Bernoulli **24**(1), 2018, 592–615
DOI: 10.3150/16-BEJ889

Local block bootstrap for inhomogeneous Poisson marked point processes

WILLIAM GARNER* and DIMITRIS N. POLITIS**

Department of Mathematics, University of California at San Diego, La Jolla, CA 92093, USA.
*E-mail: *wgarner@math.ucsd.edu; **dpolitis@ucsd.edu*

The asymptotic theory for the sample mean of a marked point process in d dimensions is established, allowing for the possibility that the underlying Poisson point process is inhomogeneous. A novel local block bootstrap method for resampling inhomogeneous Poisson marked point processes is introduced, and its consistency is proven for the sample mean and related statistics. Finite-sample simulations are carried out to complement the asymptotic results, and demonstrate the feasibility of the proposed methodology.

Keywords: resampling; sample mean; stochastic processes

1. Introduction

The literature is vast with resampling methods for data observed at regularly spaced points in time and/or space; see Lahiri [15] or Politis, Romano, and Wolf [25] for an overview. However, in many applied settings such as queuing theory, spatial statistics, mining and geostatistics, and meterology, the observations typically occur at non-lattice, irregularly spaced points.

The most common assumption for the occurrence of irregularly spaced data is that of a Poisson process which we also adopt. Another central assumption that we will adopt throughout is that the locations of our points are independent of the associated measurements – also called "marks" – at these locations. Karr [11] provides some justification for both above mentioned assumptions, and puts together a complete asymptotic theory for the sample mean of a homogeneous Poisson process with wide-sense stationary marks.

To elaborate, the mathematical set-up that we adopt goes as follows.

- Let $\{X(\mathbf{t})$ for $\mathbf{t} \in \mathbb{R}^d\}$ be a real-valued, strictly stationary random field in the continuous, d-dimensional parameter $\mathbf{t}$; denote $\mu = EX(\mathbf{t})$ and $R(\mathbf{t}) = \mathrm{Cov}(X(\mathbf{s}), X(\mathbf{s}+\mathbf{t})) = \mathrm{Cov}(X(\mathbf{0}), X(\mathbf{t}))$ which are assumed finite.
- Let $\{N(\mathbf{t})$ for $\mathbf{t} \in \mathbb{R}^d\}$ be an inhomogeneous Poisson point process with rate $\lambda(\mathbf{t})$. The point process $\{N(\mathbf{t})\}$ is assumed to be independent of the random field $\{X(\mathbf{t})\}$.
- Let $\tau_1, \tau_2, \ldots, \tau_{N(K)}$ denote the points generated by $\{N(\mathbf{t})\}$ inside the observation region K that will be assumed to be a compact, convex subset of $\mathbb{R}^d$.
- The pairs $(X(\tau_i), N(\tau_i))$ for $i = 1, \ldots, N(K)$ constitute our observed data from an inhomogeneous *Marked Point Process*.
- Let $|K|$ denote the volume of K, and $\mathrm{diam}(K)$ the supremum of the diameters of all l_∞ balls contained in K; so if K is a rectangle, $\mathrm{diam}(K)$ is its smallest dimension. All asymptotic results to be discussed in this paper will be taken under the condition $\mathrm{diam}(K) \to \infty$.

- To avoid possible pitfalls, we will also assume that the observation region K expands in a *"nested"* way as $\text{diam}(K) \to \infty$, that is, that $\text{diam}(K) < \text{diam}(K')$ has as a necessary implication that $K \subset K'$.

Consider the problem of estimation of the mean $\mu = EX(\mathbf{t})$ based on the above marked point process data. A natural estimator is the sample mean,

$$\overline{X}_K = \frac{1}{N(K)} \int_K X(\mathbf{t}) N(d\mathbf{t}).$$

Note that the denominator includes a random quantity, $N(K)$, which causes difficulties. To start with, one may adopt the convention that $0/0 = 1$. In any case, it is easier to work with the "proxy" sample mean defined as

$$\widetilde{X}_K = \frac{1}{\Lambda(K)} \int_K X(\mathbf{t}) N(d\mathbf{t}),$$

where we denote $\Lambda(K) = E[N(K)]$. In the above, $N(dt)$ acts as a counting measure so each of the above integrals is really just a sum over the observed values of N, namely the τ_i, which are themselves random, that is, we can write

$$\overline{X}_K = \frac{1}{N(K)} \sum_{i=1}^{N(K)} X(\tau_i) \quad \text{and} \quad \widetilde{X}_K = \frac{1}{\Lambda(K)} \sum_{i=1}^{N(K)} X(\tau_i).$$

Note that $\widetilde{X}_K$ is not a proper statistic unless $\Lambda(K)$ is known which is unrealistic since the rate $\lambda(\mathbf{t})$ is typically unknown. The sample mean $\overline{X}_K$ is the statistic of choice; $\widetilde{X}_K$ can serve as a "proxy" for $\overline{X}_K$ in terms of analyzing some theoretical properties. For the longest time, it was thought that $\overline{X}_K$ and $\widetilde{X}_K$ are asymptotically equivalent. However, this is only true when $\mu = 0$; see Garner and Politis [6]. Still, the asymptotic normality of $\overline{X}_K$ can be inferred from the asymptotic normality of $\widetilde{X}_K$ although the asymptotic variances are different. This is achieved in Section 2 under standard moment and mixing conditions which are weaker than Brillinger's [2] cumulant summability conditions.

In Section 3, we show how existing methods for resampling homogeneous marked point processes can be adapted to the inhomogeneous case when the dimension d is one. In Section 4, we introduce the Local Block Bootstrap (LBB) procedure for general d-dimensional inhomogeneous marked point processes data, and establish its validity for the sample mean and related statistics. Finally, in Section 5 we compare the finite-sample performance of these methods. Technical proofs are given in the appendix, and also in the supplement [5].

2. Basic asymptotic theory for an inhomogeneous Poisson process

2.1. Background in the homogeneous case

Karr [11] explored the case where N is a homogeneous Poisson process and established the fact that the quantities $\overline{X}_K$ and $\widetilde{X}_K$ defined in Section 1 are consistent and asymptotically normal at

rate $\sqrt{|K|}$ with the same asymptotic variance, where $|\cdot|$ denotes the Lebesgue measure (volume). There is an error in Karr's variance formula which was first noted by Politis, Paparoditis, and Romano [24]; the error was formally corrected by Garner and Politis [6]. We state the corrected version of Karr's theorem below as it is the springboard for our extension to the inhomogeneous setting.

In terms of notation, $\int$ is short-hand for $\int_{\mathbb{R}^d}$, while $\xrightarrow{d}$ denotes convergence in distribution and $\xrightarrow{p}$ denotes convergence in probability.

Theorem 2.1 (Karr's theorem–corrected). *Assume that* $\lambda(\mathbf{t}) = \lambda$ *is constant, and that*

$$\int \big|R(\mathbf{t})\big|\,d\mathbf{t} < \infty. \tag{2.1}$$

Also, assume $\frac{1}{\sqrt{|K|}}\int_K (X(\mathbf{t}) - \mu)\,d\mathbf{t} \xrightarrow{d} N(0, \int R(\mathbf{t})\,d\mathbf{t})$ *as* $\operatorname{diam}(K) \to \infty$. *Then, as* $\operatorname{diam}(K) \to \infty$, *we have*

$$\sqrt{|K|}(\widetilde{X}_K - \mu) \xrightarrow{d} N\big(0, \sigma^2\big) \quad \textit{and} \quad \sqrt{|K|}(\overline{X}_K - \mu) \xrightarrow{d} N\big(0, \theta^2\big),$$

where $\theta^2 = \int_K R(\mathbf{y})\,d\mathbf{y} + \frac{R(0)}{\lambda}$ *and* $\sigma^2 = \theta^2 + \frac{\mu^2}{\lambda}$.

2.2. Strong mixing coefficients

We shall assume that our random process $\{X(\mathbf{t}), \mathbf{t} \in \mathbb{R}^d\}$ satisfies a certain weak dependence condition that will be quantified in terms of mixing coefficients. Let $\rho(\cdot,\cdot)$ denote the distance in the l_∞-norm on $\mathbb{R}^d$. The strong mixing coefficients of Rosenblatt [27] are then defined as:

$$\alpha_X(k) \equiv \sup_{E_1, E_2 \subset \mathbb{R}^d} \big\{\big|P(A_1 \cap A_2) - P(A_1)P(A_2)\big| : A_i \in \mathcal{F}(E_i), i = 1, 2, \rho(E_1, E_2) \geq k\big\},$$

where $\mathcal{F}(E_i)$ is the σ-algebra generated by $\{X(\mathbf{t}), \mathbf{t} \in E_i\}$. A random field is said to be $\boldsymbol{\alpha}$-*strong* mixing provided that $\lim_{k\to\infty}\alpha_X(k) = 0$.

