HAL
open science

Median-based estimation of the intensity of a spatial point process

Jean-François Coeurjolly

To cite this version:

Jean-François Coeurjolly. Median-based estimation of the intensity of a spatial point process. Annals of the Institute of Statistical Mathematics, 2017, 69 (2), pp.303-331. 10.1007/s10463-015-0536-7 . hal-01071605v2

HAL Id: hal-01071605

https://hal.science/hal-01071605v2

Submitted on 24 May 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Median-based estimation of the intensity of a spatial point process

Jean-François Coeurjolly[1]

[1]Univ. Grenoble Alpes, F-38000 Grenoble, France,
Jean-Francois.Coeurjolly@upmf-grenoble.fr.

Abstract

This paper is concerned with a robust estimator of the intensity of a stationary spatial point process. The estimator corresponds to the median of a jittered sample of the number of points, computed from a tessellation of the observation domain. We show that this median-based estimator satisfies a Bahadur representation from which we deduce its consistency and asymptotic normality under mild assumptions on the spatial point process. Through a simulation study, we compare the new estimator, in particular, with the standard one counting the mean number of points per unit volume. The empirical study confirms the asymptotic properties established in the theoretical part and shows that the median-based estimator is more robust to outliers than standard procedures.

Keywords: Cox processes; Robust statistics; Sample quantiles; Bahadur representation.

1 Introduction

Spatial point patterns are datasets containing the random locations of some event of interest. These datasets arise in many scientific fields such as biology, epidemiology, seismology, hydrology. Spatial point processes are the stochastic models generating such data. We refer to Stoyan et al. (1995), Illian et al. (2008) or Møller and Waagepetersen (2004) for an overview on spatial point processes. These references cover practical as well as theoretical aspects. A point process $\mathbf{X}$ in $\mathbb{R}^d$ is a locally finite random subset of $\mathbb{R}^d$ meaning that the restriction to any bounded Borel set is finite. The point process $\mathbf{X}$ takes values in Ω, which contains in all locally finite subsets of $\mathbb{R}^d$. Thus, the distribution of $\mathbf{X}$ is a probability measure on an appropriate σ-algebra of Ω. The Poisson point process is the reference process to model random locations of points without interaction. Many alternative models such as Cox point processes, determinantal point processes, Gibbs point processes allow us to introduce clustering effects or to produce regular patterns (see again e.g. Møller and Waagepetersen (2004) and Lavancier et al. (2014) for an overview). If the distribution of $\mathbf{X}$ is invariant by translation, we say that $\mathbf{X}$ is stationary. We are interested

in this paper in first-order characteristics of $\mathbf{X}$, which under the assumption of stationarity, reduce to a single real parameter denoted by λ. This intensity λ measures the average number of points per unit volume.

Estimating λ is a well-known problem and has been the subject of a large literature. Based on a single realization of the point process $\mathbf{X}$ in a bounded domain W of $\mathbb{R}^d$, the natural way of estimating λ is to compute the average number of points observed by unit volume, i.e. to evaluate $N(W)/|W|$ where $N(W)$ denotes the number of points of $\mathbf{X}$ falling into the observation domain W with volume $|W|$. We denote this estimator by $\widehat{\lambda}^{\mathrm{std}}$ for this estimator. If the point process is a homogeneous Poisson point process, $\widehat{\lambda}^{\mathrm{std}}$ is also the maximum likelihood estimator. Asymptotic properties of $\widehat{\lambda}^{\mathrm{std}}$ are well established for a large class of models. In particular, as the observation window expands to $\mathbb{R}^d$, it can be shown under mild assumptions on $\mathbf{X}$ (mainly mixing conditions) that $\widehat{\lambda}^{\mathrm{std}}$ is consistent and satisfies a central limit theorem with asymptotic variance, which can be consistently estimated (see Heinrich and Prokešová (2010) and the references therein for more details). In some applications, it may be too time-consuming to count all points. In such situations, distance based methods, where mainly nearest distances between points are used, have been developed (see e.g. Byth (1982); Diggle (2003); Magnussen (2012)). Unlike the estimator $\widehat{\lambda}^{\mathrm{std}}$, those methods are quite sensitive to the model, which explains why the only practically applicable case is the Poisson process (Illian et al., 2008). Other moment-based methods include the adapted estimator proposed by Mrkvička and Molchanov (2005) or the recent Stein estimator (in the Poisson case) proposed by Clausel et al. (2015).

As outlined in particular in the book Illian et al. (2008), an important step in the statistical analysis of point patterns is the search for unusual points or unusual point configurations, i.e. the search of outliers. Two kinds of outliers appear when dealing with point pattern: first points may appear at locations where they are not expected. This situation could appear for instance when two species of plants or trees cannot be distinguished at the time of data collection. Second, it is possible that there are missing points in the pattern, i.e. areas of the observation domain where, according to the general structure of the pattern, points are expected. Illian et al. (2008) or Baddeley et al. (2005) have proposed several diagnostic tools to detect outliers and more generally to judge the quality of fit of a model. To the best of our knowledge, the works by Berndt and Stoyan (1997) and Assunção and Guttorp (1999) are the only works where robustness of estimation procedures are tackled. Assunção and Guttorp (1999) developed an M-estimator to estimate the intensity of an inhomogeneous Poisson point process. For an application in materials science, Berndt and Stoyan (1997) proposed the following methodology to estimate the intensity parameter of a stationary point process: let $\mathcal{C}$ be a typical cell of the Voronoi tessellation built from a stationary point process. It is known, see e.g. Stoyan et al. (1995); Møller (1994), that $\mathrm{E}\,|\mathcal{C}| = 1/\lambda$, whereby an estimator of λ can be deduced by evaluating the sample mean of cell areas produced by the Voronoi tessellation of the observed point pattern. Berndt and Stoyan (1997) proposed to replace the sample mean by a more robust estimator like a sample median or a trimmed-mean. Illian et al. (2008, p. 252) have suggested a slightly different procedure. Let $G = \{g_1, \ldots, g_{\#G}\}$ be a grid of $\#G$ dummy points and let $a(g, \mathbf{X})$ be the cell area of the closest point $g \in G$

in $\mathbf{X}$. Starting from the fact that $(\#G)^{-1}\sum_{g\in G} a(g,\mathbf{X})^{-1}$ is an unbiased estimator of λ (up to edge effects), then a robust estimator can be constructed by replacing the sample mean by a trimmed-mean for instance (see Section 5 for more details). The two latter procedures described, which were not supported by theoretical results, are the closest to the present work and we discuss them in Section 5.

As far as we know, no model free robust techniques supported by theoretical results have been developed. In this paper, we aim at developing a simple median-based estimator of λ. It is not so straightforward to see what a median means for a spatial point process but we may remark that if W is decomposed as a union of K non-overlapping cells C_k, then $N(W) = \sum_k N(C_k)$, which yields that $\widehat{\lambda}^{\text{std}}$ can be actually rewritten as the empirical mean of the normalized counts variables $N(C_k)/|C_k|$. We have the cornerstone to define a more robust estimator by simply replacing the sample mean by the sample median.

The classical definition of sample quantiles and their asymptotic properties for continuous distributions are nowadays well-known, see e.g. David and Nagaraja (2003). In particular, sample medians in the i.i.d. setting, computed from an absolutely continuous distribution, f, positive at the true median, Me, are consistent and satisfy a central limit theorem with asymptotic variance $1/4f(\text{Me})^2$. Such a result obviously fails for discrete distributions. In this paper, we follow a strategy introduced by Stevens (1950) and applied to count data by Machado and Santos Silva (2005) which consists in artificially imposing smoothness in the problem through jittering: i.e. we add to each count variable $N(C_k)$ a random variable U_k following a uniform distribution on $(0,1)$. Now, the random variable $N(C_k)+U_k$ admits a density and asymptotic results can be expected. To get around the problem of large sample behavior for discrete distributions, another approach could be to consider the median based on the mid-distribution, see Ma et al. (2011). The authors prove that such sample quantiles behave more favourably than the classical one and satisfy, in the i.i.d. setting, a central limit theorem even if the distribution is discrete. We leave to a future work the question of deriving asymptotic properties for the sample median based on the mid-distribution in the context of this paper.

The rest of the paper is organized as follows. Section 2 gives a short background on spatial point processes. General notation as well as the definition of our estimator are presented in Section 3. We also examine in Section 3 how far the true median of $N(C_k)+U_k$ is from the intensity $\lambda|C_k|$. Section 4 contains our main asymptotic results. General assumptions are discussed and a particular focus on Cox point processes is investigated. The main difficulty here is to establish a Bahadur representation for the jittered sample median, which can be applied to a large class of models. Section 5 presents the results of a simulation study where we compare our procedure with the standard estimator $\widehat{\lambda}^{\text{std}}$ and the estimator proposed by Berndt and Stoyan (1997). The research contained in this paper leads to a number of interesting open questions, which are mentioned in Section 6. Proofs of the results and additional comments are postponed to Appendices A and B

2 Background on spatial point processes

Let $\mathbf{X}$ be a spatial point process defined on $\mathbb{R}^d$, which we see as a random locally finite subset of $\mathbb{R}^d$. Let W be a bounded Borel set of $\mathbb{R}^d$, then the number of points in $\mathbf{X} \cap W$, denoted by $N(W)$, is finite, and a realization of $\mathbf{X} \cap W$ is of the form $\mathbf{x} = \{x_1, \ldots, x_m\} \subset W$ for some nonnegative finite integer m. If $m = 0$, then $\mathbf{x} = \emptyset$ is an empty point pattern in W. For $u \in \mathbb{R}^d$, we denote by $\|u\|$ its Euclidean norm. For further background and measure theory on spatial point processes, see e.g. Daley and Vere-Jones (2003) and Møller and Waagepetersen (2004). We assume that $\mathbf{X}$ is a stationary point process with intensity λ, which, by Campbell's theorem (see e.g. Møller and Waagepetersen (2004)), is characterized by the fact that for any real Borel function h defined on $\mathbb{R}^d$ and absolutely integrable (with respect to the Lebesgue measure on $\mathbb{R}^d$)

$$\mathrm{E} \sum_{u \in \mathbf{X}} h(u) = \lambda \int h(u) \mathrm{d}u. \tag{2.1}$$

Furthermore, for any integer $l \geq 1$, $\mathbf{X}$ is said to have an lth-order product density ρ_l if ρ_l is a non-negative Borel function on $\mathbb{R}^{dl}$ such that for all non-negative Borel functions h defined on $\mathbb{R}^{dl}$,

$$\mathrm{E} \sum_{u_1, \ldots, u_l \in \mathbf{X}}^{\neq} h(u_1, \ldots, u_l) = \int_{\mathbb{R}^d} \cdots \int_{\mathbb{R}^d} h(u_1, \ldots, u_l) \rho_l(u_1, \ldots, u_l) \, \mathrm{d}u_1 \cdots \mathrm{d}u_l, \tag{2.2}$$

where the sign $\neq$ over the summation means that $u_1, \ldots, u_l$ are pairwise distinct. Note that $\lambda = \rho_1$ and that for the homogeneous Poisson point process $\rho_l(u_1, \ldots, u_l) = \lambda^l$. If $\rho^{(2)}$ exists, then by the stationarity of $\mathbf{X}$, $\rho^{(2)}(u, v)$ depends only on $u - v$. In that case, we define the pair correlation function g as a function from $\mathbb{R}^d$ to $\mathbb{R}^+$ by $g(u - v) = \rho^{(2)}(u, v)/\lambda^2$.

In this paper, we sometimes pay attention to Cox point processes, which are defined as follows.

Definition 2.1. *Let $(\xi(s), s \in \mathbb{R}^d)$ be a non-negative locally integrable random field. Then, $\mathbf{X}$ is a Cox point process if the distribution of $\mathbf{X}$ given ξ is an inhomogeneous Poisson point process with intensity function ξ. If ξ is stationary, so is $\mathbf{X}$ and $\lambda = \mathrm{E}(\xi(s))$ for any s.*

Among often used models of stationary Cox point processes, we can cite

- Log-Gaussian Cox processes (e.g. Møller and Waagepetersen (2004)): Let Y be a stationary Gaussian process on $\mathbb{R}^d$ with mean μ and stationary covariance function $c(u) = \sigma^2 r(u)$, $u \in \mathbb{R}^d$, where $\sigma^2 > 0$ is the variance and r the correlation function. If $\mathbf{X}$ conditional on Y is a Poisson point process with intensity function $\xi = \exp(Y)$, then $\mathbf{X}$ is a (homogeneous) log-Gaussian Cox process. One example of correlation function is the Matérn correlation function (which includes the exponential correlation function) given by $r(u) = (\sqrt{2\nu}\|u\|/\phi)^\nu K_\nu(\sqrt{2\nu}\|u\|/\phi)/(2^{\nu-1}\Gamma(\nu))$ where Γ is the gamma function, K_ν is the modified Bessel function of the second kind, and ϕ and ν are non-negative parameters. In particular, the intensity of $\mathbf{X}$ equals $\lambda = e^{\mu+\sigma^2/2}$.

- Neyman-Scott processes (e.g. Møller and Waagepetersen (2004)): Let $\mathbf{C}$ be a stationary Poisson point process with intensity $\kappa > 0$, and f_σ a density function on $\mathbb{R}^d$. If $\mathbf{X}$ conditional on $\mathbf{C}$ is a Poisson point process with intensity

$$\alpha \sum_{c \in \mathbf{C}} f_\sigma(u - c)/\kappa, \quad u \in \mathbb{R}^2, \tag{2.3}$$

for some $\alpha > 0$, then $\mathbf{X}$ is a (homogeneous) Neyman-Scott process. When f_σ corresponds to the density of a uniform distribution on $B(0, \sigma^2)$ (resp. a Gaussian random variable with mean 0 and variance σ^2), we refer to $\mathbf{X}$ as the (homogeneous) Matérn Cluser (resp. Thomas) point process. In particular, the intensity of $\mathbf{X}$ equals $\lambda = \alpha\kappa$.

