete-time moving average processes. Some results and applications can be found in Fasen (2005), Fasen (2006) or Cohen and Lindner (2013).
It is easy to see that the Lévy semi-stationary processes encompass the class of Lévy processes and shot-noise processes when choosing an appropriate function h and a process σ in the first term of the right-hand side of Eq. (3.4.1) and setting the remaining terms to zero.
Those processes have gained a growing interest in the past ten years because of their ability to jointly incorporate several traits, see for instance Barndorff-Nielsen et al. (2010)

and its applications to energy spot prices or Benth et al. (2014) for path-dependent options in finance.
However, estimation methods based on a finite sample of low frequencies observations of a Lévy semi-stationary are rare or restricted to particular cases. For example, Bennedsen et al. (2014) restricts to the case of a Brownian semi-stationary process (i.e. a Lévy semi-stationary process when the two-sided Lévy process $(L_t)_{t\in\mathbb{R}}$ actually is a Brownian motion) and apply estimators developed in the context of high-frequency observations.

3.4.2 CARMA processes

The CARMA processes constitute a family of stochastic processes that were introduced by Brockwell and Davis (1991) and which can be interpreted as the convolution between a linear filter and a Lévy process. The definition given herebelow can be found in Brockwell and Schlemm (2013a).

Definition 12 (CARMA process)**.** *Let p and q be two nonnegative integers such that $0 \leq q < p$. Let also $a_1, \dots, a_p, b_0, \dots, b_p$ be nonnegative numbers such that $b_q \neq 0$ and $b_k = 0$ for $q < k < p$.*
A CARMA(p,q) [9] *process with parameters $a_1, \dots, a_p$ and $b_0, \dots, b_q$ $(Y_t)_{t\geq 0}$ is formally defined by*

$$a(D)Y_t = b(D)DL_t \,, \quad t \geq 0 \,,$$

where D denotes the differentiation operator with respect to t, $(L_t)_{t\geq 0}$ is a Lévy process such that $\mathbb{E}\left[L_1^2\right] < \infty$ and

$$\begin{aligned} a(z) &:= z^p + a_1 z^{p-1} + \cdots + a_p \,, \\ b(z) &:= b_0 + b_1 z^1 + \cdots + b_{p-1} z^{p-1} \,. \end{aligned}$$

Alternatively, the process $(Y_t)_{t\geq 0}$ can be rewritten using the following state-space representation

$$\begin{cases} Y_t = \underline{b}^{\mathrm{T}} \underline{X}_t \\ \mathrm{d}\underline{X}_t = A\underline{X}_t \mathrm{d}t + \underline{e}\mathrm{d}L_t \end{cases} , \quad t \geq 0 \,.$$

where

$$A = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ -a_p & -a_{p-1} & -a_{p-2} & \cdots & -a_1 \end{bmatrix} , \; \underline{b} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{p-2} \\ b_{p-1} \end{bmatrix} , \; \underline{e} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} .$$

It can be shown (see Brockwell (2001) for instance) that $(\underline{X}_t)_{t\geq 0}$ has the form

$$\underline{X}_t = \mathrm{e}^{At}\underline{X}_0 + \int_0^t \mathrm{e}^{As}\underline{e}\mathrm{d}L_s \,, \quad t \geq 0 \,. \tag{3.4.2}$$

[9] which stands for Continuous-time ARMA.

Note that in the unidimensional case, i.e. when $p = 1$, $q = 0$ and $A = -a_1$, the process $(X_t)_{t\geq 0}$ solution of the stochastic differential equation

$$\mathrm{d}X_t = -a_1 X_t \mathrm{d}t + \mathrm{d}L_t \ , \quad t \geq 0 \ ,$$

admits, whenever $a_1 > 0$, a stationary solution which corresponds to a Lévy semi-stationary process with an exponential impulse response. In the multidimensional case, the process $(\underline{X}_t)_{t\geq 0}$ admits a weakly stationary solution whenever the eigenvalues of the matrix A have negative real parts.
The statistical study of CARMA processes has been investigated by many authors, see for instance Brockwell and Schlemm (2013b), García et al. (2011), Brockwell et al. (2013). Among them, the paper that is the most related to our problem is Brockwell and Schlemm (2013b): assuming that the CARMA process $(Y_t)_{t\geq 0}$ is continuously observed, the authors construct estimators of the increments $(L_{k+1}-L_k)_{k\in\mathbb{N}}$ of the driving Lévy process. However, these results cannot be extended to general impulse responses h because the corresponding processes fail to be Markovian.
Indeed, from (3.4.2), it is obvious that the state-space representation process $(\underline{X}_t)_{t\geq 0}$ of a CARMA process is a Markov process. This property allows to prove sharp results concerning the dependence structure of the CARMA process $(Y_t)_{t\geq 0}$, which in turn imply both the consistency and the rates of convergence of the estimators of the driving Lévy process $(L_t)_{t\geq 0}$ developed by Brockwell and Schlemm (2013b). Indeed, the dependence structure of stationary stochastic processes is often made precise by using mixing rates, especially strong α or β mixing. We recall the definition of strong mixing while the definition of β mixing coefficients is given in Chapter 4. Given a stationary stochastic process $(Z_t)_{t\in\mathbb{R}}$ and two real numbers n, m such that $n < m$, let $\mathcal{F}_n^m$ be the σ-algebra defined by $\mathcal{F}_n^m := \sigma\left(Z_t \ , n < t < m\right)$. The α-strong mixing coefficient $\alpha(m)$ is defined as

$$\alpha(m) := \sup_{A\in\mathcal{F}_{-\infty}^0 \ , \ B\in\mathcal{F}_m^\infty} |\mathbb{P}(A\cap B) - \mathbb{P}(A)\mathbb{P}(B)| \ ,$$

and the stochastic process $(Z_t)_{t\in\mathbb{R}}$ is said to be respectively *strongly mixing* or exponentially strongly mixing if $\lim_{m\to\infty}\alpha(m) = 0$ or $|\alpha(m)| \leq \mathrm{e}^{-am}$ for some positive number a.
In Marquardt and Stelzer (2007), the following result proves that under some mild moment condition, the CARMA processes are exponentially strongly mixing.

Proposition 6 (Marquardt and Stelzer (2007), Proposition 3.34)**.** *Let $(L_t)_{t\geq 0}$ be a Lévy process, p and q be two nonnegative integers such that $0 \leq q < p$ and let $(Y_t)_{t\geq 0}$ the stationary version of the CARMA(p, q) process given by Definition (12). If there exists a positive number r such that*

$$\mathbb{E}\left[|L_1|^r\right] < \infty \ ,$$

then, the CARMA process $(Y_t)_{t\geq 0}$ is exponentially α-mixing.

4 Nonparametric estimation of an exponential shot-noise process

Abstract

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

4.1 Introduction

In this paper, we consider a nonlinear inverse problem arising in nuclear science: neutron transport or gamma spectroscopy. For the latter, a radioactive source, for instance an excited nucleus, randomly emits gamma photons according to a homogeneous Poisson point process. These high frequency radiations can be associated to high energy photons which interact with matter via three phenomena : the photoelectric absorption, the Compton scattering and the pair production (further details can be found in Knoll (1989)). When photons interact with the semiconductor detector (usually High-Purity Germanium (HPGe) detectors) arranged between two electrodes, a number of electron-holes pairs proportional to the photon transferred energy is created. Accordingly, the electrodes generate an electric current called impulse response whenever the detector is hit by a particle, with an amplitude corresponding to the transferred energy. In this context, a feature of interest is the distribution of this energy. Indeed, it can be compared to known spectra in order to identify the composition of the nuclear source. In practice, the electric current is not continuously observed but the sampling rate is typically smaller than the mean inter-arrival time of two photons. Therefore, there is a high probability that several photons are emitted between two measurements so that the energy deposited

is superimposed in the detector, a phenomenon called pile-up. Because of the pile-up, it is impossible to establish a one-to-one correspondence between a gamma ray and the associated deposited energy.
This inverse problem can be modeled as follows. The electric current generated in the detector is given by a stationary shot-noise process $\mathbf{X} = (X_t)_{t\in\mathbb{R}}$ defined by:

$$X_t = \sum_{k:T_k \le t} Y_k\, h(t - T_k)\ , \tag{4.1.1}$$

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

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

We wish to estimate the density f from a regular observation sample $X_1, \ldots, X_n$ of the shot noise (4.1.1). Note that the sampling rate is set to 1 without meaningful loss of generality. If a different sampling rate is used, e.g. we observe $X_\delta, \ldots, X_{n\delta}$ for some $\delta \neq 1$, it amounts to change λ and to scale h accordingly.
The process (4.1.1) is well defined whenever the following condition holds on the impulse response h and the density f

$$\int \min(1, |y\, h(s)|) f(y) \mathrm{d}y \mathrm{d}s < \infty\ . \tag{4.1.2}$$

As shown in Iksanov and Jurek (2003), this condition is also necessary. Moreover, the marginal distribution of $\mathbf{X}$ belongs to the class of infinitely divisible (ID) distributions and has Lévy measure ν satisfying, for all Borel sets B in $\mathbb{R} \setminus \{0\}$,

$$\nu(B) := \lambda \int_0^\infty \mathbb{P}\left(h(s)Y_1 \in B\right)\ \mathrm{d}s\ . \tag{4.1.3}$$

The ID property of the marginal distribution shows that this estimation problem is closely related to the estimation of the Lévy measure ν. This property strongly suggests to use estimators of the Lévy triplet, see for instance Neumann and Reiß (2009) and Gugushvili (2009). However, up to our best knowledge, these estimators use the increments of the corresponding Lévy process which are i.i.d. and they assume a finite Lévy Khintchine measure. In contrast, the observations are not independent and the Lévy measure of the process is infinite since from (4.1.3), we have that

$$\nu(\mathbb{R}) = \lambda \int_0^\infty \mathbb{P}\left(h(s)Y_1 \in \mathbb{R}\right) \mathrm{d}s = \infty. \tag{4.1.4}$$

In order to tackle this estimation problem, we then propose to bypass the estimation of ν and directly retrieve the density f of the marks distribution F from the empirical characteristic function of the measurements. Coarsely speaking, using (4.1.3), the Lévy-Khintchine representation provides an expression of the characteristic function φ_X

of the marginal distribution as a functional of f. The estimator is built upon replacing φ_X by its empirical version and inverting the mapping $f \mapsto \varphi_X$. A more standard marginal-based approach would be to rely on the p.d.f. of $\mathbf{X}$. However, the density of X_0 is intractable, which precludes the use of a likelihood inference method. Consequently, although shot-noise models are widespread in applications (for example, such models were used to model internet traffic Barakat et al. (2003), river streamflows Claps et al. (2005), spikes in neuroscience (Holden (1976),Hohn and Burkitt (2001)) and in signal processing (Sequeira and Gubner (1995) , Sequeira and Gubner (1997))), theoretical results on the statistical inference of shot-noise appear to be limited. Recently, Xiao and al. (Xiao and Lund (2006)) provide consistent and asymptotically normal estimators for parametric shot-noise processes with specific impulse responses.
In this contribution, we consider the particular case given by the following assumption.

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

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

Under (**SN**-2), the process $(X_t)_{t\in\mathbb{R}}$ is usually called an *exponential shot-noise*. In this case, Condition (4.1.2) becomes

$$\mathbb{E}\left[\log_+(|Y_1|)\right] < \infty\ . \tag{4.1.5}$$

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

$$\mathrm{d}X_t = -\alpha X_t \mathrm{d}t + \mathrm{d}L_t\ , \quad X_0 = x \in \mathbb{R} \tag{4.1.6}$$

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

$$L_t := \sum_{k=1}^{N_\lambda(t)} Y_k \quad \text{with} \quad N_\lambda(t) := \sum_k \mathbb{1}_{T_k \leq t}\ , \tag{4.1.7}$$

where $(T_k, Y_k)_{k\geq 0}$ satisfies (**SN**-1). The solution to the equation (4.1.6) is called a Ornstein-Uhlenbeck(O-U) process ((Sato, 1999, Chapter 17)) driven by $\mathbf{L}$ with initial condition $X_0 = x$ and rate α. Note that $\mathbf{L}$ defined by (4.1.7) has Lévy measure λF. Thus, by (Sato, 1999, Theorem 17.5), this Markov process admits a unique stationary version if (4.1.5) holds, and this stationary solution corresponds to the shot-noise process (4.1.1).

In recent works, (Brockwell and Schlemm, 2013b, Brockwell, Schlemm) exploit the integrated version of (4.1.6) to recover the Lévy process $\mathbf{L}$ and show that the increments of $\mathbf{L}$ can be represented as:

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

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

To the best of our knowledge, the paper that best fits our setting is Jongbloed et al. (2005). The authors propose a nonparametric estimation procedure from a low frequency sample of a stationary O-U process which exploits the self decomposability property of the marginal distribution. The authors construct an estimator of the so called canonical function k defined by:

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

The two main additional assumptions are that k is decreasing on $(0, \infty)$ and ν satisfies the integrability condition $\int_0^\infty (1 \wedge x)\nu(\mathrm{d}x) < \infty$. In our setting (i.e. when specifying the Lévy process to be the compound Poisson process defined in (4.1.7)), it is easily shown that these conditions hold and the canonical function and the cumulative distribution of the marks are related by the equation:

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

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

The paper is organized as follows. In Section 4.2, we introduce some preliminaries on the characteristic function of an exponential shot-noise process and provide both the inversion formula and the estimator of the density f. In particular, we derive an upper bound of convergence for our estimator over a broad class of densities under the assumption that λ/α is known. In Section 4.3, we present in details the algorithm used to perform the density estimation and illustrate our findings with a limited Monte-Carlo experiment. Section 4.4 provides error bounds for the empirical characteristic function based on discrete-time observations and exploit the β-mixing structure of the process. Finally, Section 4.5 is devoted to the proofs of the various theorems.

4.2 Main result

4.2.1 Inversion formula

As mentioned in the introduction, it is difficult to derive the probability density function of the stationary shot-noise unless the marks are distributed according to an exponential random variable and the impulse response is an exponential function. In this case, it turns out that the marginal distribution of the shot-noise is Gamma-distributed (the reader can refer to Bondesson (1992) for details). In all other cases, we can only compute the characteristic function of the marginal distribution of the stationary version of the shot-noise when treating it as a filtered point process (see for example Resnick (1992) for details). We have for every real u:

$$\varphi(u) := \mathbb{E}[\mathrm{e}^{\mathrm{i}uX_0}] = \exp\left(\lambda \int_{\mathbb{R}} \int_0^\infty (\mathrm{e}^{\mathrm{i}uyh(v)} - 1)\mathrm{d}v\, F(\mathrm{d}y)\right). \tag{4.2.1}$$

From (4.2.1), the characteristic function of X_0 can be expressed as follows:

$$\varphi_{X_0}(u) = \exp\left(\int_{\mathbb{R}} \lambda K_h(uy) F(\mathrm{d}y)\right). \tag{4.2.2}$$

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

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

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

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

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

$$\varphi_{X_0}'(u) = \lambda\, \varphi_{X_0}(u) \int_{\mathbb{R}} yK_h'(uy)F(\mathrm{d}y). \tag{4.2.3}$$

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

$$K_h(u) = \int_0^\infty \left(\mathrm{e}^{\mathrm{i}u\mathrm{e}^{-\alpha v}} - 1\right)\mathrm{d}v = \int_0^u \frac{\mathrm{e}^{\mathrm{i}s} - 1}{\alpha s}\mathrm{d}s. \tag{4.2.4}$$

With (4.2.3), we obtain that

$$\varphi_{X_0}'(u) = \varphi_{X_0}(u)\frac{\lambda}{\alpha u}\left(\varphi_{Y_0}(u) - 1\right). \tag{4.2.5}$$

Since the marginal distribution of X is infinitely divisible, we have by (Sato, 1999, Lemma 7.5.) that $\varphi_X(u)$ does not vanish. If in addition φ_Y is integrable, (4.2.5) provides a way to recover f knowing α/λ, namely, for all $x \in \mathbb{R}$,

$$f(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}xu}\varphi_{Y_0}(u)du = \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}xu}\left(1 + \frac{\alpha u}{\lambda}\frac{\varphi_{X_0}'(u)}{\varphi_{X_0}(u)}\right)\mathrm{d}u\,. \tag{4.2.6}$$

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

Remark 3. We assume in the following that the ratio α/λ appearing in the inversion formula (4.2.6) is a known constant, as it typically depends on the measurement device. Interestingly, however, an estimator of this constant can be derived from (Iksanov and Jurek, 2003, Theorem 1), where it is shown that the marginal distribution G of the stationary shot-noise is regularly varying at 0 with index λ/α, i.e. :

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

with L being slowly varying at 0. Hence it is possible to estimate α/λ by applying Hill's estimator Hill (1975) to the sample $X_1^{-1}, \cdots, X_n^{-1}$.

4.2.2 Nonparametric estimation

Let $\hat{\varphi}_n(u) := n^{-1} \sum_{j=1}^{n} \mathrm{e}^{\mathrm{i}uX_j}$ denotes the empirical characteristic function (e.c.f.) obtained from the observations and $\hat{\varphi}_n'$ its derivative. From (4.2.6), we are tempted to plug the e.c.f. of the observations to estimate the p.d.f. f. Let $(h_n)_{n\geq 0}$ and$(\kappa_n)_{n\geq 0}$ be two sequences of positive numbers such that

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

and consider the following sequence of estimators:

$$\hat{f}_n(x) := \max\left(\frac{1}{2\pi} \int_{-\frac{1}{h_n}}^{\frac{1}{h_n}} \mathrm{e}^{-\mathrm{i}xu} \left(1 + \frac{\alpha u}{\lambda} \frac{\hat{\varphi}_n'(u)}{\hat{\varphi}_n(u)} \mathbb{1}_{|\hat{\varphi}_n(u)|>\kappa_n}\right) \mathrm{d}u, 0 \right) . \tag{4.2.7}$$

Remark 4. We estimate $1/\varphi(u)$ by $\mathbb{1}_{\{|\hat{\varphi}_n(u)|\geq\kappa_n\}}/\hat{\varphi}_n(u)$ with a suitable choice of a sequence $(\kappa_n)_{n\geq 1}$ which converges to zero. The constant κ_n is chosen such that $|\hat{\varphi}_n(u) - \varphi(u)|$ remains smaller than $|\hat{\varphi}_n(u)|$ and $|\varphi(u)|$ with high probability in order to avoid large errors when inverting $\hat{\varphi}_n(u)$. In Neumann and Reiß (2009), the authors deal with the empirical characteristic function of i.i.d. random variables. In this case, the deviations of $\sqrt{n}(\hat{\varphi}_n(u) - \varphi(u))$ are bounded in probability, hence, they use $1_{\{|\hat{\varphi}_n(u)|\geq\kappa n^{-1/2}\}}/\hat{\varphi}_n(u)$ as an estimator of $1/\varphi(u)$. Here we truncate the interval of integration $\mathbb{R}$ by $\left[-h_n^{-1}, h_n^{-1}\right]$, where h_n is a bandwidth parameter. This allows us to bound the estimation error $\hat{f}_n - f$ in sup norm. The deviation of $\sqrt{n}(\hat{\varphi}_n(u)-\varphi(u))$ on $\left[-h_n^{-1}, h_n^{-1}\right]$ depends on h_n, see Theorem 4.4.1. The resulting κ_n is then taken slightly larger than $n^{-1/2}$.

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

$$\Theta(K, L, s, m) = \Big\{ f \text{ is a density s.t.} \int |y|^{4+m} f(y)\,\mathrm{d}y \leq K \ , \quad \int_{\mathbb{R}} (1 + |u|^2)^s \, |\mathcal{F} f(u)|^2 \,\mathrm{d}u \leq L^2 \Big\} , \tag{4.2.8}$$

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

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

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

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

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

Theorem 4.2.1. *Assume that the process* $\mathbf{X} = (X_t)_{t\geq 0}$ *given by* (4.1.1) *satisfies the assumptions* (**SN**-1)-(**SN**-2) *for some positive constants* λ *and* α*. Let* K, L, m *be positive constants and* $s > 1/2$*. Let* $C_{K,L,m,\lambda/\alpha}$ *be the constant defined by*

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

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

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

and define $\hat{f}_n$ *by* (4.2.7)*. Then, for* $n \geq 3$*, the density estimator* $\hat{f}_n$ *satisfies*

$$\sup_{f\in\Theta(K,L,s,m)} \mathbb{E}_f\left[\left\|f - \hat{f}_n\right\|_\infty\right] \leq Mn^{-(2s-1)/(4s+2+4\lambda/\alpha)} \log(n)^{1/2} , \tag{4.2.9}$$

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

Proof. See Section 4.5.4. □

Remark 5. The constant C in Theorem 4.2.1 might be adaptively chosen. Indeed, the well known relationship (see Daley and Vere-Jones (1988)[Chapter 6] for example) between the cumulant function of a filtered Poisson process and its intensity measure implies that the mean μ_f of X_0 is given by

$$\mu_f = \lambda\mathbb{E}_f[Y_0]\int_0^\infty e^{-\alpha s}ds = \frac{\lambda}{\alpha}\mathbb{E}_f[Y_0] \leq \frac{\lambda}{\alpha}K^{1/(4+m)} . \tag{4.2.10}$$

Since $\mathbf{X}$ is ergodic (see Section 4.4), the empirical mean $\hat{\mu}_n$ of the sample $X_1, \cdots, X_n$ converges to μ_f almost surely and thus, replacing C in Theorem 4.2.1 by $\hat{C}_n := \exp(-\hat{\mu}_n)/2$, leads to the same rate of convergence since

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

Remark 6. This theorem provides that the error in uniform norm converges at least at a polynomial rate $n^{-(2s-1)/(4s+2+8\lambda/\alpha)}\log(n)^{1/2}$ that depends both on the quantity λ/α and the smoothness coefficient s. For a given ratio λ/α, the convergence rate becomes faster as s increases and tend to behave as $n^{-1/2}$ when $s \to \infty$. On the other side, for a given smoothness parameter s, the rate of convergence decreases when the ratio λ/α tends to infinity. This can be interpreted as the consequence of the *pileup* effect that occurs whenever the intensity λ is large or the impulse response coefficient α is close to zero.