In a similar manner, Bolthausen [1] defined mixing conditions that also depend on the size of the sets considered, that is, he let

$$\begin{aligned}\alpha_X(k; l_1, l_2) \equiv \sup_{E_1, E_2 \subset \mathbb{R}^d} \big\{&\big|P(A_1 \cap A_2) - P(A_1)P(A_2)\big| : A_i \in \mathcal{F}(E_i), |E_i| \leq l_i,\\ &i = 1, 2, \rho(E_1, E_2) \geq k\big\}.\end{aligned}$$

Note that $\alpha_X(k; l_1, l_2) \leq \alpha_X(k)$, and $\alpha_X(k) = \alpha_X(k; \infty, \infty)$. Bolthausen [1] was able to prove a Central Limit theorem (CLT) using conditions on the rate of decay of $\alpha_X(k; l_1, l_2)$ with l_1, l_2 being finite (but allowed to grow with the sample size).

Politis, Paparoditis, and Romano [23,24] defined a slightly weaker notion of mixing called $\bar{\boldsymbol{\alpha}}$-*strong mixing* via the coefficients

$$\bar{\alpha}_X(k;l) \equiv \sup\Big\{\big|P(A_1 \cap A_2) - P(A_1)P(A_2)\big| : A_i \in \mathcal{F}(E_i), i = 1,2, \\ E_2 = E_1 + \mathbf{t}, |E_1| \le l, \rho(E_1, E_2) \ge k\Big\},$$

where the supremum is taken over all compact and convex sets $E_1 \subset \mathbb{R}^d$, and over all $\mathbf{t} \in \mathbb{R}^d$ such that $\rho(E_1, E_1 + \mathbf{t}) \ge k$. Also, we define $\bar{\alpha}_X(k) = \bar{\alpha}_X(k;\infty)$. Notice that $\bar{\alpha}_X(k) \le \alpha_X(k)$, so that if the random field is α-strong mixing, then it will necessarily be $\bar{\alpha}$-strong mixing. In the case where we only wish to consider sets E_1 and E_2 consisting of a single point, a condition on $\bar{\alpha}_X(k;0)$ is in order.

More discussion and references on strong mixing coefficients can be found in Doukhan [3], Roussas and Ioannides [28], and Ivanov and Leonenko [10].

2.3. The asymptotic distribution of the sample mean in the inhomogeneous case

For our theoretical results, we will need a regularity condition on the inhomogeneous Poisson process, as well as the usual mixing and moment assumptions on the process $X(\mathbf{t})$.

Assumptions.

A.2.1 $\lambda(\mathbf{t})$ is a continuous function such that $0 < \lambda_{\min} \le \lambda(\mathbf{t}) \le \lambda_{\max} < \infty$.
A.2.2 $X(\mathbf{t})$ is strictly stationary with $E[|X(\mathbf{t})|^{2+\delta}] < \infty$ for some $\delta > 0$.
A.2.3 $\bar{\alpha}_X(k;0) < \text{const.}\ k^{-d-\varepsilon}$ for some $\varepsilon > 2d/\delta$ for the δ specified in Assumption A.2.2.

A standard conditioning argument shows the following lemma.

Lemma 2.2. *$\widetilde{X}_K$ and $\overline{X}_K$ are both unbiased for μ.*

The following lemma will allow us to compare $\widetilde{X}_K$ and $\overline{X}_K$ later on.

Lemma 2.3. *Assume* A.2.1. *Then, as* $\operatorname{diam}(K) \to \infty$, $\frac{N(K)}{\Lambda(K)} \xrightarrow{p} 1$.

We can now compute the large-sample variance of $\widetilde{X}_K$. Note that by Assumption A.2.1

$$\frac{1}{\Lambda(K)} \iint_{K\,K} R(\mathbf{s}-\mathbf{t})\lambda(\mathbf{s})\lambda(\mathbf{t})\,d\mathbf{s}\,d\mathbf{t} \le \frac{(\lambda_{\max})^2|K|}{\Lambda(K)} \int_{-\infty}^{\infty} \big|R(\mathbf{u})\big|\,d\mathbf{u}.$$

By equation (2.1), the above upper bound is finite. Hence, we may define $\theta^2 = \lim_{\operatorname{diam}(K)\to\infty} \frac{1}{\Lambda(K)} \iint_{K\,K} R(\mathbf{s}-\mathbf{t})\lambda(\mathbf{s})\lambda(\mathbf{t})\,d\mathbf{s}\,d\mathbf{t}$ assuming the limit exists.

Theorem 2.4. *Assuming* A.2.1 *and equation* (2.1), *we have that*

$$\operatorname{Var}\Big(\sqrt{\big[\Lambda(K)\big]}\widetilde{X}_K\Big) \to \theta^2 + R(\mathbf{0}) + \mu^2 \qquad \text{as } \operatorname{diam}(K) \to \infty.$$

The proof of Theorem 2.4 is immediate noting that

$$E\bigg[\iint_{K\,K} X(\mathbf{t})N(d\mathbf{t})X(\mathbf{s})N(d\mathbf{s})\bigg] = \iint_{K\,K} E\big[X(\mathbf{t})X(\mathbf{s})\big]E\big[N(d\mathbf{t})N(d\mathbf{s})\big]$$

and $E[X(\mathbf{t})X(\mathbf{s})] = R(\mathbf{s}-\mathbf{t}) + \mu^2$; here, we used the stationarity of $X(\cdot)$ and the independence of $X(\cdot)$ and $N(\cdot)$. Since $N(d\mathbf{t})$ is a Poisson$(\lambda(\mathbf{t})\,d\mathbf{t})$ random variable, we have that

$$E\big[N(d\mathbf{t})N(d\mathbf{s})\big] = \begin{cases} \lambda(\mathbf{t})\lambda(\mathbf{s})\,d\mathbf{t}\,d\mathbf{s}, & \text{if } \mathbf{t} \neq \mathbf{s}, \\ \big(\lambda(\mathbf{t})\,d\mathbf{t}\big)^2 + \lambda(\mathbf{t})\,d\mathbf{t}, & \text{if } \mathbf{t} = \mathbf{s}, \end{cases}$$

from which it follows that

$$\begin{aligned} &E\bigg[\iint_{K\,K} X(\mathbf{t})N(d\mathbf{t})X(\mathbf{s})N(d\mathbf{s})\bigg] \\ &\quad= \iint_{K\,K} \big(R(\mathbf{s}-\mathbf{t}) + \mu^2\big)\lambda(\mathbf{t})\lambda(\mathbf{s})\,d\mathbf{t}\,d\mathbf{s} + \int_K \big(R(\mathbf{0}) + \mu^2\big)\lambda(\mathbf{t})\,d\mathbf{t} \end{aligned}$$

as required.

Our final goal for this section is to establish a CLT for the sample mean of our marked point process. We will do this under the high-level condition that a CLT holds true for the underlying continuous-time process, that is, that

$$\frac{1}{\sqrt{\Lambda(K)}} \int_K \big(X(\mathbf{t}) - \mu\big)\lambda(\mathbf{t})\,d\mathbf{t} \xrightarrow{d} N\big(0, \theta^2\big) \tag{2.2}$$

as $\operatorname{diam}(K) \to \infty$, where θ^2 is the quantity appearing in Theorem 2.4. Note that the above continuous-time CLT is guaranteed under standard mixing and moment conditions as the following lemma shows.

Lemma 2.5. *If there exists a* $\delta > 0$ *such that Assumptions* A.2.2 *and* A.2.3 *hold*, *then equations* (2.1) *and* (2.2) *hold true*.

Different sufficient conditions for equations (2.1) and (2.2) can also be found in Rozanov [29]. As mentioned before, Bolthausen [1] proved a CLT for a discrete-parameter random field using conditions on the rate of decay of $\alpha_X(k; l_1, l_2)$ with l_1, l_2 being finite but allowed to grow with the sample size; it is conjectured that an identical result would hold for the continuous-parameter random field under consideration.

As mentioned in the Introduction, $\overline{X}_K$ and $\widetilde{X}_K$ are asymptotically equivalent when $\mu = 0$; this is the subject of the following lemma whose proof is analogous to the one given in Garner and Politis [6].

Lemma 2.6. *Assume equations* (2.1) *and* (2.2). *If* $\mu = 0$, *then*

$$\sqrt{\Lambda(K)}(\overline{X}_K - \widetilde{X}_K) \xrightarrow{p} 0 \qquad \text{as } \operatorname{diam}(K) \to \infty.$$

Our main result in this section now follows.

Theorem 2.7. *Assuming* A.2.1, *and equations* (2.1) *and* (2.2), *then*

$$\sqrt{\Lambda(K)}(\widetilde{X}_K - \mu) \xrightarrow{d} N\big(0, \sigma^2\big) \quad \text{and} \quad \sqrt{\Lambda(K)}(\overline{X}_K - \mu) \xrightarrow{d} N\big(0, \phi^2\big),$$

as $\operatorname{diam}(K) \to \infty$, *where* $\phi^2 = \theta^2 + R(0)$ *and* $\sigma^2 = \phi^2 + \mu^2$.

Theorem 2.7 is the extension of Theorem 2.1 to the inhomogeneous case. The reason we are able to extend results from the homogeneous setting to the inhomogeneous setting is due to the generality of Campbell's theorem [12] (stated below) which applies to both homogeneous and inhomogeneous point processes alike.