3 Median-based estimator of λ

For any real-valued random variable Y, we denote by $F_Y(\cdot)$ its cdf, by $F_Y^{-1}(p)$ its quantile of order $p \in (0, 1)$, by $\mathrm{Me}_Y = F_Y^{-1}(1/2)$ its theoretical median. Based on a sample $\mathbf{Y} = (Y_1, \ldots, Y_n)$ of n identically distributed random variables we denote by $\widehat{F}(\cdot; \mathbf{Y})$ the empirical cdf, by $\widehat{F}^{-1}(p; \mathbf{Y})$ the sample quantile of order p given by

$$\widehat{F}^{-1}(p; \mathbf{Y}) = \inf\{x \in \mathbb{R} : p \leq \widehat{F}(x; \mathbf{Y})\}. \tag{3.1}$$

The sample median is simply denoted by $\widehat{\mathrm{Me}}(\mathbf{Y}) = \widehat{F}^{-1}(1/2; \mathbf{Y})$.

We will study the large-sample behavior of estimators of the intensity λ. Specifically, we consider a region W_n assumed to increase to $\mathbb{R}^d$ as $n \to \infty$. We assume that the domain of observation W_n can be decomposed as $W_n = \cup_{k \in \mathcal{K}_n} C_{n,k}$ where the cells $C_{n,k}$ are non-overlapping and equally sized with volume $c_n = |C_{n,k}|$ and where $\mathcal{K}_n$ is a subset of $\mathbb{Z}^d$ with cardinality $k_n = |\mathcal{K}_n|$. More details on W_n, c_n and k_n will be provided in the appropriate Section 4 when we present asymptotic results. Finally, for any random variable Y or any random vector $\mathbf{Y}$, we denote by $\check{Y} = Y/c_n$ and $\check{\mathbf{Y}} = \mathbf{Y}/c_n$.

The classical estimator of the intensity λ is given by $\widehat{\lambda}^{\mathrm{std}} = N(W_n)/|W_n|$. In order to define a more robust estimator, we can note that

$$\widehat{\lambda}^{\mathrm{std}} = \frac{1}{k_n} \sum_{k \in \mathcal{K}_n} \frac{N(C_{n,k})}{c_n} \tag{3.2}$$

since $|W_n| = k_n c_n$, i.e. $\widehat{\lambda}^{\mathrm{std}}$ is nothing else that the sample mean of intensity estimators computed in cells $C_{n,k}$. The strategy adopted in this paper is to replace the sample mean by the sample median, which is known to be more robust to outliers. As underlined in the introduction, quantile estimators based on count data or more generally on discrete data can cause some troubles in the asymptotic theory. The problems come from the fact that, in the continuous case, the asymptotic variance of the sample median involves the probability distribution function at the true median.

To overcome the problem of discontinuity of the counts variables $N(C_{n,k})$, we follow a well-known technique (e.g. Machado and Santos Silva (2005)) which introduces smoothness in the problem. Let $(U_k, k \in \mathcal{K}_n)$ be a collection of independent

and identically distributed random variables, distributed as $U \sim \mathcal{U}([0,1])$. Then, for any $k \in \mathcal{K}_n$, we define

$$Z_{n,k} = N(C_{n,k}) + U_k \quad \text{and} \quad \mathbf{Z} = (Z_{n,k},\ k \in \mathcal{K}_n). \tag{3.3}$$

Since $\mathbf{X}$ is stationary, the variables $Z_{n,k}$ are identically distributed and we let $Z \sim Z_{n,k}$. The jittering effect shows up right away: the cdf of Z is given for any $t \geq 0$ by

$$F_Z(t) = P(N(C_{n,0}) \leq \lfloor t \rfloor - 1) + P(N(C_{n,0}) = \lfloor t \rfloor)\,(t - \lfloor t \rfloor),$$

which is a continuously differentiable function whereby we deduce that Z admits a density f_Z at t given by

$$f_Z(t) = P(N(C_{n,0}) = \lfloor t \rfloor). \tag{3.4}$$

We define the jittered estimator of the intensity λ by

$$\widehat{\lambda}^J = \widehat{\mathrm{Me}}(\check{\mathbf{Z}}) = \frac{\widehat{\mathrm{Me}}(\mathbf{Z})}{c_n} \tag{3.5}$$

where the sample median is defined by (3.1). Since it is expected that $\widehat{\mathrm{Me}}(\check{\mathbf{Z}})$ is close to $\mathrm{Me}_{\check{Z}} = \mathrm{Me}_Z/c_n$, we need to understand how far $\mathrm{Me}_{\check{Z}}$ is from λ. Using the definition of the median we can prove the following result.

Proposition 3.1. *Assume that the pair correlation function of the stationary point process* $\mathbf{X}$ *exists for* $u, v \in \mathbb{R}^d$ *and satisfies* $\int_{\mathbb{R}^d} |g(w) - 1| \mathrm{d}w < \infty$*, then for any* $\varepsilon > 0$ *we have for* n *sufficiently large*

$$|\mathrm{Me}_{\check{Z}} - \lambda| \leq \frac{1}{c_n}\left(\frac{1}{2} + \sqrt{\frac{1}{12}}\right) + (1+\varepsilon)\sqrt{\frac{\sigma}{c_n}} = \mathcal{O}(c_n^{-1/2}) \tag{3.6}$$

where $\sigma^2 = \lambda + \lambda^2 \int_{\mathbb{R}^d} (g(w) - 1) \mathrm{d}w$.

The assumption $\int |g(w) - 1| \mathrm{d}w < \infty$ is quite standard when we deal with asymptotics for spatial point processes, see e.g. Guan and Loh (2007) or Heinrich and Prokešová (2010). It ensures that for any sequence of regular domains Δ_n, $|\Delta_n|^{-1} \operatorname{Var}(N(\Delta_n)) \to \sigma^2$ as $n \to \infty$. We refer the reader to these papers and to Section 4.1 for a discussion on this assumption.

Proof. By the previous remark, ${c_n}^{-1} \operatorname{Var}(N(C_{n,k})) \to \sigma^2$ as $n \to \infty$. Since for any continuous random variable Y with finite first two moments, $|\mathrm{Me}_Y - \mathrm{E}(Y)| \leq \sqrt{\operatorname{Var}(Y)}$ and since $\mathrm{E}(Z) = \lambda c_n + 1/2$, then for any $\varepsilon > 0$ we have for n sufficiently large

$$|\mathrm{Me}_Z - \lambda c_n| \leq \frac{1}{2} + \sqrt{\frac{1}{12} + (1+\varepsilon)^2 \sigma^2 c_n}.$$

Dividing both sides of the inequality by c_n leads to (3.6). □

Let $\Pi \sim \mathcal{P}(\nu)$ be a Poisson random variable with mean $\nu > 0$. Several results are known for the theoretical median of Π, see e.g. Adell and Jodrá (2005). For instance, when ν is an integer $\mathrm{Me}_\Pi = \nu$ and for non integer ν, $-\log 2 \leq \mathrm{Me}_\Pi \leq 1/3$ (see Figure 1). Based on this, we can obtain a sharper inequality than (3.6) for Poisson and Cox point processes.

Proposition 3.2. *Let* $\mathbf{X}$ *be a stationary Cox point process with latent random field* ξ*, then*

$$\lambda c_n - \log 2 \leq \mathrm{Me}_{N(C_{n,0})} \leq \lambda c_n + \frac{1}{3} \quad \textit{and} \quad |\mathrm{Me}_Z - \lambda c_n| \leq \frac{4}{3}. \tag{3.7}$$

A reformulation of (3.7) is of course $\mathrm{Me}_{\check{Z}} - \lambda = \mathcal{O}(c_n^{-1})$.

Proof. Given ξ, for any $k \in \mathcal{K}_n$, $N(C_{n,k})$ follows a Poisson distribution with mean $\int_{C_{n,k}} \xi(s)\mathrm{d}s$. Denote by $\mathrm{Me}_{N(C_{n,k})|\xi}$ the median of $N(C_{n,k})$ given ξ defined by

$$\mathrm{Me}_{N(C_{n,k})|\xi} = \inf\left\{z \in \mathbb{R} : F_{N(C_{n,k})|\xi}(z) \geq 1/2\right\}$$

where $F_{N(C_{n,k})|\xi}$ is the cumulative distribution function of $N(C_{n,k})$ given ξ. From the property of the median of a Poisson distribution, we have for any $k \in \mathcal{K}_n$

$$\int_{C_{n,k}} \xi(s)\mathrm{d}s - \log 2 \leq \mathrm{Me}_{N_{n,k}|\xi} \leq \int_{C_{n,k}} \xi(s)\mathrm{d}s + \frac{1}{3}.$$

Since $\mathrm{E}\int_{C_{n,k}} \xi(s)\mathrm{d}s = \lambda c_n$, the first result is deduced by taking the expectation of each term of the previous inequality. Since $N(C_{n,0}) \leq Z \leq N(C_{n,0}) + 1$, $\mathrm{Me}_{N(C_{n,0})} \leq \mathrm{Me}_Z \leq \mathrm{Me}_{N(C_{n,0})} + 1$ which leads to the second result. □

4 Asymptotic results

We state in this section our main results and the general assumptions required to obtain them. Proofs of the different results presented here, as well as auxiliary results, are presented in Appendix A.

4.1 General assumptions and discussion

We recall the classical definition of mixing coefficients (see e.g. Politis et al. (1998)): for $j, k \in \mathbb{N} \cup \{\infty\}$ and $m \geq 1$, define

$$\begin{aligned}\alpha_{j,k}(m) = \sup\{&|P(A \cap B) - P(A)P(B)| : A \in \mathcal{F}(\Lambda_1),\, B \in \mathcal{F}(\Lambda_2),\\ &\Lambda_1 \in \mathcal{B}(\mathbb{R}^d),\, \Lambda_2 \in \mathcal{B}(\mathbb{R}^d),\, |\Lambda_1| \leq j,\, |\Lambda_2| \leq k,\, d(\Lambda_1, \Lambda_2) \geq m\}\end{aligned} \tag{4.1}$$

where $\mathcal{F}(\Lambda_i)$ is the σ-algebra generated by $\mathbf{X} \cap \Lambda_i$, $i = 1, 2$, $d(\Lambda_1, \Lambda_2)$ is the minimal distance between the sets Λ_1 and Λ_2, and $\mathcal{B}(\mathbb{R}^d)$ denotes the class of Borel sets in $\mathbb{R}^d$.

We require the following assumptions to prove our asymptotic results.

(i) For any $n \geq 1$, we assume that $W_n = \cup_{k \in \mathcal{K}_n} C_{n,k}$ where $\mathcal{K}_n$ is a subset of $\mathbb{Z}^d$ with cardinality $k_n = |\mathcal{K}_n|$ and where the cells $C_{n,k}$ are equally sized and non-overlapping cubes with volume c_n, defined by

$$C_{n,k} = \left\{u = (u_1, \ldots, u_d)^\top \in \mathbb{R}^d : c_n^{1/d}(k_l - 1/2) \leq u_l \leq c_n^{1/d}(k_l + 1/2),\ l = 1 \ldots, d\right\}.$$

We assume that $0 \in \mathcal{K}_n$ and that there exists $0 < \eta' < \eta$ such that as $n \to \infty$

$$k_n \to \infty, \quad c_n \to \infty, \quad \frac{k_n}{c_n^{\eta'/2\wedge(1-2\ell)}} \to 0$$

where ℓ is given by Assumption (ii) and η by Assumption (iv).
(ii)

(ii-1) $\mathrm{Me}_Z - \lambda c_n = \mathcal{O}({c_n}^{\ell})$ with $0 \leq \ell < 1/2$.

(ii-2) $\forall t_n = \lambda c_n + \mathcal{O}(\sqrt{c_n/k_n})$, $\mathrm{P}(N(C_{n,0}) = \lfloor t_n \rfloor)/\mathrm{P}(N(C_{n,0}) = \lfloor \lambda c_n \rfloor) \to 1$.

(ii-3) There exist $\underline{\kappa}, \overline{\kappa} > 0$ such that for n large enough, $\underline{\kappa} \leq \sqrt{c_n} f_Z(\mathrm{Me}_Z) \leq \overline{\kappa}$.

(iii) $\mathbf{X}$ has a pair correlation function g satisfying $\int_{\mathbb{R}^d} |g(w) - 1| \mathrm{d}w < \infty$.

(iv) There exists $\eta > 0$ such that

$$\alpha(m) = \sup_{p \geq 1} \frac{\alpha_{p,p}(m)}{p} = \mathcal{O}(m^{-d(1+\eta)}) \quad \text{and} \quad \alpha_{2,\infty}(m) = \mathcal{O}(m^{-d(1+\eta)})$$

where $\alpha_{j,k}(m)$ for $j, k \in \mathbb{N} \cup \{\infty\}$ is defined by (4.1).

Now, we discuss the different assumptions. The last statement of Assumption (i) is required to control the dependency between the variables $Z_{n,k}$ through the control of the mixing coefficients and to ensure that asymptotically $|W_n|^{1/2}(\widehat{\lambda}^J - \lambda)$ behaves as $|W_n|^{1/2}(\widehat{\lambda}^J - \mathrm{Me}_{\check{Z}})$. We note that if $\mathbf{X}$ is a stationary Cox point process, Proposition 3.2 yields that (ii-1) is satisfied for $\ell = 0$. So if $\eta > 2$, Assumption (i) can be rewritten as $c_n \to \infty$, $k_n \to \infty$ and $k_n/c_n \to 0$ as $n \to \infty$.