Remark 7. Based on the previous theorem, one might wonder whether the rates of convergence are optimal. According to similar but not identical problems (Neumann and Reiß (2009),Gugushvili (2009)) in which authors estimate in a nonparametric fashion a Lévy triplet (with finite activity) based on a low frequency sample of the associated process, the optimal rates of convergence are identical to ours. Our estimation procedure lies on stationary but dependent infinitely divisible random variables associated to an infinite Lévy measure so that these results do not apply here. However we believe that the rates obtained in Theorem 4.2.1 are also optimal in this dependent context. The proof of this conjecture is left for future work.

4

4.3 Experimental results

The estimation procedure based on the estimator $\hat{f}_n$ given by (4.2.7) can be made time-efficient and thus well suited to a very large dataset. In nuclear applications, it is usual to deal with several million of observations while the intensity of the time-arrival point process can reach several thousand of occurrences per second. Typically, the shot-noise process in nuclear applications corresponding to the electric current is discretely observed for three minutes at a sampling rate of 10Mhz and the mean number of arrivals between two observations lies between 10 and 100. Such large values for the intensity and the number of sampled points motivate us to present a practical way to compute the estimator (4.2.7).

Practical computation of the estimator

4

In Section 4.2, we have defined the estimator of mark's density by (4.2.7). Although it theoretically converges to the true density of shot-noise marks, the evaluation of the empirical characteristic function and its derivative based on observations $X_1, \cdots, X_n$ might be time-consuming when the sample size n is large. To circumvent this issue, we propose to compute the empirical characteristic function using the fast fourier transform of an appropriate histogram of the vector $X_1, \cdots, X_n$. More precisely, for a strictly positive fixed h, we consider the grid $G = \{hl : \lfloor \min_{k\leq n}(X_k)/h \rfloor \leq l \leq \lceil \max_{k\leq n}(X_k)/h \rceil\}$ and compute the normalized histogram H of the sample sequence $(X_l)_{1\leq l\leq n}$ with respect to the grid G defined by

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

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

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

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

$$\hat{\varphi}_{h,n}(u) := \sum_{l=m_n}^{M_n} H(l)\mathrm{e}^{\mathrm{i}uh(l+1/2)} \,. \tag{4.3.1}$$

For any real u, we have the following upper bounds

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

and

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

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

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

The advantage of using $\hat{\varphi}_{h,n}$ is that $\hat{\varphi}_{h,n}(u)$ and $\hat{\varphi}'_{h,n}(u)$ can be evaluated on a regular grid using the fast Fourier Transform algorithm.
The last step in the numerical computation of the estimator (4.2.7) consists in evaluating the quantity

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

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

Numerical results

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

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

Furthermore, in order to fit with nuclear science applications, where detectors have a time resolution of 10 Mhz, corresponding to a sampling time $\Delta = 10^{-7}$ seconds. The parameters of the experiment are set to $\alpha = 8.10^8$, $\lambda = 10^9$. Moreover, in nuclear spectrometry, the bandwidth h_n is directly related to the known precision of the measuring instrument. For the following numerical experiment, we set it to 2.5 which is in range with the detector resolution as described in Knoll (1989), Chapter 4. Figure 4.1 below shows a simulated sample path of such a shot-noise with its associated marked point process.

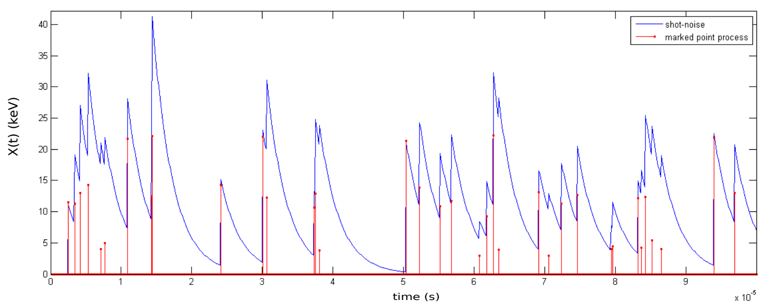


Figure 4.1 : Sample path of a simulated exponential shot-noise process

As shown in Figure 4.2 below, our estimator $\hat{f}_n$ defined by (4.2.7) well retrieves the three modes of the Gaussian mixture as well as the corresponding variance from a sample of size 10^5, which corresponds to a signal observed for one hundredth second. Current estimators used in nuclear spectrometry for similar data requires much longer measurements (up to 10 seconds). The reason is that these estimators do not consider observations where pile-up is suspected to occur, thus throwing away a large part of the available information.

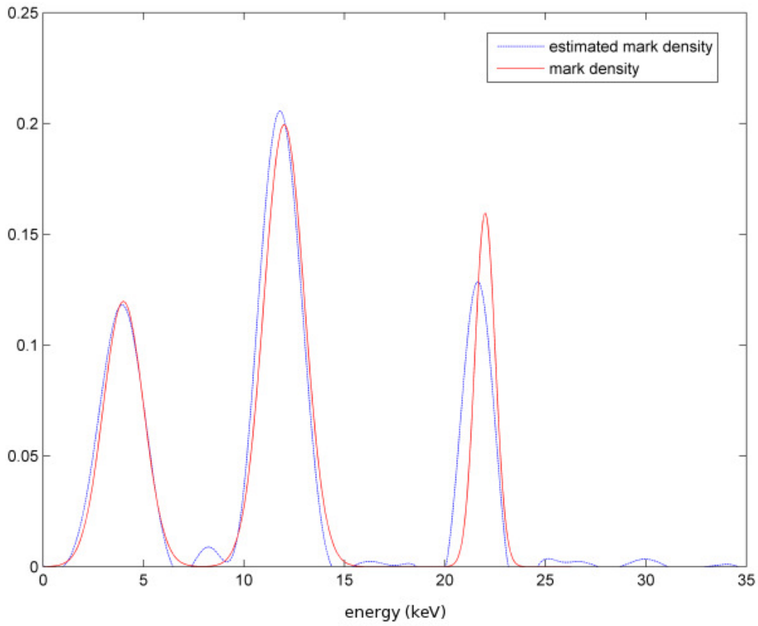


Figure 4.2 : Illustration of the nonparametric estimator of a Gaussian mixture model.

Moreover, an estimation of the risk $\mathbb{E}_f\left[\|f - \hat{f}_n\|_\infty\right]$ is provided in Table 4.1 for the shot-noise configuration described at the beginning of this section and for three different sample sizes.

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

Table 4.1 : Mean ℓ^∞ error $\mathbb{E}_f\left[\|f - \hat{f}_n\|_\infty\right]$ for different sample sizes n with f given by the Gaussian mixture defined above. The displayed values are obtained by averaging $\|f - \hat{f}_n\|_\infty$ over 100 Monte Carlo simulations. The variance of the error over the Monte Carlo runs is indicated between parentheses to assess the precision of the given estimates.

4.4 Error bounds for the empirical characteristic function and its derivatives

To derive Theorem 4.2.1, since our constructed estimator involves the empirical characteristic function and its derivative, we rely on deviation bounds for

$$\mathbb{E}_f\left[\sup_{u\in[-h^{-1},h^{-1}]}\left|\hat{\varphi}_n^{(k)}(u)-\varphi^{(k)}(u)\right|\right] \quad , \quad k=0,1 \quad , \quad h>0\,, \tag{4.4.1}$$

which are uniform over $f \in \Theta(K, L, s, m)$, where the smoothness class $\Theta(K, L, s, m)$ is defined by (4.2.8). Here, $\varphi^{(k)}$ and $\hat{\varphi}_n^{(k)}$ respectively denote the k-th derivative of the characteristic function and its empirical counterpart associated to the sample $X_1, \dots, X_n$. These bounds are of independent interest and therefore are stated in this separate section. Upper bounds of the empirical characteristic function deviations have been derived in the case of i.i.d. samples: Gugushvili (2009)[Theorem 2.2.] provides upper bounds of (4.4.1) for i.i.d. infinitely divisible random variables, based on general deviation bounds for the empirical process of i.i.d. samples found in van der Vaart and Wellner (1996). Here we are concerned with a dependent sample $X_1, \cdots, X_n$ and we rely instead on Doukhan et al. (1995). We obtain upper bounds with the same rate of convergence as in the i.i.d. case but depending on the β-mixing coefficients, see Theorem 4.4.1. An additional difficulty in the non-parametric setting that we consider is to derive upper bounds that are uniform over smoothness classes for the density f, and thus to carefully examine how the β coefficients depend on f, see Theorem 4.4.2.

Let us first recall the definition of β-mixing coefficient (also called absolutely regular or completely regular coefficient) as introduced by Volkonskii and Rozanov Volkonskii and Rozanov (1959). For $\mathcal{A}, \mathcal{B}$ two σ-algebras of Ω, the coefficient $\beta(\mathcal{A}, \mathcal{B})$ is defined by

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

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

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

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

We first state a result essentially following from Doukhan et al. (1995) which specify how the β coefficients allows us to derive bounds on the estimation of the characteristic function and its derivatives.

Theorem 4.4.1. *Let k be a non-negative integer and $X_1, \cdots, X_n$ a sample of a stationary β-mixing process. Suppose that there exists $C \geq 1$ and $\rho \in (0, 1)$ such that $\beta_n \leq C\rho^n$ for all $n \geq 1$. Let $r > 1$ and suppose that $\mathbb{E}\left[|X_1|^{2(k+1)r}\right] < \infty$.*

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

$$\mathbb{E}\left[\sup_{u\in\left[-h^{-1},h^{-1}\right]}\left|\hat{\varphi}_n^{(k)}(u)-\varphi^{(k)}(u)\right|\right] \\ \leq A\frac{\max\left(\mathbb{E}\left[|X_1|^{2kr}\right]^{1/2r},\mathbb{E}\left[|X_1|^{2(k+1)r}\right]^{1/2r}\right)\left(1+\sqrt{\log(1+h^{-1})}\right)}{n^{1/2}}. \quad (4.4.2)$$

Proof. The proof is deferred to Section 4.5.3 □

It turns out that the stationary exponential shot-noise process $\mathbf{X}$ defined by (4.1.1) is exponentially β-mixing if the first absolute moment of the marks is finite, see Masuda (2004)[Theorem 4.3] for a slightly more general condition. However, in order to obtain a uniform bound of the risk of our estimator $\hat{f}_n$ over a smoothness class, a more precise result is needed. In the sequel, we add a superscript f to the β-mixing sequence to make explicit the dependence with respect to the mark's density f. The following theorem provides a geometric bound for the β-mixing coefficients of the shot noise which is uniform over the class $\Theta(K, L, s, m)$.

Theorem 4.4.2. *Let $X_1, \cdots, X_n$ be a sample of the stationary shot-noise process given by* (4.1.1) *satisfying* (**SN**-1)-(**SN**-2)*. Let $K, L, m > 0$ and $s > 1/2$. Then there exist two constants $C > 0$ and $\rho \in (0, 1)$ only depending on λ, α, K, L, s and m such that, for all $n \geq 1$,*

$$\sup_{f\in\Theta(K,L,s,m)} \beta_n^f \leq C\rho^n < \infty\,. \quad (4.4.3)$$

Proof. See Section 4.5.2. □

As a corollary of Theorems 4.4.1 and 4.4.2, we obtain error bounds for the empirical characteristic function when dealing with observations $X_1, \cdots, X_n$ of the stationary shot-noise process given by (4.1.1).

Corollary 1. *Let $X_1, \cdots, X_n$ be a sample of the stationary shot-noise process given by* (4.1.1) *satisfying* (**SN**-1)-(**SN**-2)*. Let $K, L, m > 0$ and $s > 1/2$ and let k be an integer such that $0 \leq k < 1 + n/2$. Then there exists a constant B only depending on k, λ, α, K, L, s and m such that for all and $n \geq 1$,*

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

This result can be compared to Gugushvili (2009)[Theorem 2.2]. Note however that although our sample has infinitely divisible marginal distributions, it is not independent and the Lévy measure is not integrable.

4.5 Proofs

4.5.1 Preliminary results on the exponential shot noise

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

Definition 13 (Geometric drift condition)**.** *A Markov Kernel P satisfies a geometric drift condition (called $D(V, \mu, b)$) if there exists a measurable function $V : \mathbb{R} \to [1, \infty[$ and constants $(\mu, b) \in (0, 1) \times \mathbb{R}_+$ such that*

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

Definition 14 (Doeblin set)**.** *A set C is called a (m, ϵ)-Doeblin set if there exists a positive integer m , a positive number ϵ and a probability measure ν on $\mathbb{R}$ such that, for any x in C and A in $\mathcal{B}(\mathbb{R})$*

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

The following proposition is borrowed from Tuominen and Tweedie (1994) and relates explicitly the geometrical drift condition to the convergence in V-norm (denoted by $|\cdot|_V$) to the stationary distribution.

Proposition 7. *Let P be a Markov kernel satisfying the drift condition $D(V, \mu, b)$. Assume moreover that for some $d > 2b(1-\mu) - 1$, $m \in \mathbb{N} - \{0\}$ and $\epsilon \in (0, 1)$, the level set $\{V \leq d\}$ is an (m, ϵ)-Doeblin set. Then P admits a unique invariant measure π and P is V-geometrically ergodic, that is, for any $0 < u < \epsilon/\left(b_m + \mu^m d - 1 + \epsilon\right)_+ \vee 1$, $n \in \mathbb{N}$ and $x \in \mathbb{R}$,*

$$||P^n(x, \cdot) - \pi||_V \leq c(u)[\pi(V) + V(x)]\rho^{\lfloor n/m \rfloor}(u) \tag{4.5.1}$$

where

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

In order to apply such a result to the sample $(X_1, \cdots, X_n)$ of the exponential shot-noise defined by (4.1.1), observe that it is the sample of a Markov chain which satisfies the autoregression equation $X_{i+1} = \mathrm{e}^{-\alpha}X_i + W_{i+1}$, where the sequence of innovations $(W_i)_{i\in\mathbb{Z}}$ is made up of i.i.d. random variables distributed as

$$W_0 := \sum_{k=1}^{N_\lambda([0,1])} Y_k \mathrm{e}^{-\alpha U_k}\ . \tag{4.5.2}$$

where

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

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

Proposition 8 (Uniform Geometric drift condition)**.** *Let $K, L, m > 0$ and $s > 1/2$ and let $f \in \Theta(K, L, s, m)$. Then the Markov kernel Q_f satisfies the drift condition $D(V, \mu, b)$, where*

$$V : x \to 1 + |x| \quad , \quad \mu = \mathrm{e}^{-\alpha} \quad , \quad b = 1 + \lambda K^{1/(4+m)} - \mathrm{e}^{-\alpha} \ . \tag{4.5.3}$$

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

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

□

Remark 8. A similar result holds for the functions $V_i : x \to 1 + |x|^i$ where $i \in \{1, \cdots, \lfloor 4 + m \rfloor\}$.

Proposition 9 (Doeblin set)**.** *Let $l > 1$, $K, L, m > 0$ and $s > 1/2$ and define V as in (4.5.3). There exists $\epsilon > 0$ only depending on l, α, λ, $K, L, m > 0$ and s such that, for all $f \in \Theta(K, L, s, m)$, the Markov kernel Q_f admits $\{V \leq l\}$ as an $(1, \epsilon)$-Doeblin set.*

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

$$\check{f}(v) = \frac{1}{\alpha v} \int_v^{v\mathrm{e}^\alpha} f(y) \mathrm{d}y \ . \tag{4.5.4}$$

The distribution of W_0 is thus given by the infinite mixture

$$\mathrm{e}^{-\lambda}\delta_0(\mathrm{d}\xi) + \sum_{k=1}^{\infty} \frac{\lambda^k}{k!} \mathrm{e}^{-\lambda} \check{f}^{*k}(\xi)\mathrm{d}\xi := \mathrm{e}^{-\lambda}\delta_0(\mathrm{d}\xi) + (1 - \mathrm{e}^{-\lambda})\tilde{g}_f(\xi)\mathrm{d}\xi \ , \tag{4.5.5}$$

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

$$Q_f(x, A) = \mathrm{e}^{-\lambda} \mathbb{1}_A(\mathrm{e}^{-\alpha} x) + \int_A (1 - \mathrm{e}^{-\lambda}) \tilde{g}_f(\xi + \mathrm{e}^{-\alpha} x) \, \mathrm{d}\xi \ .$$

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

$$\int_A \tilde{g}_f(\xi + \mathrm{e}^{-\alpha}x)\mathrm{d}\xi \geq (1-\mathrm{e}^{-\lambda})^{-1}\epsilon\, \nu(A)\ .$$

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

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

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

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

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

$$\inf_{f\in\Theta(K,L,s,m)} \inf_{c'(f)\leq\xi\leq d'(f)} [d'(f)-c'(f)-2\mathrm{e}^{-\alpha}(l-1)]\,\check{f}^{*k(f)}(\xi) > 0\ . \tag{4.5.6}$$

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

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

Finally, Lemma 4 and the previous bound yield (4.5.6), which concludes the proof. □

4.5.2 Proof of Theorem 4.4.2

As explained in Section 4.5.1, $(X_i)_{i\geq0}$ is a stationary V-geometrically ergodic Markov chain with Markov kernel denoted by Q_f. By Davydov (1973), the β-coefficient of the stationary Markov chain $(X_i)_{i\geq0}$ can be expressed for all $n \geq 1$ and $f \in \Theta(K, L, s, m)$ as

$$\beta_n^f = \int_{\mathbb{R}} \|Q_f^n(x,\cdot) - \pi_f\|_{TV}\,\pi_f(dx)\ ,$$

where π_f is the invariant marginal distribution and $|\cdot|_{TV}$ denotes the total variation norm, i.e. the V-norm with $V = 1$. Combining Propositions 8, 9 and 7, we can find constants $C > 0$ and $\rho \in (0, 1)$ only depending on λ, α, K, L, s and m such that

$$\|Q_f^n(x,\cdot) - \pi_f\|_{TV} \leq \|Q_f^n(x,\cdot) - \pi_f\|_V \leq C\left(2 + \mathbb{E}_f[|X_1|] + |x|\right)\rho^n\ ,$$

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

$$\begin{aligned}\beta_n^f &\leq 2C\,(1+\mathbb{E}_f[|X_1|])\rho^n \\ &\leq 2C\,(1+\lambda K^{1/(4+m)})\rho^n\ ,\end{aligned} \tag{4.5.7}$$

which concludes the proof.

4.5.3 Proof of Theorem 1

In Doukhan et al. (1995), the authors establish a Donsker invariance principle for the process $\{Z_n(f), f \in \mathcal{F}\}$ where $Z_n := n^{-1/2}\sum_{i=1}^n(\delta_{X_i} - P)$ is the normalized centered empirical process associated to a stationary sequence of β-mixing random variables $(X_1, \cdots, X_n)$ with marginal distribution P and $\mathcal{F}$ is a class of functions satisfying an entropy condition. To be more precise, suppose that the sequence $(X_i)_{i\geq 1}$ is β-mixing with $\sum_{n\in\mathbb{N}}\beta_n < \infty$. The mixing rate function β is defined by $\beta(t) = \beta_{\lfloor t\rfloor}$ if $t \geq 1$ and $\beta(t) = 0$ otherwise while its càdlàg inverse β^{-1} is defined by:

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

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

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

The space $\mathcal{L}_{2,\beta}$ is defined as the class of functions f such that $\|f\|_{2,\beta} < \infty$. In Doukhan et al. (1995), the authors proved that $\left(\mathcal{L}_{2,\beta}, \|\cdot\|_{2,\beta}\right)$ is a normed subspace of $\mathcal{L}_2$. A useful and trivial result from the definition of the norm $\mathcal{L}_{2,\beta}(P)$ provides the following relation:

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

For any real $r > 1$, another useful (less trivial) result in Doukhan et al. (1995) states that under the condition

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

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

$$\|f\|_{2,\beta} \leq \|f\|_{2r}\sqrt{1 + r\sum_{n\geq 0}\beta_n\, n^{r/(r-1)}} , \tag{4.5.8}$$

where $\|f\|_{2r} = \mathbb{E}\left[|f(X_0)|^{2r}\right]^{1/2r}$ denote the usual L^{2r}-norm.
Now, we can state a result directly adapted from Doukhan et al. (1995)[Theorem 3] that will serve our goal to prove Theorem 4.2.1. For the sake of self-consistency, we recall that, given a metric space of real-valued funtions $(E, \|\cdot\|)$ and two functions l, u , the bracket $[l, u]$ represents the set of all functions f such that $l \leq f \leq u$. For any positive ϵ, $[l, u]$ is called an ϵ-bracket if $\|l - u\| < \epsilon$.

Theorem 4.5.1. *Suppose that the sequence $(X_i)_{i\geq 1}$ is exponentially β-mixing and that there exists $C \geq 1$ and $\rho \in (0, 1)$ such that $\beta_n \leq C\rho^n$ for all $n \geq 1$. Let $\sigma > 0$ and let $\mathcal{F} \subset \mathcal{L}_{2,\beta}$ be a class of functions such that for every f in $\mathcal{F}$, $\|f\|_{2,\beta} \leq \sigma$. Define*

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

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

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

(DMR2) $\phi(1) < \infty$.

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

$$\mathbb{E}\left[\sup_{f\in\mathcal{F}} |Z_n(f)|\right] \leq A\phi(\sigma)\left(1 + \frac{\|F\|_{2r}}{\sigma\sqrt{1-r^{-1}}}\right) . \tag{4.5.9}$$

Having this result at hand, we now remark that (4.4.1), for a fixed integer k, can be rewritten as

$$\mathbb{E}\left[\sup_{u\in[-h^{-1},h^{-1}]} \left|\hat{\varphi}_n^{(k)}(u) - \varphi^{(k)}(u)\right|\right] = n^{-1/2}\mathbb{E}\left[\sup_{f\in\mathcal{F}_h^k} |Z_n(f)|\right], \tag{4.5.10}$$

where

$$\mathcal{F}_h^k := \{f_u : x \to (ix)^k e^{iux}, u \in \left[-h^{-1}, h^{-1}\right]\}. \tag{4.5.11}$$

The proof of Theorem 4.4.1 based on an application of the previous theorem is as follows.

Proof. We apply Theorem 4.5.1 for a fixed integer k, $\mathcal{F} = \mathcal{F}_h^k$, $F = F_k$ and $r = (4+m)/4$ where $F_k : x \to |x|^k$.