Theorem 2.8 (Campbell's theorem [12]). *Let* N *be a Poisson process on* K *with rate* $\lambda(\mathbf{t})$, *and let the function* $f(\mathbf{t}) : K \to \mathbb{R}$ *be measurable. Then the sum* $S = \sum_{i=1}^{N(K)} f(\tau_i)$ *is absolutely convergent in probability if and only if* $\int_K \min(|f(\mathbf{t})|, 1)\lambda(\mathbf{t})\, d\mathbf{t} < \infty$. *If this condition holds, then*

$$E\big[e^{\theta S}\big] = \exp\left\{\int_K \big(e^{\theta f(\mathbf{t})} - 1\big)\lambda(\mathbf{t})\, d\mathbf{t}\right\}$$

for any complex θ *for which the above integral converges.*

Moreover, $E[S] = \int_K f(\mathbf{t})\lambda(\mathbf{t})\, d\mathbf{t}$ *in the sense that the expectation exists if and only if the integral converges, and they are equal. If the expected value exists, then* $\operatorname{Var}[S] = \int_K [f(\mathbf{t})]^2\lambda(\mathbf{t})\, d\mathbf{t}$, *be that finite or infinite.*

Remark 2.9. Lahiri [14] has proven a CLT for irregularly sampled random processes that is closely related to our Theorem 2.7. Lahiri's [14] results are quite general, and in particular allow for the possibility of partially "in-fill" asymptotics. However, the assumptions and sampling set-up of Lahiri's [14] CLT deviate from the classical set-up of a random process sampled by an independent Poisson point process over an expanding domain that we adopt here.

To elaborate, Lahiri [14] assumes that a (nonrandom) number, say n, of sampling points $\tau_1, \ldots, \tau_n$ are generated by an arbitrary probability density, say $f_L(\cdot)$, over a finite domain, say K_L, that can be then "inflated", that is, scaled up, so that it matches our expanding domain K. Consequently, equation (3.4) of Lahiri [14] gives an asymptotic variance for the sample mean that depends on the autocovariance function $R(\cdot)$ as well as the quantity $\int f_L^2(\tau)\, d\tau$. Not surprisingly, the quantity ϕ^2 appearing in our Theorem 2.7 depends instead on $R(\cdot)$ and the rate $\lambda(\cdot)$.

Hence, to use the CLT of Lahiri [14] in practice, one has to estimate $R(\cdot)$ and $f_L^2(\cdot)$ from the data at hand; by contrast, to employ Theorem 2.7 one has to estimate $R(\cdot)$ and $\lambda(\cdot)$. Both of these procedures are quite cumbersome for a task that is as simple as computing a standard error for

the mean. It is for this purpose that the local block bootstrap is developed in Section 4 that can yield tests and confidence intervals for the mean, side-stepping the cumbersome task of analytic estimation of the asymptotic variance.

3. Overview of existing resampling methods and extensions

3.1. Existing resampling methods in the homogeneous case

The closest related works are by Politis *et al.* [23,24] for resampling homogeneous marked point processes and by Politis and Sherman [26] for estimating moments from marked point processes where the sampling is done with a homogeneous point process. Consequently, it would seem natural to try to extend this approach to the inhomogeneous setting. As we shall use this method shortly, we record it below.

Politis *et al.* [24] propose a block resampling scheme that would fill a block with *any* block of equal width from the entire window. We state their circular block bootstrap algorithm below. (Note: This algorithm performs toroidal wrapping of the data. A similar algorithm is also presented in the paper that does not wrap, but instead changes the probabilities near the boundary.)

1. Begin by imagining that $K = \{\mathbf{t} = (t_1, \ldots, t_d) : 0 \le t_i \le K_i, i = 1, \ldots, d\}$ is a rectangular set in $\mathbb{R}^d$ that is "wrapped around" on a compact torus; in other words, we interpret the index $\mathbf{t}$ as being modulo K. To give meaning to the latter notation, we define $\mathbf{t}$ (modulo K) as the vector whose ith coordinate is t_i (modulo K_i). With this definition, we have data $X(\mathbf{t})$ even if $t \notin K$.
2. Let $c = c(K)$ be a number in $(0, 1)$ depending on K such that $c \to 0$ but $c \min_i K_i \to \infty$. Define a scaled-down replica of K by $B = \{c\mathbf{t} : \mathbf{t} \in K\}$. B has the same shape as K but smaller dimensions. Also, define the displaced sets $B + \mathbf{y}$ and let $L = \lfloor 1/c \rfloor$.
3. Generate random points $Y_1, Y_2, \ldots, Y_L$ independent and identically distributed from a uniform distribution on K and define

$$\widetilde{X}^* \equiv \frac{1}{L} \sum_{i=1}^{L} \frac{1}{\lambda |B|} \int_{B+Y_i} X(\mathbf{t}) N(d\mathbf{t}),$$

   where λ is the rate of the homogeneous Poisson point process and

$$\overline{X}^* \equiv \frac{1}{L} \sum_{i=1}^{L} \frac{1}{N(B + Y_i)} \int_{B+Y_i} X(\mathbf{t}) N(d\mathbf{t}).$$

4. Approximate $P(\sqrt{|K|}(\overline{X}_K - \mu) \le x)$ with $P^*(\sqrt{|K|}(\overline{X}^* - E^*\overline{X}^*) \le x)$, where P^* is the resampling probability mechanism that is understood to be conditional on the marked point process data that were actually observed.

Remark 3.1. The above algorithm is tailor-made for homogeneous marked point processes. The literature on resampling for possibly inhomogeneous marked point processes is scarce. To our

knowledge, the only exception is the bootstrap algorithm of Lahiri and Zhu [16] that is intimately related to the sampling set-up of Lahiri [14]. The Lahiri and Zhu [16] bootstrap algorithm can handle an arbitrary degree of inhomogeneity but requires a certain degree of "in-fill" asymptotics in order to work, i.e., requires the availability of obtaining more and more "dense" data around any spatial point of interest. As discussed in Remark 2.9, this set-up is quite different from the classical, expanding domain set-up of sampling via an independent Poisson process that we have adopted here.

3.2. Proposed extensions of existings methods: Case $d = 1$

The above mentioned circular block bootstrap has been shown to be applicable for stationary marked process data. As such, its direct application to inhomogeneous data is not recommended. However, in the case of $d = 1$, it is a fact that an inhomogeneous Poisson process $X(t)$ with rate $\lambda(t)$ can be *transformed* into a homogeneous Poisson process $Y(u(t))$ with rate 1 via the time transformation $u(t) = \Lambda(t)$. Here (and elsewhere when $d = 1$) we make the simplifying assumption that the observation region K is simply the interval $[0, K]$; therefore, the notation $\Lambda(K)$ can be used as short for $\Lambda([0, K])$, and the notation $\Lambda(t)$ is short for $\Lambda([0, t])$.

Consequently, after this transformation, we can appeal to the Politis, Paparoditis, and Romano [24] resampling method described in the previous subsection. We consider this method to be a "circular block bootstrap after transformation to homogeneity". It can be broken down into steps as follows:

1. Identify the cumulative intensity function $\Lambda(t)$.
2. Use the fact that an inhomogeneous Poisson process $N(t)$ with rate $\lambda(t)$ can be transformed to a homogeneous Poisson process $Y(u)$ with rate 1 via the time transformation $u(t) = \Lambda(t)$.
3. Apply the circular block bootstrap of Politis *et al.* [24] to the homogeneous marked point process $Y(u(t))$; in so doing, create bootstrap pseudo-data Y^*.
4. Use the inverse of $\Lambda(t)$ to map back the Y^* pseudo-data to the inhomogeneous setting, that is, to create X^* pseudo-data.
5. Recompute the statistic of interest on the X^* pseudo-data.

Note that if the statistic of interest is the sample mean, Step 4 is superfluous as the average of the X^* pseudo-data is tantamount to the average of the Y^* pseudo-data.

In practice, the cumulative intensity function, $\Lambda(t)$ will not be known, and must be estimated from the data at hand. Law and Kelton [17] suggest a non-parametric procedure for estimating $\lambda(t)$ with a piecewise-constant function. Their method is to divide the interval $[0, K]$ into non-overlapping pieces (of size $2w$) on which the intensity is assumed to be (fairly) constant and estimate a single rate for each interval. The estimated intensity is $\hat{\lambda}(t) = \frac{\#\text{ of points}}{2w}$. The question remains as to what constitutes the optimal choice of w which is similar to the difficult problem of choosing the knots in the context of smoothing splines.

Lewis and Shedler [19] suggest a general nonparametric kernel estimate of the form

$$\hat{\lambda}(t; K) = \frac{1}{h(K)} \sum_{j=1}^{N(K)} W\left(\frac{t - \tau_j}{h(K)}\right),$$

where $N(K)$ is the number of observations on $[0, K]$, $\tau_1, \ldots, \tau_{N(K)}$ are the observed values, $W(\cdot)$ is a bounded non-negative integrable kernel function with $\int_{-\infty}^{\infty} W(u)\,du = 1$, and $h(K)$ is a positive bandwidth which tends to zero as $K \to \infty$ but in such a way that $Kh(K) \to \infty$. Choosing the bandwidth optimally in practice is an unavoidable (and difficult) problem associated with all kernel smoothing methods.