Regarding Assumption (ii), Proposition 4.1 stated below shows it can be simplified for a large class of Cox point processes. We underline that Assumptions (i), (ii-1)-(ii-2) imply the existence of $\overline{\kappa} < \infty$ such that $\sqrt{c_n} f_Z(\mathrm{Me}_Z) \leq \overline{\kappa}$, so (ii-3) could actually be simplified.

Assumption (iii) is classical when dealing with asymptotics of intensity estimates, see e.g. Heinrich and Prokešová (2010). For isotropic pair correlation functions, i.e. $g(w) = \overline{g}(\|w\|)$ for $\overline{g} : \mathbb{R}^+ \to \mathbb{R}$, Assumption (iii) is fulfilled when $\overline{g}(r) = 0$ for $r \geq R$ or when $\overline{g}(r) = \mathcal{O}(r^{-d-\gamma})$ for some $\gamma > d$. This includes the Matérn cluster and Thomas processes and the log-Gaussian Cox process with Matérn-Whittle covariance functions.

Assumption (iv) is also quite standard and has been discussed a lot in the literature: Guan and Loh (2007); Guan et al. (2007); Prokešová and Jensen (2013) discussed the first part of (iv) while the second one was commented in Waagepetersen and Guan (2009); Coeurjolly and Møller (2014). Both of them are satisfied for the Matérn Cluster and Thomas processes and for log Gaussian Cox processes with correlation function decaying fast enough to zero.

We point out that it is not so common to use both the mixing coefficients $\alpha(m)$ and $\alpha_{2,\infty}(m)$. As detailed in the proof of Theorem 4.2, the first one is used to control the dependence between the random variables $Z_{n,k}$ for $k \in \mathcal{K}_n$ and derive a central limit theorem using the blocking technique developed by Ibragimov and Linnik (1971) which is pertinent and appropriate here since the cells $C_{n,k}$ are increasing. The second mixing coefficient is necessary to apply a multivariate central limit theorem inside the cell $C_{n,0}$. We prove in particular that

$\mathrm{P}(N(C_{n,0}) \leq \lambda c_n, N(C_{n,0}^-) \leq \lambda c_n) \to 1/2$ as $n \to \infty$ where $C_{n,0}^-$ is a "small" erosion of $C_{n,0}$ (see the proof of Step 1 of Theorem 4.2 for more details).

The next result shows the simplifications we can obtain for Cox point processes.

Proposition 4.1. *Let* $\mathbf{X}$ *be a stationary Cox point process with latent random field* $(\xi(s), s \in \mathbb{R}^d)$ *satisfying the Assumptions (iii)-(iv). Assume there exists* $\delta > 2/\eta$, *where* η *is given by Assumption (iv), such that* $\mathrm{E}(|\xi(0)|^{2+\delta}) < \infty$. *Let* $t_n = \lambda c_n + \mathcal{O}(\sqrt{c_n/k_n})$ *and* $T_n = \lfloor t_n \rfloor^{-1} \int_{C_{n,0}} \xi(s)\mathrm{d}s$. *We also assume that the sequence of random variables* $(B_n)_n$ *defined by* $\log(B_n) = \lfloor t_n \rfloor (\log(T_n) - (T_n - 1) + (T_n - 1)^2/2)$ *is uniformly integrable. Then, Assumption (ii) holds (with* $\ell = 0$*) and as* $n \to \infty$

$$\sqrt{c_n}\mathrm{P}(N(C_{n,0}) = \lfloor \lambda c_n \rfloor) \to \left(2\pi\sigma^2\right)^{-1/2} \tag{4.2}$$

where $\sigma^2 = \lambda + \lambda^2 \int_{\mathbb{R}^d} (g(w) - 1)\mathrm{d}w$.

4.2 Results

In this section, we present asymptotic results for $\widehat{F}(\cdot; \mathbf{Z})$, the empirical cumulative distribution function based on $\mathbf{Z}$ and for the median-based estimator $\widehat{\lambda}^J$.

Theorem 4.2. *Under the Assumptions (i)-(iv), we have the following two statements.*
(a) Let $(a_n)_{n\geq 1}$ *be a sequence of real numbers satisfying* $\lambda c_n = a_n + o(\sqrt{c_n})$, *then as* $n \to \infty$

$$\sqrt{k_n}\left(\widehat{F}(\lambda c_n + a_n; \mathbf{Z}) - F_Z(\lambda c_n + a_n)\right) \to \mathcal{N}(0, 1/4)$$

in distribution.
(b) As $n \to \infty$

$$\sqrt{k_n}\left(\widehat{F}(\mathrm{Me}_Z; \mathbf{Z}) - 1/2\right) \to \mathcal{N}(0, 1/4)$$

in distribution.

If $\mathbf{Z}$ corresponds to a sample of n i.i.d. random variables, $\sqrt{n}(\widehat{F}(p; \mathbf{Z}) - p)$ tends to a Gaussian random variable with mean zero and variance $p(1-p)$. Hence, we recover the same result as in our dependency setting.

The next result establishes a Bahadur representation for the sample median, leading to its asymptotic normality. The notation $X_n = o_{\mathrm{P}}(v_n^{-1})$ for a sequence of random variables X_n and a sequence of positive real numbers v_n means that $v_n X_n$ tends to 0 in probability as $n \to \infty$.

Theorem 4.3. *Under the assumptions (i)-(iv), we have the following two statements.*
(a) As $n \to \infty$

$$\widehat{\mathrm{Me}}(\mathbf{Z}) - \mathrm{Me}_Z = \frac{1/2 - \widehat{F}(\mathrm{Me}_Z; \mathbf{Z})}{f_Z(\mathrm{Me}_Z)} + o_{\mathrm{P}}\left(\sqrt{\frac{c_n}{k_n}}\right). \tag{4.3}$$

(b) Let $s_n = \sqrt{c_n}\mathrm{P}(N(C_{n,0}) = \lfloor \lambda c_n \rfloor)$, *then as* $n \to \infty$

$$|W_n|^{1/2} s_n \left(\widehat{\lambda}^J - \lambda\right) \to \mathcal{N}(0, 1/4) \tag{4.4}$$

in distribution.

We deduce the following Corollary given without proof for Cox point processes.

Corollary 4.4. Under the Assumption (i) and the Assumptions of Proposition 4.1, we have as $n \to \infty$

$$|W_n|^{1/2}\left(\widehat{\lambda}^J - \lambda\right) \to \mathcal{N}\left(0, \pi\sigma^2/2\right)$$

in distribution, where $\sigma^2 = \lambda + \lambda^2 \int_{\mathbb{R}^d}(g(w) - 1)\mathrm{d}w$.

As detailed after Proposition 3.1, σ^2 corresponds to the asymptotic variance of $|W_n|^{-1}N(W_n)$. Actually, if we denote by $\widehat{\lambda}^{\mathrm{std}}$ the standard estimator of λ given by $\widehat{\lambda}^{\mathrm{std}} = |W_n|^{-1}N(W_n)$ then with quite similar assumptions, it has been proved, see e.g. Heinrich and Prokešová (2010), that $|W_n|^{1/2}(\widehat{\lambda}^{\mathrm{std}} - \lambda) \to \mathcal{N}(0, \sigma^2)$. It is worth noting that the two estimators $\widehat{\lambda}^{\mathrm{std}}$ and $\widehat{\lambda}^J$ only differ from their asymptotic variance and that the ratio of the asymptotic variances is equal to $\pi/2$. When we estimate the location of a Gaussian sample using the sample mean or the sample median, it is remarkable that the ratio of the asymptotic variances is also $\pi/2$.

Finally, let us add that on the basis of Corollary 4.4, an asymptotic confidence interval of λ can be constructed using a consistent estimator of σ^2. By the previous remark, we can use the kernel-based estimator proposed by Heinrich and Prokešová (2010) (or any other estimator presented in the mentioned paper), which precisely estimates the asymptotic variance of $\widehat{\lambda}^{\mathrm{std}}$, i.e. σ^2.

5 Simulation study

We present in this section a simulation study where, in particular, we intend to compare the median-based estimator defined by (3.5) with the standard moment-based estimator $\widehat{\lambda}^{\mathrm{std}} = N(W)/|W|$. In the end of this section, we also investigate the robust estimator proposed by Berndt and Stoyan (1997).

We focus on the planar case $d = 2$. Three models of spatial point processes are considered:

- Poisson point processes (referred to as POISSON) with intensity λ.
- Log-Gaussian Cox Processes (referred to as LGCP) point processes with exponential covariance function. We fixed the variance to 0.5 and ϕ to 0.02. The parameter μ is fixed by the relation $\mu = \log\lambda - \sigma^2/2$ (see Section 2 for details).
- Poisson hard-core (referred to as PHC) with parameter β and hard core R. The resulting process consists in a conditional Poisson point process with intensity β where we have conditioned on the hard core condition which consists in prohibiting points from being closer than distance R apart. We fixed β to 200 and R to 0.05.

The POISSON model is used as a benchmark. The LGCP model enters into the class of Cox point processes for which we focused a lot in our asymptotic results. We have also considered the Thomas model in a separate simulation study and have obtained quite similar results to the LGCP case. The PHC model is used as an example of repulsive point process model. Note that for fixed β, R the intensity parameter λ of

the process is not explicit. In our simulation study, we fixed $\lambda = 100$ for the POISSON and LGCP models. With the parameters $\beta = 200$ and $R = 0.05$, we estimated the intensity of a PHC to $\lambda \simeq 86$ using 10000 Monte-Carlo replications. The simulations have been performed using the `R` package `spatstat` (Baddeley and Turner, 2005).

To illustrate the performances of (3.5) we generated the point processes on the domain of observation $W_n = [-n, n]^2$ for different values of n and considered the three following settings: let $\mathbf{y}$ be a realization from one of the three models described above, generated on W_n and with m points. The observed point pattern is denoted by $\mathbf{x}$ and is obtained as follows.

(A) Pure case: no modification is considered, $\mathbf{x} = \mathbf{y}$.

(B) A few points are added: in a sub-square Δ_n with side-length $n/5$ included in W_n and randomly chosen, we generated a point process $\mathbf{y}^{\text{add}}$ of $n^{\text{add}} = \rho\, m$ uniform points in Δ_n. We chose $\rho = 0.05$ or 0.1. Then, we defined $\mathbf{x} = \mathbf{y} \cup \mathbf{y}^{\text{add}}$.

(C) A few points are deleted: let $\Delta_n = \cup_{q=1}^{4} \Delta_n^q$ where the Δ_n^q's are the four equally sized squares included in W_n, located in each corner of W_n. The volume of Δ_n is chosen such that $\mathrm{E}(N(\Delta_n)) = \rho\, \mathrm{E}(N(W_n)) = \rho\lambda|W_n|$ and we chose either $\rho = 0.05$ or 0.1. Then, we define $\mathbf{x} = \mathbf{y} \setminus (\mathbf{y} \cap \Delta_n)$, .i.e. $\mathbf{x}$ is the initial configuration thinned by 5% or 10% of its points.

We conducted a Monte-Carlo simulation and generated 1000 replications of the models POISSON, LGCP, PHC and for the three different settings (A)-(C). The observation windows for which we report the empirical results hereafter are $n = 1, 2$. Regarding the setting (C), we placed the squares in which points are thinned at the corners of W_n. By stationarity, the empirical results are the same if we decide to choose them randomly. For each replication, we evaluated $\widehat{\lambda}^{\text{std}}$ and $\widehat{\lambda}^{J}$ for different number of non-overlapping and equally sized cells k_n. More precisely, we chose $k_n = 9, 16, 25, 36, 49$.

Empirical means and standard deviations related to the case (A) are reported in Table 1. We can check that the standard estimator is of course unbiased and that the standard deviation decreases by a factor close to 2, which is equal to $\sqrt{|W_2|/|W_1|}$. The median-based estimator is not theoretically unbiased but the bias is clearly not important and tends to decrease when the observation window grows up. Similarly, the rate of convergence of the empirical standard deviation is not too far from the expected value 2. We also computed separately $\widehat{\mathrm{Var}}(\widehat{\lambda}^{\text{std}})/\widehat{\mathrm{Var}}(\widehat{\lambda}^{J})$ for each value of k_n and n and found interesting that these ratios are not too far from $\pi/2$. Finally, we underline that the choice of the number of cells k_n has little influence on the performances. When $n = 1$, a too large value of k_n seems to increase the bias, especially for the LGCP and PHC model. The differences are however reduced when $n = 2$. We also note that the empirical variance is almost the same whatever the value of k_n.

Tables 2 and 3 are respectively related to the settings (B) and (C) (described above). Both these contaminations (B) or (C) can affect significantly the bias of the estimator. In both tables, we report the bias of the different estimators and the gain (in percent) in terms of mean squared error of the median based-estimator

with respect to the standard one, i.e. for each model and each value of ρ, n, k_n, we computed

$$\widehat{\text{Gain}} = \left(\frac{\widehat{\text{MSE}}(\widehat{\lambda}^{\text{std}}) - \widehat{\text{MSE}}(\widehat{\lambda}^{J})}{\widehat{\text{MSE}}(\widehat{\lambda}^{\text{std}})} \right) \times 100\% \tag{5.1}$$

where $\widehat{\text{MSE}}$ is the empirical mean squared error based on the 1000 replications. Thus a positive (resp. negative) empirical gain means that the median-based estimator is more efficient (resp. less efficient) than the standard procedure.