Assumption (DMR1): Let k be a fixed integer. On the one hand, the function F_k is an envelope function of the class $\mathcal{F}_h^k$ and on the other hand, for any real $r > 1$, from (4.5.8), we have

$$\|F_k\|_{2,\beta} \leq \mathbb{E}\left[|X_1|^{2kr}\right]^{1/2r} \sqrt{1 + r\sum_{n\geq 0} \beta_n\, n^{r/(r-1)}} := \sigma_{r,k} < \infty . \tag{4.5.12}$$

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

$$\left|(\mathrm{i}x)^k \mathrm{e}^{\mathrm{i}sx} - (\mathrm{i}x)^k \mathrm{e}^{\mathrm{i}tx}\right| \leq |s-t|\,|x|^{k+1} \tag{4.5.13}$$

A direct application of van der Vaart and Wellner (1996)[Theorem 2.7.11] for the classes $\mathcal{F}_h^k$ gives for any $\epsilon > 0$:

$$\begin{aligned} N_{[\,]}\left(2\epsilon\|F_{k+1}\|_{2,\beta}, \mathcal{F}_h^k, \|\cdot\|_{2,\beta}\right) &\leq N\left(\epsilon, \left[-h^{-1}, h^{-1}\right], |\cdot|\right) \\ &\leq 1 + \frac{2h^{-1}}{\epsilon} \end{aligned} \tag{4.5.14}$$

where N and $N_{[\,]}$ are respectively called the covering numbers and bracketing number (these numbers respectively represent the minimum number of balls and brackets of a given size necessary to cover a space with respect to a given norm). From (4.5.14), it follows that for any $\sigma > 0$, we have

$$\begin{aligned}
\phi(\sigma) &= \int_0^{\sigma} \sqrt{1 + \log\left(N_{[\,]}\left(u, \mathcal{F}_h^k, \|\cdot\|_{2,\beta}\right)\right)} \mathrm{d}u \\
&\leq \int_0^{\sigma} \sqrt{1 + \log\left(1 + \frac{4\|F_{k+1}\|_{2,\beta} h^{-1}}{u}\right)} \mathrm{d}u \qquad (4.5.15)\\
&\leq \int_0^{\sigma} \left(1 + \frac{2\|F_{k+1}\|_{2,\beta}^{1/2} h^{-1/2}}{u^{1/2}}\right) \mathrm{d}u \\
&= \sigma + 4\sqrt{\sigma}\,\|F_{k+1}\|_{2,\beta}^{1/2} h^{-1/2} < \infty \qquad (4.5.16)
\end{aligned}$$

because we supposed $F_{k+1} \in \mathcal{L}_{2r}$ and $\beta_n \leq C\rho^n$ which, from (4.5.8), implies that $\|F_{k+1}\|_{2,\beta} < \infty$.

Conclusion of the proof The application of Theorem 4.5.1 gives

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

where $\tilde{A} = A(1 + 1/\sqrt{1 - r^{-1}})$ since, from (4.5.12), we have $\|F_k\|_{2r}/\sigma_{r,k} \leq 1$.
Set $c_{r,\bar{\beta}} := \sqrt{1 + r\sum_{n\geq 0}\beta_n\, n^{r/(r-1)}}$. From (4.5.15) and (4.5.12), we can write

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

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

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

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

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

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

4.5.4 Proof of Theorem 4.2.1

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

$$f_n^0(x) := \max\left(0, \frac{1}{2\pi}\int_{-h_n^{-1}}^{h_n^{-1}} \mathrm{e}^{-\mathrm{i}xu}\varphi_{Y_0}(u)\mathrm{d}u\right). \qquad (4.5.17)$$

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

$$f(x) = \max\left(0, \frac{1}{2\pi}\int_{\mathbb{R}} \mathrm{e}^{-\mathrm{i}xu}\varphi_{Y_0}(u)\mathrm{d}u\right) . \tag{4.5.18}$$

We decompose the error in infinite norm as

$$\|f - \hat{f}_n\|_\infty \leq \|f - f_n^0\|_\infty + \|f_n^0 - \hat{f}_n\|_\infty . \tag{4.5.19}$$

From (4.5.17) and (4.5.18), we get

$$\begin{aligned}
\|f - f_n^0\|_\infty &\leq \frac{1}{\pi}\int_{h_n^{-1}}^{\infty} \left|\varphi_{Y_0}(u)\right| \mathrm{d}u \\
&= \frac{1}{\pi}\int_{h_n^{-1}}^{\infty} \left|u^{-s}u^{s}\varphi_{Y_0}(u)\right| \mathrm{d}u \\
&\leq \frac{1}{\pi}\left(\int_{h_n^{-1}}^{\infty} |u|^{-2s}\,\mathrm{d}u\right)^{1/2}\left(\int_{h_n^{-1}}^{\infty} \left|u^{s}\varphi_{Y_0}(u)\right|^2 \mathrm{d}u\right)^{1/2} \\
&\leq \frac{L}{\pi}\frac{h_n^{s-1/2}}{2s-1} = \frac{L}{(2s-1)\pi} n^{-(2s-1)/(4s+2+4\lambda/\alpha)} .
\end{aligned} \tag{4.5.20}$$

where we used the Cauchy-Schwartz inequality and the assumption that $f \in \Theta(K, L, s, m)$. We conclude with a bound of the term involving $\|f_n^0 - \hat{f}_n\|_\infty$ in (4.5.19). To this end, the following inequality will be useful. Using (4.2.5) and the mean-value theorem, we have

$$\sup_{u\in\mathbb{R}} \left|\frac{\varphi'_{X_0}(u)}{\varphi_{X_0}(u)}\right| \leq \frac{\lambda}{\alpha}\sup_{u\in\mathbb{R}} \left|\varphi'_{Y_0}(u)\right| \leq \frac{\lambda}{\alpha}\mathbb{E}_f\left[|Y_0|\right] \leq \frac{\lambda}{\alpha}K^{1/(4+m)} . \tag{4.5.21}$$

By (4.2.5), we can bound the term $\|f_n^0 - \hat{f}_n\|_\infty$ by

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

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

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

Thus, using (4.5.21), we get

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

The two terms on the right hand side can be bounded using Corollary 1 with $r=(4+m)/4$. It gives

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

and

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

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

$$P \lesssim Q \Longleftrightarrow \text{for all } n\geq 3,\quad \sup_{f\in\Theta(K,L,s,m)}\frac{P}{Q}<\infty\,. \tag{4.5.22}$$

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

$$\begin{aligned}\mathbb{E}_f\left[A_{n,1}\right] &\lesssim \frac{1+\sqrt{\log(1+h_n^{-1})}}{\kappa_n\, n^{1/2}h_n}\\ &\lesssim \frac{1+\sqrt{\log(1+h_n^{-1})}}{n^{1/2}h_n^{\lambda/\alpha+1}}\\ &\lesssim n^{-(2s-1)/(4s+2+4\lambda/\alpha)}\log(n)^{1/2}\,,\end{aligned} \tag{4.5.23}$$

where we used the fact that $\kappa_n^{-1}\leq 2C^{-1}h_n^{\lambda/\alpha}$ for any integer n.
We now bound $A_{n,2}$. From (4.5.21), remark that

$$\begin{aligned}\mathbb{E}_f\left[A_{n,2}\right] &\leq \frac{2}{\pi h_n}K^{1/(4+m)}\mathbb{P}_f\left(\exists u\in\left[-h_n^{-1},h_n^{-1}\right],|\hat{\varphi}_n(u)|\leq\kappa_n\right)\\ &\leq \frac{2}{\pi h_n}K^{1/(4+m)}\mathbb{P}_f\left(\inf_{|u|\leq h_n^{-1}}|\hat{\varphi}_n(u)|\leq\kappa_n\right)\,,\end{aligned} \tag{4.5.24}$$

From Lemma (5), we have

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

It follows that

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

Since $0 < C < C_{K,L,m,\lambda/\alpha}$, applying Corollary 1 combined to the Markov's inequality, and using (4.5.24), we get

$$\begin{aligned}
\mathbb{E}_f\left[A_{n,2}\right] &\lesssim \frac{1 + \sqrt{\log\left(1 + h_n^{-1}\right)}}{n^{1/2}\kappa_n h_n} \\
&\lesssim n^{-(2s-1)/(4s+2+4\lambda/\alpha)} \log(n)^{1/2} \ , \qquad (4.5.25)
\end{aligned}$$

Equations (4.5.20), (4.5.23) and (4.5.25) imply (4.2.9) and the proof is concluded. □

4.6 Useful lemmas

4

The following classical embedding will be useful.

Lemma 1 (Sobolev embedding). *Let $K, L, m > 0$ and $s > 1/2$. Let $f \in \Theta(K, L, s, m)$ defined in (4.2.8). Then, for any $\gamma \in (0, (s-1/2) \wedge 1)$, there is a constant $C > 0$ depending on L, s and γ such that, for every real numbers x, y,*

$$|f(x) - f(y)| \le C\,|x - y|^\gamma \ , \qquad (4.6.1)$$

where

$$C = \frac{3}{2\pi} L \left(\int_{\mathbb{R}} \frac{|\xi|^{2\gamma}}{(1 + |\xi|^2)^s} \mathrm{d}\xi\right)^{1/2} . \qquad (4.6.2)$$

The following result is used in the proof of Proposition 9.

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

$$\inf_{a\le x\le a+\Delta} f(x) \ge \frac{1}{16}(2K)^{-1/(4+m)} ,$$

where $T_K = (2K)^{-1/(\gamma(4+m))}$, $\Delta = (2K)^{-1/(\gamma(4+m))}(16C)^{-1/\gamma}$ with C defined by (4.6.2).

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

$$\sup_{|x|\le T} f(x) \ge (2T)^{-1}\left(1 - T^{-(4+m)}K\right) ,$$

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

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

and on the other side

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

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

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

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

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

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

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

□

4

Lemma 3. *Let $K, L, m, \alpha > 0$ and $s > 1/2$. Let $\gamma \in (0, (s-1/2) \wedge 1)$ and $f \in \Theta(K, L, s, m)$. Define $T_K = (2K)^{-1/(4+m)}$, $\Delta = (2K)^{-1/(\gamma(4+m))}(16C)^{-1/\gamma}$ with C defined by (4.6.2) and let δ be a positive number satisfying*

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

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

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

Proof. From Lemma 2, we have

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

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

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

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

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

which concludes the proof. □

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

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

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

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

$$\left(f \star \tilde{f}\right)(x) \geq \min(1,\delta)\, \epsilon\tilde{\epsilon}\; \mathbb{1}_{[a+c+\delta, b+d-\delta]}(x)\ . \tag{4.6.3}$$

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

$$f^{\star n}(x) \geq \left(\min\left(1, \frac{b-a}{2n}\right)\right)^{n-1} \epsilon^n \mathbb{1}_{[na+(b-a)/2, nb-(b-a)/2]}(x)\ . \tag{4.6.4}$$

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

Lemma 5. *Assume that the process X is given by (4.1.1) under (SN-1)-(SN-2) with some positive constant α and λ. Let K, L, m and s be positive constants. Then for all $f \in \Theta(K, L, s, m)$ and $u \in \mathbb{R}$, we have*

$$\left|\varphi_{X_0}(u)\right| \geq C_{K,L,m,\lambda/\alpha}\, (1+|u|)^{-\lambda/\alpha} \tag{4.6.5}$$

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

Proof. From (4.2.2) and (4.2.4), we have for all $u \in \mathbb{R}$,

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

If follows that

$$\begin{aligned}
\left|\varphi_{X_0}(u)\right| &= \exp\left(\frac{\lambda}{\alpha} \int_{\mathbb{R}} \left(\int_0^{ux} \frac{\cos(v)-1}{v} \mathrm{d}v\right) f(x)\mathrm{d}x\right) \\
&= \exp\left(-\frac{\lambda}{\alpha} \int_0^{|u|} \frac{1-\mathrm{Re}\left(\varphi_{Y_0}(z)\right)}{z} \mathrm{d}z\right)\ .
\end{aligned} \tag{4.6.6}$$

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

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

We thus get that

$$\int_0^1 \frac{1-\mathrm{Re}\left(\varphi_{Y_0}(z)\right)}{z}\mathrm{d}z \leq K^{1/(4+m)} . \tag{4.6.7}$$

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

$$\begin{aligned}\int_1^{|u|} \frac{1-\mathrm{Re}\left(\varphi_{Y_0}(z)\right)}{z}\mathrm{d}z &\leq \log|u| + \int_1^{|u|} z^{-1}\left|\mathrm{Re}\left(\varphi_{Y_0}(z)\right)\right| \mathrm{d}z \\ &\leq \log|u| + L , \end{aligned} \tag{4.6.8}$$

where we use the Cauchy-Schwartz inequality and

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

Inserting (4.6.7) and (4.6.8) in (4.6.6), we get the result. □

4

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

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

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

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

As $u \to 0$, $\sqrt{\log(1+u)}$ is equivalent to $\sqrt{u}$.
As $u \to \infty$, the Karamata's Theorem (see Resnick (2007)[Theorem 0.6]) applied to the function $u \to \sqrt{\log(1+u)}u^{-2}$, which is regularly varying with index -2, gives that

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

which concludes the proof. □

5 Nonparametric estimation of a shot-noise process

Abstract

We consider a nonlinear inverse problem that occurs in γ-spectroscopy. A radioactive source randomly emits photons that interact with a detector. An acquisition system then converts each interaction into a pulse of short duration whose shape is proportional to the deposited energy. In this paper, we model this signal using a shot-noise process and assume the shape of the electric current is known. Based on low frequency observations of the shot-noise process, we then introduce a new nonparametric procedure in order to estimate both the intensity and the probability density function of the underlying marked point process. The algorithm developed in this article is time-efficient and suited for large datasets commonly used in nuclear science. We illustrate the performance of our method on simulated and real datasets.

5.1 Introduction

We consider 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 Knoll (1989)). 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. 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 physicians, we assume, in this article, that we exactly know the impulse response. The measured electric current can be modeled by a shot-noise process $\mathbf{X} := (X_t)_{t\geq 0}$ defined by

$$X_t := \sum_{k:T_k\leq t} Y_k\, h(t-T_k)\ , \qquad (5.1.1)$$

where we suppose that

(**SN**-3) $\sum_k \delta_{T_k,Y_k}$ is a Poisson point process with times $T_k \in \mathbb{R}$ arriving homogeneously with a finite intensity $\lambda > 0$ and independent i.i.d. marks $Y_k \in \mathbb{R}_+^*$ with probability density function (p.d.f.) f.

(**SN**-4) the impulse response h of the detector is causal, integrable and satisfies

$$|h(t)| \leq C\mathrm{e}^{-\alpha t}\ ,$$

for some positive constants C and α.

For statistical estimation procedures, we will additionally assume that

(**SN**-5) f is compactly supported in $[0,1]$ and belongs to the Sobolev space $\mathcal{W}^2$ defined by

$$\mathcal{W}^2 := \{f : [0,1] \to \mathbb{R} : \quad f, f' \text{ absolutely continuous}\ ,$$
$$f(1) = f'(1) = 0 \text{ and } \int_0^1 |f''(t)|^2\,\mathrm{d}t < \infty\}. \qquad (5.1.2)$$

In addition, let $\mathcal{W}_+^2$ denote the subset of nonnegative functions that belong to $\mathcal{W}^2$.

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

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

Shot-noise processes constitute natural models to study dynamical systems which can be represented as the convolution between a filter (an *impulse response* in this article) and a marked point process. Those stochastic processes were initially introduced in physics (see Campbell (1909b), Campbell (1909a), Schottky (1918)), electrical engineering (see Rice (1977),Sequeira and Gubner (1997), Sequeira and Gubner (1995)) and then in the last decades to model several phenomena such as teletraffic data by Barakat et al. (2003), finance by Samorodnitsky (1996) , river streamflows Claps et al. (2005), neuroscience (Hohn and Burkitt (2001),Holden (1976)), astrophysics Chaplin et al. (2013) or computed tomography Taguchi et al. (2010). In this article, we lay the emphasis on two nonlinear inverse problems occurring in nuclear science, namely neutron transport and gamma spectroscopy: for the latter, we describe the experimental set-up and model in Section 5.4. Based on a regular sample of observations $X_\delta, \ldots, X_{n\delta}$ where n is an integer that tends to infinity and the knowledge of the impulse response of the system, we aim to estimate the function called *flow* function

$$f_\lambda := \lambda f \ . \tag{5.1.3}$$

As we will explain it in Section 5.2, the problem of recovering f_λ from the observations $X_1, \ldots, X_n$ is *ill-posed*. Recall that a problem is *ill-posed* in the Hadamard sense Hadamard (1902) whenever one of the three conditions fails to be valid:

1. A solution exists.
2. The solution is unique.
3. The operator A^{-1} is continuous.

To our knowledge, this problem has paid little attention (other than through ad-hoc problems) mainly because classical estimation techniques fail to work. For instance, no closed form of likelihood even the marginal one, if it exists, is available.
In Xiao and Lund (2006), assuming the impulse response to be of the form $h = h_{\theta_0}$ where θ_0 belongs to an open set of an Euclidean space, moment-type conditions are used to both estimate the product $\lambda\mathbb{E}[Y_0]$ and θ_0 when observing a discrete-time version of (5.1.1). On the one hand, this technique does not provide enough information about λf and on the other hand, the assumption that h is "interval-similar", that is to say

$$\exists (c_n)_{n\geq 0}, \quad \forall n \geq 0, \quad h_{|[n,n+1]} = c_n \, h_{|[0,1]}$$

is too specific and thus not relevant for our field of applications compared to assumption (**SN**-4).
Another way of looking at this problem consists in rewriting (5.1.1) as

$$X_t = \int_{-\infty}^{t} h(t-s)\mathrm{d}L_s$$

where $(L_t)_{t\in\mathbb{R}}$ is a two-sided Lévy process such that L_1 has the same distribution as $Y_1 + \cdots + Y_{N_\lambda}$ where $(Y_i)_{i\geq 0}$ are given by (**SN**-3) and N_λ is a Poisson random variable with parameter λ independent of the random variables $(Y_i)_{i\in\mathbb{Z}}$. In this form, the shot-noise process $\mathbf{X}$ can be seen as a continuous moving average process. In particular, the subclass of Lévy-driven CARMA processes has received recent attention in the statistical inference domain, see Brockwell et al. (2007), Brockwell and Schlemm (2013a) for instance. However, even tough these models allow to consider broader driving Lévy processes with continuous parts (in contrast with pure-jump parts), the impulse responses are constrained to be, in the unidimensional case, of the specific form (see Brockwell and Schlemm (2013a)[Section 3])

$$h(t) = b\mathrm{e}^{-at}\mathbb{1}_{t\geq 0}, \quad a, b \in \mathbb{R}_+ \ .$$

Consequently, we choose to address the estimation problem from a different perspective. Coarsely speaking, we exploit the infinitely divisible property (see Sato (1999) e.g.) of the shot-noise process $\mathbf{X}$ given by (5.1.1) which provides an analytical expression of the characteristic function φ_{X_0} of the marginal distribution of X_0 and then derive a functional of the form

$$g(u) = K_h [f_\lambda] (u), \quad u \in \mathbb{R} , \tag{5.1.4}$$

where

- K_h is a bounded linear operator only depending on h acting from an Hilbert space to another Hilbert space,
- g depends on the characteristic function φ_{X_0} and its derivative,
- we only have access to a noisy version $\hat{g}_n$ of the function g.

It is unclear, for a given function h, whether the operator K_h is compact or not. In many cases, classical techniques to recover f_λ from the observations $\hat{g}_n$ based on the spectral theorem decomposition of the self-adjoint operator $K_h^\star K_h$ or on the singular value decomposition K_h when compact are not explicit.

In order to bypass this issue, we propose to sample both the function g and its estimator $\hat{g}_n$ and to estimate f_λ using two nonparametric procedures. Coarsely speaking, the problem of interest becomes the estimation of f_λ from

$$\hat{g}_n(u_i) = K_h [f_\lambda] (u_i) + \frac{\epsilon_n(u_i)}{\sqrt{n}}, \quad i \in \{1, \dots, N\} . \tag{5.1.5}$$

where N is an integer (the choice is detailed in Section 5.3) and $(\epsilon(u_1), \dots, \epsilon(u_N))$ are error terms.

When equipped with the norm

$$\|f\|_{\mathcal{W}^2} = \sqrt{\int_0^1 (f''(t))^2 \, \mathrm{d}t} , \tag{5.1.6}$$

the space $\mathcal{W}^2$ is a Reproducing Kernel Hilbert space: the representer theorem then provides a first estimator of the function f_λ. However, as explained in Section 5.2, the basis functions are not easily interpretable and not suited to handle positivity constraints on the function f_λ.

As a consequence, we introduce a second estimation technique using B-splines (see for instance De Boor (1978)) which allow to efficiently handle these constraints. Moreover, B-splines basis functions have local support which allows us to construct an estimator of f_λ that carries the local regularity of the targeted function. Finally, the minimization algorithm related to B-splines is a quadratic programming (QP) problem which leads to an easy and fast implementation.

The paper is organized as follows. In Section 5.2, we define the operator K_h and derive two estimators of f_λ. In Section 5.3, we present in details the procedure used to perform the density estimation and methods to handle the choice of the various hyperparameters. In Section 5.4, we illustrate the efficiency of our estimators on simulated datasets and

show the results obtained when dealing with gamma-ray spectroscopy datasets. Finally, Section 5.5 is devoted to the proofs of the main theorems.
In the following, we set the sampling rate δ to 1 without loss of generality since it only requires to modify λ and to scale h accordingly. Also, the random variables X and Y will respectively represent generic copies of X_0 and Y_0.