Both methods give a reasonable approximation to the true intensity function $\lambda(t)$. Once we have an estimate for $\lambda(t)$, we can estimate $\Lambda(t) = \int_0^t \lambda(s)\,ds$, and use the time transformation $u(t) = \Lambda(t)$ as required above. It should be noted that the methods we used above are not exhaustive. Many other techniques for estimating $\lambda(t)$ and $\Lambda(t)$ exist. For example, Leemis [18] provides another nonparametric technique for estimating $\Lambda(t)$ that does not require the user to specify any parameters or weighting functions.

4. Local block bootstrap for an inhomogeneous Poisson process

In Section 2, we established the asymptotic normality of the sample mean. In order to construct confidence intervals for the mean μ, though, the asymptotic variance would need to be explicitly estimated. Bootstrap methods, on the other hand, should be able to capture the asymptotic variance and distribution automatically.

To address this situation, we introduce the Local Block Bootstrap (LBB) algorithm in Section 4.1 that, in principle, can handle an arbitrary dimension d. The LBB is able to yield confidence intervals for the sample mean and related statistics without the need for explicit estimation of the asymptotic variance. Alternatively, the LBB method may be used to provide an explicit estimate of the asymptotic variance to be used in connection with the asymptotic normality result established in Section 2.

For conciseness, in this paper we only explicitly address the case of a stationary random field sampled via an inhomogeneous Poisson process. Nevertheless, the LBB methodogy is quite general and is expected to be valid under more general situations, for example, when the random field $X(\mathbf{t})$ is only locally stationary; see, for example, Paparoditis and Politis [21] for a discussion in the discrete-time case.

4.1. Local block bootstrap (LBB) algorithm

Suppose we have observations $\{\tau_i, X(\tau_i), i = 1, \ldots, N(K)\}$ from an inhomogeneous Poisson process with intensity function $\lambda(\mathbf{t})$. K can be any compact, convex set in $\mathbb{R}^d$, though for the sake of simplicity, we suppose that K is a d-dimensional rectangular "box" given by $[0, K_1] \times [0, K_2] \times \cdots \times [0, K_d]$. We shall employ a local block bootstrap method to resample such data in a spirit analogous to the LBB for discrete time of Paparoditis and Politis [21].

We resample the data in blocks (just like a block bootstrap), but when filling a particular block, we only consider blocks that are in a "local" neighborhood of the original block. The blocks are compact sets of rectangular shape in $\mathbb{R}^d$; their size is a parameter while the proximity to the original block is determined by a second parameter. Define the "bottom-left" vertex of a block as the point with ith coordinate being the minimum of the ith coordinates of all points in the block

for $i = 1, \ldots, d$; the term "bottom-left" refers to $\mathbb{R}^2$ but the above definition holds true for any dimension d.

The LBB algorithm can be described as follows:

1. For each dimension $i = 1, 2, \ldots, d$, select a block size b_i and a bandwidth parameter h_i such that $b_i << h_i << K_i$; both b_i and h_i are real-valued and positive. For the asymptotic results that follow, we will require that $b_i \to \infty$ but $b_i^{5/2}/K_i \to 0$ as $K_i \to \infty$, and that $h_i = \Omega(b_i^{\alpha})$ for some $\alpha > 2$; the notation $h = \Omega(g)$ is used to denote that there are positive constants C_1, C_2 such that $C_1 \leq h/g \leq C_2$.
2. Consider a blank template identical to the real-world d-dimensional "box" $[0, K_1] \times [0, K_2] \times \cdots \times [0, K_d]$; this template is to be "filled" with bootstrap data. We break this template into L (non-overlapping) blocks where each block has size $b_1 \times b_2 \times \cdots \times b_d$; the total number of blocks to be "filled" is $L = \prod_{i=1}^{d} K_i/b_i$ (assumed to be an integer for simplicity).
3. Order the L blocks in some way, and label them according to their "bottom-left" vertex, that is, label the jth block by its "bottom-left" vertex $\mathbf{c}_j$.
4. Generate i.i.d. random vectors $\mathbf{d}_1, \ldots, \mathbf{d}_L$ where the components of $\mathbf{d}_j$ are i.i.d. Uniform$[-h_i, h_i]$ random variables.
5. For $j = 1, \ldots, L$, create a bootstrap version of the block $\mathbf{c}_j$ by "filling" it with the points and associated marks found in the real-world block of size $b_1 \times b_2 \times \cdots \times b_d$ that has $\mathbf{c}_j + \mathbf{d}_j$ as its "bottom-left" vertex.

   Note: It is possible that parts of the block may lie outside of K. To correct this, imagine that the real-world observations extend beyond K by "wrapping it around" on itself as in the well-known circular block bootstrap; in other words, our calculations are done modulo K.
6. The above steps define a realization of the bootstrap marked point process that has marks denoted by $X^*(\cdot)$ and point process denoted by N^*. The generation of the bootstrap marked point process is governed by a probability mechanism which we will denote by P^*, with moments denoted by E^*, Var*, etc. This generation is done conditionally on the real-world marked point process data observed; thus P^* is really a conditional probability.
7. Having generated the bootstrap marked point process, we can now re-compute our statistic of interest on the bootstrap data. In principle, the LBB idea can accommodate a variety of statistics; for simplicity, we focus on the aforementioned sample mean statistics that can be computed as $\widetilde{X}^* = \frac{1}{\Lambda^*(K)} \sum_{i=1}^{N(K)} W_i X(\tau_i)$ and $\overline{X}^* = \frac{1}{N^*(K)} \sum_{i=1}^{N(K)} W_i X(\tau_i)$, respectively where W_i is the number of times that $X(\tau_i)$ occurs in the bootstrap data, and $\Lambda^*(K) = E^*[N^*(K)]$.
8. Let $P^*(\sqrt{\Lambda^*(K)}(\widetilde{X}^* - E^*\widetilde{X}^*) \leq x)$ and $P^*(\sqrt{N^*(K)}(\overline{X}^* - E^*\overline{X}^*) \leq x)$ denote the conditional (given the marked point process data) distribution functions of the bootstrap sample means. These will be used to approximate the real-world distributions $P(\sqrt{\Lambda(K)}(\widetilde{X}_K - \mu) \leq x)$ and $P(\sqrt{N(K)}(\overline{X}_K - \mu) \leq x)$, respectively.

Note that we will not concern ourselves with the trivial matter of divisibility, and issues like K_i/b_i being an integer. The reason for this is that for a practical application with a finite sample, we can modify the window size K_i and obtain perfect divisibility. As for the asymptotic case, we can always ignore truncations which are clearly of negligible order.

Remark 4.1. The LBB algorithm as described above is expected to be valid *verbatim* even when the random field $\{X(\mathbf{t})\}$ is only *locally stationary* extending the discrete sampling case studied by Paparoditis and Politis [21]. However, in order to keep the notation and proofs manageable, we will continue our exposition under the simplifying assumption that $\{X(\mathbf{t})\}$ is stationary, and the nonstationarity is only present in the Poisson point process.

For our bootstrap asymptotic results, we need to impose some restrictions on the process $X(\mathbf{t})$, the mixing coefficients, and the parameters.

Assumptions.

A.4.1 $\lambda(\mathbf{t})$ is a continuous function such that $0 < \lambda_{\min} \leq \lambda(\mathbf{t}) \leq \lambda_{\max} < \infty$.

A.4.2 $\{X(\mathbf{t})\}$ is strictly stationary with $E[|X(\mathbf{t})|^{6+\delta}] < \infty$ for some $\delta > 0$.

A.4.3 $\bar{\alpha}_X(k; \mathbf{0}) <$ const. $k^{-d-\varepsilon}$ for some $\varepsilon > \max\{\frac{2d}{\delta}, \frac{8+\delta}{4+\delta}\}$ for the δ specified in Assumption A.4.2.

A.4.4 $\int |Q(\mathbf{u}, \mathbf{v}, \mathbf{v} - \mathbf{w})|\, d\mathbf{v} \leq C_Q$ for all $\mathbf{u}, \mathbf{w}$; here, C_Q is some finite number that may depend on the fourth-order cumulant $Q(\mathbf{s}, \mathbf{t}, \mathbf{u})$ of the random field $\{X(\mathbf{t})\}$ defined as $Q(\mathbf{s}, \mathbf{t}, \mathbf{u}) = E[X(\mathbf{0})X(\mathbf{s})X(\mathbf{t})X(\mathbf{u})] - R(\mathbf{s})R(\mathbf{u} - \mathbf{t}) - R(\mathbf{t})R(\mathbf{u} - \mathbf{s}) - R(\mathbf{u})R(\mathbf{t} - \mathbf{s})$.

A.4.5 $h_i / K_i \to 0$ as $K_i \to \infty$ for $i = 1, 2, \ldots, d$.

A.4.6 $h_i = \Omega(b_i^{\alpha})$ for some $\alpha > 2$, and $b_i \to \infty$ but $b_i^{5/2} / K_i \to 0$ as $K_i \to \infty$ for $i = 1, 2, \ldots, d$.

Note that Assumption A.4.6 would imply Assumption A.4.5 under the additional assumption that $\alpha \leq 5/2$; the latter, however, is not required, and thus we do not impose it here.