The standard estimator, based only on the global number of points, is of course not robust to perturbations. It is clearly seen that $\widehat{\lambda}^{\text{std}}$ has a positive bias when we add points (setting (B)) and a negative one when we delete points (setting (C)). This bias is obviously all the more important as ρ (the ratio of points added or deleted) increases. Unlike this, the median-based estimator shows its advantages. When points are added (setting (B)), the estimator $\widehat{\lambda}^{J}$ remains much more stable and is more efficient in terms of MSE except when $\rho = 0.05$ and $n = 1$ for the three models where the empirical results do not lead to clear conclusions. When n and/or ρ increases the gains in percent are quite important and it is worth noticing that results do not fluctuate that much with the choice of the number of cells k_n. We also mention that, in a separate simulation study not shown here, we tried to add a clustered point process or repulsive point process, instead of adding uniform points. The empirical results remained almost unchanged.

Comments regarding Table 3 (setting (C)) are very similar. The results vary a lot when $n = 1$ and $\rho = 0.05$ but as soon as one of this parameter increases, the bias of the median-based estimators are clearly reduced, which implies they outperform significantly the standard estimator.

	Emprical mean (Standard Deviation)					
	$\widehat{\lambda}^{\text{std}}$	$\widehat{\lambda}^{J}$				
		$k_n = 9$	16	25	36	49
POISSON						
$n = 1$	99.6 (4.9)	100.5 (5.9)	101 (5.7)	101.8 (6)	102.8 (6.1)	104 (6)
$n = 2$	99.9 (2.5)	100.2 (3)	100.3 (3)	100.5 (3.1)	100.5 (3.1)	101 (3)
LGCP						
$n = 1$	100.3 (5.5)	101.2 (6.6)	101.8 (6.4)	102.3 (6.6)	103 (6.8)	104.2 (6.8)
$n = 2$	100.1 (2.7)	100.5 (3.2)	100.4 (3.2)	100.6 (3.3)	100.7 (3.3)	101 (3.3)
PHC						
$n = 1$	86 (3)	87.3 (4)	87.7 (3.9)	88.9 (4.1)	90.1 (4.1)	91.7 (4.2)
$n = 2$	86 (1.6)	86.3 (2)	86.4 (2)	86.7 (2)	87 (2)	87.4 (2.1)

Table 1: Empirical means and standard deviations between brackets of estimates of the intensity λ for different models of spatial point processes (POISSON, LGCP, PHC). The empirical results are based on 1000 replications simulated on $[-n, n]^2$ for $n = 1, 2$. The second and third columns correspond to the standard estimator $\widehat{\lambda}^{\text{std}} = N(W_n)/|W_n|$, while the following ones correspond to the median-based estimator (3.5) for different number of cells k_n. The intensity λ equals 100 for the models POISSON,LGCP and (approximately) 86 for the model PHC.

	Bias (Gain of MSE %)					
	std	$k_n = 9$	16	25	36	49
$\rho = 0.05$						
POISSON, $n = 1$	5.4 (0)	4.2 (-14)	4.5 (-9)	4.8 (-19)	6.3 (-53)	7.1 (-75)
$n = 2$	5 (0)	1.9 (50)	1.8 (51)	2.2 (47)	2.3 (47)	2.7 (43)
LGCP, $n = 1$	5 (0)	3.2 (-9)	3.6 (-6)	4.6 (-26)	5.3 (-35)	6.4 (-64)
$n = 2$	5.2 (0)	2.2 (44)	2.2 (49)	2.2 (46)	2.5 (47)	2.8 (39)
PHC, $n = 1$	5 (0)	3.3 (14)	3.8 (5)	4.9 (-19)	6 (-53)	7.6 (-110)
$n = 2$	5 (0)	1.4 (70)	1.6 (72)	1.8 (70)	2.1 (66)	2.5 (57)
$\rho = 0.1$						
POISSON, $n = 1$	10.1 (0)	4.7 (44)	5 (45)	5.7 (35)	6.9 (29)	7.8 (17)
$n = 2$	10.1 (0)	2.6 (79)	2.2 (84)	2.5 (83)	2.7 (83)	2.8 (83)
LGCP, $n = 1$	9.8 (0)	4.4 (38)	5 (41)	5.7 (30)	6.3 (28)	7.2 (21)
$n = 2$	9.8 (0)	2.2 (81)	2.3 (79)	2.2 (82)	2.2 (81)	2.5 (81)
PHC, $n = 1$	10 (0)	3.9 (64)	4.1 (66)	5.3 (57)	6.5 (44)	8 (24)
$n = 2$	10 (0)	1.8 (89)	1.7 (92)	1.9 (91)	2.1 (91)	2.5 (89)

Table 2: Bias and empirical gains in percent between brackets, see (5.1), for the standard and median based estimators for different values of k_n. The empirical results are based on 1000 replications generated on $[-n, n]^2$ for $n = 1, 2$ for the models POISSON, LGCP, PHC where 5% or 10% of points are added to each configuration. This corresponds to the case (B) described in details above.

	Bias (Gain of MSE %)					
	std	$k_n = 9$	16	25	36	49
$\rho = 0.05$						
POISSON, $n = 1$	-5 (0)	-3.9 (-13)	-2.8 (2)	-2 (8)	-0.7 (13)	0.6 (11)
$n = 2$	-4.9 (0)	-4.2 (6)	-3 (34)	-2 (46)	-1 (59)	-0.4 (62)
LGCP, $n = 1$	-5 (0)	-3.9 (0)	-3.1 (-2)	-2.1 (2)	-1.1 (17)	0.1 (18)
$n = 2$	-5 (0)	-4.3 (3)	-3.2 (27)	-2 (46)	-1.2 (60)	-0.6 (62)
PHC, $n = 1$	-4.2 (0)	-2.7 (11)	-1.6 (26)	0 (25)	1.6 (27)	3.4 (-5)
$n = 2$	-4.3 (0)	-3.3 (22)	-1.8 (59)	-0.8 (72)	0 (78)	0.7 (75)
$\rho = 0.1$						
POISSON, $n = 1$	-10 (0)	-8.6 (1)	-6 (32)	-3.3 (53)	-1.1 (63)	-1.4 (65)
$n = 2$	-10 (0)	-7.2 (34)	-3.1 (81)	-1.9 (88)	-1 (88)	-2.4 (86)
LGCP, $n = 1$	-10.4 (0)	-9.3 (2)	-6.4 (35)	-4 (50)	-2.4 (59)	-2.7 (58)
$n = 2$	-10 (0)	-7.6 (23)	-3.7 (73)	-2.3 (82)	-1.4 (86)	-2.7 (81)
PHC, $n = 1$	-8.5 (0)	-6.7 (19)	-3.1 (62)	-0.7 (74)	1.8 (72)	1.9 (74)
$n = 2$	-8.6 (0)	-4.8 (55)	-1.8 (88)	-0.7 (92)	0 (94)	-1 (91)

Table 3: Bias and empirical gains in percent between brackets, see (5.1), for the standard and median based estimators for different values of k_n. The empirical results are based on 1000 replications generated on $[-n, n]^2$ for $n = 1, 2$ for the models POISSON, LGCP, PHC where 5% or 10% of points are deleted to each configuration. This corresponds to the case (C) described in details above.

Finally, we compared our estimator with the one proposed by Berndt and Stoyan

(1997). In particular, we used the version presented by Illian et al. (2008, p. 252). Let G be a grid of $\#G$ dummy points, and for any $g \in G$, let $a(g, \mathbf{X})$ be the Voronoi cell area of the cell corresponding to the closest point of g in $\mathbf{X}$, then $(\#G)^{-1} \sum_g a(g, \mathbf{X})^{-1}$ is (up to edge effects) an unbiased estimator of λ. Illian et al. (2008), proposed to replace the previous sample mean by a sample median or sample trimmed-mean. We investigated the latter estimator, referred to as the Voronoi estimator in the following, in a shorter simulation study. We considered only the POISSON model (similar results were observed for the LGCP and PHC models), fixed the grid G to a regular grid of $\#G = 200^2$ dummy points (the results were very stable to that parameter) and used a symmetrically trimmed mean with a fraction of $f = 0.025, 0.05$ or 0.1 observations trimmed from each end. The `spatstat R` package was used to compute the Voronoi tessellation and cell areas. As suggested by the authors, to correct border effects, we removed all border cells from the analysis. Empirical means and standard deviations for the Voronoi estimator in the settings (A)-(C), described above, are reported in Table 4.

In the setting (A), the Voronoi estimator has a bias comparable to the one of the median-based estimator (see Table 1) when $f = 0.025$ and $f = 0.05$. The bias is surprisingly very large when $f = 0.1$. When extra points are observed (setting (B)) or omitted (setting (C)), the results are less ambiguous: the Voronoi estimator remains strongly biased, even more than with the standard estimator $\widehat{\lambda}^{\text{std}}$ in a few cases when $f = 0.025$ or $f = 0.1$. The choice $f = 0.05$ seems to offer a better compromise. We computed the gains for the Voronoi estimator as we did for the median-based estimator in Tables 2-3 using (5.1). The results, not reported, show that the median-based estimator is more efficient in the settings (A), (B) and (C) when $f = 0.025$ and $f = 0.1$ and slightly more efficient when $f = 0.05$.

From a computational point of view, the Voronoi estimator is more expensive to evaluate. For instance it takes 6 seconds in average to evaluate the Voronoi estimator when $n = 2$ while it takes approximately 0.03 second to evaluate the median-based estimator (for all the values of $k_n = 9, 16, 25, 36, 49$). This precludes from using the Voronoi estimator (at least in that form) for very large point pattern. Asymptotic properties were not the focus of the paper by Berndt and Stoyan (1997) or the book by Illian et al. (2008). But another argument in support of our approach is that we believe that deriving asymptotic results for the Voronoi estimator (consistency, asymptotic variance, central limit theorem) looks more awkward.

		(A)	(B)		(C)	
			$\rho = 0.05$	$\rho = 0.1$	$\rho = 0.05$	$\rho = 0.1$
$f = 0.025$						
	$n = 1$	98.4 (5.8)	104.7 (5.7)	109.7 (5.9)	96 (5.4)	94.9 (5.7)
	$n = 2$	98 (2.6)	104.6 (2.6)	107.3 (2.8)	95.4 (3)	91.1 (3.1)
$f = 0.05$						
	$n = 1$	99.3 (5.5)	103.4 (5.5)	104.9 (5.7)	94.4 (4.9)	92 (5.4)
	$n = 2$	98.8 (2.7)	101.8 (2.5)	103.8 (2.7)	94 (2.5)	91.9 (3)
$f = 0.1$						
	$n = 1$	94.1 (5.1)	97.1 (5)	98.5 (5.1)	91.5 (5.1)	88.9 (5.5)
	$n = 2$	94.5 (2.3)	97.4 (2.4)	97.1 (2.7)	88.9 (2.3)	87.9 (2.7)

Table 4: Empirical means and standard deviations between brackets of estimates of the intensity $\lambda = 100$ based on 1000 replications of the POISSON on $[-n, n]^2$ for $n = 1, 2$ and in the settings (A), (B) and (C) described in the text. The estimator considered in this table is the Voronoi cell estimator proposed by Illian et al. (2008, p.252).

6 Conclusion

In this paper, we propose a median-based estimator of the intensity parameter of a stationary spatial point process. We prove asymptotic properties of this estimator as the observation window expands to $\mathbb{R}^d$. In particular for a large class of models, we show that the estimator $\widehat{\lambda}^J$ satisfies a central limit theory, which allows us to derive asymptotic confidence intervals.

As a general conclusion of the simulation study, it turns out that the estimator $\widehat{\lambda}^J$ confirms expected asymptotic properties and improves the robustness property of the standard procedure. Even if the choice of the tuning parameter k_n has a moderate influence on the empirical results when the observation window is large or when the point pattern is strongly contaminated, it is an open question to propose a data-driven procedure to select the number of cells k_n.

In this paper, we did not aim at detecting outliers or detecting areas where problems are suspected (abundance or lack of points). If the assumption of stationarity seems valid, a large difference between the median-based estimator and a classical estimator of the intensity parameter might allow the user to reconsider the observation window in a second step.

The research contained in this paper leads to interesting open issues: (i) It could be worth continuing the comparison between the Voronoi estimator and our approach. This would require to propose a data-driven procedure to fix the tuning parameter of the fraction of deleted observations, to investigate a more evolved border correction and to derive asymptotic properties. (ii) In the setting (B), we considered outliers as extra points added in a small subsquare. Extra points could be uniformly distributed on the observation domain. Such a situation would be closer to the application in image analysis investigated by Berndt and Stoyan (1997). A similar problem was also considered by Redenbach et al. (2015). (iii) A quantile based estimator is the most natural way of defining a robust estimator. A more advanced technique would consist in extending our theory to M-estimators. (iv) Another step

could be to tackle the problem of robust estimators for second-order characteristics like the K, F or G functions. (v) Finally, extending the methodology and results to the estimation of the intensity of inhomogeneous spatial point processes constitutes also an interesting perspective.

Acknowledgements

The author would like to thank sincerely Michaela Prokešová for discussing the initial idea of this paper, Frédéric Lavancier for discussions on an earlier version of the manuscript and Jérôme Lelong for a careful reading. The author is very grateful to the associate editor and the referee in particular for pointing out interesting references, which enriched the simulation study. The research of the author is partially funded by Persyval-lab EA Oculo-Nimbus.

A Proofs

In all the proofs, κ denotes a generic constant which may vary from line to line. For $k = (k_1, \ldots, k_d)^\top \in \mathbb{Z}^d$, we denote $|k| = \max(|k_1|, \ldots, |k_d|)$.