5.2 Functional Inverse Problem and Estimator

5.2.1 Functional operator

Statistical inference methods for shot-noise processes were initiated in Macchi and Picinbono (1972). More recent contributions include Trigano et al. (2015a), Trigano et al. (2015b) and Ilhe et al. (2015). The low sampling rate of the signal in nuclear applications 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 the latter is infinitely divisible, we can derive the analytical expression of the characteristic function φ of the shot-noise marginal. For every real u, we have (see e.g. Daley and Vere-Jones (1988) for details)

$$\varphi(u) = \exp\left(\lambda \int_0^{\infty} \left[\varphi_Y\left(uh(s)\right) - 1\right] \mathrm{d}s\right) . \tag{5.2.1}$$

In addition, under (**SN**-3) and (**SN**-4), we have that

$$\lambda \int_0^{\infty} \int_0^{\infty} |xh(s)|\, f(x) \mathrm{d}s \mathrm{d}x < \infty . \tag{5.2.2}$$

Using Fubini-Tonelli theorem, φ defined by (5.2.1) can be written as

$$\varphi(u) = \exp\left(\int_0^{\infty} \left(\int_0^{\infty} \left[\mathrm{e}^{\mathrm{i}uxh(s)} - 1\right] \mathrm{d}s\right) \lambda f(x) \mathrm{d}x\right)$$

Using (5.2.2), the characteristic function φ of X is differentiable and we get that

$$g(u) := \frac{\varphi'(u)}{\varphi(u)} = \int_0^1 \int_0^{\infty} \mathrm{i}xh(s)\mathrm{e}^{\mathrm{i}uxh(s)} f_\lambda(x) \mathrm{d}x \mathrm{d}s =: K_h[f_\lambda](u), \quad u \in \mathbb{R} . \tag{5.2.3}$$

In practice, we only observe an empirical version $\hat{g}_n$ of the function g in Equation (5.2.3) which is defined by

$$\hat{g}_n(u) := \frac{\hat{\varphi}_n'(u)}{\hat{\varphi}_n(u)} \mathbb{1}_{\{|\hat{\varphi}_n(u)| \geq \mathrm{e}^{-\hat{\sigma}_n^2 u^2/2}/2\}} , \tag{5.2.4}$$

where $\hat{\varphi}_n$ is the empirical characteristic function and $\hat{\sigma}_n^2$ represents the empirical variance associated to the sample of observations. The indicator of the set

$$\{|\hat{\varphi}_n(u)| \geq \mathrm{e}^{-\hat{\sigma}_n^2 u^2/2}/2\}$$

5

is a safeguard against instabilities in the ratio $\hat{\varphi}_n'/\hat{\varphi}_n$ when $\hat{\varphi}_n(u)$ is small. According to Lemma 12, we have

$$\mathbb{1}_{\{|\hat{\varphi}_n(u)|\geq e^{-\hat{\sigma}_n^2 u^2/2}/2\}} \underset{n\to\infty}{\to} 1 \text{ in probability }.$$

Is this paper, the idea behind the estimation of f_λ is to replace g in Equation (5.2.3) by $\hat{g}_n$ and to minimize over $\mathcal{W}_+^2$ a functional of the form

$$\xi \to \; F\left(K_h[\xi], \hat{g}_n\right) + \mu P(\xi)\;, \tag{5.2.5}$$

where F is measuring the goodness of $K_h[\xi]$ with $\hat{g}_n$, P is a penalty term and μ is some positive number called regularization parameter or *penalty weigth*. In order to justify this approach, let us review some well-known facts regarding the linear operator K_h. Setting

$$k_h(u,x) := \int_0^\infty \mathrm{i}xh(s)e^{\mathrm{i}uxh(s)}\mathrm{d}s \quad \text{ for } u\in\mathbb{R}, x\in[0,1]\;, \tag{5.2.6}$$

Equation (5.2.3) rewrites as

$$g(u) = K_h[f_\lambda](u) = \int_0^1 k(u,x)f_\lambda(x)\mathrm{d}x, \quad u\in\mathbb{R}\;. \tag{5.2.7}$$

The functional equation $g = K_h\left[f_\lambda\right]$ corresponds to a Fredholm integral equation of the first kind. In Wahba (1973), convergence rates are provided when the operator is compact and are linked to the eigenvalues resulting from the singular value decomposition of the operator. It is not clear to determine for which impulse response functions h the operator K_h is compact (See Section 5.8 for a first result). However, the operator K_h can be described by exploiting the RKHS property of the functional space $(\mathcal{W}^2, \|\cdot\|_{\mathcal{W}^2})$. Its reproducing kernel is given by

$$R(x,y) = \int_0^1 (u-x)_+\,(u-y)_+\,\mathrm{d}u\;, \quad x,y\in[0,1]\;. \tag{5.2.8}$$

Then, Nashed and Wahba (1974) and Wahba (1990)[Chapter 8] show that $\{K_h[f] \;:\; f\in\mathcal{W}^2\} = \mathcal{H}_Q$, where $\mathcal{H}_Q$ corresponds to the Reproducing Kernel Hilbert space with continuous kernel Q defined by

$$Q(s,t) := \int_0^1\int_0^1 k_h(s,x)R(x,y)k_h(t,y)\mathrm{d}s\mathrm{d}t, \quad t,s\in\mathbb{R}\;.$$

In the next section, we detail our estimation procedure taking into consideration the asymptotic properties of the estimated function $\hat{g}_n$ defined in (5.2.4). Then, we first exploit the reproducing property of the space $\mathcal{W}^2$ to construct a first estimator of the *flow* function f_λ and propose another method using B-splines functions that enjoys a more interpretable solution and allows us to handle the positivity constraints of the function f_λ.

5.2.2 Estimators

In the light of the previous arguments, we thus propose to use regularized functional regression techniques in order to construct estimators of the flow function f_λ from the observations $X_1, \ldots, X_n$ of the shot-noise process given by (5.1.1). For reasons related to the complexity and the time efficiency of the numerical estimation procedure, we only consider a finite number of equations of the form

$$\hat{g}_n(u_i) = K_h\left[f_\lambda\right](u_i) + \frac{\epsilon_n(u_i)}{\sqrt{n}}, \quad i \in \{1, \ldots, N\} . \tag{5.2.9}$$

where N is an integer and $\epsilon_n(u_1), \ldots, \epsilon_n(u_N)$ are correlated error terms due to the approximation of g by $\hat{g}_n$ defined by (5.2.4) in order to estimate the function f_λ. The choice of the evaluation grid $\underline{u}$ is postponed and will be discussed in Section 5.3 dedicated to the practical procedure. Following Nashed and Wahba (1974), we define F and P respectively by

$$P(\xi) := \|\xi\|^2_{\mathcal{W}^2} , \tag{5.2.10}$$

and

$$F\left(K_h[\xi], \hat{g}_n\right) = \tilde{F}\left(K_h[\xi](u_1), \ldots, K_h[\xi](u_N), \hat{g}_n(u_1), \ldots, \hat{g}_n(u_N)\right) , \tag{5.2.11}$$

where $\tilde{F} : \mathbb{C}^{2N} \to \mathbb{R}_+$ has to be chosen adequately.
A usual choice for $\tilde{F}$ (see Wahba (1977)) is the mean squared-error defined for all $x_{1:N}, y_{1:N} \in \mathbb{C}^N$ by

$$\|x_{1:N} - y_{1:N}\|^2 := \sum_{k=1}^{N} |x_i - y_i|^2 .$$

Such a choice is particularly adapted in the particular case where the function g can be estimated by a function $\hat{g}_n$ such that the approximation error terms $\epsilon_n(u_i) = \sqrt{n}\left(\hat{g}_n(u_i) - g(u_i)\right)$ for $i = 1, \ldots, n$ in (5.2.9) are centered, uncorrelated and with constant variance. However, such an assumption fails when using the estimator $\hat{g}_n$ defined by (5.2.4). Indeed, the distribution of $\epsilon(u_1), \ldots, \epsilon(u_N)$ is quite involved and depends on the unknown parameter f_λ. We thus rely on the following Gaussian approximation that holds as $n \to \infty$.

Theorem 5.2.1. *Let $\lambda > 0$ and $X_1, \cdots, X_n$ be a sample of the stationary shot-noise process given by* (5.1.1) *under Assumptions* (**SN**-3), (**SN**-4) *and such that $\mathbb{E}\left[|Y_1|^4\right] < \infty$ for some $\eta > 0$. Let φ and φ' respectively denote the characteristic function of X_1 and its derivative. Then, for any positive number K, the stochastic process $(\epsilon_n(u))_{u \in [-K,K]}$ converges weakly to a centered multivariate complex-valued Gaussian process denoted by $\left(\check{Z}_u\right)_{u \in [-K,K]}$. where $(Z_u)_{u \in [-K,K]}$ is a centered multivariate complex-valued Gaussian process with covariance function denoted C.*

Proof. The proof is deferred to Section 5.5. □

5

From this result holds the following corollary.

Corollary 2. *Let $\lambda > 0$ and $X_1, \cdots, X_n$ be a sample of the stationary shot-noise process given by (5.1.1) under Assumptions (SN-3), (SN-4) and such that $\mathbb{E}\left[|Y_1|^4\right] < \infty$. Let φ and φ' respectively denote the characteristic function of X_1 and its derivative. Then, for any positive integer N and any choice of distinct real numbers $u_1, \ldots, u_N$, the vector*

$$(\epsilon_n(u_1), \ldots, \epsilon_n(u_N)) \ , \tag{5.2.12}$$

is asymptotically normal with mean zero, covariance matrix C and relation matrix Γ defined by

$$C(i,j) = \check{C}(u_i, u_j) \ , \quad i, j = 1, \ldots, N \ ,$$

and

$$\Gamma(i,j) = \check{\Gamma}(u_i, u_j) \ , \quad i, j = 1, \ldots, N \ ,$$

where $\check{C}$ and $\check{\Gamma}$ are respectively the asymptotic covariance and relation functions of $(\epsilon_n(u))_{u \in [-K,K]}$.

As a consequence of this theorem and this corollary, it is preferable to define more general goodness of fit terms $\tilde{F}$ that can take account of the covariance structure of the errors terms $(\epsilon(u_1), \ldots, \epsilon_n(u_N))$. Given a grid $\underline{u}$ of N points and an symmetric matrix W of size $2N$, we set for all $x, y \in \mathbb{C}^N$,

$$\tilde{F}_{W,\underline{u}}(x,y) := \left\| W^{1/2} \begin{bmatrix} \mathrm{Re}(x-y) \\ \mathrm{Im}(x-y) \end{bmatrix} \right\|_2^2 .$$

The choice of W depends on the correlation structure of $(\epsilon_n(u))_{u \in \mathbb{R}}$. In standard generalized least square estimation, the optimal choice is to take W as the inverse of the covariance matrix of

$$(\mathrm{Re}(\epsilon_n(u_1)), \ldots, \mathrm{Re}(\epsilon_n(u_N)), \mathrm{Im}(\epsilon_n(u_1)), \ldots, \mathrm{Im}(\epsilon_n(u_N)))$$

which is given by

$$\Omega = \begin{bmatrix} \frac{1}{2}\mathrm{Re}(C+\Gamma) & \frac{1}{2}\mathrm{Im}(C+\Gamma) \\ \frac{1}{2}\mathrm{Im}(C+\Gamma)^T & \frac{1}{2}\mathrm{Re}(C-\Gamma) \end{bmatrix} \tag{5.2.13}$$

However, the explicit form of Ω is unknown since it depends on the flow function f_λ associated to the shot-noise observations $X_1, \ldots, X_n$. In Section 5.3, we detail two ways to estimate this covariance matrix using bootstrap techniques.
We now present how to construct two different estimators of f_λ.

5.2.2.1 Nonparametric estimation using Mercer's theorem

The minimization problem given by (5.2.5) can be rewritten

$$\hat{f}_{n,\mu,W} := \underset{\xi \in \mathcal{W}^2, \mathrm{Re}(\xi) \geq 0, \mathrm{Im}(\xi)=0}{\operatorname{argmin}} \|W^{1/2} \cdot \begin{bmatrix} \mathrm{Re}\left(K_{\underline{u}}[\xi] - \hat{g}_n(\underline{u})\right) \\ \mathrm{Im}\left(K_{\underline{u}}[\xi] - \hat{g}_n(\underline{u})\right) \end{bmatrix} \|_2^2 + \mu \|\xi\|_{\mathcal{W}^2}^2 . \tag{5.2.14}$$

The following proposition establishes that the solution of the above minimization problem is a linear combination of a finite number of functions.

Theorem 5.2.2 (Heckman et al. (2012), Corollary 3.1.). *Let $\left(\mathcal{W}^2, \|\cdot\|_{\mathcal{W}^2}\right)$ denote the Sobolev space of order 2 given by (SN-3), R its reproducing kernel given by 5.2.8, K_h be the operator defined by (5.2.3). Let also $X_1, \ldots, X_n$ be n observations of the shot-noise process given by (5.1.1) under Assumptions (SN-3), (SN-4) and (SN-5). Finally, let N be a positive integer, $0 \leq u_1 < \cdots < u_N$ be positive numbers and $\hat{\Omega}_n$ be a positive definite estimator of the covariance matrix Ω given by 5.2.13. Then, the solution of the minimization problem over $\mathcal{W}^2_+$*

$$\underset{\xi \in \mathcal{W}^2, \xi \geq 0}{argmin} \left\| \hat{\Omega}_n^{-1/2} \cdot \begin{bmatrix} Re\left(K_{\underline{u}}[\xi] - \hat{g}_n(\underline{u})\right) \\ Im\left(K_{\underline{u}}[\xi] - \hat{g}_n(\underline{u})\right) \end{bmatrix} \right\|_2^2 + \mu \|\xi\|^2_{\mathcal{W}^2} ,$$

where

$$K_{\underline{u}}[\xi] := (K_h[\xi](u_1), \ldots, K_h[\xi](u_N)) \quad \text{and} \quad \hat{g}_n(\underline{u}) := (\hat{g}_n(u_1), \ldots, \hat{g}_n(u_N)) , \tag{5.2.15}$$

can be written as

$$f_a(x) := \sum_{k=1}^{2N} a_k \eta_k(x) ,$$

where $a_1, \ldots, a_{2N}$ are real numbers and $\eta_{\underline{u}} := (\eta_1, \ldots, \eta_N, \eta_{N+1}, \ldots, \eta_{2N})$ are the functions defined by

$$\eta_i(x) := K_h[R(\cdot, x)](u_i) , \quad x \in [0,1], i \in \{1, \ldots, N\} . \tag{5.2.16}$$

Furthermore, the solution $f_{\hat{a}_n}$ of the minimization problem

$$\underset{\xi \in \mathcal{W}^2, \xi \geq 0}{argmin} \left\| \hat{\Omega}_n^{-1/2} \cdot \begin{bmatrix} Re\left(K_{\underline{u}}[\xi] - \hat{g}_n(\underline{u})\right) \\ Im\left(K_{\underline{u}}[\xi] - \hat{g}_n(\underline{u})\right) \end{bmatrix} \right\|_2^2 + \mu \|\xi\|^2_{\mathcal{W}^2}$$

is obtained for the vector $\hat{a}_n(\mu)$ in $\mathbb{R}^{2N}$ that is solution of the quadratic problem

$$\begin{aligned} &\underset{a \in \mathbb{R}^{2N}}{minimize}\ a^\star \{G \hat{\Omega}_n^{-1} G + \mu G\} a - a^\star G \hat{g}_n(\underline{u}) - \hat{g}_n(\underline{u})^\star G a \\ &\text{subject to } a^{\mathrm{T}} \eta_{\underline{u}} \geq 0 . \end{aligned} \tag{5.2.17}$$

where G is the positive definite matrix defined by

$$G_{i,j} := < \eta_i, \eta_j >_{\mathcal{W}^2} , \quad 1 \leq i, j \leq 2N .$$

Proof. See Section 5.7. □

This result takes advantage of the Riesz Representation Theorem and uses the Representer theorem to reduce the minimization problem (5.2.5) over a infinite-dimensional space into the minimization problem over a finite-dimensional one.

5.2.2.2 B-splines estimation

We will see in Section 5.4 that the estimator obtained in Subsection 5.2.2.1 does not perform well both on simulated signals and real datasets. This is essentially due to the fact that the space of functions $(\eta_i)_{1\leq i\leq N}$ is not adapted to reconstruct functions belonging to $\mathcal{W}^2$ and because the decomposition over $(\eta_i)_{1\leq i\leq N}$ does not allow to handle efficiently positivity constraints. To tackle this issue, we now introduce another estimator.
Rather than relying on the reproducing property of the space $\mathcal{W}^2$ to reduce the minimization problem into a finite-dimensional one, we propose to approximate the space $\mathcal{W}^2$ by approaching it with finite-dimensional spaces. We propose to approximate the functional space $\mathcal{W}^2$ using a B-spline basis whose definition is recalled hereafter.

Definition 15. *Let k be a positive integer and q be a non-negative integer. Also, let $\underline{t} := (t_j)_{0\leq j\leq 2q+k}$ be the sequence of real numbers defined by*

$$t_j := \begin{cases} 0 & \text{if } j \in \{0,\dots,q\} \\ (j-q)/k & \text{if } j \in \{q+1,\dots,q+k-1\} \\ 1 & \text{if } j \in \{q+k,\dots,2q+k\} \end{cases} .$$

The B-spline functions of order 0 with $k-1$ equispaced interior knots are given by

$$B_j^0(t) := \mathbb{1}_{[t_{j-1},t_j)}(t) , \quad j \in \{1,\dots,2q+k\} .$$

For any positive integer s such that $s \leq q$, the B-spline basis functions of order s are given by the recursive equations

$$B_j^s(t) := w_{j,s-1}(t) B_j^{s-1}(t) + \Big(1 - w_{j+1,s-1}(t)\Big) B_{j+1}^{s-1}(t), \quad j \in \{1,\dots,2q+k\} , \tag{5.2.18}$$

where the functions $w_{i,j}$ are defined for $(i,j) \in \{\} \times \{0,\dots,q\}$ by

$$w_{i,j}(t) := \begin{cases} \frac{t-t_i}{t_{i+j}-t_i} & \text{if } t_{i+j} \neq t_i \\ 0 & \text{else} \end{cases} , \quad t \in \mathbb{R} .$$

Then, the B-spline space $\mathbb{S}_k^q$ (respectively the positive B-spline space $\mathbb{S}_{k,+}^q$) or order q with $k-1$ equispaced interior knots is defined as the vector space (respectively the cone) spanned by the $k+q$ nonzero functions $(B_j^q)_{1\leq j\leq k+q}$ defined by (5.2.18). In the following, we say that f is a B-spline (respectively a positive B-spline) of order q with $k-1$ equispaced interior knots whenever f belongs to the space $\mathbb{S}_k^q$ (respectively $\mathbb{S}_{k,+}^q$).

We will use the following features offered by B-spline bases. Firstly, there exists a simple relation between the derivatives of B-splines of order q and B-splines of order $q-1$ for any integer $q > 0$. Indeed, let v be a function belonging to $\mathbb{S}_k^q$ for k and q two positive integers. Then

$$v(t) = \sum_{j=1}^{k+q} \omega_j \, B_{k,j}^q(t) := \underline{\omega}^T \underline{B}_k^q(t) , \quad \forall t \in [0,1] ,$$

5

for some $\underline{\omega} = (\omega_1, \ldots, \omega_{k+q}) \in \mathbb{R}^{k+q}$. From De Boor (1978), we have

$$v'(t) = \underline{\omega}^{(1)^T} \underline{B}_k^{q-1}(t) \quad \text{with } \underline{\omega}^{(1)} = D^1_{k,q}\underline{\omega} \,, \tag{5.2.19}$$

where $D^1_{k,q}$ is a matrix only depending on k and q. Such a result can be generalized by recursion and we can express the m derivative of v with $m < q-1$ as

$$v^{(m)}(t) = \underline{\omega}^{(m)^T} \underline{B}_k^{q-m}(t) \quad \text{with } \underline{\omega}^{(m)} = D^m_{k,q}\underline{\omega} \,. \tag{5.2.20}$$

Secondly, for a given non-negative integer q and a positive integer k, the B-splines functions $(B^q_j)_{1\le j\le k+q}$ are compactly supported in the interval

$$[t_j, t_{j+q+1}) \quad \text{that satisfies} \quad 0 \le t_{j+q+1} - t_j \le (q+1)/k \,,$$

where $\underline{t}$ is the knots vector given in Definition 15. As we will detail in Section 5.4, this property is appealing in the context of particle counting systems where the flow functions f_λ usually correspond to a mixture of *singular peak distributions* and a *continuum near zero*.

Finally, the space $\mathcal{W}^2$ can be approximated by the spaces $\mathbb{S}^q_k$. On the one hand, for two positive integers k and q, the B-spline functions of order q are $q-1$ times continuously differentiable. Since we choose a penalty term P given by (5.2.10) that involves derivatives up to order 2, we need to parametrize using B-splines with order $q \ge 3$. In fact, we exactly set $q = 3$ to reduce the size of the supports. On the other hand, any positive function in $\mathcal{W}^2$ can be approximated by a function belonging to $\mathbb{S}^3_{k,+}$, provided that k is large enough.

In addition, the following lemma provides another writing of the norm of a function in the space $\mathcal{S}^3_k$.

Lemma 7. *Let k be a positive integer and f be a function belonging to the space $\mathcal{S}^3_k$. By definition, there exists a vector ω in $\mathbb{R}^{k+3}$ such that $f = \omega^{\mathrm{T}}\underline{B}^3_k$. Let $\|\cdot\|_{\mathcal{W}^2}$ be the norm of the space $\mathcal{W}^2$ introduced in (5.1.6). Then, we have*

$$\|f\|^2_{\mathcal{W}^2} = \omega^{\mathrm{T}} M_k \omega \,, \tag{5.2.21}$$

where M_k is a matrix of size $k+3$ with coefficients given by

$$M_k(i,j) = \int_0^1 (B^3_i)''(u)(B^3_j)''(u)\mathrm{d}u \,, \quad 1 \le i,j \le k+3 \,.$$

Combining the three previous arguments provides a way to control the local regularity of functions by deriving a minimization procedure that penalizes certain coefficients of the B-splines basis while being explicitely solvable. The proofs of all the above properties are available in De Boor (1978).