Remark 4.2. Assumptions A.4.1, A.4.2 and A.4.3 are almost identical (albeit slightly stronger) to Assumptions A.2.1, A.2.2 and A.2.3 required for the real-world CLT. Assumption A.4.4 is new but can be shown to hold under some stronger moment and mixing condition. For example, in the case $d = 1$, if $\mu = 0$ and $\int t^2 \alpha_X(t)^{\delta/(6+\delta)}\, dt < \infty$ is satisfied, then Assumption A.4.2 together with condition (A1) used in the proof of Theorem 3.3 in Künsch [13] implies that Assumption A.4.4 holds; see also Theorem 17.2.3 of Ibragimov and Linnik [9].

Remark 4.3. In the LBB algorithm, reference was made to $\widetilde{X}^*$ and $\overline{X}^*$. Define $\widetilde{\widetilde{X}}^*$ and $\overline{\overline{X}}^*$ similarly but with denominators of $\Lambda(K)$ and $N(K)$ respectively, that is, let $\widetilde{\widetilde{X}}^* = \widetilde{X}^*[\Lambda^*(K)/\Lambda(K)]$ and $\overline{\overline{X}}^* = \overline{X}^*[N^*(K)/N(K)]$. Note that $N(K)$ is not random in the bootstrap world. Since the division by a random quantity complicates matters, we shall use these modified estimators to prove our bootstrap asymptotic results. Lemma 4.4 below, coupled with Slutsky's lemma [30], allows for this simplification.

Lemma 4.4. *Assume* A.4.1. *Then*, *as* $\operatorname{diam}(K) \to \infty$, $\frac{\Lambda^*(K)}{\Lambda(K)} \xrightarrow{p} 1$.

4.2. Consistency of the local block bootstrap for the sample mean

For simplicity of exposition, we focus on the case $d = 1$ throughout Section 4.2. However, all consistency results are valid for a general $d \geq 1$ under analogous conditions – see Section 4.3 for details.

So in the present section, the observation region K is the interval $[0, K]$; that is, K denotes what was previously denoted by K_1. A key element in showing consistency of the LBB is computing the variance $\mathrm{Var}^*[\sqrt{\Lambda^*(K)}\widetilde{X}^*]$. Recall that $\widetilde{X}^* = \frac{1}{\Lambda^*(K)} \sum_{i=1}^{N(K)} W_i X(\tau_i)$, where W_i is the number of times that $X(\tau_i)$ occurs in the resampled data.

We can think of W_i as a sum of Bernoulli random variables, where the sum runs over all of the blocks. That is, suppose we have L blocks. Then, $W_i = \sum_{j=1}^{L} Y_{ij}$ where $Y_{ij} \sim \mathrm{Bernoulli}(p_{ij})$; here, p_{ij} represents the probability that τ_i is contained in block j. Recall that τ_i denotes the position of the mark of the ith datapoint, while $X(\tau_i)$ is the mark at τ_i. Due to the toroidal wrapping in the LBB algorithm, the following lemma is immediate.

Lemma 4.5. $E^*[W_i] = 1$.

Next, we consider $\mathrm{Cov}^*(W_i, W_k)$. Note that

$$\mathrm{Cov}^*(W_i, W_k) = \mathrm{Cov}^*\left(\sum_{j=1}^{L} Y_{ij}, \sum_{m=1}^{L} Y_{km}\right) = \sum_{j=1}^{L}\sum_{m=1}^{L} \mathrm{Cov}^*(Y_{ij}, Y_{km}).$$

Thus, we have reduced the problem above to computing $\mathrm{Cov}^*(Y_{ij}, Y_{km})$. Since the covariance of two indicator random variables 1_A and 1_B is given by $P^*(A \cap B) - P^*(A)P^*(B)$, we are left computing the joint probabilities of Y_{ij} and Y_{km}.

Definition 4.6. We say that two points τ_i and τ_k are *sufficiently close* to each other provided that $|\tau_k - \tau_i| < b$. Otherwise, we say that τ_i and τ_k are *far apart*.

Definition 4.7. Given an interval $[a_1, a_2]$, where $a_1, a_2 \in \mathbb{R}$, if $|\tau_i - a_1| > h$ and $|\tau_i - a_2| > h$, then we say that τ_i is *removed from the boundary* of $[a_1, a_2]$ by h units.

Claim 4.8. *If τ_i and τ_k are far apart, then the probability that a particular LBB block contains both τ_i and τ_k will be zero. On the other hand, if τ_i and τ_k are sufficiently close, then the probability that a particular block contains both τ_i and τ_k will be positive.*

The reason is because if τ_i and τ_k are far apart, then both cannot appear in the same block. Moreover, if, for example, the probability for τ_i is non-zero, then necessarily the probability for τ_j will be zero.

By construction, the probability that block j contains τ_i and block m contains τ_k are independent, because of the independence of the resampling of the blocks.

Thus, $P^*(Y_{ij}Y_{km}) = P^*(Y_{ij})P^*(Y_{km})$ for all $j \neq m$, so we have that

$$\mathrm{Cov}^*(Y_{ij}, Y_{km}) = P^*(Y_{ij}Y_{km}) - P^*(Y_{ij})P^*(Y_{km}) = 0.$$

Hence, $\mathrm{Cov}^*(W_i, W_k) = \sum_{j=1}^{L} \mathrm{Cov}^*(Y_{ij}, Y_{kj})$, and by Claim 4.8, we only need to consider the case where τ_i and τ_k are sufficiently close, that is, when τ_i and τ_k are within b units of each other. Moreover, we only need to consider those pairs τ_i and τ_k that are removed from the boundary by h units. The reason can be seen via Claim 4.9.

Claim 4.9. *Assuming* A.4.1 *and* A.4.5, *the expected number of τ_i that are within h units of the boundary is $O(h)$.*

Recall that if $N \sim \text{Poisson}(\lambda)$, then on a window of size w, we would expect to see $w\lambda$ observations in that window. Under Assumption A.4.1, we have that the expected number of points within h units of the boundary would be between $2\lambda_{\min}h$ and $2\lambda_{\max}h$. Thus, the number of observed points within h units of the boundary is $O(h)$.

Since $K/h \to \infty$ (by Assumption A.4.5), the expected number of points that are within h units of the boundary is asymptotically negligible.

Lemma 4.10. *For τ_i and τ_k sufficiently close, removed from the boundary by h units, we have that*

$$\mathrm{Cov}^*(W_i, W_k) = 1 - \frac{|\tau_i - \tau_k|}{b} - \frac{b}{2h} + O\left(\frac{b^2}{h^2}\right),$$

where the $O(\cdot)$ term is uniform as it does not depend on i and k.

Therefore, we have that

$$\begin{aligned}
&\mathrm{Var}^*\big[\sqrt{\Lambda^*(K)}\widetilde{X}^*\big] \\
&= \frac{1}{\Lambda^*(K)}\mathrm{Var}^*\left[\sum_{i=1}^{N(K)} W_i X(\tau_i)\right] \\
&= \frac{1}{\Lambda^*(K)}\sum_{i=1}^{N(K)}\sum_{k=1}^{N(K)} \mathrm{Cov}^*(W_i, W_k)X(\tau_i)X(\tau_k) \\
&= \frac{1}{\Lambda^*(K)}\sum_{i\neq k}\left(1 - \frac{|\tau_i - \tau_k|}{b}\right)1_{\{|\tau_i-\tau_k|<b\}}X(\tau_i)X(\tau_k) \\
&\quad + \frac{1}{\Lambda^*(K)}\sum_{i=1}^{N(K)} X^2(\tau_i) \\
&\quad + \frac{1}{\Lambda^*(K)}\sum_{i=1}^{N(K)}\sum_{k=1}^{N(K)} O\left(\frac{b}{h}\right)1_{\{|\tau_i-\tau_k|<b\}}X(\tau_i)X(\tau_k).
\end{aligned} \tag{4.1}$$

The asympotics of the three terms in the variance formula (4.1) are the subject of the following three lemmas whose proofs are given in the supplement [5].

Lemma 4.11. *Assuming* A.4.1, A.4.5, *and* A.4.6, *we have that*

$$\frac{1}{\Lambda^*(K)} \sum_{i=1}^{N(K)} \sum_{k=1}^{N(K)} O\left(\frac{b}{h}\right) 1_{\{|\tau_i - \tau_k| < b\}} X(\tau_i) X(\tau_k) \xrightarrow{p} 0.$$

Lemma 4.12. *Let* $\hat{R}(0) = \frac{1}{\Lambda^*(K)} \sum_{i=1}^{N(K)} X^2(\tau_i)$. *Then, under Assumptions* A.4.1–A.4.3,

i. $E[\hat{R}(0)] = R(0) + \mu^2$;
ii. $\operatorname{Var}[\hat{R}(0)] \to 0$ *as* $K \to \infty$.

Lemma 4.13. *Suppose Assumptions* A.4.1–A.4.4, *and* A.4.6 *hold.*
Let $\hat{\theta}^2 = \frac{1}{\Lambda^*(K)} \sum_{i \neq k} (1 - \frac{|\tau_i - \tau_k|}{b}) 1_{\{|\tau_i - \tau_k| < b\}} X(\tau_i) X(\tau_k)$. *As* $K \to \infty$,

i. $E[\hat{\theta}^2] \xrightarrow{p} \theta^2$;
ii. $\operatorname{Var}[\hat{\theta}^2] \to 0$;

where $\theta^2 = \lim_{|K| \to \infty} \frac{1}{\Lambda^*(K)} \int_K \int_K R(s-t) \lambda(s) \lambda(t)\, ds\, dt$.