A.1 Auxiliary results

We present in this section an auxiliary lemma providing a control of the covariance of counting variables and a general central limit theorem adapted to our context.

Lemma A.1. *We define $\mathcal{C}_\tau$ the cube centered at 0 with volume $\tau^d c_n$, i.e.*

$$\mathcal{C}_\tau = \left\{ u = (u_1, \ldots, u_d)^\top \in \mathbb{R}^d : |u_l| \leq \tau c_n^{1/d}/2, l = 1, \ldots, d \right\}.$$

Under Assumption (iii), we have the following three statements.
(a) For any $\tau \in (0, 1]$

$$\mathrm{Var}(N(\mathcal{C}_\tau)) \sim |\mathcal{C}_\tau| \left(\lambda + \lambda^2 \int_{\mathbb{R}^d} (g(w) - 1) \mathrm{d}w \right)$$

as $n \to \infty$.
(b) Let $\varepsilon \in (0, 1)$ then

$$\mathrm{Cov}(N(\mathcal{C}_{1-\varepsilon}), N(\mathcal{C}_1)) \sim \lambda |\mathcal{C}_{1-\varepsilon}| + \lambda^2 |\mathcal{C}_{1-\varepsilon/2}| \int_{\mathbb{R}^d} (g(w) - 1) \mathrm{d}w$$

as $n \to \infty$.
(c) Let $(\varepsilon_n)_{n \geq 1}$ be a sequence of real numbers such that $\varepsilon_n \to 0$ and $c_n^{1/d} \varepsilon_n \to \infty$ as $n \to \infty$, then

$$\begin{aligned} \mathrm{Var}(N(\mathcal{C}_{1-\varepsilon_n})) &\sim \mathrm{Var}(N(\mathcal{C}_1)) \sim \mathrm{Cov}(N(\mathcal{C}_{1-\varepsilon_n}), N(\mathcal{C}_1)) \\ &\sim c_n \left(\lambda + \lambda^2 \int_{\mathbb{R}^d} (g(w) - 1) \mathrm{d}w \right) \end{aligned}$$

as $n \to \infty$.

Proof. (a) is a classical result, see e.g. Heinrich and Prokešová (2010). As we need to refer to specific equations, we report the proof here. By Campbell's Theorem and since $\mathbf{X}$ admits a pair correlation function

$$\begin{aligned}
\mathrm{Var}(N(\mathcal{C}_\tau)) &= \lambda|\mathcal{C}_\tau| + \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \mathbf{1}(u\in\mathcal{C}_\tau)\mathbf{1}(v\in\mathcal{C}_\tau)(g(u-v)-1)\mathrm{d}u\mathrm{d}v \\
&= \lambda|\mathcal{C}_\tau| + \lambda^2\int_{\mathbb{R}^d} |\mathcal{C}_\tau\cap(\mathcal{C}_\tau)_{-w}|(g(w)-1)\mathrm{d}w \\
&= \lambda|\mathcal{C}_\tau| + \lambda^2\int_{\mathcal{C}_{2\tau}} |\mathcal{C}_\tau\cap(\mathcal{C}_\tau)_{-w}|(g(w)-1)\mathrm{d}w \qquad \text{(A.1)} \\
&= \lambda|\mathcal{C}_\tau| + \lambda^2\int_{\mathcal{C}_{2\tau}} \prod_{l=1}^d(\tau c_n^{1/d}-|w_l|)(g((w_1,\dots,w_d)^\top)-1)\mathrm{d}w_1\dots\mathrm{d}w_d \qquad \text{(A.2)} \\
&\sim |\mathcal{C}_\tau|\left(\lambda+\lambda^2\int_{\mathbb{R}^d}(g(w)-1)\mathrm{d}w\right)
\end{aligned}$$

by Assumption (iii).
(b) For brevity, let K_ε denote the covariance to evaluate. Following (a) we have

$$\begin{aligned}
K_\varepsilon &= \lambda|\mathcal{C}_{1-\varepsilon}\cap\mathcal{C}_1| + \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\mathbf{1}(u\in\mathcal{C}_{1-\varepsilon})\mathbf{1}(v\in\mathcal{C}_1)(g(u-v)-1)\mathrm{d}u\mathrm{d}v \\
&= \lambda|\mathcal{C}_{1-\varepsilon}| + \lambda^2\int_{\mathbb{R}^d}|\mathcal{C}_{1-\varepsilon}\cap(\mathcal{C}_1)_{-w}|(g(w)-1)\mathrm{d}w.
\end{aligned}$$

Let $w=(w_1,\dots,w_d)^\top$. We can check that

$$|\mathcal{C}_{1-\varepsilon}\cap(\mathcal{C}_1)_{-w}| = \begin{cases} 0 & \text{if } w\in\mathbb{R}^d\setminus\mathcal{C}_{2-\varepsilon} \\ \prod_{l=1}^d\left((1-\frac{\varepsilon}{2})c_n^{1/d}-|w_l|\right) & \text{if } w\in\mathcal{C}_{2-\varepsilon}\end{cases}$$

whereby we deduce using (A.1)-(A.2) and Assumption (iii) that

$$\begin{aligned}
K_\varepsilon &= \lambda|\mathcal{C}_{1-\varepsilon}| + \lambda^2\int_{\mathcal{C}_{2-\varepsilon}}\prod_{l=1}^d\Big((1-\varepsilon/2)c_n^{1/d}-|w_l|\Big)\Big(g((w_1,\dots,w_d)^\top)-1\Big)\mathrm{d}w_1\dots\mathrm{d}w_d \\
&= \lambda|\mathcal{C}_{1-\varepsilon}| + \lambda^2\int_{\mathcal{C}_{2-\varepsilon}}|\mathcal{C}_{1-\varepsilon/2}\cap(C_{1-\varepsilon/2})_{-w}|(g(w)-1)\mathrm{d}w \\
&\sim \lambda|\mathcal{C}_{1-\varepsilon}| + \lambda^2|\mathcal{C}_{1-\varepsilon/2}|\int_{\mathbb{R}^d}(g(w)-1)\mathrm{d}w
\end{aligned}$$

as $n\to\infty$.
(c) The assumptions on the sequence (ε_n) allow us to apply (a)-(b) which leads to the result since $|\mathcal{C}_1|\sim|\mathcal{C}_{1-\varepsilon_n}|\sim|\mathcal{C}_{1-\varepsilon_n/2}|\sim c_n$ as $n\to\infty$. □

Now we present a central limit theorem for stationary random fields with asymptotic covariance matrix not necessarily positive definite. It is very close to Guyon (1991, Theorem 3.3.1) and to Karáczony (2006, Theorem 1) but we were not able to find it in the following form in the literature.

For two square matrices A, B, $A\geq B$ (resp. $A>B$) means that $A-B$ is a positive (resp. positive definite) matrix. Finally, $\|A\|$ stands for the Frobenius norm of A given by $\|A\| = \mathrm{Tr}(A^\top A)^{1/2}$.

Theorem A.2. *Let $(X_k, k \in \mathbb{Z}^d)$ be a stationary random field in a measurable space S. Let $\mathcal{K}_n \subset \mathbb{Z}^d$ with $k_n = |\mathcal{K}_n| \to \infty$ as $n \to \infty$. For any $n \geq 1$ and $k \in \mathcal{K}_n$, we define $Y_{n,k} = f_{n,k}(X_k)$ where $f_{n,k} : S \to \mathbb{R}^p$ for some $p \geq 1$ is a measurable function. We denote by $S_n = \sum_{k \in \mathcal{K}_n} Y_{n,k}$ and by $\Sigma_n = \mathrm{Var}(S_n)$ and assume that for any $n \geq 1$, $k \in \mathcal{K}_n$, $\mathrm{E}\, Y_{n,k} = 0$. We also assume that*
(I) $\sup_{n \geq 1} \sup_{k \in \mathcal{K}_n} \|Y_{n,k}\|_\infty < \infty$.
(II)There exists $\eta > 0$ such that $\alpha_{2,\infty}(m) = \mathcal{O}(m^{-d(1+\eta)})$.
(III) There exists $\Sigma \geq 0$ a (p,p) matrix with rank $1 \leq r \leq p$ such that $k_n^{-1}\Sigma_n \to \Sigma$ as $n \to \infty$.
Then, $k_n^{-1/2} S_n \to \mathcal{N}(0, \Sigma)$ in distribution as $n \to \infty$.

We present Theorem A.2 for bounded random vectors and with only one mixing coefficient, namely $\alpha_{2,\infty}$. It can obviously be generalized along similar lines as in Guyon (1991, Theorem 3.3.1).

Proof. Assume $\Sigma > 0$, then for n large enough $k_n^{-1}\Sigma_n \geq \Sigma/2 > 0$, which combined with Assumptions (I)-(II) allows us to apply Karáczony (2006, Theorem 1) to conclude the result.

The end of the proof follows the same arguments as the proof of a central limit theorem for triangular arrays of conditionally centered random fields obtained by Coeurjolly and Lavancier (2013, Theorem 2). If Σ is not positive definite, we can find an orthonormal basis $(h_1, \ldots, h_p)$ of $\mathbb{R}^p$ where the f_i's are eigenvectors of Σ. We let $(f_1, \ldots, f_r)$ be the basis of the image of Σ and $(f_{r+1}, \ldots, f_p)$ be the basis of its kernel. Let also H_{Im} (resp. H_{Ker}) be the matrix formed by the column vectors of $(f_1, \ldots, f_r)$ (resp. $(f_{r+1}, \ldots, f_p)$). Similarly for $v \in \mathbb{R}^p$, we denote by v_j its jth coordinate in the basis of $(f_1, \ldots, f_p)$, $v_{Im} = (v_1, \ldots, v_r)$ and $v_{Ker} = (v_{r+1}, \ldots, v_p)$. Using the Cramer-Wold device, we need to prove that for any $v \in \mathbb{R}^p$, $v^\top k_n^{-1/2} S_n$ converges towards a Gaussian random variable. We have

$$v^\top k_n^{-1/2} S_n = v_{Im}^\top H_{Im}^\top k_n^{-1/2} S_n + v_{Ker}^\top H_{Ker}^\top k_n^{-1/2} S_n.$$

Let $S'_n = \sum_k Y'_{n,k}$ where $Y'_{n,k} = H_{Im}^\top Y_{n,k}$. The random variables $Y'_{n,k}$ are bounded variables for any $n \geq 1$ and $k \in \mathcal{K}_n$. By assumption (III), $k_n^{-1} \mathrm{Var}(S'_n) \to H_{Im}^\top \Sigma H_{Im}$ which is a positive definite matrix since $r \geq 1$. Therefore from the first part of the proof, $v_{Im}^\top H_{Im}^\top k_n^{-1/2} S_n$ tends to a Gaussian random variable in distribution as $n \to \infty$. By Slutsky's Lemma (see e.g. Van der Vaart (2000)), the proof will be done if $v_{Ker}^\top H_{Ker}^\top k_n^{-1/2} S_n$ tends to 0 in probability as $n \to \infty$. Since, $H_{Ker}^\top \Sigma H_{Ker} = 0$, the expected convergence follows from

$$\begin{aligned}
\mathrm{Var}(v_{Ker}^\top H_{Ker}^\top k_n^{-1/2} S_n) &= v_{Ker}^\top H_{Ker}^\top k_n^{-1} \Sigma_n H_{Ker} v_{Ker} \\
&= v_{Ker}^\top H_{Ker}^\top (k_n^{-1}\Sigma_n - \Sigma) H_{Ker} v_{Ker} \\
&\leq \|v_{Ker}\| \, \|H_{Ker}\| \, \|k_n^{-1}\Sigma_n - \Sigma\|
\end{aligned}$$

which tends to 0 by Assumption (III). □

A.2 Proof of Proposition 4.1

Proof. Assumption (ii-1) corresponds to Proposition 3.2.