We can now propose the explicit definition of the second estimator. Given a grid of N points $\{u_1, \ldots, u_N\}$, a positive definite estimator $\hat{\Omega}_n$ of the matrix Ω and a positive integer k , the estimator $\hat{f}_{n,\lambda}(\underline{u}, \mu, \hat{\Omega}_n, k)$ is defined as the solution of the minimization problem

$$\underset{\xi\in\mathbb{S}^3_{k,+}}{\operatorname{argmin}} \left\| \hat{\Omega}_n^{-1/2} \cdot \begin{bmatrix} \mathrm{Re}\left(K_{\underline{u}}[\xi] - \hat{g}_n(\underline{u})\right) \\ \mathrm{Im}\left(K_{\underline{u}}[\xi] - \hat{g}_n(\underline{u})\right) \end{bmatrix} \right\|_2^2 + \mu\|\xi\|^2_{\mathcal{W}^2} \,, \tag{5.2.22}$$

where $\hat{g}_n(\underline{u})$ and $K_{\underline{u}}[\xi]$ are defined by (5.2.15).
For any positive integer k, let $\underline{KB}_k$ denote the multivalued function defined by

$$\underline{KB}_k := \left(KB_1^3, \ldots, KB_{k+3}^3\right) ,$$

where we set

$$KB_j^3 := K_h[B_j^3] , \quad j = 1, \ldots, k+3 .$$

For $j \in \{1, \ldots, k+3$, let also rKB_j^3 and iKB_j^3 respectively denote the real and imaginary parts of the function KB_j^3.
In fact, using Lemma 7, it turns out that the solution $\hat{f}_{n,\lambda}(\underline{u}, \mu, \hat{\Omega}_n, k)$ can be written in the form

$$\hat{f}_{n,\lambda}(\underline{u}, \hat{\Omega}_n, k, \mu) = \underline{\hat{\omega}}_{n,\lambda}^T \underline{B}_k^3$$

where $\underline{\hat{\omega}}_{n,\lambda}$ is the solution of quadratic programming problem over $\mathbb{R}_+^{k+3}$

$$\left\| \hat{\Omega}_n^{-1/2} \cdot \begin{bmatrix} \mathrm{Re}\left(\underline{KB}_k(\underline{u})\ \underline{\omega} - \hat{g}_n(\underline{u})\right) \\ \mathrm{Im}\left(\underline{KB}_k(\underline{u})\ \underline{\omega} - \hat{g}_n(\underline{u})\right) \end{bmatrix} \right\|_2^2 + \mu \underline{\omega}^T M_k \underline{\omega} , \tag{5.2.23}$$

where M_k is the symmetric matrix given by Lemma 7 and $\underline{KB}_k(\underline{u})$ is a matrix of size $N \times p$ whose ith line is given by $\underline{KB}_k(u_i)$.

5

5.3 Practical procedure

In this section, we address the practical problem of choosing the several parameters of the optimization procedure given by (5.2.23) based on the sole knowledge of the impulse response function h and a finite sample $X_1, \ldots, X_n$ of the shot-noise process given by (5.1.1).
The sequence of estimators $\left(\hat{f}_{n,\lambda}(\underline{u}, \mu, \hat{\Omega}_n, k)\right)_{n \geq 1}$ obtained by the minimization problem (5.2.3) relies on multiple hyper-parameters.

5.3.1 Choice of N and $\underline{u}$

The choice of the evaluation grid of the function $\hat{g}_n$ given by (5.2.4) directly influences the covariance matrix Ω and the number of knots k. We discuss in this section how to select them from the observations $X_1, \ldots, X_n$ of the shot-noise process. As it is tempting to choose a large integer N and a fine grid $\underline{u}$, papers tackling ad-hoc problems (Carrasco et al. (2002),Yu (2004)) suggest that such a choice would necessary lead to singular estimators $\hat{\Omega}_n$ of the covariance matrix of the normalized errors vector in Corollary 2.
Since an optimal selection of the evaluation grid still remains an open question in several fields, we address the sub-problem of selecting a grid $\{u_1, \ldots, u_N\}$ in the particular case where

$$u_i := \frac{2\pi(i-1)}{N\Delta} \quad \text{for } i = 1, \ldots, N$$

and some positive number Δ. We impose this choice because it both reduces the number of hyperparameters to N and Δ and allows to compute a quick estimator of the estimator vector $\hat{g}_n(\underline{u})$. We detail below how to compute these values in two steps.
Firstly, although it theoretically converges to the true density of shot-noise marks, the evaluation of the empirical characteristic function and its derivative based on observations $X_1, \cdots, X_n$ might be time-consuming when the sample size n is large. To circumvent this issue, we propose to compute the empirical characteristic function using the Fast Fourier Transform (abbreviated FFT in the following) of an appropriate histogram of the vector $X_1, \cdots, X_n$. More precisely, for a strictly positive fixed Δ, we consider the grid $G = \{\Delta l : \lfloor \min_{k\leq n}(X_k)/\Delta \rfloor \leq l \leq \lceil \max_{k\leq n}(X_k)/\Delta \rceil\}$ and compute the normalized histogram H of the sample sequence $(X_l)_{1\leq l\leq n}$ with respect to the grid G defined by

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

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

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

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

$$\hat{\varphi}_{\Delta,n}(u) := \sum_{l=m_n}^{M_n} H(l) e^{iu\Delta(l+1/2)} . \tag{5.3.1}$$

For any real u, we have the following upper bounds

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

and

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

showing that the approximations are close to the true functions for small values of Δ and u. Then, the N-point Fast Fourier Transform of the vectors $H = (H(l))_{l\in[[m_n,M_n]]}$ and $\tilde{H} = (H(l)\Delta\cdot(l+1/2))_{l\in[[m_n,M_n]]}$ respectively provide, for small values of Δ, estimators of $\hat{\varphi}_n(\underline{u})$ and $\hat{\varphi}'_n(\underline{u})$.

5.3.2 Estimation of the covariance operator $\hat{\Omega}_n$

In order to implement the estimation procedure, we need to construct a positive definite and consistent estimator of the covariance matrix Ω defined in Corollary 2. In this

section, we present three different procedures to construct estimators $\hat{\Omega}_n$ of the covariance function. The first two methods are based on resampling methods introduced in Efron (1979) while the last one is purely empirical.
For all the estimators $\hat{\Omega}_n$ of Ω derived in the next lines, we have no guarantee that the matrix $\hat{\Omega}_n$ is nonsingular and thus we need to consider a regularized version. For any positive sequence $(\alpha_n)_{n\in\mathbb{N}}$, we define the Tikhonov approximation of the generalized inverse of Ω by

$$\left(\hat{\Omega}_n^{\alpha_n}\right)^{-1} := \left(\hat{\Omega}_n + \alpha_n I_N\right)^{-1} .$$

Parametric bootstrap:
We can obtain a first estimator of the covariance matrix by a two-step procedure. First, one can construct an estimator $\tilde{\omega}_n$ defined as the solution of the minimization problem

$$\tilde{\omega}_n := \underset{\omega\in\mathbb{R}_+^{k+3}}{\operatorname{argmin}} \|\underline{KB}_k(\underline{u})\ \omega - \hat{g}_n(\underline{u})\|_2^2 . \tag{5.3.2}$$

We thus use this estimator to simulate N (which depends on the number of initial observations n) independent shot-noise processes $\mathrm{X}^{(i)}$ which satisfy Assumption (**SN**-3), (**SN**-4) and whose intensity is given by $\tilde{\lambda}_n := \sum_i \tilde{\omega}_n(i)$ and the marks density is set to $(\tilde{\omega}_n)^T \underline{B}_k^3/\lambda_n$. We then extract M samples of observations $(X_1^{(i)}, \ldots, X_n^{(i)})$, $1 \leq i \leq M$ and we estimate the covariance matrix Ω by

$$\hat{\Omega}_n(i,j) := M^{-1} \sum_{k=1}^{M} \left(\tilde{g}_{n,k}(u_i) - \tilde{\omega}_n^T \underline{KB}_k(u_i)\right) \overline{\left(\tilde{g}_{n,k}(u_j) - \tilde{\omega}_n^T \underline{KB}_k(u_j)\right)} . \tag{5.3.3}$$

Nonparametric bootstrap (moving block procedure):
Unlike the parametric bootstrap, this procedure makes no assumptions on how the statistics of interest depend on the parameter to estimate. In particular, we focus our attention to the moving block bootstrap procedure. Given a sample $(X_1, \ldots, X_n)$, the main idea is to construct new time series by concatenating smaller blocks drawn from the initial sequence of observations as illustrated in Figure 5.1. We briefly review the method based on random resampling and with replacement introduced in Kunsch (1989) as follows:

1. Let $(X_1, \ldots, X_n)$ be a sample of observations arising from a real-valued stationary process $(X_t)_{t\geq 0}$ depending on a parameter θ_0, possibly infinite dimensional. We want to estimate a statistic- in our case Ω given in Corollary 2- that depends on the unknown parameter.

2. Let $l := l_n$ be an integer such that $1 < l < n$ and $b := \lfloor n/l_n \rfloor$. We then draw b blocks $(\mathbf{X}_1^\star, \ldots, \mathbf{X}_b^\star)$ from the random distribution

$$F_n^l := (n-l+1)^{-1} \sum_{i=0}^{n-l} \delta_{(X_{i+1},\ldots,X_{i+l})} .$$

3. Construct a sequence $\mathbf{Y}_1^\star := \left(Y_{1,1}^\star, \ldots, Y_{1,bl}^\star\right)$ such that

$$Y_{1,kb+j}^\star := X_{k+1}^\star(j), \quad 0 \leq k \leq b-1, \quad 1 \leq j \leq l .$$

5

4. Repeat steps 2 and 3 $B := B_n$ times to construct $\mathbf{Y}_i^\star$ for $2 \le i \le B$.

5. Compute the estimator $\tilde{\Omega}_n$ of Ω defined by

$$\tilde{\Omega}_n(i,j) := B_n^{-1} \sum_{k=1}^{B_n} \left(\breve{g}_{n,k}(u_i) - B_n^{-1} \sum_{k=1}^{B_n} \breve{g}_{n,k}(u_i) \right) \left(\breve{g}_{n,k}(u_j) - B_n^{-1} \sum_{k=1}^{B_n} \breve{g}_{n,k}(u_j) \right) . \quad (5.3.4)$$

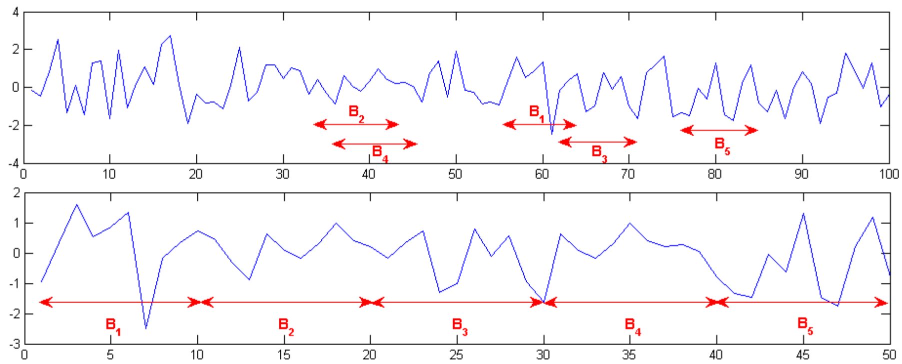


Figure 5.1 : Moving-Block Bootstrap example

Remark 9. There are several slight modifications (random lengths of blocks, versions without replacement) of the moving-block procedure gathered in Lahiri (1999) and more recently Politis and White (2004) where the asymptotic efficiencies of the different methods are compared.

5.3.3 Cross-validation choice of μ

We propose to choose the penalty coefficient μ in the optimization problem (5.2.23) in an adaptive fashion. However, (5.2.23) corresponds to a non-negative ridge regression problem and, up to our knowledge, there is no cross-validation methods to select an optimal regularization coefficient when dealing with positivity constraints. Instead we follow Wood (2000) that introduces a practical and efficient selection procedure in the unconstrained scenario. Hence, given a discrete grid of M penalty terms $\{\mu_1, \dots, \mu_M\}$, we choose the regularizing parameter μ that minimizes some criterion in relation to the quadratic unconstrained problem

$$\min_{\omega \in \mathbb{R}^p} \left\| W^{1/2}(y - X\omega) \right\|_2^2 + \mu\, \omega^T S \omega \,,$$

where, for n and p two non-negative integers, $W \in \mathbb{R}^{n\times n}$ is a weight matrix with positive eigenvalues, $y \in \mathbb{R}^n$ a vector of data to be fitted, $\omega \in \mathbb{R}^p$ a parameter vector associated to some model and $X \in \mathbb{R}^{n\times p}$ the design matrix.
The smoothing selection boils down to choose the real number $\mu^\star \in \{\mu_1, \ldots, \mu_M\}$ that minimizes the generalized cross-validation score

$$\mathrm{GCV}(\mu) := \frac{\|W^{1/2}\{y - A_\mu y\}\|_2^2}{\left(n - \mathrm{Tr}(A_\mu)\right)^2} \ ,$$

where A_μ is the *influence matrix* defined by

$$A_\mu := X\left(X^T W X + \mu S\right)^{-1} X^\star W \ .$$

Furthermore, the algorithmic method detailed in Wood (2000) is computationally economical since it involves between $O(p^3)$ and $O(Mp^3)$ operations. Theoretical justifications of this method can be found in the earlier work Craven and Wahba (1978).
We then use $\mu^\star$ as the regularization parameter of our algorithm and illustrate in Section 5.4 the performance of this choice by comparing the integrated squared error

$$\|\hat{f}_{n,\lambda}(\underline{u}, \hat{\Omega}_n, k, \mu^\star) - f_\lambda\|_2^2$$

given by this procedure with the optimal integrated squared error

$$\underset{\mu\in\{\mu_1,\ldots,\mu_M\}}{\operatorname{argmin}} \|\hat{f}_{n,\lambda}(\underline{u}, \hat{\Omega}_n, k, \mu) - f_\lambda\|_2^2 \ .$$

Applying it to the shot-noise framework, the generalized cross-validation score becomes

$$\mathrm{GCV}(\mu) = \frac{\|\hat{\Omega}_n^{-1/2}\{\hat{g}_n(\underline{u}) - \hat{A}_{n,\mu}\hat{g}_n(\underline{u})\}\|_2^2}{\left(N - \mathrm{Tr}(\hat{A}_{n,\mu})\right)^2} \ , \tag{5.3.5}$$

with

$$\hat{A}_{n,\mu} := \underline{KB}_k \cdot \left(\underline{KB}_k^\star \hat{\Omega}_n^{-1} \underline{KB}_k + \mu M_k\right)^{-1} \cdot \underline{KB}_k^\star \hat{\Omega}_n^{-1} \ ,$$

where $\hat{\Omega}_n$ is a positive definite estimator of the asymptotic covariance matrix Ω given in Theorem 5.2.1 and the matrices $\underline{KB}_k$ and M_k are defined in (5.2.23). In our applications, we select a grid of 100 penalization terms that are logarithmically equispaced between 10^{-8} and 10^8.

5.3.4 Algorithms

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

Algorithm 1: Kernel estimator

Inputs : Shot-noise observations $X_1, \cdots, X_n$, impulse response function h

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 $dH_i = \Delta_n \, iH_i$
2. Choice of $N := 2^{\lfloor \log(n) \rfloor}$ and computation of $X\phi = \mathrm{FFT}(H, N)$, $dX\phi = \mathrm{FFT}(dH, N)$.
   For $k = 1, \cdots, N$, we set $\hat{g}_n\left(\frac{2\pi(k-1)}{N\Delta}\right) := \frac{dX\phi_k}{X\phi_k}$
3. Computation of the basis functions $\eta_1, \ldots, \eta_{2N}$ defined by (5.2.16).
4. Estimation of $\hat{\Omega}_n$ by bootstrap.
5. Estimation of $\hat{a}_n(\mu)$ for $\mu \in \{10^{-8}, \ldots, 10^{8}\}$ via (5.2.17) and choice by the cross-validation score using (5.3.5).

Algorithm 2: B-splines estimator

Inputs : Shot-noise observations $X_1, \cdots, X_n$, impulse response function h

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 $dH_i = \Delta_n \, iH_i$
2. Choice of $N := 2^{\lfloor \log(n) \rfloor}$ and computation of $X\phi = \mathrm{FFT}(H, N)$, $dX\phi = \mathrm{FFT}(dH, N)$.
   For $k = 1, \cdots, N$, we set $\hat{g}_n\left(\frac{2\pi(k-1)}{N\Delta}\right) := \frac{dX\phi_k}{X\phi_k}$
3. Choice of the knots' number $k = N$, the degree $q = 3$ and computation of the design matrix $(K_h[B_j^3](2\pi(i-1)/N\Delta))_{i,j}$ for $i \in \{1, \ldots, N\}$, $j \in \{1, \ldots k+3\}$
4. Estimation of $\hat{\Omega}_n$ by bootstrap.
5. Estimation of $\hat{\omega}_n(\alpha)$ for $\alpha \in \{2^{-15}, \ldots, 2^{15}\}$ via (5.2.23) and choice by the cross-validation score using (5.3.5).

5

5.4 Numerical Resuts and Applications

In this section, we illustrate the efficiency of Algorithms 1 and 2 both on simulated and real data sets. Usually, in nuclear applications, the words intensity and impulse response are respectively replaced by *counting rate* and *pulse* .

Existing methods: a quick state of the art

Up until now, the nuclear science experts use a pretty natural method to estimate the gamma ray spectra from the observations $X_1, \ldots, X_n$. After a preprocessing step, it basically consists in establishing an histogram of the local maxima of the observations. This idea stems from the fact that, at low counting rates λ, the probability that two pulses overlap becomes close to zero provided that the support of the pulse is short enough.

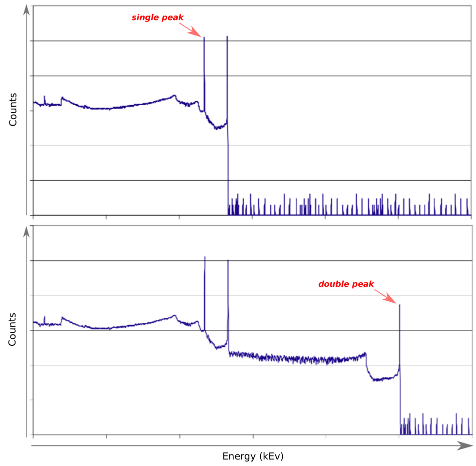


Figure 5.2 : Illustration of a double peak artefact. Top: incident spectrum. Bottom: estimated spectrum using the local maxima counting method.

5

Then, measuring the local maxima of the signal provide a good approximation of the spectrum f_λ. However, this argument fail to work when the intensity λ of the source increases. For instance, the resulting spectrum exhibits artifacts like sum peak, that is to say, we observe "artificial" peaks appearing at multiples of a present mono-energetic ray, as shown in Figure 5.2.

5.4.1 Simulated data

In this section, we illustrate the performances of Algorithms 1 and 2. Let $(X_t)_{t\in\mathbb{R}}$ be the shot-noise process given by (5.1.1) with intensity and impulse response illustrated by Figure 5.3 that are given by

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

and where the i.i.d. marks $(Y_k)_k$ follow 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 a Gaussian random variable with mean μ and variance σ^2. An example of the simulated shot-noise process is given by Figure 5.4. Figure 5.5 respectively displays the densities obtained using Algorithm 2 for several number of observations and a boxplot of the integrated squared error $\hat{f}_{n,\lambda}$ and f_λ for 100 independent runs of Algorithm 2 and three different sample sizes.
The results of Algorithm 1 on the same dataset are illustrated in Figure 5.6. The B-spline method described by Algorithm 2 achieves better performances for two reasons. The first one is related to the basis functions. Recall that the cubic B-splines spaces $\mathcal{S}_k^3$ are defined as the set of linear combinations of the basis functions $(B_j^3)_{1\leq j\leq k+3}$ defined

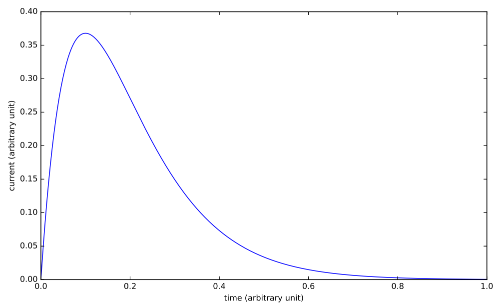


Figure 5.3 : Impulse response used in numerical applications

5

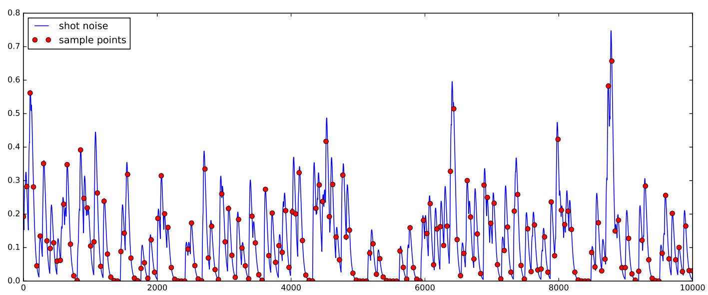


Figure 5.4 : Sample path of the simulated shot-noise process

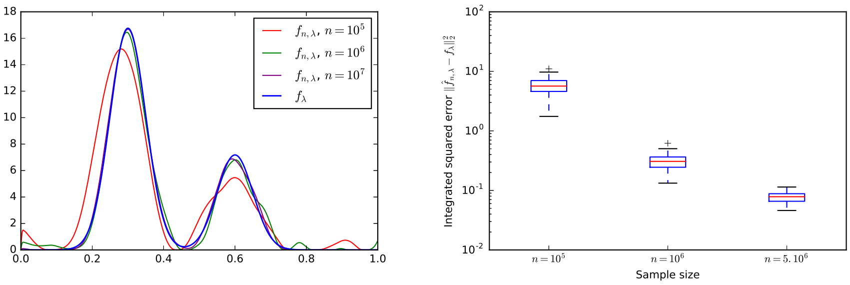


Figure 5.5 : Left: Estimation of f_λ for $n \in \{10^5, 10^6, 10^7\}$ using Algorithm 2. Right: Boxplot of the Integrated Squared Error $\|\hat{f}_{n,\lambda} - f_\lambda\|_2^2$ for $n \in \{10^5, 10^6, 5.10^6\}$ and 100 independent runs of Algorithm 2 over independent input data sets

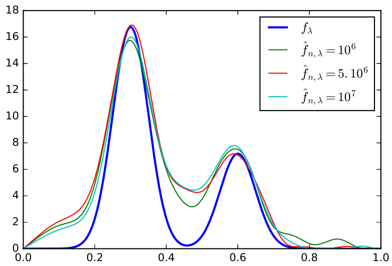


Figure 5.6 : Estimation of f_λ for $n \in \{10^5, 10^6, 10^7\}$ using Algorithm 1

by 15. In Schumaker (2007)[Theorem 4.18], it is proved that these functions are linearly independent while we can not say much about the basis functions $(\eta_1, \ldots, \eta_{2N})$ defined by Theorem 5.2.2. In addition, the analytical expression of the functions $(\eta_1, \ldots, \eta_{2N})$ is explicit but very involved since it intricately depends on the impulse response function h. The second reason is computational. The decomposition basis $(\eta_1, \ldots, \eta_{2N})$ obtained using the representer's theorem is complex-valued while we are looking to construct an estimator of the positive-valued function f_λ. The solution of the optimization problem (5.2.17) needs to satisfy the continuum of following constraint

$$\alpha^{\mathrm{T}} \eta_{\underline{u}}(x) \geq 0 \quad \text{for all } x \in [0, 1] \ . \tag{5.4.1}$$

In practice, we only consider a finite number of constraints that are given by the evaluation of (5.4.1) on the finite grid of points $\{1/N,\ 2/N, \ldots, 1\}$ where N is chosen according to Step 2 of Algorithm 1. Another choice would consists in relaxing all the constraints and taking the positive part of the real part of the solution of the problem (5.2.17). Either way, this algorithm seems less efficient than Algorithm 2 using B-splines.