Lemma 4.14 below puts it all together; its proof is in the supplement [5].

Lemma 4.14. *If Assumptions* A.4.1–A.4.6 *hold, then* $\operatorname{Var}^*[\sqrt{\Lambda^*(K)} \widetilde{X}^*] \xrightarrow{p} \theta^2 + R(0) + \mu^2$ *as* $K \to \infty$.

Having established that the bootstrap variance tends (asymptotically) to the true variance as given by Theorem 2.4, we are ready to state our main bootstrap theorems.

Theorem 4.15. *Suppose Assumptions* A.4.1–A.4.6 *as well as equations* (2.1) *and* (2.2) *hold true. Then, as* $K \to \infty$, *we have the following*:

i. $E^*[\widetilde{X}^*] = \widetilde{X}_K$;
ii. $\frac{\operatorname{Var}^*[\widetilde{X}^*]}{\operatorname{Var}[\widetilde{X}_K]} \xrightarrow{p} 1$;
iii. $\sup_x |P^*(\sqrt{\Lambda(K)}(\widetilde{X}^* - \widetilde{X}_K) \leq x) - P(\sqrt{\Lambda(K)}(\widetilde{X}_K - \mu) \leq x)| \xrightarrow{p} 0$.

It was mentioned that $\widetilde{X}_K$ is an auxiliary variable as it is not a proper statistic in the usual situation where $\Lambda(\cdot)$ is not known. The statistic of interest is the sample mean $\overline{X}_K$ that is covered in the following theorem.

Theorem 4.16. *Suppose Assumptions* A.4.1–A.4.6 *as well as equations* (2.1) *and* (2.2) *hold true. Then, as* $K \to \infty$, *we have*

$$\sup_x \left| P^*\left(\sqrt{N(K)}\left(\overline{X}^* - \overline{X}_K\right) \leq x\right) - P\left(\sqrt{N(K)}(\overline{X}_K - \mu) \leq x\right) \right| \xrightarrow{p} 0.$$

Theorem 4.16 is our main theoretical result showing asymptotic validity of the LBB for the sample mean of a stationary random process sampled via an independent, inhomogeneous Poisson process. The theorem ensures that the LBB bootstrap is consistent in terms of estimating the asymptotic distribution of the sample mean $\overline{X}_K$. Interestingly, the mixing conditions involved in proving the bootstrap CLT are only conditions on $\bar{\alpha}_X(k;0)$, that is, pointwise asymptotic independence. The usual conditions on either $\bar{\alpha}_X(k;l_1)$ or $\alpha_X(k;l_1,l_2)$ with l_1, l_2 growing with sample size are hidden in the higher-level adoption of equations (2.1) and (2.2).

4.3. Extension to multivariate random fields and smooth functions of means

We now revert to the d-dimensional case to present some further results. Recall that, up to now, the marked point process set-up involved the real-valued random field $\{X(\mathbf{t})$ for $\mathbf{t} \in \mathbb{R}^d\}$ sampled according to the point process $\{N(\mathbf{t})\}$. The case of a multivariate random field, say $\{Y(\mathbf{t})$ for $\mathbf{t} \in \mathbb{R}^d\}$ where $Y(\mathbf{t})$ takes values in $\mathbb{R}^q$ is similar. The sample mean statistics are defined in the same way, that is,

$$\overline{Y}_K = \frac{1}{N(K)} \sum_{i=1}^{N(K)} Y(\tau_i) \quad \text{and} \quad \widetilde{Y}_K = \frac{1}{\Lambda(K)} \sum_{i=1}^{N(K)} Y(\tau_i).$$

Then, by the Cramér–Wold device, we can obtain multivariate analogs of both the real-world and bootstrap central limit theorems; these are given in the two theorems below. To describe them, let $Y(\mathbf{t}) = (Y^{(1)}(\mathbf{t}), \ldots, Y^{(q)}(\mathbf{t}))'$ be a strictly stationary random field that is independent of the Poisson point process $\{N(\mathbf{t})\}$; also let $\mu_Y = EY(\mathbf{t})$.

Theorem 4.17. *Assume that each coordinate of $Y(\mathbf{t})$ satisfies the assumptions of Theorem 2.7, that is, letting $X(\mathbf{t}) = Y^{(j)}(\mathbf{t})$ – for any j – the assumptions of Theorem 2.7 are satisfied. Then,*

$$\sqrt{\Lambda(K)}(\widetilde{Y}_K - \mu_Y) \xrightarrow{d} N(0, \Sigma) \quad \textit{and} \quad \sqrt{\Lambda(K)}(\overline{Y}_K - \mu_Y) \xrightarrow{d} N(0, \Phi)$$

as $\operatorname{diam}(K) \to \infty$, *where Φ and Σ are appropriate matrices.*

The LBB can generate the bootstrap marked point process that has multivariate marks $Y^*(\cdot)$ and point process N^* in an identical fashion as described in the LBB algorithm of Section 4.1. The bootstrap theorem below is helpful in estimating the distribution of the sample mean statistics, and by-passes the need to explicitly estimate the large sample covariance matrices Φ and Σ appearing in Theorem 4.17. To state it, let $Y(\mathbf{t}) = (Y^{(1)}(\mathbf{t}), \ldots, Y^{(q)}(\mathbf{t}))'$, and consider the modified assumptions:

A.4.2′ $\{Y(\mathbf{t})\}$ is strictly stationary with $E[|Y^{(i)}(\mathbf{t})|^{6+\delta}] < \infty$ for some $\delta > 0$ and for all i.

A.4.3′ $\bar{\alpha}_Y(k;\mathbf{0}) <$ const. $k^{-d-\varepsilon}$ for some $\varepsilon > \max\{\frac{2d}{\delta}, \frac{8+\delta}{4+\delta}\}$ for the δ specified in Assumption A.4.2′.

A.4.4′ $\int |Q^{(i)}(\mathbf{u}, \mathbf{v}, \mathbf{v}-\mathbf{w})|\, d\mathbf{v} \le C_Q^{(i)}$ for all $\mathbf{u}, \mathbf{w}$, and for each i; here, $C_Q^{(i)}$ is some finite number that may depend on the fourth-order cumulant $Q^{(i)}(\mathbf{s}, \mathbf{t}, \mathbf{u})$ of the univariate random field $\{Y^{(i)}(\mathbf{t})\}$.

Theorem 4.18. *Assume that Assumptions* A.4.1, A.4.2′, A.4.3′, A.4.4′, A.4.5, A.4.6, *as well as equations* (2.1) *and* (2.2) *hold true. Then, as* $\operatorname{diam}(K) \to \infty$, *we have*

$$\sup_{x \in \mathbb{R}^q} \left| P^*\left(\sqrt{N(K)}\left(\overline{Y}^* - \overline{Y}_K\right) \le x\right) - P\left(\sqrt{N(K)}(\overline{Y}_K - \mu_Y) \le x\right)\right| \xrightarrow{p} 0,$$

where the above inequality sign ($\le$) *is to be interpreted coordinate-wise.*

In the case $d = 1$, Theorem 4.18 follows immediately from Theorem 4.16 via the Cramér–Wold device. In the general case $d \ge 1$, Theorem 4.18 follows from a d-dimensional analog of Theorem 4.16; see Chapter 7 of Garner [4] for details.

Remark 4.19. The above multivariate bootstrap limit theorem allows us to extend the applicability of the LBB to the smooth function of means model developed by Hall [7] and Lahiri [15]. To elaborate, let $H : \mathbb{R}^q \to \mathbb{R}^r$ be a smooth (differentiable) function such that $H'(\mu_Y) \ne 0$. Under the assumptions of Theorem 4.18, the multivariate δ-method immediately implies

$$\sup_{x \in \mathbb{R}^r} \left| P^*\left(\sqrt{N(K)}\left(H\left(\overline{Y}^*\right) - H(\overline{Y}_K)\right) \le x\right) - P\left(\sqrt{N(K)}\left(H(\overline{Y}_K) - H(\mu_Y)\right) \le x\right)\right| \xrightarrow{p} 0.$$

Many statistics of interest can be put in the form $H(\overline{Y}_K)$ with an appropriate choice for the function H and the random field $Y(\mathbf{t})$; see Lahiri [15] for examples.