Assumptions (ii-2) and (ii-3). By definition of $\mathbf{X}$,

$$\sqrt{2\pi\lambda c_n}\mathrm{P}(N(C_{n,0}) = \lfloor t_n \rfloor \mid \xi) = \frac{\left(\int_{C_{n,0}} \xi(s)\mathrm{d}s\right)^{\lfloor t_n \rfloor} e^{-\int_{C_{n,0}} \xi(s)\mathrm{d}s}}{\lfloor t_n \rfloor^{\lfloor t_n \rfloor} e^{-\lfloor t_n \rfloor}} v_n$$

where

$$v_n = \sqrt{\frac{\lambda c_n}{\lfloor t_n \rfloor}} \frac{\sqrt{2\pi}\lfloor t_n \rfloor^{\lfloor t_n \rfloor + 1/2} e^{-\lfloor t_n \rfloor}}{\lfloor t_n \rfloor!}.$$

Since $t_n/(\lambda c_n) \to 1$ as $n \to \infty$, then using Stirling's Formula we obviously have $v_n \to 1$ as $n \to \infty$. Now using the notation $T_n = \lfloor t_n \rfloor^{-1} \int_{C_{n,0}} \xi(s)\mathrm{d}s$, we rewrite the first equation as follows

$$(v_n)^{-1}\sqrt{2\pi\lambda c_n}\mathrm{P}(N(C_{n,0}) = \lfloor t_n \rfloor \mid \xi) = T_n^{\lfloor t_n \rfloor} e^{\lfloor t_n \rfloor (1-T_n)} = A_n B_n$$

where A_n and B_n are defined by

$$A_n = e^{-\lfloor t_n \rfloor (T_n - 1)^2/2} \quad \text{and } B_n = e^{\lfloor t_n \rfloor \left(\log T_n - (T_n - 1) + (T_n - 1)^2/2\right)}.$$

Let η be given by Assumption (iv). Since $\mathrm{E}\,|\xi(0)|^{2+\delta} < \infty$ for some $\delta > 2/\eta$, we are ensured that $\alpha_{2,\infty} = \mathcal{O}(m^{-\nu})$ for some $\nu > d(2+\delta)/\delta$. Therefore, we can apply Guyon (1991, Theorem 3.3.1) and show that there exists $\tau > 0$ such that $\sqrt{\lambda c_n}(I_n - 1) \to \mathcal{N}(0, \tau^2)$ in distribution where $I_n = (\lambda c_n)^{-1} \int_{C_{n,0}} \xi(s)\mathrm{d}s$. To compute τ^2, we observe that using the definition of a Cox point process

$$\begin{aligned}\mathrm{Var}(N(C_{n,0})) &= \mathrm{E}\left(\mathrm{Var}(N(C_{n,0}) \mid \xi)\right) + \mathrm{Var}\left(\mathrm{E}\left(N(C_{n,0})\right) \mid \xi\right) \\ &= \lambda c_n + \mathrm{Var} \int_{C_{n,0}} \xi(s)\mathrm{d}s.\end{aligned}$$

We use Assumption (iii) and Lemma A.1 (a) to deduce that as $n \to \infty$

$$\mathrm{Var} \int_{C_{n,0}} \xi(s)\mathrm{d}s \sim \lambda^2 c_n \int_{\mathbb{R}^d} (g(w) - 1)\mathrm{d}w$$

which leads to $\mathrm{Var}(\sqrt{\lambda c_n} I_n) \to \tau^2 = \lambda \int_{\mathbb{R}^d} (g(w) - 1)\mathrm{d}w$ as $n \to \infty$. From the definition of t_n and Slutsky's Lemma, it can be shown that $\sqrt{\lfloor t_n \rfloor}(T_n - I_n) \to 0$ in probability which leads to $T_n \to 1$ in probability and $\sqrt{\lfloor t_n \rfloor}(T_n - 1) \to \mathcal{N}(0, \tau^2)$ in distribution. We deduce that $A_n \to A = e^{-\tau^2 L^2/2}$ in distribution, where $L \sim \mathcal{N}(0,1)$, which, by the uniform integrability of the sequence $(A_n)_n$, leads to $A_n \to A$ in L^1. Now a Taylor expansion shows that there exists $\widetilde{T}_n \in (0 \wedge (T_n - 1), 0 \vee (T_n - 1))$ such that

$$|\log(B_n)| = \lfloor t_n \rfloor |T_n - 1| \frac{\widetilde{T}_n^2}{1 + \widetilde{T}_n} \leq \lfloor t_n \rfloor (T_n - 1)^2 \frac{|T_n - 1|}{\widetilde{T}_n + 1}.$$

It is clear that $\widetilde{T}_n$ tends to 0 in probability as $n \to \infty$, which yields that $\log(B_n) \to 0$ and $B_n \to 1$ in probability by Slutsky's Lemma. Again, the uniform integrability assumption of the sequence $(B_n)_n$ implies that $B_n \to 1$ in L^1. Since

$$|A_n B_n - A| \leq A_n |B_n - 1| + |A_n - A| \leq |B_n - 1| + |A_n - A|$$

we conclude that $A_nB_n \to A$ in L^1 as $n \to \infty$. In other words as $n \to \infty$

$$\sqrt{2\pi\lambda c_n}\mathrm{P}(N(C_{n,0}) = \lfloor t_n \rfloor) \sim v_n^{-1}\,\mathrm{E}\left(\sqrt{2\pi\lambda c_n}\mathrm{P}(N(C_{n,0}) = \lfloor t_n \rfloor \mid \xi)\right) \to \mathrm{E}(A).$$

Using the definition of the moment generating function of a χ_1^2 distribution, we have $\mathrm{E}(A) = (1+\tau^2)^{-1/2}$ whereby we deduce that

$$\sqrt{c_n}\mathrm{P}(N(C_{n,0}) = \lfloor t_n \rfloor) \to \left(2\pi\lambda(1+\tau^2)\right)^{-1/2} = \left(2\pi\sigma^2\right)^{-1/2}$$

with $\sigma^2 = \lambda + \lambda^2 \int_{\mathbb{R}^d}(g(w)-1)\mathrm{d}w$. □

A.3 Proof of Theorem 4.2

Proof. We focus only on (a) as (b) follows from (a), Slutsky's Lemma and Assumption (ii-2). Let $t_n = \lambda c_n + a_n$. By definition

$$\widehat{F}(t_n;\mathbf{Z}) - F_Z(t_n) = \frac{1}{k_n}\sum_{k\in\mathcal{K}_n}\left(\mathbf{1}(Z_{n,k} \le t_n) - \mathrm{P}(Z_{n,k} \le t_n)\right). \tag{A.3}$$

Let $(\varepsilon_n)_{n\ge1}$ be a sequence of real numbers such that $\varepsilon_n \to 0$ and $\varepsilon_n c_n^{1/d} \to \infty$ as $n \to \infty$.We denote by $Z_{n,k}^- = N(C_{n,k}^-) + U_k$ where $C_{n,k}^-$ is the erosion of the cell $C_{n,k}$ by a closed ball with radius $\varepsilon_n c_n^{1/d}$. Two cells $C_{n,k}^-$ and $C_{n,k'}^-$ for $k, k' \in \mathcal{K}_n$ $(k \ne k')$ are therefore at distance greater than $2\varepsilon_n c_n^{1/d}$. To prove Theorem 4.2 (a), we use the blocking technique introduced by Ibragimov and Linnik (1971) and applied to spatial point processes by Guan and Loh (2007); Guan et al. (2007) and Prokešová and Jensen (2013). To this end, we need additional notation. For any $n \ge 1$ and $k \in \mathcal{K}_n$, let $t_n^- = \lambda|C_{n,k}^-| + 1/2 = \lambda(1-\varepsilon_n)^d c_n + 1/2$ and let $(\widetilde{Z}_{n,k}^-, k \in \mathcal{K}_n)$ be a collection of independent random variables such that $\widetilde{Z}_{n,k}^- \overset{d}{=} Z_{n,k}^-$. We decompose the sum in (A.3) as follows

$$\sum_{k\in\mathcal{K}_n}\left(\mathbf{1}(Z_{n,k} \le t_n) - \mathrm{P}(Z_{n,k} \le t_n)\right) = D_n + S_n^- + \widetilde{S}_n^- \tag{A.4}$$

where

$$\begin{aligned}
D_n &= \sum_{k\in\mathcal{K}_n} D_{n,k} = \sum_{k\in\mathcal{K}_n}\left\{\mathbf{1}(Z_{n,k} \le t_n) - \mathrm{P}(Z_{n,k} \le t_n) - \mathbf{1}(Z_{n,k}^- \le t_n^-) + \mathrm{P}(Z_{n,k}^- \le t_n^-)\right\}\\
S_n^- &= \sum_{k\in\mathcal{K}_n}\mathbf{1}(Z_{n,k}^- \le t_n^-) - \mathrm{P}(Z_{n,k}^- \le t_n^-)\\
\widetilde{S}_n^- &= \sum_{k\in\mathcal{K}_n}\mathbf{1}(\widetilde{Z}_{n,k}^- \le t_n^-) - \mathrm{P}(\widetilde{Z}_{n,k}^- \le t_n^-).
\end{aligned}$$

We split the proof into three steps. As $n \to \infty$, we prove that
Step 1. $D_n/\sqrt{k_n} \to 0$ in probability.
Step 2. for any $u \in \mathbb{R}$, $\phi_n^-(u) - \widetilde{\phi}_n^-(u) \to 0$ as $n \to \infty$ where $i = \sqrt{-1}$, $\phi_n^-(u) = \mathrm{E}(e^{iuS_n^-/\sqrt{k_n}})$ and $\widetilde{\phi}_n^-(u) = \mathrm{E}(e^{iu\widetilde{S}_n^-/\sqrt{k_n}})$, which will imply that $(S_n^- - \widetilde{S}_n^-)/\sqrt{k_n} \to 0$.

Step 3. $\widetilde{S}_n^-/\sqrt{k_n} \to \mathcal{N}(0,1/4)$ in distribution.

The conclusion will follow directly from Steps 1-3, (A.3)-(A.4) and Slutsky's Lemma.

Step 1. To achieve this step, we will prove that $k_n^{-1}\,\mathrm{E}(D_n^2) = k_n^{-1}\,\mathrm{Var}(D_n) \to 0$ as $n \to \infty$. We have

$$\frac{1}{k_n}\,\mathrm{Var}(D_n) = \frac{1}{k_n}\sum_{\substack{k,k'\in\mathcal{K}_n\\ |k-k'|\le 1}} \mathrm{Cov}(D_{n,k}, D_{n,k'}) + \frac{1}{k_n}\sum_{\substack{k,k'\in\mathcal{K}_n\\ |k-k'|>1}} \mathrm{Cov}(D_{n,k}, D_{n,k'}).$$

Let $k, k' \in \mathcal{K}_n$ with $k \neq k'$, Assumption (i) asserts that $d(C_{n,k}, C_{n,k'}) = |k-k'-1|c_n^{1/d}$. Since $D_{n,k} \in \mathcal{F}(C_{n,k})$ and $D_{n,k'} \in \mathcal{F}(C_{n,k'})$, we have from Zhengyan and Chuanrong (1996, Lemma 2.1)

$$\begin{aligned}\mathrm{Cov}(D_{n,k}, D_{n,k'}) &\le 4\alpha_{c_n,c_n}(|k-k'-1|c_n^{1/d})\\ &\le 4c_n\alpha(|k-k'-1|c_n^{1/d}) = \mathcal{O}(|k-k'-1|^{-d(1+\eta)}c_n^{-\eta}).\end{aligned}$$

Since the series $\sum_{k\in\mathbb{Z}^d\setminus\{0\}} |k|^{-d(1+\eta)}$ converges, it is clear that

$$\frac{1}{k_n}\sum_{\substack{k,k'\in\mathcal{K}_n\\ |k-k'|>1}} \mathrm{Cov}(D_{n,k}, D_{n,k'}) = \mathcal{O}(c_n^{-\eta}), \tag{A.5}$$

which tends to 0 as $n \to \infty$. Since the variables $D_{n,k}$ are identically distributed, we get from the Cauchy-Schwarz inequality

$$\begin{aligned}\Big|\frac{1}{k_n}\sum_{\substack{k,k'\in\mathcal{K}_n\\ |k-k'|\le 1}} \mathrm{Cov}(D_{n,k}, D_{n,k'})\Big| &\le \frac{1}{k_n}\sum_{\substack{k,k'\in\mathcal{K}_n\\ |k-k'|\le 1}} \sqrt{\mathrm{Var}(D_{n,k})\,\mathrm{Var}(D_{n,k'})}\\ &\le \mathrm{Var}(D_{n,0})\frac{1}{k_n}\sum_{\substack{k,k'\in\mathcal{K}_n\\ |k-k'|\le 1}} 1\\ &\le 3^d\,\mathrm{Var}(D_{n,0}).\end{aligned}$$

Thus, Step 1 is achieved once we prove that $\mathrm{Var}(D_{n,0}) \to 0$ as $n \to \infty$. A straightforward calculation yields that

$$\begin{aligned}\mathrm{Var}(D_{n,0}) = \mathrm{P}(Z_{n,0} \le t_n)(1-\mathrm{P}(Z_{n,0} \le t_n)) + \mathrm{P}(Z_{n,0}^- \le t_n^-)(1-\mathrm{P}(Z_{n,0}^- \le t_n^-))\\ + 2\mathrm{P}(Z_{n,0} \le t_n)\mathrm{P}(Z_{n,0}^- \le t_n^-) - 2\mathrm{P}\big(Z_{n,0} \le t_n, Z_{n,0}^- \le t_n^-\big).\end{aligned}$$

Let Δ_j be the unit cube centered at $j \in \mathbb{Z}^d$ and let $\mathcal{J}_n = \{j \in \mathbb{Z}^d : \Delta_j \cap C_{n,0} \neq \emptyset\}$. We denote by $Y_{n,j}$ the random vector

$$Y_{n,j} = \left(\frac{U_0}{j_n} + \mathbf{1}(u \in C_{n,0} \cap \Delta_j), \frac{U_0}{j_n} + \mathbf{1}(u \in C_{n,0}^- \cap \Delta_j)\right)^\top$$

where $j_n = |\mathcal{J}_n|$ satisfies $j_n \sim c_n$ as $n \to \infty$. We have $(Z_{n,0}, Z_{n,0}^-)^\top = \sum_{j\in\mathcal{J}_n} Y_{n,j}$ and we note that $\sup_{n\geq 1}\sup_{j\in\mathcal{J}_n}\|Y_{n,j}\|_\infty < \infty$. Since $\varepsilon_n \to 0$ and $c_n^{1/d}\varepsilon_n \to \infty$ as $n \to \infty$, we can apply Lemma A.1 (c) to derive

$$\mathrm{Var}(Z_{n,0}) \sim \mathrm{Var}(Z_{n,0}^-) \sim \mathrm{Cov}(Z_{n,0}, Z_{n,0}^-) \sim \frac{1}{12} + \sigma^2 c_n$$

where $\sigma^2 = \lambda + \lambda^2 \int_{\mathbb{R}^d}(g(w)-1)\mathrm{d}w$. In other words,

$$j_n^{-1}\,\mathrm{Var}((Z_{n,0}, Z_{n,0}^-)^\top) \to \Sigma = \sigma^2 \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \tag{A.6}$$

which is a matrix with rank 1. By combining this with Assumption (iv), we can apply Theorem A.2 to get as $n \to \infty$

$$c_n^{-1/2}\left(Z_{n,0} - \mathrm{E}(Z_{n,0}), Z_{n,0}^- - \mathrm{E}(Z_{n,0}^-)\right)^\top \to \mathcal{N}(0,\Sigma)$$

in distribution. Since $t_n^- = \mathrm{E}(Z_{n,0}^-)$ and $\mathrm{E}(Z_{n,0}) - t_n = 1/2 - a_n = o(c_n^{1/2})$ by definition of t_n, an application of Slutsky's Lemma yields that

$$c_n^{-1/2}\left(Z_{n,0} - t_n, Z_{n,0}^- - t_n^-)\right)^\top \to \mathcal{N}(0,\Sigma)$$

in distribution as $n \to \infty$ whereby we deduce that

$$\mathrm{P}(Z_{n,0} \leq t_n) \to 1/2 \quad \text{and} \quad \mathrm{P}(Z_{n,0}^- \leq t_n^-) \to 1/2. \tag{A.7}$$

Rose and Smith (1996) proved that if $U = (U_1, U_2)^\top$ follows a bivariate normal distribution with mean 0, variance 1 and correlation ρ, $\mathrm{P}(U_1 \leq 0, U_2 \leq 0) = 1/4 + \sin^{-1}(\rho)/2\pi$ which equals to $1/2$ when $\rho = 1$. From (A.6), this shows that $\mathrm{P}(Z_{n,0} \leq t_n, Z_{n,0}^- \leq t_n^-) \to 1/2$ as $n \to \infty$. As a consequence, $\mathrm{Var}(D_{n,0}) \to 0$ which combined with (A.5) leads to $k_n^{-1}\,\mathrm{Var}(D_n) \to 0$ as $n \to \infty$.