5.4.2 Real datasets

In this section, we provide an application of Algorithm 2 on real datasets provided by the CEA Saclay. Ideally, we would like to estimate, as precisely as possible, the function f_λ. However, practitioners establish precise gamma ray spectra of a given radionuclide ${}^{A}_{Z}X$ where X is a chemical element, Z the number of protons and A the number of nucleons, by using a radioactive source only containing atoms ${}^{A}_{Z}X$ with a low counting rate λ in order to avoid pileup. For instance, we illustrate in Figure 5.7 a gamma-ray spectrum (for energies between 0 and 100 keV) that is obtained by computing the histogram of the local maxima of a shot-noise process whose the source is exclusively made of Americium 241. The horizontal axis is expressed in terms of channels that corresponds to energy intervals while the vertical axis, in logarithmic scale, corresponds to the number of recorded energies that lie in the associated channel. In addition, the color code has to be understood as follows: red peaks correspond to the main photoelectric peaks, blue peaks to the minor photoelectric peaks and yellow peaks to artifacts that are either escape peaks, multiple peaks of photoelectric peaks or detected photons that follow from an α or β disintegration reaction.

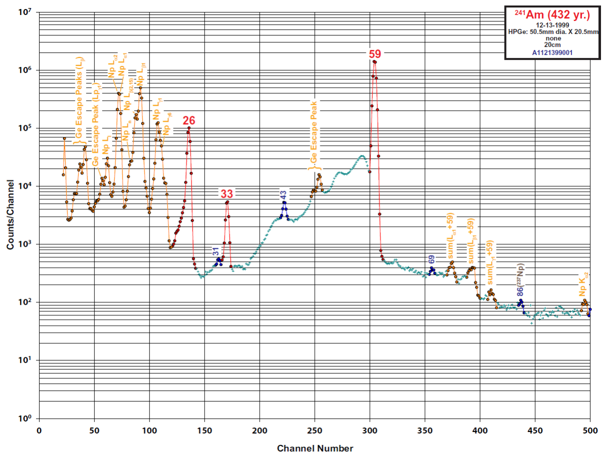


Figure 5.7 : Gamma-ray spectrum of ^{241}Am taken from the website www.radiochemistry.org

The data available from the CEA Saclay laboratory correspond to gamma radiation measurements of a mixture of Americium 241 and Plutonium 239 with an approximate counting rate of 400 000 counts per second.
In Figure 5.8, we illustrate the estimator obtained by computing the histogram of the local maxima of the measured signal described in Chapter 1. Note that this estimator exhibits multiple peaks at abscissa points 119,14 keV (2.59,57) and 178,71 keV (3.59,57). The estimator computed using Algorithm 2 is displayed in blue in Figure 5.9. This estimator is not as precise as the gamma-ray spectrum presented in Figure 5.7 as illustrated by Figure 5.10 but is able to identify several photoelectric peaks of Americium 241. Furthermore, it removes the multiple peaks that appear when using the method based

on local maxima. In conclusion, the results obtained using Algorithm 2 are encouraging. Some limitations of the performances of Algorithm 2 could come from the additional electric noise that has not been modeled as well as the assumption about the impulse response. Indeed, for theoretical reasons, we assumed that the shape of the impulse response does not depend on the incoming photon.

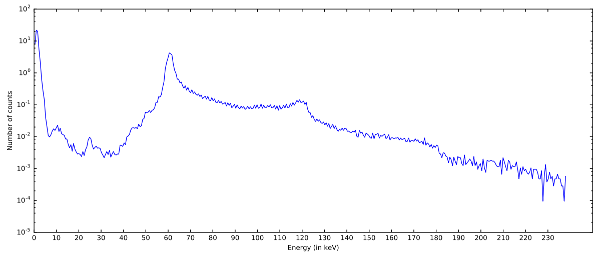


Figure 5.8 : Estimator of the marks' density using the local maxima method

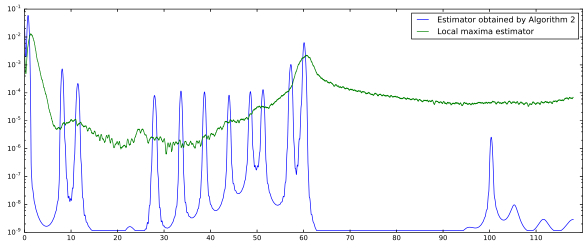


Figure 5.9 : Estimator of f_λ using Algorithm 2

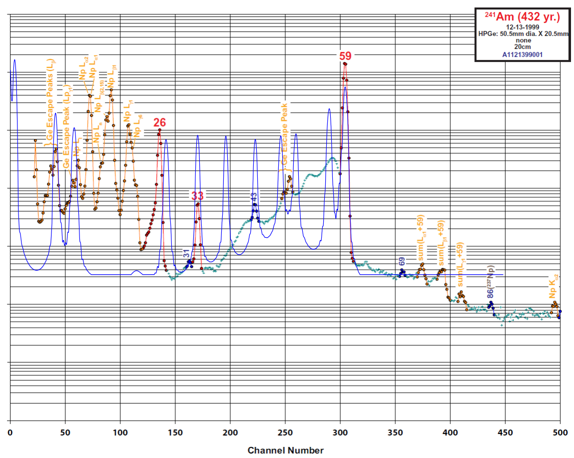


Figure 5.10 : Comparison between the gamma-ray spectrum of Figure 5.7 and the estimator of f_λ using Algorithm 2 (in blue)

5.5 Proof of Theorem 5.2.1

Proof. Let $(\mathcal{E}_n)_n$ be the stochastic processes defined for any positive integer n and all $u \in \mathbb{R}$ by

$$\mathcal{E}_n(u) := n^{-1/2} \sum_{k=1}^{n} \left(\mathrm{i}X_k \mathrm{e}^{\mathrm{i}uX_k} - \mathbb{E}\left[\mathrm{i}X_0 \mathrm{e}^{\mathrm{i}uX_0}\right] \right) = \sqrt{n}\left(\hat{\varphi}_n'(u) - \varphi'(u)\right) . \quad (5.5.1)$$

The proof of Theorem 5.2.1 is based on the following result.

Proposition 10. *Let $(X_t)_{t\in\mathbb{R}}$ be the stationary shot-noise process defined by (5.1.1)] under Assumptions (**SN**-3), (**SN**-4) and $\mathbb{E}\left[|Y_0|^4\right] < \infty$. Then, for any positive number K, we have the weak convergence*

$$\mathcal{E}_n \Rightarrow Z \quad \text{in } \mathrm{C}([-K, K]) , \quad (5.5.2)$$

where $(\mathcal{E}_n)_n$ is defined by (5.5.1) and $Z = (Z_u)_{u\in\mathbb{R}}$ is a Gaussian process with mean zero and covariance function γ defined by (5.6.4).

Proof. See Section 5.5.1.1. □

From this result, we can conclude the proof of Theorem 5.2.1 using the following steps. Let K be a positive integer. From the continuity of the mapping

$$\begin{aligned} T : (\mathrm{C}([-K, K]), \|\cdot\|_\infty) &\to \left(\mathrm{C}([-K, K])^2, \|\cdot\|_{2,\infty}\right) \\ f &\to \left(f, \left[u \to 1 + \int_0^u f(t)\mathrm{d}t\right]\right) , \end{aligned}$$

where $\|\cdot\|_{2,\infty}$ is the norm defined on $\mathrm{C}([-K, K])^2$ by

$$\|(f, g)\|_{2,\infty} := \|f\|_\infty + \|g\|_\infty , \quad f, g \in \mathrm{C}\left([-K, K]\right) ,$$

we have that

$$\sqrt{n}\left(\hat{\varphi}_n'(u) - \varphi'(u), \hat{\varphi}_n(u) - \varphi(u)\right) = T(\mathcal{E}_n) \Rightarrow \left(Z, \bar{Z}\right) \quad \text{in } \mathrm{C}\left([-K, K]\right) ,$$

where $\bar{Z}$ is defined by $\bar{Z}_u := 1 + \int_0^u Z_t \mathrm{d}t$.
Next, we apply the delta method using the mapping

$$\begin{aligned} S : (\mathrm{C}([-K, K]) \times \mathrm{C}^*([-K, K]), \|\cdot\|_{2,\infty}) &\to (\mathrm{C}([-K, K]), \|\cdot\|_\infty) \\ (f, g) &\to f/g , \end{aligned}$$

where $\mathrm{C}^\star([-K, K])$ denotes the set of functions in $\mathrm{C}([-K, K])$ valued in $\mathbb{R}^* = \mathbb{R} - \{0\}$. Let us show that S is differentiable at any $(f, g) \in \mathrm{C}([-K, K]) \times \mathrm{C}^\star([-K, K])$.For all $h_1, h_2 \in \mathrm{C}([-K, K])$ such that $\|h_2\|_\infty \le \inf_{x\in[-K,K]} |g(x)|/2$, we have

$$S(f + h_1, g + h_2) = S(f, g) + \frac{h_1 g - h_2 f}{g^2} + r(h) ,$$

with $r(h) = \frac{h_2(h_1 g - h_2 f)}{g^2}$ so that $\|r(h)\|_\infty = o\left(\|h_1\|_\infty + \|h_2\|_\infty\right)$.
Now, since $\varphi' \in \mathrm{C}([-K, K])$ and $\varphi \in \mathrm{C}^\star([-K, K])$ (see Sato (1999) for instance), the δ-method gives that

$$\sqrt{n}\left(\frac{\hat{\varphi}_n'}{\hat{\varphi}_n} - \frac{\varphi'}{\varphi}\right) \Rightarrow \frac{Z}{\varphi} - \bar{Z}\frac{\varphi'}{\varphi^2} =: \check{Z} \quad \text{in } \mathrm{C}([-K, K])\,.$$

Then, we conclude the proof using Lemma 12. □

5.5.1 Proof of Proposition 10

The proof of Proposition 10 follows the two usual steps. We first consider the convergence in the finite-dimensional sense of $\mathcal{E}_n$. Then, we show that it is tight in $(\mathrm{C}([-K, K]), \|\cdot\|_\infty)$.

5.5.1.1 Convergence of finite-dimensional laws distributions

5

We use the classical m-dependent approximation approach (see Brockwell and Davis (1991)). Let $(X_t^{(m)})_{t\in\mathbb{R}}$ and $(R_t^{(m)})_{t\in\mathbb{R}}$ be the stochastic processes defined for any positive integer m by

$$X_t^{(m)} := \sum_{T_k \le t} Y_k\ h(t - T_k)\mathbb{1}_{\{t - T_k \le m\}}\ , \tag{5.5.3}$$

and

$$R_t^{(m)} := \sum_{T_k < t} Y_k\ h(t - T_k)\mathbb{1}_{\{t - T_k \ge m\}}\ , \quad t \in \mathbb{R}\,.$$

For any positive integers n and m, let also $(\mathcal{E}_{n,m}(u))_{u\in\mathbb{R}}$ be the stochastic processes defined by

$$\mathcal{E}_{n,m}(u) := n^{-1/2}\left(\sum_{k=1}^{n} \mathrm{i}X_k^{(m)}\mathrm{e}^{\mathrm{i}uX_k^{(m)}} - \mathbb{E}\left[\mathrm{i}X_0^{(m)}\mathrm{e}^{\mathrm{i}uX_0^{(m)}}\right]\right)\ , \quad u \in \mathbb{R}\,. \tag{5.5.4}$$

Note that the sequence of random variables $\left(X_k^{(m)}(u)\right)_{k\in\mathbb{N}}$ is m-dependent. Hence, for a fixed m, Brockwell and Davis (1991)[Theorem 6.4.2.] implies that the finite-dimensional laws of $(\mathcal{E}_{n,m}(u))_{u\in\mathbb{R}}$ converge to the finite-dimensional laws of a centered Gaussian process $\left(Z_u^{(m)}\right)_{u\in\mathbb{R}}$ with covariance function γ_m given by Lemma 13. If, in addition, one can prove that, for all u, v,

$$\lim_{m\to\infty} \gamma_m(u, v) = \gamma(u, v) \tag{5.5.5}$$

and

$$\lim_{m\to\infty} \limsup_{n} \mathbb{E}\left[\left|\mathcal{E}_{n,m}(u) - \mathcal{E}_n(u)\right|^2\right] = 0\,, \tag{5.5.6}$$

then, Brockwell and Davis (1991) gives that $\mathcal{E}_n$ converges in the sense of finite-dimensional distributions to a centered Gaussian process $(Z_u)_{u\in\mathbb{R}}$ with covariance function γ.

The proof of (5.5.5) follows from Lemma 13. We now prove (5.5.6). For any real number u, we have

$$\begin{aligned}&\mathbb{E}\left[\left|\mathcal{E}_{n,m}(u)-\mathcal{E}_n(u)\right|^2\right]\\&\leq n^{-1}\sum_{k\in\mathbb{Z}}\mathrm{Cov}\left(X_0\mathrm{e}^{\mathrm{i}uX_0}-X_0^{(m)}\mathrm{e}^{\mathrm{i}uX_0^{(m)}},X_k\mathrm{e}^{\mathrm{i}uX_k}-X_k^{(m)}\mathrm{e}^{\mathrm{i}uX_k^{(m)}}\right)\ ,\end{aligned}$$

which tends to zero as $n\to\infty$ by Lemma 10.

5.5.1.2 Tightness

It remains to prove the tightness of the sequence of stochastic processes $(\mathcal{E}_n)_{n\geq 1}$ in $(\mathrm{C}([-K,K]),\|\cdot\|_\infty)$. A sufficient condition (see Kallenberg (2002)[Corollary 16.9]) is to show that

- $\mathcal{E}_n$ is tight in $\mathrm{C}([-K,K])$,
- There exists positive constants a,b,C such that

$$\sup_{n\geq 1}\ \mathbb{E}\left[|\mathcal{E}_n(s)-\mathcal{E}_n(s)|^a\right]\leq C\,|t-s|^{1+b}\ ,\quad -K\leq s,t\leq K\ .$$

Before that, let us define weighted Lipschitz seminorms.

Definition 16 (Weighted Lipschitz seminorm)**.** *Let k be a positive integer. We call $\mathcal{L}^k$ the space of complex-valued functions δ defined on the real line that satisfy $|\delta|_{\mathcal{L}^k}<\infty$ where*

$$|\delta|_{\mathcal{L}^k}:=\sup_{x,y\in\mathbb{R}}\frac{|\delta(x)-\delta(y)|}{(1+|x|^k+|y|^k)\,|x-y|}\ .$$

Then, we have the following result.

Proposition 11 (Burkhölder's inequality for shot-noise processes)**.** *Let (p,q,k) be three nonnegative integers that jointly satisfy*

$$q \text{ is even}\ ,\ p\geq 1\ ,\ (k+1)p\leq q \text{ and } \mathbb{E}\left[Y^q\right]<\infty\ . \tag{5.5.7}$$

Let δ be a complex-valued function defined on $\mathbb{R}$ such that $|\delta|_{\mathcal{L}^k}<\infty$. Then, there exists a constant $C_{p,q,k}$ only depending on (p,q,k) and (λ,C_0,α) in (SN-3)-(SN-4) such that, for every integer n, we have

$$\mathbb{E}\left[\left|\sum_{j=1}^{n}\left(\delta(X_j)-\mathbb{E}\left[\delta(X_j)\right]\right)\right|^{2p}\right]\leq C_{p,q,k}n^p\,|\delta|_{\mathcal{L}^k}^{2p}\ . \tag{5.5.8}$$

Proof. The proof is a direct application of Rio (1999)[Theorem 2.5.] using Lemma 8. □

Since $\mathcal{E}_n(0) = \sqrt{n}\sum_{k=1}^{n}(X_k - \mathbb{E}[X_1])$, Condition (i) follows immediately as soon as $\mathbb{E}\left[Y^2\right] < \infty$.
We now show that Condition (ii) holds.
For $u, v \in [-K, K]$, let $\delta_{u,v}$ be the function defined on $\mathbb{R}$ by

$$\delta_{u,v}(x) := \mathrm{i}x\mathrm{e}^{\mathrm{i}ux} - \mathrm{i}x\mathrm{e}^{\mathrm{i}vx}\,,\, x \in \mathbb{R}\,.$$

For $u, v \in [-K, K]$, one easily gets that $\left|\delta_{u,v}\right|_{\mathcal{L}^2} \le K/2|u-v|$. Let us detail the last inequality. We write for all $u \le v \in [-K, K]$ and $x, y \in \mathbb{R}$,

$$\begin{aligned}\left|\delta_{u,v}(x) - \delta_{u,v}(y)\right| &= \left|\int_u^v \left(x^2\mathrm{e}^{\mathrm{i}sx} - y^2\mathrm{e}^{\mathrm{i}sy}\right)\mathrm{d}s\right| \\ &\le \int_u^v \left|x^2\mathrm{e}^{\mathrm{i}sx} - y^2\mathrm{e}^{\mathrm{i}sy}\right|\mathrm{d}s \\ &\le K/2\left(x^2 + y^2\right)|u - v||x - y|\,,\end{aligned}$$

since $x \to x^2\mathrm{e}^{\mathrm{i}sx}$ has a derivative bounded by Kx^2 for a given s.
Applying Proposition 11 to the function $\delta_{u,v}$ with $(p, q, k) = (1, 4, 2)$ leads to

$$\begin{aligned}\mathbb{E}\left[|\mathcal{E}_n(u) - \mathcal{E}_n(v)|^2\right] &= n^{-1}\mathbb{E}\left[\left|\sum_{k=1}^{n}\left(\delta_{u,v}(X_k) - \mathbb{E}\left[\delta_{u,v}(X_k)\right]\right)\right|^2\right] \\ &\le C_{1,4,2}K^2/2|u - v|^2\,.\end{aligned}$$

We thus obtain Condition (ii) which concludes the proof of the tightness of the stochastic processes $(\mathcal{E}_n)_{n\ge1}$ in $\mathrm{C}([-K, K])$.

5.5.2 Lemmas

This section gathers technical lemmas involved in the proof of Theorem 5.2.1. For any $p \ge 1$ and any random variable X, let $\|X\|_p := \mathbb{E}\left[|X|^p\right]^{1/p}$.
In order to prove the Burkhölder's inequality in Proposition 11, we need the following lemma.