Remark 4.20. *Any* choice of the window/block size parameters h_i and b_i (as functions of their respective K_i) will result into a consistent LBB procedure as long as h_i and b_i satisfy Assumptions A.4.5 and A.4.6. However, the question of *optimal* choice of these parameters remains open at this point both at the theoretical as well as the empirical level. To elaborate, optimality depends on the criterion being optimized. In the case of stationary and regularly sampled time series, Hall *et al.* [8] show that choosing the block size, that is, b_i, proportional to the 4th root of sample size is optimal for estimating the distribution of the sample mean as is the subject of Theorems 4.16 and 4.18. Nevertheless, to optimally estimate the variance of the sample mean via a standard block bootstrap, the optimal block size is proportional to the 3rd root of sample size. Recall that in our "locally" homogenous context, homogeneity of the Marked Point Process can be thought to hold (at least approximately) over a window of size h_i; hence, we conjecture that choosing $b_i \sim h_i^{1/\alpha}$ with $\alpha \in [3, 4]$ as in Assumption A.4.6 may give reasonable results. Then, one may let $h_i \sim K_i^{\beta}$ for some $\beta \in (0, 1)$. However, β and the size of the window h_i are intimately related to the degree of inhomogeneity of the Marked Point Process. For example, if the process is almost homogeneous throughout K, one may take β close to 1; otherwise, β should not be taken close to 1. In the context of Theorems 4.16 and 4.18, the inhomogeneity is only due to the changing intensity so the optimal β may well have to do with the derivative of $\lambda(\cdot)$. Future work will shed light on these important aspects, and help construct viable data-based rules for optimally choosing the window/block size parameters in practice.

5. Simulations

In the simulations presented in this section, we focus on the case $d = 1$; simulations in the case $d = 2$ are given in Garner [4].

For the case at hand where $d = 1$, we employ accept-reject algorithms specified by Lewis and Shedler [20] to generate one-dimensional inhomogeneous Poisson process data. Pasupathy [22] is another excellent reference for generating (in)homogeneous Poisson processes data. We consider five different intensity functions to generate the points.

1. $\lambda_1(t) = 2 + \sin \frac{2\pi}{250} t$ for $0 \le t \le 1000$.
2. $\lambda_2(t) = 1 + \frac{1}{500} t$ for $0 \le t \le 1000$.
3. $\lambda_3(t) = 3 - \frac{1}{250}|t - 500|$ for $0 \le t \le 1000$.
4. $\lambda_4(t) = 1 + \frac{3}{1\,000\,000} t^2$ for $0 \le t \le 1000$.
5. $\lambda_5(t) = 4 - 2(\frac{t-500}{450})^6$ for $0 \le t \le 1000$.

Once we have the points, we use two different covariance functions to generate associated marks. First, we consider $R(t) = \exp(-|t|)$ and second, we consider $R(t) = \exp(-|t|/3)$.

We shall consider two methods using the local block bootstrap: with toroidal wrapping and without. Also, we will consider two competing, exisiting methods that will first transform the inhomogeneous Poisson process into a homogeneous Poisson process and then resample that using the block bootstrap discussed by Politis, Paparoditis, and Romano [24]. Finally, we compare that with another transformation method that uses the actual intensity function to see the effect of estimating $\lambda(t)$. All approaches require a block size, b, for which we consider three different choices ($b = 2, 5, 10$).

The transformation methods are given as follows:

Method 1: Estimate $\lambda(t)$ as a piecewise constant function on intervals of length $2w$. Estimate $\Lambda(t)$ by finding the area under these rectangles.

Method 2: Estimate $\lambda(t)$ with a local average over a window $[-w, w]$. Construct a fine grid (with points every 0.1) and use this as an approximation to $\lambda(t)$. Again, by considering the area under the curve, estimate $\Lambda(t)$.

Method 3: Use the true $\lambda(t)$ to integrate the function and obtain an exact expression for $\Lambda(t)$, that is, $\Lambda(t) = \int_0^t \lambda(s)\,ds$. The additional parameter w is not needed here as we (unrealistically) use the actual intensity function.

The local block bootstrap methods require a parameter h to denote the proximity of the resampled block. (We cast a net on the interval $[-h, h]$ to shift the block, and in the case of no wrapping, we reduced the size accordingly.) We choose h so that the requirement of $h = \Omega(b^\alpha)$ for some $\alpha > 5/2$ would be plausible. Since the parameter w in the transformation methods plays a similar role as h, we lump these together in a single column, entitled "Band", for which we consider values of 10, 20, and 40.

In the table below, we present the empirical coverage probabilities of 95% equal-tailed confidence intervals for the various methods considered. Each model was simulated 1000 times. Regarding block sizes, $b = 2$ seems to be the best choice for all methods. When considering Table 1, we see that an increase in the block size results in a mild reduction (around 5%) in coverage probabilities. In Table 2, however, increases in b result in more noticeable reductions which are likely attributable to the increased dependence among the marks.

Table 1. Comparison of resampling methods with $R(t) = \exp(-|t|)$

			LBB method		Transformation method		
$\lambda_i(t)$	b	Band	Wrap	No wrap	Method 1	Method 2	Method 3
1	2	10	95.6%	94.8%	87.9%	83.7%	95.0%
	5	20	91.3%	91.6%	78.2%	81.4%	90.7%
	10	40	93.7%	92.8%	74.5%	74.4%	87.0%
2	2	10	95.6%	95.2%	85.4%	73.7%	94.9%
	5	20	93.1%	92.8%	71.2%	82.7%	93.1%
	10	40	87.2%	88.4%	67.3%	67.9%	92.0%
3	2	10	94.4%	95.7%	95.1%	93.7%	93.5%
	5	20	88.9%	91.2%	86.5%	88.6%	86.1%
	10	40	92.4%	93.5%	86.4%	85.7%	84.8%
4	2	10	94.4%	95.2%	93.1%	94.4%	93.5%
	5	20	89.8%	90.8%	88.1%	87.8%	88.9%
	10	40	85.0%	83.8%	83.8%	87.5%	85.5%
5	2	10	88.8%	89.6%	57.8%	61.5%	65.6%
	5	20	80.7%	85.0%	58.1%	61.4%	66.9%
	10	40	76.7%	78.7%	58.7%	59.8%	66.3%

Looking closer at Table 1, we see that in the case of $\lambda_1(t)$ and $\lambda_2(t)$ the coverage of the LBB methods is 10 percentage points closer to the nominal coverage as compared to the coverage of the two averaging transformation methods (Methods 1 and 2), and comparable to the exact transformation method (Method 3). For $\lambda_3(t)$ and $\lambda_4(t)$, both sets of methods yield similar results, with the local block bootstrap performing slightly better. For $\lambda_5(t)$, though, there is a significant improvement (20–30%) using the local block bootstrap; this may be due to the sharp declines in the intensity.

Similar results hold in Table 2 when examining $\lambda_1(t)$ and $\lambda_2(t)$. However, *significant* differences appear for $\lambda_3(t)$ and $\lambda_4(t)$, with the local block bootstrap maintaining its coverage probabilities, but the transformation method failing to work. (The differences are 30–60%!) The reason for this difference is not clear and the models were re-run with similar results; this will be the subject of future study. The coverage with $\lambda_5(t)$ drops for all methods and the effects are more noticeable when considering a larger b which is more in line with expectations.

To summarize the findings from the simulation:

- For the transformation methods, a small choice for block size ($b = 2$) and a moderate choice for local window size ($w = 10$) produced the best results; this may be attributed to the fast (exponential) decay of $R(t)$. As noted, future work may involve the determination of b and w from the data itself. As the block size increased, there was a notable difference in coverage when the dependence of the marks was increased.
- For the local block bootstrap methods, it was again observed that a small block size led to the best results, and that increasing h led to a reduction in coverage. Notably, choosing h to be too large would reduce our local block bootstrap to a regular block bootstrap that loses all information regarding spatial changes of intensity. Future work may entail the study of

Table 2. Comparison of resampling methods with $R(t) = \exp(-|t|/3)$

			LBB method		Transformation method		
$\lambda_i(t)$	b	Band	Wrap	No wrap	Method 1	Method 2	Method 3
1	2	10	93.2%	94.2%	87.8%	84.2%	88.3%
	5	20	80.5%	83.7%	74.9%	73.4%	78.3%
	10	40	74.2%	72.1%	65.6%	62.6%	66.9%
2	2	10	93.2%	94.9%	92.6%	90.2%	87.4%
	5	20	85.5%	87.0%	80.7%	81.7%	77.5%
	10	40	74.2%	73.9%	67.7%	70.6%	72.0%
3	2	10	92.2%	94.1%	50.3%	45.8%	47.2%
	5	20	82.3%	81.7%	46.5%	46.1%	47.1%
	10	40	73.6%	80.7%	45.7%	45.6%	43.9%
4	2	10	94.1%	94.1%	36.4%	38.4%	36.6%
	5	20	82.3%	83.8%	35.2%	33.3%	35.5%
	10	40	71.9%	72.4%	32.7%	34.9%	31.3%
5	2	10	84.4%	86.2%	53.4%	51.4%	51.6%
	5	20	72.7%	70.8%	53.6%	52.2%	53.5%
	10	40	66.1%	64.4%	48.7%	49.7%	50.4%

data-based determination of b and h. Notably, the LBB method that uses toroidal wrapping performs similarly to LBB without wrapping but the latter may have some finite-sample advantages as it avoids the edge effects that wrapping imposes.

- Comparing the two approaches, it appears that the local block bootstrap consistently performed as well – if not better – than the transformation methods.