Step 2. This step is the core of the blocking technique. Let h denote a bijection from $\mathcal{K}_n$ to $\{1,\ldots,k_n\}$. Let $j \in \{1,\ldots,k_n\}$ and $V_j = e^{iu(\mathbf{1}(Z_{n,h^{-1}(j)} \leq t_n^-) - \mathrm{P}(Z_{n,h^{-1}(j)} \leq t_n^-))/\sqrt{k_n}}$. Then

$$\phi_n^-(u) = \mathrm{E}\prod_{j=1}^{k_n} V_j \quad \text{and} \quad \widetilde{\phi}_n^-(u) = \prod_{j=1}^{k_n}\mathrm{E}(V_j).$$

and

$$|\phi_n^-(u) - \widetilde{\phi}_n^-(u)| \leq \sum_{j=1}^{k_n-1}\Big|\,\mathrm{E}\,\big(\prod_{s=1}^{j+1} V_s\big) - \mathrm{E}\,\big(\prod_{s=1}^{j} V_s\big)\,\mathrm{E}(V_{j+1})\Big|.$$

Let $j \in \{1,\ldots,k_n-1\}$ and $A_j = \prod_{s=1}^j V_s$. Clearly, $A_j \in \mathcal{F}(\cup_{s=1}^j C_{n,h^{-1}(s)}^-)$ and $V_{j+1} \in \mathcal{F}(C_{n,h^{-1}(j+1)}^-)$, $|\cup_{s=1}^j C_{n,h^{-1}(s)}^-| = j(1-\varepsilon_n)^d c_n$, $|C_{n,h^{-1}(j+1)}^-| = (1-\varepsilon_n)^d c_n$ and $d(\cup_{s=1}^j C_{n,h^{-1}(s)}^-, C_{n,h^{-1}(j+1)}^-) \geq 2\varepsilon_n c_n^{1/d}$. Since A_j and V_{j+1} are bounded random

variables, we have the following upper-bound on their covariance by means of the strong mixing coefficient, see Zhengyan and Chuanrong (1996, Lemma 2.1)

$$\begin{aligned}\mathrm{Cov}(A_j, V_{j+1}) &\leq 4\alpha_{j(1-\varepsilon_n)^d c_n,(1-\varepsilon_n)^d c_n}(2\varepsilon_n c_n^{1/d}) \\ &\leq 4jc_n \sup_p \frac{\alpha_{p,p}(2\varepsilon_n c_n^{1/d})}{p} \\ &\leq 4c_n k_n \mathcal{O}(\varepsilon_n^{-d(1+\eta)} c_n^{-(1+\eta)}) = \mathcal{O}(k_n \varepsilon_n^{-d(1+\eta)} c_n^{-\eta})\end{aligned}$$

whereby we deduce that $|\phi_n^-(u) - \widetilde{\phi}_n^-(u)| = \mathcal{O}(k_n^2 c_n^{-\eta} \varepsilon_n^{-d(1+\eta)})$. Now we can fix the sequence $(\varepsilon_n)_{n\geq 1}$. Specifically, we set $\varepsilon_n = c_n^{(\eta'-\eta)/d(1+\eta)}$ for some $0 < \eta' < \eta$. This choice ensures that $\varepsilon_n \to 0$, $c_n^{1/d}\varepsilon_n = c_n^{(1+\eta')/d(1+\eta)} \to \infty$ and yields that $|\phi_n^-(u) - \widetilde{\phi}_n^-(u)| = \mathcal{O}(k_n^2/c_n^{\eta'})$ which tends to 0 as $n \to \infty$ by Assumption (i).

Step 3. Since $\widetilde{Z}_{n,k}^- \stackrel{d}{=} Z_{n,k}^-$ and since $\mathrm{P}(Z_{n,k}^- \leq t_n^-) \to 1/2$ as $n \to \infty$ from Step 2, we deduce that

$$\begin{aligned}\mathrm{Var}(\widetilde{S}_n^-) &= \sum_{k\in\mathcal{K}_n} \mathrm{P}(\widetilde{Z}_{n,k}^- \leq t_n^-)(1 - \mathrm{P}(\widetilde{Z}_{n,k}^- \leq t_n^-)) \\ &= k_n \mathrm{P}(\widetilde{Z}_{n,0}^- \leq t_n^-)(1 - \mathrm{P}(\widetilde{Z}_{n,0}^- \leq t_n^-)) \sim k_n/4\end{aligned}$$

as $n \to \infty$. Since $(\mathbf{1}(\widetilde{Z}_{n,k}^- \leq t_n^-), k \in \mathcal{K}_n)$ is a collection of bounded and independent random variables, Step 3 follows from an application of Lyapounov Theorem. □

A.4 Proof of Theorem 4.3

Proof. (a) Let us define for any $t \geq 0$

$$A_n = \sqrt{\frac{k_n}{c_n}}\left(\widehat{\mathrm{Me}}(\mathbf{Z}) - \mathrm{Me}_Z\right) \quad \text{and} \quad B_n(t) = \sqrt{\frac{k_n}{c_n}}\left(\frac{F_Z(t) - \widehat{F}(t;\mathbf{Z})}{f_Z(\mathrm{Me}_Z)}\right).$$

We have to prove that $A_n - B_n(\mathrm{Me}_Z)$ converges in probability to 0 as $n \to \infty$. The proof is based on the application of Ghosh (1971, Lemma 1) which consists in satisfying the two following conditions:
(I) for all $\delta > 0$, there exists $\varepsilon = \varepsilon(\delta)$ such that $\mathrm{P}(|B_n(\mathrm{Me}_Z)| > \varepsilon) < \delta$.
(II) for all $y \in \mathbb{R}$ and $\varepsilon > 0$

$$\lim_{n\to\infty} \mathrm{P}(A_n \leq y, B_n(\mathrm{Me}_Z) \geq y + \varepsilon) = \lim_{n\to\infty} \mathrm{P}(A_n \geq y + \varepsilon, B_n(\mathrm{Me}_Z) \leq y) = 0.$$

In particular (I) is fulfilled if we prove that $\mathrm{Var}\, B_n(\mathrm{Me}_Z) = \mathcal{O}(1)$. The proof of Theorem 4.2 shows that $\mathrm{Var}\, \widehat{F}(\mathrm{Me}_Z;\mathbf{Z}) = \mathcal{O}(k_n^{-1})$ as $n \to \infty$. By Assumption (ii-3), we obtain

$$\mathrm{Var}\, B_n(\mathrm{Me}_Z) = \frac{1}{c_n f_Z(\mathrm{Me}_Z)^2}\, \mathrm{Var}(\sqrt{k_n}\widehat{F}(\mathrm{Me}_Z;\mathbf{Z})) = \mathcal{O}(1).$$

(II) Let $y \in \mathbb{R}$ (and without loss of generality, assume $y \geq 0$). By definition of the sample median, we have

$$\begin{aligned}\{A_n \leq y\} &= \left\{\widehat{\mathrm{Me}}(\mathbf{Z}) \leq \mathrm{Me}_Z + y\sqrt{c_n/k_n}\right\} \\ &= \left\{\frac{1}{2} \leq \widehat{F}\left(\mathrm{Me}_Z + y\sqrt{c_n/k_n}\right)\right\} \\ &= \left\{B_n\left(\mathrm{Me}_Z + y\sqrt{c_n/k_n}\right) \leq y_n\right\}\end{aligned}$$

where

$$y_n = \sqrt{k_n/c_n}\,\frac{1}{f_Z(\mathrm{Me}_Z)}\left(F_Z\left(\mathrm{Me}_Z + y\sqrt{c_n/k_n}\right) - F_Z(\mathrm{Me}_Z)\right).$$

Now, we intend to prove that as $n \to \infty$, $y_n \to y$ and $\widetilde{B}_n = B_n(\mathrm{Me}_Z + y\sqrt{c_n/k_n}) - B_n(\mathrm{Me}_Z) \to 0$ in probability. First, since Z admits a density everywhere, there exists $\tau_n \in (\mathrm{Me}_Z, \mathrm{Me}_Z + y\sqrt{c_n/k_n})$ such that $y_n = y\ f_Z(\tau_n)/f_Z(\mathrm{Me}_Z)$. From (3.4)

$$\frac{f_Z(\tau_n)}{f_Z(\mathrm{Me}_Z)} = \frac{\mathrm{P}(N(C_{n,0}) = \lfloor \tau_n \rfloor)}{\mathrm{P}(N(C_{n,0}) = \lfloor \mathrm{Me}_Z \rfloor)},$$

which tends to 1 by Assumptions (ii-1)-(ii-2) and implies the convergence of y_n towards y. Second, we show that $\mathrm{Var}(\widetilde{B}_n) \to 0$ as $n \to \infty$ by decomposing the variance as follows. Let $\widetilde{B}_{n,k} = \mathbf{1}(\mathrm{Me}_Z \leq Z_{n,k} \leq \mathrm{Me}_Z + y\sqrt{c_n/k_n}) - \mathrm{P}(\mathrm{Me}_Z \leq Z_{n,k} \leq \mathrm{Me}_Z + y\sqrt{c_n/k_n})$

$$\begin{aligned}\mathrm{Var}(\widetilde{B}_n) &= \frac{1}{c_n f_Z(\mathrm{Me}_Z)^2}\,\frac{1}{k_n}\sum_{k,k' \in \mathcal{K}_n} \mathrm{Cov}(\widetilde{B}_{n,k}, \widetilde{B}_{n,k'}) \\ &\leq \frac{\kappa}{k_n}\sum_{\substack{k,k' \in \mathcal{K}_n \\ |k-k'| \leq 1}} |\,\mathrm{Cov}(\widetilde{B}_{n,k}, \widetilde{B}_{n,k'})| + \frac{\kappa}{k_n}\sum_{\substack{k,k' \in \mathcal{K}_n \\ |k-k'| > 1}} |\,\mathrm{Cov}(\widetilde{B}_{n,k}, \widetilde{B}_{n,k'})|. \qquad \text{(A.8)}\end{aligned}$$

We follow the proof of Step 1 of Theorem 4.2. For any $k, k' \in \mathcal{K}_n$ $k \neq k'$, $\mathrm{Cov}(\widetilde{B}_{n,k}, \widetilde{B}_{n,k'}) = \mathcal{O}(|k-k'-1|^{-d(1+\eta)} c_n^{-\eta})$. So

$$\frac{1}{k_n}\sum_{\substack{k,k' \in \mathcal{K}_n \\ |k-k'| > 1}} |\,\mathrm{Cov}(\widetilde{B}_{n,k}, \widetilde{B}_{n,k'})| = \mathcal{O}(c_n^{-\eta})$$

which tends to 0 as $n \to \infty$. The first double sum in (A.8) is upper-bounded by $3^d \kappa\, \mathrm{Var}(\widetilde{B}_{n,0})$ and

$$\mathrm{Var}(\widetilde{B}_{n,0}) = \mathrm{P}(\mathrm{Me}_Z \leq Z_{n,0} \leq \mathrm{Me}_Z + y\sqrt{c_n/k_n})\big(1 - \mathrm{P}(\mathrm{Me}_Z \leq Z_{n,0} \leq \mathrm{Me}_Z + y\sqrt{c_n/k_n})\big).$$

By Assumption (i)-(ii), $\mathrm{Me}_Z = \lambda c_n + o(\sqrt{c_n})$ and $\mathrm{Me}_Z + y\sqrt{c_n/k_n} = \lambda c_n + o(\sqrt{c_n})$ for every $y \in \mathbb{R}$. So we can apply (A.7) which leads to $\mathrm{P}(Z_{n,0} \geq \mathrm{Me}_Z) \to 1/2$, $\mathrm{P}(Z_{n,0} \leq \mathrm{Me}_Z + y\sqrt{c_n/k_n}) \to 1/2$ and finally to $\mathrm{Var}(\widetilde{B}_{n,0}) \to 0$ and $\widetilde{B}_n \to 0$ in probability as $n \to \infty$.