Lemma 8. *Let $N = \sum_{k\in\mathbb{Z}}\delta_{T_k,Y_k}$ be the marked point process given by Assumption (**SN**-3) and let $\mathcal{F}_0$ be the σ-algebra defined by*

$$\mathcal{F}_0 := \sigma\left(N(A \times B)\,, A \in \mathcal{B}((-\infty, 0])\ ,\ B \in \mathcal{B}(\mathbb{R})\right)\,.$$

*Let $(X_t)_{t\in\mathbb{R}}$ be the stationary shot-noise process (5.1.1) under Assumptions (**SN**-3) and (**SN**-4). Let (p, q, k) be a triplet of nonnegative integers such that (5.5.7) holds.*
Then, we have for all $s > 0$ and $\delta \in \mathcal{L}^k$ (see Definition 16)

$$\|\mathbb{E}\left[\delta(X_s) \mid \mathcal{F}_0\right] - \mathbb{E}\left[\delta(X_s)\right]\|_p \le K_{p,q,k}\,|\delta|_{\mathcal{L}^k}\,\mathrm{e}^{-\alpha s}\,,$$

*where $K_{p,q,k}$ is a constant only depending on (p, q, k) and (λ, C_0, α) in Assumptions (**SN**-3)-(**SN**-4).*

Proof. For any positive s, let $X_s^{<0}$ and $X_s^{>0}$ be the random variables defined by

$$X_s^{<0} := \sum_{T_k \leq 0} Y_k \, h(s - T_k) \quad \text{and} \quad X_s^{>0} := \sum_{T_k > 0} Y_k \, h(s - T_k) \, ,$$

so that $X_s = X_s^{<0} + X_s^{>0}$. Set, for any positive number s,

$$\bar{\delta}(X_s) := \delta(X_s) - \mathbb{E}\left[\delta(X_s)\right] \, .$$

Then, we have

$$\left|\mathbb{E}\left[\bar{\delta}(X_s) \mid \mathcal{F}_0\right]\right| \leq \mathbb{E}\left[\left|\delta(X_s) - \delta(X_s^{>0})\right| \mid \mathcal{F}_0\right] + \mathbb{E}\left[\left|\delta(X_s) - \delta(X_s^{>0})\right|\right] \, .$$

This with Jensen's inequality and $\delta \in \mathcal{L}^k$ yields that

$$\begin{aligned}
\|\mathbb{E}\left[\bar{\delta}(X_s) \mid \mathcal{F}_0\right]\|_p &\leq 2\|\delta(X_s) - \delta(X_s^{>0})\|_p \\
&\leq 2\,|\delta|_{\mathcal{L}^k}\, \mathbb{E}\left[\left(1 + |X_s|^k + \left|X_s^{>0}\right|^k\right)^p \left|X_s^{<0}\right|^p\right]^{1/p} \\
&\leq 2.3^{p-1}\,|\delta|_{\mathcal{L}^k}\, \mathbb{E}\left[\left(1 + |X_s|^{kp} + \left|X_s^{>0}\right|^{kp}\right) \left|X_s^{<0}\right|^p\right]^{1/p} \\
&\leq 2.3^{p-1}\,|\delta|_{\mathcal{L}^k}\left(\|X_s^{<0}\|_p + \|X_s^{<0} X_s^k\|_p + \|X_s^{<0}(X_s^{>0})^k\|_p\right) \, .
\end{aligned}$$

Then, using Hölder's inequality, we have that

$$\|X_s^{<0} X_s^k\|_p \leq \|X_s^{<0}\|_{(k+1)p} \|X_s\|_{(k+1)p}^k$$

and

$$\|X_s^{<0}(X_s^{>0})^k\|_p \leq \|X_s^{<0}\|_{(k+1)p} \|X_s^{>0}\|_{(k+1)p}^k \, .$$

From Jensen's inequality and Lemma 9, we have that

$$\|X_s^{<0}\|_p \leq \|X_s^{<0}\|_{(k+1)p} \leq D_k C_0 (1 + \lambda)(1 + 1/\alpha) \|Y_0\|_{(k+1)p} \mathrm{e}^{-\alpha s}$$

and

$$\|X_s^{>0}\|_{(k+1)p} \leq \|X_s\|_{(k+1)p} \leq D_k C_0 (1 + \lambda)(1 + 1/\alpha) \|Y_0\|_{(k+1)p} \, .$$

Hence, the proof is concluded by setting

$$K_{p,q,k} = 2.3^{p-1} D_k C_0 (1 + \lambda)(1 + 1/\alpha) \|Y_0\|_q \, .$$

□

5.6 AUXILIARY RESULTS

Lemma 9 (Upper bounds of the moments of a shot-noise process)**.** *Let $(X_t)_{t\in\mathbb{R}}$ be the stationary shot-noise process given by (5.1.1) under Assumptions (SN-3) and (SN-4). For any positive s, let $X_s^{<0}$ denote the random variable*

$$X_s^{<0} := \sum_{T_k \le 0} Y_k \, h(s - T_k) ,$$

and $X_0^{(m)}$ is defined by (5.5.3).
Let k be a positive integer. If $\mathbb{E}\left[|Y_0|^k\right] < \infty$, then there exists a constant D_k only depending on k such that, for every positive integer m and real number s, we have

(i) $\|X_0\|_k \le D_k C_0 (1+\lambda)^k (1+1/\alpha)^k \|Y_0\|_k$,

(ii) $\|X_0 - X_0^{(m)}\|_k \le D_k C_0 (1+\lambda)^k (1+1/\alpha)^k \|Y_0\|_k \mathrm{e}^{-\alpha m}$,

(iii) $\|X_t^{<0}\|_k \le D_k C_0 (1+\lambda)^k (1+1/\alpha)^k \|Y_0\|_k \mathrm{e}^{-\alpha t}$.

where λ, C_0 and α are the constants appearing in Assumptions (SN-3)-(SN-4).

Proof. The cumulant of order k of the shot-noise marginal X_0 (see for instance Daley and Vere-Jones (1988)) is given by

$$\kappa_k = \lambda \mathbb{E}\left[|Y_0|^k\right] \int_0^\infty h^k(u)\mathrm{d}u \le \frac{\lambda}{k\alpha} \mathbb{E}\left[|Y_0|^k\right] C^k .$$

Then, (i) follows by the usual relation between moments and cumulants (see e.g. citer Rosenblatt) and

$$\left| \int_0^\infty h^j(t)\mathrm{d}t \right| \le \frac{C_0^j}{\alpha} \qquad j \ge 1 .$$

The proof of (ii) is similar noting that

$$\left| \int_m^\infty h^j(t)\mathrm{d}t \right| \le \frac{C_0^j}{\alpha} \mathrm{e}^{-\alpha m} \qquad j \ge 1 .$$

Finally, (iii) is the same as (ii) since $X_t^{<0}$ has the same law as $X_0 - X_0^{(t)}$.

□

Lemma 10. *Let $(X_t)_{t\in\mathbb{R}}$ be the stationary shot-noise process defined by (5.1.1)] under Assumptions (SN-3) and (SN-4) such that $\mathbb{E}\left[Y_0^4\right] < \infty$. For any integer m, let $(X_t^{(m)})_{t\in\mathbb{R}}$ be the truncated shot-noise process given by (5.5.3). Then, we have*

$$\lim_{m\to\infty} \sum_{k\in\mathbb{N}} \left| Cov\left(X_0 \mathrm{e}^{\mathrm{i}uX_0} - X_0^{(m)} \mathrm{e}^{\mathrm{i}uX_0^{(m)}}, X_k \mathrm{e}^{\mathrm{i}uX_k} - X_k^{(m)} \mathrm{e}^{\mathrm{i}uX_k^{(m)}} \right) \right| = 0 .$$

Proof. Let u be a real number and k, m two nonnegative numbers. We have

$$\begin{aligned}
&\left|\mathrm{Cov}\left(X_0 e^{iuX_0} - X_0^{(m)} e^{iuX_0^{(m)}}, X_k^{(m)} e^{iuX_k^{(m)}}\right)\right| \\
&\leq \left\|X_0 e^{iuX_0} - X_0^{(m)} e^{iuX_0^{(m)}}\right\|_2 \cdot \left\|\mathbb{E}\left[X_k^{(m)} e^{iuX_k^{(m)}} - \mathbb{E}\left[X_k^{(m)} e^{iuX_k^{(m)}}\right] \mid \mathcal{F}_0\right]\right\|_2 \\
&\leq \left(\|X_0\|_2 + \|X_0^{(m)}\|_2\right) C_{2,2,C_0,\alpha,\lambda} e^{-\alpha k} \qquad (5.6.1) \\
&\leq \frac{C_{2,2,C_0,\alpha,\lambda} C_0 \sqrt{\lambda}}{\sqrt{\alpha}} e^{-\alpha k} ,
\end{aligned}$$

where the inequality (5.6.1) is an application of Lemma 8 with $(p, q, k) = (2, 4, 1)$ and $\delta = \delta_u$. Then, by dominated convergence, it is sufficient to prove that, for every nonnegative integer, we have

$$\lim_{m\to\infty} \mathrm{Cov}\left(X_0 e^{iuX_0} - X_0^{(m)} e^{iuX_0^{(m)}}, X_k e^{iuX_k} - X_k^{(m)} e^{iuX_k^{(m)}}\right) = 0 .$$

Since

$$\begin{aligned}
&\left|\mathrm{Cov}\left(X_0 e^{iuX_0} - X_0^{(m)} e^{iuX_0^{(m)}}, X_k e^{iuX_k} - X_k^{(m)} e^{iuX_k^{(m)}}\right)\right| \\
&\leq \|X_0 e^{iuX_0} - X_0^{(m)} e^{iuX_0^{(m)}}\|_2 \cdot \|X_k e^{iuX_k} - X_k^{(m)} e^{iuX_k^{(m)}}\|_2 ,
\end{aligned}$$

and

$$\begin{aligned}
&\|X_k e^{iuX_k} - X_k^{(m)} e^{iuX_k^{(m)}}\|_2 \leq \|X_k - X_k^{(m)}\|_2 + \|X_k^{(m)}\left(e^{iuX_k^m} - e^{iuX_k}\right)\|_2 \\
&\leq \|X_k - X_k^{(m)}\|_2 + |u| \|X_k^{(m)}\|_4 \|X_k - X_k^{(m)}\|_4 ,
\end{aligned}$$

the result follows by Lemma 9. □

Lemma 11. *Assume that the stationary shot-noise process $(X_t)_{t\in\mathbb{R}}$ is given by (5.1.1) under Assumptions* (**SN**-3) *and* (**SN**-4) *and satisfies $\mathbb{E}\left[Y_0^2\right] < \infty$. Then, for all real u, we have*

$$\left|\varphi_{X_0}(u)\right| \geq \exp\left(-\frac{\lambda \mathbb{E}\left[Y_0^2\right] \|h\|_2^2}{2} u^2\right) . \qquad (5.6.2)$$

Proof. For every $u \in \mathbb{R}$, we have

$$\begin{aligned}
\left|\varphi_{X_0}(u)\right| &= \exp\left(\mathrm{Re}\left(\lambda \int_0^\infty \int_0^\infty \left(e^{iuxh(s)} - 1\right) f(x) dx ds\right)\right) \\
&= \exp\left(\lambda \int_0^\infty \int_0^\infty (\cos(uxh(s)) - 1) f(x) dx ds\right) \\
&\geq \exp\left(-\lambda \int_0^\infty \int_0^\infty u^2 x^2 h^2(s) f(x) dx ds\right) \\
&= \exp\left(-\frac{\lambda \mathbb{E}\left[Y_0^2\right] \|h\|_2^2}{2} u^2\right) .
\end{aligned}$$

□

Lemma 12. *Let $(X_t)_{t\in\mathbb{R}}$ be the stationary shot-noise process defined by (5.1.1) under Assumptions (SN-3) and (SN-4) such that $\mathbb{E}\left[Y_0^4\right] < \infty$. For any nonnegative integer n, let also $\hat{\sigma}_n^2$ and $\hat{\varphi}_n$ respectively denote the empirical variance and the empirical characteristic function associated to the observations $X_1, \dots, X_n$.*
Then, for every positive number K, we have

$$\left(\mathbb{1}_{\{|\hat{\varphi}_n(u)|\geq \frac{1}{2}\mathrm{e}^{-\hat{\sigma}_n^2 u^2/2}\}}\right)_{u\in[-K,K]} \Rightarrow 1 \quad \text{in } \mathrm{C}([-K,K]) \, .$$

Proof. To prove this lemma, note that it is sufficient to show that

$$\lim_{n\to\infty} \mathbb{P}\left(\exists u \in [-K,K] \; : \; |\hat{\varphi}_n(u)| \leq \mathrm{e}^{-\hat{\sigma}_n^2 u^2/2}/2\right) = 0 \, .$$

Observe now that

$$\begin{aligned}
&\mathbb{P}\left(\exists u \in [-K,K] \; : \; |\hat{\varphi}_n(u)| \leq \frac{1}{2}\mathrm{e}^{-\hat{\sigma}_n^2 u^2/2}\right) \\
&\mathbb{P}\left(\inf_{u\in[-K,K]} |\hat{\varphi}_n(u)| \leq \frac{1}{2}\mathrm{e}^{-\hat{\sigma}_n^2 K^2/2}\right) \\
&\leq \mathbb{P}\left(\sup_{u\in[-K,K]} |\hat{\varphi}_n(u) - \varphi(u)| \geq \inf_{u\in[-K,K]} |\varphi(u)| - \frac{1}{2}\mathrm{e}^{-\hat{\sigma}_n^2 K^2/2}\right) \, .
\end{aligned}$$

Let ϵ be a positive number such that $\epsilon < 2\log(2)/K^2$. By Lemma 11, we have

$$\inf_{u\in[-K,K]} |\varphi(u)| \geq \mathrm{e}^{-\sigma^2 K^2/2} \, .$$

It follows that

$$\begin{aligned}
&\mathbb{P}\left(\sup_{u\in[-K,K]} |\hat{\varphi}_n(u) - \varphi(u)| \geq \inf_{u\in[-K,K]} |\varphi(u)| - \frac{1}{2}\mathrm{e}^{-\hat{\sigma}_n^2 K^2/2}\right) \\
&\leq \mathbb{P}\left(\sup_{u\in[-K,K]} |\hat{\varphi}_n(u) - \varphi(u)| \geq \mathrm{e}^{-\sigma^2 K^2/2} - \frac{1}{2}\mathrm{e}^{-\hat{\sigma}_n^2 K^2/2} \, , \; \left|\hat{\sigma}_n^2 - \sigma^2\right| \leq \epsilon\right) \\
&+ \mathbb{P}\left(\left|\hat{\sigma}_n^2 - \sigma^2\right| > \epsilon\right) \, .
\end{aligned}$$

The first term in the right-hand side of the above inequality can be bounded by

$$\mathbb{P}\left(\sup_{u\in[-K,K]} |\hat{\varphi}_n(u) - \varphi(u)| \geq \mathrm{e}^{-\sigma^2 K^2/2}\left(1 - \mathrm{e}^{\epsilon K^2/2}/2\right)\right) \, ,$$

which tends to zero when $n \to \infty$ as a consequence of Proposition 10. Then, the proof is concluded since, by Proposition 11, we have $\hat{\sigma}_n^2 \xrightarrow{\mathbb{P}} \sigma^2$. □

Lemma 13. *Let $(X_t)_{t\in\mathbb{R}}$ be the stationary shot-noise process defined by (5.1.1) under Assumptions (SN-3) and (SN-4) such that $\mathbb{E}\left[Y_0^4\right] < \infty$.*

For any nonnegative integers k and m, let $c_k^{(m)}$ *and* c_k *be the functions respectively defined on* $\mathbb{R}^2$ *by*

$$c_k^{(m)}(u,v) := Cov\left(\mathrm{e}^{\mathrm{i}uX_0^{(m)}}\ ,\ \mathrm{e}^{\mathrm{i}vX_k^{(m)}}\right) \ \text{ and } \ c_k(u,v) := Cov\left(\mathrm{e}^{\mathrm{i}uX_0}\ ,\ \mathrm{e}^{\mathrm{i}vX_k}\right)\ .$$

Then, the covariance function γ_m *of the stochastic processes* $(\mathcal{E}_{n,m})_{n\geq 1}$ *given by* (5.5.4) *for any positive m is given by*

$$\gamma_m(u,v) = \sum_{k=-m}^{m} \partial_u\partial_v c_k^{(m)}(u,v)\ . \tag{5.6.3}$$

and the covariance of the Gaussian process Z is given by

$$\gamma_m(u,v) = \sum_{k=-m}^{m} \partial_u\partial_v c_k(u,v)\ . \tag{5.6.4}$$

In addition, the Gaussian process $\bar{\bar{Z}} := (Z_u, \bar{Z}_u)_{u\in\mathbb{R}}$ *defined in Proposition 10 has a covariance matrix* Σ *given by*

$$\Sigma(u,v) = Cov\left(\bar{\bar{Z}}(u), \bar{\bar{Z}}(v)\right) = \begin{bmatrix} \partial_u\partial_v C(u,v) & \partial_u C(u,v) \\ \partial_v C(u,v) & C(u,v)\end{bmatrix}\ , \quad u,v\in\mathbb{R}\ ,$$

where $C(u,v) := \sum_{k\in\mathbb{Z}} c_k(u,v)$ *for every real numbers* u,v.

Proof. Let k,m be nonnegative integers and $\epsilon_1,\epsilon_2\in\{0,1\}$. Since $\mathbb{E}\left[Y_0^4\right] < \infty$, we have from Fubini-Tonelli theorem's that

$$\partial_u^{\epsilon_1}\partial_v^{\epsilon_2} c_k^{(m)}(u,v) = \mathrm{Cov}\left(\left(X_0^{(m)}\right)^{\epsilon_1}\mathrm{e}^{\mathrm{i}uX_0^{(m)}}, \left(X_k^{(m)}\right)^{\epsilon_2}\mathrm{e}^{\mathrm{i}vX_k^{(m)}}\right)\ .$$

From Lemma 8, we get

$$\begin{aligned}
&\left|\partial_u^{\epsilon_1}\partial_v^{\epsilon_2} c_k^{(m)}(u,v)\right| \\
&\leq \left\|\left(X_0^{(m)}\right)^{\epsilon_1}\mathrm{e}^{\mathrm{i}uX_0^{(m)}} - \mathbb{E}\left[\left(X_0^{(m)}\right)^{\epsilon_1}\mathrm{e}^{\mathrm{i}uX_0^{(m)}}\right]\right\|_2 \\
&\qquad \cdot \left\|\mathbb{E}\left[\left(X_k^{(m)}\right)^{\epsilon_2}\mathrm{e}^{\mathrm{i}vX_k^{(m)}} - \mathbb{E}\left[\left(X_k^{(m)}\right)^{\epsilon_2}\mathrm{e}^{\mathrm{i}vX_k^{(m)}}\right] \mid \mathcal{F}_0\right]\right\|_2 \\
&\leq 2\|X_0^{\epsilon_1}\|_2 K_{1,2,1,C_0,\alpha,\lambda}(1+|v|)\mathrm{e}^{-\alpha k}\ .
\end{aligned}$$

If one can show that there exists some number E only depending on C_0, λ and K such that

$$\sup_{k\geq 1}\ \sup_{u,v\in[-K,K]} \left|c_k^{(m)}(u,v) - c_k(u,v)\right| \leq E\mathrm{e}^{-\alpha m}\ , \tag{5.6.5}$$

then, by the dominated convergence theorem, we have

$$\lim_{m\to\infty}\sum_{k=-m}^{m}\partial_u^{\epsilon_1}\partial_v^{\epsilon_2}c_k^{(m)}(u,v) = \sum_{k\in\mathbb{Z}}\partial_u^{\epsilon_1}\partial_v^{\epsilon_2}c_k(u,v)\ , \quad u,v\in\mathbb{R}\ ,$$

so that

$$\lim_{m\to\infty}\gamma_m(u,v) = \lim_{n\to\infty}\ \sqrt{n}\mathrm{Cov}\left(\hat{\varphi}_n'(u), \hat{\varphi}_n'(v)\right) =: \gamma(u,v)\ .$$

The proof of (5.6.5) is similar to the one of Lemma 10. By Fubini-Tonelli, it follows that for all u,v, we have $\gamma(u,v) = \partial_u\partial_v C(u,v)$ and the proof is concluded. □

5.7 Proof of Theorem 5.2.2

Proof. Let $\mathcal{G}_{2N}$ be the closure in $\mathcal{W}^2$ of all the linear combination of functions $\eta_1, \ldots, \eta_{2N}$ and $\mathcal{G}_{2N}^{\perp}$ its orthogonal complement. For any function $\xi \in \mathcal{W}^2$, one can find complex numbers $a_1, \ldots, a_{2N}$ such that

$$\xi - \xi_a := \xi - \sum_{k=1}^{2N} a_k \eta_k \in \mathcal{G}_{2N}^{\perp} \ ,$$

which implies that

$$K_{\underline{u}}[\xi] = K_{\underline{u}}[\xi_a] + K_{\underline{u}}[\xi - \xi_a] = K_{\underline{u}}[\xi_a] \ .$$

Since we have, by Pythagorean's theorem, that

$$\|\xi\|_{\mathcal{W}^2}^2 = \|\xi_a\|_{\mathcal{W}^2}^2 + \|\xi - \xi_a\|_{\mathcal{W}^2}^2 \ ,$$

it follows that

$$\|\hat{\Omega}_n^{1/2} \cdot \{\hat{g}_n(\underline{u}) - K_{\underline{u}}[\xi]\}\|_2^2 + \mu\|\xi\|_{\mathcal{W}^2}^2 \geq \|\hat{\Omega}_n^{1/2} \cdot \{\hat{g}_n(\underline{u}) - K_{\underline{u}}[\xi_a]\}\|_2^2 + \mu\|\xi_a\|_{\mathcal{W}^2}^2 \ ,$$

and the proof is concluded. □

5

5.8 Compacity of the operator K_h

In Section 5.2, we claim that it is difficult to derive necessary and sufficient conditions on the function h such that the operator K_h defined by (5.2.3) is compact as a mapping from $L^2([0,1], \mathbb{R})$ to $L^2(\mathbb{R}, \mathbb{C})$. The following proposition provides a sufficient condition.

Proposition 12. *Suppose that h is a causal function satisfying Assumption (**SN**-4). Let k_h be the function defined on $\mathbb{R} \times [0,1]$ by (5.2.6). Let K_h be the operator defined by*

$$K_h[f](t) := \int_0^1 k_h(t,x) f(x) \mathrm{d}x \ , \quad \text{for all } f \in L^2([0,1]) \text{ and } t \in \mathbb{R} \ .$$

If k_h satisfies

$$\int_0^1 \int_0^\infty |k_h(t,x)|^2 \, \mathrm{d}t \mathrm{d}t < \infty \ ,$$

then K_h is a compact operator from $L^2([0,1])$ to $L^2(\mathbb{R}, \mathbb{C})$.

Proof. By Young (1988)[Theorem 8.8], K_h is Hilbert-Schmidt hence compact operator. □

The following lemma provides a sufficient condition on the impulse response function h that ensures that the kernel k_h is square integrable.

Lemma 14. *Suppose that $h : \mathbb{R}_+ \to \mathbb{R}$ is a function that satisfies, for some $\alpha > 0$,*

$$D_\alpha := \sup_{u \in \mathbb{R}} (1 + |u|)^\alpha \left| \int_{\mathbb{R}} h(s) \mathrm{e}^{\mathrm{i}uh(s)} \mathrm{d}s \right| < \infty \,. \tag{5.8.1}$$

Then, as $t \to \infty$, we have

$$\sup_{f \,:\, \|f\|_{L^2} \leq 1} |K_h[f](t)| = O(|t|^{-\beta}) \,,$$

for some $\beta > 1$ and K_h denotes the operator defined by (5.2.7).

Proof. We have for all real t that $|K_h[f](t)| \leq \left(\int_0^1 |k_h(t, x)|^2 \, \mathrm{d}s \right)^{1/2}$ where k_h is defined in (5.2.6). Observe now that

$$\begin{aligned}
\int_0^1 |k_h(t, x)|^2 \, \mathrm{d}x &\leq \int_0^1 |x|^2 \cdot \left| \int_0^\infty h(s) \mathrm{e}^{\mathrm{i}th(s)} \mathrm{d}s \right|^2 \mathrm{d}x \\
&\leq D_\alpha \int_0^1 |x|^2 \, (1 + |tx|)^{-2\alpha} \, \mathrm{d}x \\
&\leq D_\alpha \left(\int_0^t (1 + |u|)^{-2\alpha} \mathrm{d}u \right) t^{-3} \\
&\underset{t \to \infty}{=} O(t^{-2\beta})
\end{aligned}$$

with

$$\begin{cases} 2\beta = \begin{cases} 3 & \text{if } \alpha > 1/2 \\ 2\alpha + 2 & \text{if } \alpha < 1/2 \end{cases} \\ 2\beta \in (0, 3) \qquad\quad \text{if } \alpha = 1/2 \end{cases}$$

In all cases we can choose $\beta > 1$. □

We give a particular case for checking (5.8.1). If $h \in L^1(\mathbb{R}_+, \mathbb{R})$ is a continuously differentiable function and satisfies $h'(t) \neq 0$ for all positive t, then

$$\int_0^\infty h(s) \mathrm{e}^{\mathrm{i}uh(s)} \mathrm{d}s = \int_0^{\|h\|_\infty} y \mathrm{e}^{\mathrm{i}uy} \frac{\mathrm{d}y}{\left| h' \circ h^{-1}(y) \right|} \,.$$

Defining $H(y) := \frac{y}{|h' \circ h^{-1}(y)|} \mathbb{1}_{[0, \|h\|_\infty]}(y)$, we only need to prove that

$$\hat{H}(u) := \int_{\mathbb{R}} H(y) \mathrm{e}^{\mathrm{i}uy} \mathrm{d}y \underset{|u| \to \infty}{=} O\left(|u|^{-\alpha}\right)$$

for some positive α.

Remark 10. If h is C^2 on $[0, \infty)$ and compactly supported, the function H is continuous on $\mathbb{R}$, supported on $[0, \|h\|_\infty]$ and C^1 on $[0, \|h\|_\infty)$. Then, $\left| \hat{H}(u) \right| \underset{|u| \to \infty}{=} O(|u|^{-1})$.

5

6 Conclusion and perspectives

Conclusion

In this thesis, we have developed new statistical methods in order to estimate the elements of shot-noise processes and we have seen that the inference problem of shot-noise processes is related to the estimation in a C^1-sense of the characteristic function of the marginal of the shot-noise process. Indeed, we have shown that there exists a linear operator that links the function $\varphi'_{X_0}/\varphi_{X_0}$ to the function f_λ and allow us to construct estimators of the flow function f_λ and, consequently, of the intensity λ and the marks' density f. Moreover, the estimators constructed in Chapters 4 and 5 can be used whatever the intensity λ of the shot-noise process but their associated rates of convergence worsen for high counting rates which is related to the presence of pileup. Moreover, it is important to note that the results presented for simulated datasets perform well in presence of pileup, where the actual methods developed in the statistical literature fail to work.

Perspectives

Even if the estimator introduced in Chapter 2 well retrieves most of the photoelectric peaks of the underlying radioactive source – and thus allows to identify the radionuclide present in the source – the performances of our algorithm on the real datasets are not as precise as the gamma-ray spectra obtained by spectroscopists at low count rates. We present some suggestions for improving the algorithms as well as the theoretical studies tackled in this manuscript.

Rates of convergence

In this manuscript, we provide upper bounds rates of convergence for the exponential shot-noise processes in Chapter 4. However, we have not shown that these polynomial rates of convergence are optimal in the minimax sense. A lower bound on the minimax risk could be studied in order to determine if our estimator achieves the best attainable rate of convergence.

In Chapter 5, no rates of convergence are provided because the principal purpose of this chapter consists in constructing an efficient estimation algorithm to infer the statistical

elements of the shot-noise processes. Nonetheless, we believe that a more theoretical treatment could be done in at least two different ways.
Firstly, instead of considering a finite number of equations defined by (5.2.11), we could consider, for some increasing sequence of positive numbers $(K_n)_n$, all the equations of the form

$$\hat{g}_n(u) = K_h[f_\lambda](u) + \epsilon_n(u) \;, \quad u \in [-K_n, K_n] \;.$$

Additionally, the data attachment term would be replaced by

$$F(\xi, \hat{g}_n) = \int_{-K_n}^{K_n} \int_{-K_n}^{K_n} \hat{C}_n^{-1}(u, v)\, (\hat{g}_n(u) - K_h[\xi](u))\, \overline{(\hat{g}_n(v) - K_h[\xi](v))} \mathrm{d}u\mathrm{d}v \;,$$

where $\hat{C}_n$ is a positive-definite estimator of the covariance operator C given by Theorem 5.2.1. A similar problem is tackled in Carrasco and Florens (2002) and discussed in Yu (2004).
The second method is suggested by Cardot (2002) which tackles a similar linear inverse problem and obtains optimal rates of convergence under the assumption that the operator is compact and satisfies additional restrictive properties. It would be interesting to study if these results can be extended to non-compact operators using the methods introduced in Mair and Ruymgaart (1996) that rely on to the spectral theorem of self-adjoint operators taking values in a Hilbert space.

A more general class of shot-noise processes

In this manuscript, we only consider shot-noise processes with a deterministic impulse response h so that each gamma particle that interacts with the detector generates the same kind of electric current multiplied by the incident photon energy. It could be interesting to soften this assumption and allow each particle to generate several forms of electric current as done in Trigano et al. (2015a). Indeed, the algorithm 2 introduced in Chapter 5 assumes that the impulse response is known so that one can evaluate the action of the operator K_h. In real applications of γ-spectroscopy, the impulse response is estimated when calibrating the multichannel analyzer and this step can induce errors.

Optimal choice of the impulse response

Experts in γ-spectroscopy instrumentation systems can choose a several number of impulse responses by tuning the preamplifier and the amplifier options. A recurrent question that arises is the following: which impulse response could be chosen in order to construct the best statistical estimators of the intensity and the marks density function from the discrete observations of the associated shot-noise process? This remains an open-ended question and suggests to derive rates of convergence that would explicitly depend on the kind of impulse responses as in Chapter 4.

Bibliography

Bibliography

Andrieu, C., Barat, É., and Doucet, A. (2001). Bayesian deconvolution of noisy filtered point processes. *Signal Processing, IEEE Transactions on*, 49(1):134–146. 12, 21, 38, 46

Applebaum, D. (2009). *Lévy processes and stochastic calculus*. Cambridge university press. 64, 69

Bar-David, I. and Nemirovsky, A. (1972). Level crossings of nondifferentiable shot processes. *Information Theory, IEEE Transactions on*, 18(1):27–34. 56

Barakat, C., Thiran, P., Iannaccone, G., Diot, C., and Owezarski, P. (2003). Modeling internet backbone traffic at the flow level. *Signal Processing, IEEE Transactions on*, 51(8):2111–2124. 79, 102

Barndorff-Nielsen, O. E., Benth, F. E., and Veraart, A. (2010). Modelling energy spot prices by lévy semistationary processes. *CREATES Research Paper*, 18. 74

Bennedsen, M., Lunde, A., and Pakkanen, M. S. (2014). Discretization of lévy semistationary processes with application to estimation. *arXiv preprint arXiv:1407.2754*. 75

Benth, F. E., Eyjolfsson, H., and Veraart, A. E. (2014). Approximating lévy semistationary processes via fourier methods in the context of power markets. *SIAM Journal on Financial Mathematics*, 5(1):71–98. 75

Biermé, H., Desolneux, A., et al. (2012). Crossings of smooth shot noise processes. *The Annals of Applied Probability*, 22(6):2240–2281. 56

Bondesson, L. (1992). Shot-noise distributions. In *Generalized Gamma Convolutions and Related Classes of Distributions and Densities*, pages 152–159. Springer. 64, 80

Brockwell, P. and Davis, R. (1991). Time series: theory and methods. *Springer series in statistics*. 75, 124

Brockwell, P. J. (2001). Lévy-driven carma processes. *Annals of the Institute of Statistical Mathematics*, 53(1):113–124. 75

Brockwell, P. J., Davis, R. A., and Yang, Y. (2007). Estimation for nonnegative lévy-driven ornstein-uhlenbeck processes. *Journal of Applied Probability*, pages 977–989. 103

Brockwell, P. J., Ferrazzano, V., and Klüppelberg, C. (2013). High-frequency sampling and kernel estimation for continuous-time moving average processes. *Journal of Time Series Analysis*, 34(3):385–404. 76

Brockwell, P. J. and Schlemm, E. (2013a). Parametric estimation of the driving lévy process of multivariate carma processes from discrete observations. *Journal of Multivariate Analysis*, 115:217–251. 75, 103

Brockwell, P. J. and Schlemm, E. (2013b). Parametric estimation of the driving lévy process of multivariate carma processes from discrete observations. *Journal of Multivariate Analysis*, 115:217–251. 76, 79

Campbell, N. (1909a). Discontinuities in light emission. In *Proc. Cambridge Phil. Soc*, volume 15, page 3. 53, 102

Campbell, N. (1909b). The study of discontinuous phenomena. In *Proc. Camb. Phil. Soc*, volume 15, page 310. 53, 102

Cardot, H. (2002). Spatially adaptive splines for statistical linear inverse problems. *Journal of Multivariate Analysis*, 81(1):100–119. 136

Carrasco, M., Chernov, M., Ghysels, E., and Florens, J.-P. (2002). Efficient estimation of jump diffusions and general dynamic models with a continuum of moment conditions. *Available at SSRN 338961*. 112

Carrasco, M. and Florens, J.-P. (2002). Efficient gmm estimation using the empirical characteristic function. *IDEI Working Paper*, 140. 136

Chaplin, V., Bhat, N., Briggs, M. S., and Connaughton, V. (2013). 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. 102

Claps, P., Giordano, A., and Laio, F. (2005). Advances in shot noise modeling of daily streamflows. *Advances in water resources*, 28(9):992–1000. 79, 102

Cohen, S. and Lindner, A. (2013). A central limit theorem for the sample autocorrelations of a lévy driven continuous time moving average process. *Journal of Statistical Planning and Inference*, 143(8):1295–1306. 74

Craven, P. and Wahba, G. (1978). Smoothing noisy data with spline functions. *Numerische Mathematik*, 31(4):377–403. 116

Daley, D. (1971). The definition of a multi-dimensional generalization of shot noise. *Journal of Applied Probability*, pages 128–135. 55

Daley, D. J. and Vere-Jones, D. (1988). *An introduction to the theory of point processes*, volume 2. Springer. 53, 54, 60, 61, 83, 105, 128

Davydov, Y. A. (1973). Mixing conditions for markov chains. *Teoriya Veroyatnostei i ee Primeneniya*, 18(2):321–338. 91

De Boor, C. (1978). A practical guide to splines. *Mathematics of Computation*. 104, 111

Doukhan, P., Massart, P., and Rio, E. (1995). Invariance principles for absolutely regular empirical processes. In *Annales de l'IHP Probabilités et statistiques*, volume 31, pages 393–427. Elsevier. 87, 92

Efron, B. (1979). Bootstrap methods: another look at the jackknife. *The Annals of Statistics*, pages 1–26. 114

Elter, Z., Bakkali, M., Jammes, C., and Pázsit, I. (2016). Performance of higher order campbell methods, part i: review and numerical convergence study. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 821:66–72. 18, 42, 66

Fan, J. (1991). On the optimal rates of convergence for nonparametric deconvolution problems. *The Annals of Statistics*, pages 1257–1272. 74

Fasen, V. (2005). Extremes of regularly varying lévy-driven mixed moving average processes. *Advances in applied probability*, pages 993–1014. 74

Fasen, V. (2006). Extremes of subexponential lévy driven moving average processes. *Stochastic processes and their applications*, 116(7):1066–1087. 74

García, I., Klüppelberg, C., and Müller, G. (2011). Estimation of stable carma models with an application to electricity spot prices. *Statistical Modelling*, 11(5):447–470. 76

Gilbert, E. and Pollak, H. (1960). Amplitude distribution of shot noise. *Bell System Technical Journal*, 39(2):333–350. 55

Gubner, J. A. (1996). Computation of shot-noise probability distributions and densities. *SIAM Journal on Scientific Computing*, 17(3):750–761. 55

Gugushvili, S. (2009). Nonparametric estimation of the characteristic triplet of a discretely observed lévy process. *Journal of Nonparametric Statistics*, 21(3):321–343. 66, 71, 78, 83, 87, 88

Hadamard, J. (1902). Sur les problèmes aux dérivées partielles et leur signification physique. *Princeton university bulletin*, 13(49-52):28. 103

Heckman, N. et al. (2012). The theory and application of penalized methods or reproducing kernel hilbert spaces made easy. *Statistics Surveys*, 6:113–141. 109

Heinrich, L. and Schmidt, V. (1985). Normal convergence of multidimensional shot noise and rates of this convergence. *Advances in Applied Probability*, pages 709–730. 55

Helfenstein, U. and Steiner, M. (1990). Random drug excess. *European journal of drug metabolism and pharmacokinetics*, 15(4):309–315. 56

Hill, B. M. (1975). A simple general approach to inference about the tail of a distribution. *The annals of statistics*, 3(5):1163–1174. 81

Hohn, N. and Burkitt, A. N. (2001). Shot noise in the leaky integrate-and-fire neuron. *Physical Review E*, 63(3):031902. 79, 102

Holden, A. V. (1976). *Models of the stochastic activity of neurones*. Springer-Verlag. 79, 102

Iksanov, A. M. and Jurek, Z. J. (2003). Shot noise distributions and selfdecomposability. *Stochastic Analysis and Applications*, 21(3):593–609. 63, 78, 81

Ilhe, P., Moulines, E., Roueff, F., and Antoine, S. (2015). Nonparametric estimation of the mark's density of an exponential shot-noise process. *Electronic Journal of Statistics*. 105

Jongbloed, G., van der Meulen, F. H., and van der Vaart, A. W. (2005). Nonparametric inference for lévy-driven ornstein-uhlenbeck processes. *Bernoulli*, 11(5):759–791. 80

Kallenberg, O. (2002). *Foundations of Modern Probability*. Springer Science & Business Media. 125

Knoll, G. F. (1989). *Radiation detection and measurement*, volume 3. Wiley New York. 77, 85, 101

Kunsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, pages 1217–1241. 114

Lahiri, S. N. (1999). Theoretical comparisons of block bootstrap methods. *Annals of Statistics*, pages 386–404. 115

Lapierre, A. B. and Fortet, R. (1953). *Théorie des fonctions aléatoires*. Masson. 55

Leadbetter, M. et al. (1966). On crossings of levels and curves by a wide class of stochastic processes. *The Annals of Mathematical Statistics*, 37(1):260–267. 56

Lopatin, M., Moskovitch, N., Trigano, T., and Sepulcre, Y. (2012). Pileup attenuation for spectroscopic signals using a sparse reconstruction. In *Electrical & Electronics Engineers in Israel (IEEEI), 2012 IEEE 27th Convention of*, pages 1–5. IEEE. 19, 45

Lowen, S. B. and Teich, M. C. (1990). Power-law shot noise. *Information Theory, IEEE Transactions on*, 36(6):1302–1318. 55

Macchi, O. and Picinbono, B. C. (1972). Estimation and detection of weak optical signals. *Information Theory, IEEE Transactions on*, 18(5):562–573. 105

Mair, B. A. and Ruymgaart, F. H. (1996). Statistical inverse estimation in hilbert scales. *SIAM Journal on Applied Mathematics*, 56(5):1424–1444. 136

Marquardt, T. and Stelzer, R. (2007). Multivariate carma processes. *Stochastic Processes and their Applications*, 117(1):96–120. 76

Masuda, H. (2004). On multidimensional ornstein-uhlenbeck processes driven by a general lévy process. *Bernoulli*, 10(1):97–120. 88

Nashed, M. Z. and Wahba, G. (1974). Generalized inverses in reproducing kernel spaces: An approach to regularization of linear operator equations. *SIAM Journal on Mathematical Analysis*, 5(6):974–987. 106, 107

Neumann, M. H. and Reiß, M. (2009). Nonparametric estimation for lévy processes from low-frequency observations. *Bernoulli*, 15(1):223–248. 66, 71, 72, 73, 78, 82, 83

Papoulis, A. (1971). High density shot noise and gaussianity. *Journal of Applied Probability*, pages 118–127. 54

Pirard, B. (2006). *Etude et validation d'un spectromètre gamma pour la mesure de la composition chimique des surfaces planétaires*. PhD thesis, Université Catholique de Louvain. 6, 8, 32, 34

Politis, D. N. and White, H. (2004). Automatic block-length selection for the dependent bootstrap. *Econometric Reviews*, 23(1):53–70. 115

Reisch, M. (2012). *High-frequency bipolar transistors*, volume 11. Springer Science & Business Media. 51

Reiss, R.-D. (2012). *A course on point processes*. Springer Science & Business Media. 56, 57, 58

Resnick, S. I. (1992). *Adventures in stochastic processes*. Springer. 65, 80

Resnick, S. I. (2007). *Extreme values, regular variation, and point processes*. Springer. 100

Rice, J. (1977). On generalized shot noise. *Advances in Applied Probability*, pages 553–565. 102

Rice, S. O. (1944). Mathematical analysis of random noise. *Bell System Technical Journal*, 23(3):282–332. 54, 56

Rio, E. (1999). *Théorie asymptotique des processus aléatoires faiblement dépendants*, volume 31. Springer Science & Business Media. 125

Roberts, J. (1966). On the response of a simple oscillator to random impulses. *Journal of Sound and Vibration*, 4(1):51–61. 55

Samorodnitsky, G. (1996). A class of shot noise models for financial applications. In *Athens conference on applied probability and time series analysis*, pages 332–353. Springer. 102

Sato, K.-I. (1999). *Lévy processes and infinitely divisible distributions*. Cambridge university press. 69, 74, 79, 81, 103, 124

Schottky, W. (1918). Über spontane stromschwankungen in verschiedenen elektrizitätsleitern. *Annalen der Physik*, 362(23):541–567. 54, 102

Schumaker, L. (2007). *Spline functions: basic theory*. Cambridge University Press. 120

Sequeira, R. E. and Gubner, J. A. (1995). Intensity estimation from shot-noise data. *Signal Processing, IEEE Transactions on*, 43(6):1527–1531. 79, 102

Sequeira, R. E. and Gubner, J. A. (1997). Blind intensity estimation from shot-noise data. *Signal Processing, IEEE Transactions on*, 45(2):421–433. 79, 102

Taguchi, K., Frey, E. C., Wang, X., Iwanczyk, J. S., and Barber, W. C. (2010). An analytical model of the effects of pulse pileup on the energy spectrum recorded by energy resolved photon counting x-ray detectors. *Medical physics*, 37(8):3957–3969. 55, 102

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 267–288. 20, 46

Trabs, M. (2014). On infinitely divisible distributions with polynomially decaying characteristic functions. *Statistics & Probability Letters*, 94:56–62. 68, 74

Trigano, T. (2005). *Traitement Statistique du Signal Spectrométrique: étude du désempilement de spectre en énergie pour la spectrométrie gamma*. PhD thesis, Télécom ParisTech. 18, 19, 44

Trigano, T., Barat, E., Dautremer, T., and Montagu, T. (2015a). Fast digital filtering of spectrometric data for pile-up correction. *Signal Processing Letters, IEEE*, 22(7):973–977. 105, 136

Trigano, T., Gildin, I., and Sepulcre, Y. (2015b). Pileup correction algorithm using an iterated sparse reconstruction method. *Signal Processing Letters, IEEE*, 22(9):1392–1395. 12, 19, 38, 45, 105

Tsoulfanidis, N. (2013). *Measurement and detection of radiation*. CRC press. 2, 28

Tuominen, P. and Tweedie, R. L. (1994). Subgeometric rates of convergence of f-ergodic markov chains. *Advances in Applied Probability*, pages 775–798. 89

van der Vaart, A. W. and Wellner, J. A. (1996). *Weak Convergence and Empirical Processes*. Springer. 87, 93

Volkonskii, V. and Rozanov, Y. A. (1959). Some limit theorems for random functions. i. *Theory of Probability & Its Applications*, 4(2):178–197. 87

Wahba, G. (1973). Convergence rates of certain approximate solutions to fredholm integral equations of the first kind. *Journal of Approximation Theory*, 7(2):167–185. 106

Wahba, G. (1977). Practical approximate solutions to linear operator equations when the data are noisy. *SIAM Journal on Numerical Analysis*, 14(4):651–667. 107

Wahba, G. (1990). *Spline models for observational data*, volume 59. Siam. 106

Wood, S. N. (2000). Modelling and smoothing parameter estimation with multiple quadratic penalties. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 62(2):413–428. 115, 116

Xiao, Y. and Lund, R. (2006). Inference for shot noise. *Statistical inference for stochastic processes*, 9(1):77–96. 79, 103

Young, N. (1988). *An introduction to Hilbert space*. Cambridge university press. 132

Yu, J. (2004). Empirical characteristic function estimation and its applications. *Econometric reviews*, 23(2):93–123. 112, 136

Statistical estimation of a shot-noise process

Paul Ilhe

RESUME : Dans le cadre de la spectrométrie gamma, cette thèse introduit de nouveaux estimateurs non-paramétriques de l'intensité et de la densité des marques d'un processus shot-noise à partir d'un nombre fini d'observations du processus échantillonné à basse fréquence. Les méthodes proposées utilisent une relation non linéaire reliant la fonction caractéristique de la loi marginale du processus à la densité des marques. Elles sont particulièrement rapides et possèdent l'avantage d'être efficaces mêmes pour des intensités élevées. Les performances de ces méthodes sont étudiées quantitativement et illustrées à la fois pour des données simulées et réelles provenant du CEA Saclay. En particulier, les estimateurs de la densité des marques permettent de corriger les artefacts de pics multiples.

MOTS-CLEFS : processus shot-noise, problème inverse non-linéaire, désempilement, estimation non-parametrique, spectrométrie gamma.

ABSTRACT : In the context of gamma-spectroscopy, this thesis introduces new nonparametric estimators of the intensity and the mark's density of a shot-noise process based on a finite sample of low-frequency observations of this stochastic process. The methods developed exploit a nonlinear functional equation linking the characteristic function of the marginal law of the shot-noise with the mark's density function. They are particularly time-efficient and perform well even for processes with high intensity. The performances of the methods are quantitatively studied and illustrations are provided both on simulated datasets and real datasets stemming from the CEA. In particular, our methods corrects the multiple peak artefacts that arises with classical techniques.

KEY-WORDS : shot-noise processes, nonlinear inverse problem, pileup correction, nonparametric estimation, gamma spectroscopy.