Now we can conclude. For all $\varepsilon > 0$, there exists $n_0(\varepsilon)$ such that for all $n \geq n_0(\varepsilon)$, $y_n \leq y + \varepsilon/2$. Therefore for $n \geq n_0(\varepsilon)$

$$\begin{aligned}
\mathrm{P}(A_n \leq y,\ B_n(\mathrm{Me}_Z) \geq y+\varepsilon) &= \mathrm{P}(B_n(\mathrm{Me}_Z + y\sqrt{c_n/k_n}) \leq y_n,\ B_n(\mathrm{Me}_Z) \geq y+\varepsilon) \\
&\leq \mathrm{P}(B_n(\mathrm{Me}_Z + y\sqrt{c_n/k_n}) \leq y+\varepsilon/2, B_n(\mathrm{Me}_Z) \geq y+\varepsilon) \\
&\leq \mathrm{P}\left(\left|B_n(\mathrm{Me}_Z + y\sqrt{c_n/k_n}) - B_n(\mathrm{Me}_Z)\right| \geq \varepsilon/2\right) \\
&\leq \mathrm{P}(|\widetilde{B}_n| \geq \varepsilon/2)
\end{aligned}$$

which tends to 0 as $n \to \infty$ and (II) is proved.

(b) It is sufficient to combine Theorem 4.2 (b) and Theorem 4.3 (a). From Slutsky's Lemma and by Assumptions (ii-2)-(ii-3), the following convergence in distribution holds as $n \to \infty$

$$\sqrt{k_n/c_n}s_n\left(\widehat{\mathrm{Me}}(\mathbf{Z}) - \mathrm{Me}_Z\right) \to \mathcal{N}(0, 1/4)$$

where $s_n = \sqrt{c_n}\mathrm{P}(N(C_{n,0}) = \lfloor \lambda c_n \rfloor)$. Since $\widehat{\mathrm{Me}}(\mathbf{Z}) = c_n\widehat{\mathrm{Me}}(\check{\mathbf{Z}})$, $\mathrm{Me}_Z = c_n\mathrm{Me}_{\check{Z}}$ and $|W_n| = k_n c_n$, this can be rewritten as

$$|W_n|^{1/2}s_n\left(\widehat{\mathrm{Me}}(\check{\mathbf{Z}}) - \mathrm{Me}_{\check{Z}}\right) \to \mathcal{N}(0, 1/4).$$

From (B.1) and by Assumptions (i)-(ii), $\mathrm{Me}_{\check{Z}} = \lambda + \mathcal{O}(c_n^{\ell-1})$ and $\sqrt{k_n c_n}c_n^{\ell-1} \to 0$ as $n \to \infty$. Hence, a last application of Slutsky's Lemma concludes the proof. □

B Additional comments

B.1 The way of jittering a sample of counts

We could think of slightly generalizing (3.3) and introduce the jittering effect as

$$Z_{n,k} = N(C_{n,k}) + \varphi^{-1}(U_k)$$

for any $k \in \mathcal{K}_n$, where $\varphi : [0,1] \to [0,1]$ is a continuously differentiable increasing function. The cumulative distribution function of Z would be in that case

$$F_Z(t) = P(N(C_{n,0}) \leq \lfloor t \rfloor - 1) + P(N(C_{n,0}) = \lfloor t \rfloor)\,\varphi(t - \lfloor t \rfloor)$$

and for any $t \notin \mathbb{N}$, Z would admit a density f_Z at t given by

$$f_Z(t) = P(N(C_{n,0}) = \lfloor t \rfloor)\,\varphi'(t - \lfloor t \rfloor).$$

When $t \in \mathbb{N}$, since $(F_Z(t+h) - F_Z(t))/h$ tends to $\mathrm{P}(N(C_{n,0}) = \lfloor t \rfloor)\varphi'(0)$ when $h \to 0^+$ and to $\mathrm{P}(N(C_{n,0}) = \lfloor t \rfloor)\varphi'(1)$ when $h \to 0^-$, Z would also admit a density at t if we add the condition $\varphi'(0) = \varphi'(1)$. However, our Theorem 4.3 requires another assumption. Namely, we need to assume that for any $t_n = \lambda c_n + \mathcal{O}(\sqrt{c_n/k_n})$, $f_Z(t_n)/f_Z(\lambda c_n)$ tends to 1. To this end, we would have to combine Assumption (ii-2), with an assumption like $\inf_t \varphi'(t) = \sup_t \varphi'(t)$. This explains why we focused on the case $\varphi(t) = t$ in Section 3 and in the presentation of our asymptotic results in Section 4.

B.2 Rule of thumb under the Poisson case

In this section, we want to examine the value of the true median of Z under the Poisson case. Even if this is useless we also had a look at different functions φ. Figure 1 presents the true median of Π and $Z = \Pi + \varphi^{-1}(U)$ where Π follows a Poisson distribution with mean ν and where U is a uniform random variable on $[0, 1]$. We considered the cases $\varphi(t) = \sqrt{t}, t, t^2$ and examined the true median minus ν in terms of ν. First, we recover a result obtained by Adell and Jodrá (2005): when ν is an integer, the median of Π equals ν and for other values of ν, it lies in the interval $[\nu - \log(2), \nu + 1/3]$. It is worth observing that the choice $\varphi(t) = t$ leads us to conjecture that when ν is large Me_Z is very close to $\nu + 1/3$.

So, we could use the rule of thumb derived under the Poisson case and modify the jittered estimator (3.5) as follows

$$\widehat{\lambda}^{J,2} = \widehat{\lambda}^{J} - \frac{1}{3c_n} = \mathrm{Me}_{\check{\mathbf{Z}}} - \frac{1}{3c_n}. \tag{B.1}$$

Since $|W_n|^{1/2}/c_n = \sqrt{k_n/c_n} \to 0$ by Assumption (i), this produces no differences asymptotically: $\widehat{\lambda}^{J,2}$ has the same behaviour as $\widehat{\lambda}^{J}$ and satisfies the central limit theorem given by (4.4) or Corollary 4.4. We compared $\widehat{\lambda}^{J}$ and $\widehat{\lambda}^{J,2}$ in the framework of the simulation study presented in Section 5. The evidence of better empirical results was unclear which explains why we did not present $\widehat{\lambda}^{J,2}$ before and kept $\widehat{\lambda}^{J}$ in the simulation study.

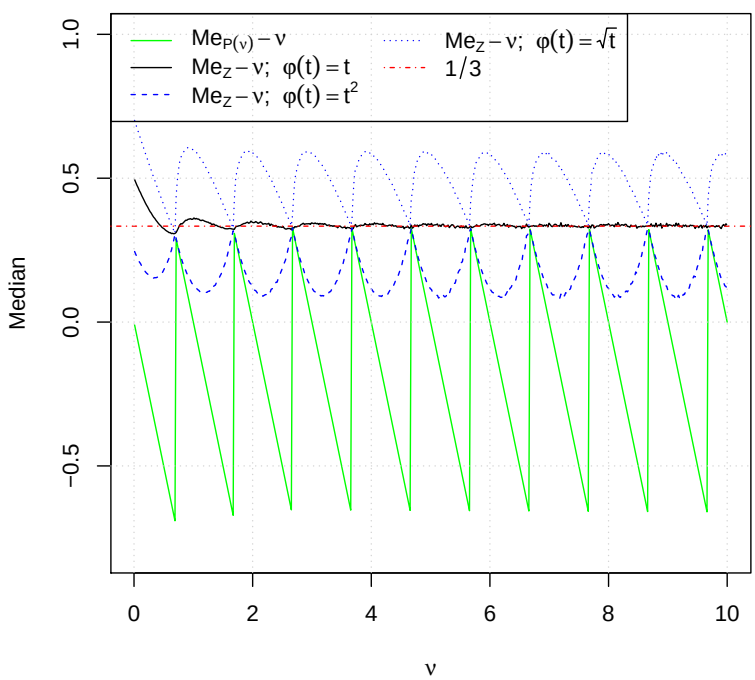

Figure 1: Sample medians based on 10^6 replications of Π or $\Pi + \varphi^{-1}(U)$ random variables where Π (resp. U) follows a Poisson distribution with mean ν (resp. uniform distribution on [0,1]) and where $\varphi = (t) = t, t^2, \sqrt{(t)}$.

References

J.A. Adell and P. Jodrá. The median of the Poisson distribution. *Metrika*, 61(3):337–346, 2005.

R. Assunção and Guttorp. Robustness for inhomogeneous Poisson point processes. *Annals of the Institute of Statistical Mathematics*, 51:657–678, 1999.

A. Baddeley and R. Turner. Spatstat: an R package for analyzing spatial point patterns. *Journal of Statistical Software*, 12:1–42, 2005.

A. Baddeley, R. Turner, J. Møller, and M. Hazelton. Residual analysis for spatial point processes (with discussion). *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67(5):617–666, 2005.

S. Berndt and D. Stoyan. Automatic determination of dendritic arm spacing in directionally solidified matters. *Materials Research and Advanced Techniques*, 88:758–763, 1997.

K. Byth. On robust distance-based intensity estimators. *Biometrics*, pages 127–135, 1982.

M. Clausel, J.-F. Coeurjolly, and J. Lelong. Stein estimation of the intensity of a spatial homogeneous Poisson point process. *arXiv preprint arXiv:1407.4372*, 2015.

J.-F. Coeurjolly and F. Lavancier. Residuals and goodness-of-fit tests for stationary marked Gibbs point processes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 75(2):247–276, 2013.

J.-F. Coeurjolly and J. Møller. Variational approach to estimate the intensity of spatial point processes. *Bernoulli*, 20(3):1097–1125, 2014.

D. J. Daley and D. Vere-Jones. *An Introduction to the Theory of Point Processes. Volume I: Elementary Theory and Methods.* Springer-Verlag, New York, second edition, 2003.

H.A. David and H.N. Nagaraja. *Order statistics (3rd ed.).* Hoboken: Wiley, 2003.

P. J. Diggle. *Statistical Analysis of Spatial Point Patterns.* Arnold, London, second edition, 2003.

J.K. Ghosh. A new proof of the Bahadur representation of quantiles and an application. *Annals of Mathematical Statistics*, 42(6):1957–1961, 1971.

Y. Guan and J. M. Loh. A thinned block bootstrap procedure for modeling inhomogeneous spatial point patterns. *Journal of the American Statistical Association*, 102:1377–1386, 2007.

Y. Guan, M. Sherman, and J. A. Calvin. On asymptotic properties of the mark variogram estimator of a marked point process. *Journal of statistical planning and inference*, 137 (1):148–161, 2007.

X. Guyon. *Random Fields on a Network.* Springer-Verlag, New York, 1991.

L. Heinrich and M. Prokešová. On estimating the asymptotic variance of stationary point processes. *Methodology and Computing in Applied Probability*, 12(3):451–471, 2010.

I. A. Ibragimov and Y. V. Linnik. *Independent and Stationary Sequences of Random Variables*. Wolters-Noordhoff, Groningen, 1971.

J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan. *Statistical Analysis and Modelling of Spatial Point Patterns*. Statistics in Practice. Wiley, Chichester, 2008.

Z. Karáczony. A central limit theorem for mixing random fields. *Miskolc Mathematical Notes*, 7:147–160, 2006.

F. Lavancier, J. Møller, and E. Rubak. Determinantal point process models and statistical inference. *Journal of the Royal Statistical Society: Series B*, 2014. doi: 10.1111/rssb.12096.

Y. Ma, M. Genton, and E. Parzen. Asymptotic properties of sample quantiles of discrete distributions. *Annals of the Institute of Statistical Mathematics*, 63(2):227–243, 2011.

J.A.F. Machado and J.M.C. Santos Silva. Quantiles for counts. *Journal of the American Statistical Association*, 100(472):1226–1237, 2005.

S. Magnussen. Fixed-count density estimation with virtual plots. *Spatial Statistics*, 2: 33–46, 2012.

J. Møller. *Lectures on random Voronoi tessellations*. Springer, 1994.

J. Møller and R. P. Waagepetersen. *Statistical Inference and Simulation for Spatial Point Processes*. Chapman and Hall/CRC, Boca Raton, 2004.

T. Mrkvička and I. Molchanov. Optimisation of linear unbiased intensity estimators for point processes. *Annals of the Institute of Statistical Mathematics*, 57(1):71–81, 2005.

D.N. Politis, E. Paparoditis, and J.P. Romano. Large sample inference for irregularly spaced dependent observations based on subsampling. *Sankhyā: The Indian Journal of Statistics, Series A*, pages 274–292, 1998.

M. Prokešová and E.B.V. Jensen. Asymptotic Palm likelihood theory for stationary point processes. *Annals of the Institute of Statistical Mathematics*, 65(2):387–412, 2013.

C. Redenbach, A. Särkkä, and M. Sormani. Classification of points in superpositions of strauss and poisson processes. *Spatial Statistics*, 12:81–95, 2015.

C. Rose and M.D. Smith. The multivariate normal distribution. *Mathematica Journal*, 6 (1):32–37, 1996.

WL Stevens. Fiducial limits of the parameter of a discontinuous distribution. *Biometrika*, 37(1-2):117–129, 1950.

D. Stoyan, W. S. Kendall, and J. Mecke. *Stochastic Geometry and Its Applications*. Wiley, Chichester, second edition, 1995.

A.W. Van der Vaart. *Asymptotic statistics*, volume 3. Cambridge university press, 2000.

R. Waagepetersen and Y. Guan. Two-step estimation for inhomogeneous spatial point processes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology*, 71: 685–702, 2009.

L. Zhengyan and L. Chuanrong. *Limit theory for mixing dependent random variables*, volume 378. Springer, 1996.