Appendix: Technical proofs

Proof of Lemma 2.5. The proof is similar to that of Theorem 1.7.1 of Ivanov and Leonenko [10]. We only need to verify the finiteness of $\int |R(\mathbf{t})|\,d\mathbf{t}$ which ensures that the limiting variance of $\frac{1}{\sqrt{\Lambda(K)}}\int_K (X(\mathbf{t}) - \mu)\lambda(\mathbf{t})\,d\mathbf{t}$ is well defined. Note that Assumption A.2.2 implies $|\mathrm{Cov}(X(\mathbf{0}), X(\mathbf{t}))| \leq \text{const.}\ \bar{\alpha}_X(\rho(\mathbf{0},\mathbf{t});0)^{1-2/(2+\delta)}$ where ρ denotes l_∞ distance; see, for example, Roussas and Ioannides [28]. Denoting $\mathbf{t} = (t_1, \ldots, t_d)'$, we have

$$\begin{aligned}\int \big|R(\mathbf{t})\big|\,d\mathbf{t} &= O\bigg(\int \bar{\alpha}_X\Big(\max_i |t_i|;0\Big)^{1-2/(2+\delta)}\,d\mathbf{t}\bigg)\\ &= O\bigg(\int_1^\infty y^{d-1}\Big(\frac{1}{y^{d+\varepsilon}}\Big)^{1-2/(2+\delta)}\,dy\bigg)\\ &< \infty \qquad \text{by Assumption A.2.3.}\end{aligned}$$

□

Proof of Theorem 2.7. We will proceed to show that the characteristic function of $\sqrt{\Lambda(K)}(\widetilde{X}_K - \mu)$ converges to that of a $N(0, \sigma^2)$ random variable.

$$\begin{aligned}
&E\big[\exp\big(i\alpha\sqrt{\Lambda(K)}(\widetilde{X}_K - \mu)\big)\big] \\
&\quad = E\big[E\big[\exp\big(i\alpha\sqrt{\Lambda(K)}(\widetilde{X}_K - \mu)\big)|X\big]\big] \\
&\quad = E\bigg[\exp\bigg(\int_K \bigg(\exp\bigg(\frac{i\alpha X(\mathbf{t})}{\sqrt{\Lambda(K)}}\bigg) - 1\bigg)\lambda(\mathbf{t})\,d\mathbf{t} - \frac{i\alpha\mu}{\sqrt{\Lambda(K)}}\int_K \lambda(\mathbf{t})\,d\mathbf{t}\bigg)\bigg].
\end{aligned}$$

The last line follows by applying Campbell's theorem to the first term (since $X(\mathbf{t})$ is a deterministic function of $\mathbf{t}$ after we condition on it). Note that

$$\begin{aligned}
\int_K \big(e^{\frac{i\alpha}{\sqrt{\Lambda(K)}}X(\mathbf{t})} - 1\big)\lambda(\mathbf{t})\,d\mathbf{t} &\approx \int_K \frac{i\alpha}{\sqrt{\Lambda(K)}}X(\mathbf{t})\lambda(\mathbf{t})\,d\mathbf{t} + \int_K \frac{1}{2}\bigg(\frac{i\alpha}{\sqrt{\Lambda(K)}}X(\mathbf{t})\bigg)^2\lambda(\mathbf{t})\,d\mathbf{t} \\
&= \frac{i\alpha}{\sqrt{\Lambda(K)}}\int_K X(\mathbf{t})\lambda(\mathbf{t})\,d\mathbf{t} - \frac{\alpha^2}{2\Lambda(K)}\int_K X^2(\mathbf{t})\lambda(\mathbf{t})\,d\mathbf{t};
\end{aligned}$$

in the above, the Taylor Series expansion $e^x = 1 + x + x^2/2 + O(x^3)$ was used, written as $e^x - 1 \approx x + x^2/2$, and noting that the error term converges in probability to zero even after integration. Therefore,

$$\begin{aligned}
&\int_K \bigg(\exp\bigg(\frac{i\alpha X(\mathbf{t})}{\sqrt{\Lambda(K)}}\bigg) - 1\bigg)\lambda(\mathbf{t})\,d\mathbf{t} - \frac{i\alpha\mu}{\sqrt{\Lambda(K)}}\int_K \lambda(\mathbf{t})\,d\mathbf{t} \\
&\quad \approx \frac{i\alpha}{\sqrt{\Lambda(K)}}\int_K \big[X(\mathbf{t}) - \mu\big]\lambda(\mathbf{t})\,d\mathbf{t} - \frac{\alpha^2}{2\Lambda(K)}\int_K X^2(\mathbf{t})\lambda(\mathbf{t})\,d\mathbf{t}.
\end{aligned}$$

Recall that, by assumption, $\frac{1}{\sqrt{\Lambda(K)}}\int_K (X(\mathbf{t}) - \mu)\lambda(\mathbf{t})\,d\mathbf{t} \xrightarrow{d} N(0, \theta^2)$. Hence, the characteristic function converges to that of a normal, that is,

$$\exp\bigg(\frac{i\alpha}{\sqrt{\Lambda(K)}}\int_K \big(X(\mathbf{t}) - \mu\big)\lambda(\mathbf{t})\,d\mathbf{t}\bigg) \to e^{-\frac{\alpha^2}{2}\theta^2}.$$

By the law of large numbers, we have almost surely,

$$\frac{1}{\Lambda(K)}\int_K X^2(\mathbf{t})\lambda(\mathbf{t})\,d\mathbf{t} \to E\big[X(\mathbf{t})^2\big] = R(\mathbf{0}) + \mu^2.$$

And hence,

$$\begin{aligned}
&E\bigg[\exp\bigg(\frac{i\alpha}{\sqrt{\Lambda(K)}}\int_K \big[X(\mathbf{t}) - \mu\big]\lambda(\mathbf{t})\,d\mathbf{t} - \frac{\alpha^2}{2\Lambda(K)}\int_K X^2(\mathbf{t})\lambda(\mathbf{t})\,d\mathbf{t}\bigg)\bigg] \\
&\quad \to \exp\bigg(-\frac{\alpha^2}{2}\theta^2 - \frac{\alpha^2}{2}\big(R(\mathbf{0}) + \mu^2\big)\bigg).
\end{aligned}$$

Thus, $\sqrt{\Lambda(K)}(\widetilde{X}_K - \mu) \xrightarrow{d} N(0, \sigma^2)$, where $\sigma^2 = \theta^2 + R(\mathbf{0}) + \mu^2$. The asymptotic normality of $\overline{X}_K$ now follows from Lemma 2.6 using a centering argument as in the proof of the CLT of Garner and Politis [6]. □

Proof of Lemma 4.4. Recall the definition $\Lambda(K) = E[N(K)]$. The bootstrap analog is given by $\Lambda^*(K) = E^*[N^*(K)]$, that is, $\Lambda^*(K)$ represents the expected number of points in our resampled data.

Suppose we are given data $X(\tau_1), \ldots, X(\tau_{N(K)})$ from an inhomogeneous marked point processes. We can express the number of points in the resampled data as $N^*(K) = \sum_{i=1}^{N(K)} W_i$, where $W_i = \sum_{j=1}^{L} Y_{ij}$, with L being the total number of blocks and each $Y_{ij} \sim \text{Bernoulli}(p_{ij})$, where p_{ij} is the probability that block j will contain τ_i.

By Lemma 4.5, $E^*[W_i] = 1$, so $\Lambda^*(K) = N(K)$. Thus, we need only show that $\frac{N(K)}{E[N(K)]} = \frac{N(K)}{\Lambda(K)} \xrightarrow{p} 1$; but this was already established in Lemma 2.3. □

Proof of Theorem 4.15. Proof of (i): Recall that $\overset{\approx}{X}{}^* = \frac{1}{\Lambda(K)} \sum_{i=1}^{N(K)} W_i X(\tau_i)$ and, by Lemma 4.5, $E^*[W_i] = 1$. In the bootstrap world, the value of $N(K)$ as well as the value of the τ_i's is given; the only randomness comes from the W_i. Thus,

$$\begin{aligned} E^*\big[\overset{\approx}{X}{}^*\big] &= \frac{1}{\Lambda(K)} \sum_{i=1}^{N(K)} E^*[W_i] X(\tau_i) \\ &= \frac{1}{\Lambda(K)} \sum_{i=1}^{N(K)} X(\tau_i) \\ &= \frac{1}{\Lambda(K)} \int_K X(t) N(dt) \\ &= \tilde{X}_K \end{aligned}$$

Proof of (ii): Note that

$$\frac{\text{Var}^*[\overset{\approx}{X}{}^*]}{\text{Var}[\widetilde{X}_K]} = \frac{\text{Var}^*[\sqrt{\Lambda(K)}\overset{\approx}{X}{}^*]}{\text{Var}[\sqrt{\Lambda(K)}\widetilde{X}_K]}.$$

But by Theorem 2.7 and Lemma 4.14 that both the numerator and denominator tend to $\theta^2 + R(0) + \mu^2$ as $K \to \infty$. Thus, we have that their ratio tends to 1 in probability.

Proof of (iii): From Theorem 2.7, we have that

$$\sqrt{\Lambda(K)}(\widetilde{X}_K - \mu) \xrightarrow{d} N\big(0, \sigma^2\big).$$

As discussed in Remark 4.3, all that remains to be shown is that

$$\frac{\sqrt{\Lambda(K)}(\overset{\approx}{X}^* - \widetilde{X}_K)}{\sqrt{\mathrm{Var}^*[\sqrt{\Lambda(K)}\overset{\approx}{X}^*]}} \xrightarrow{d} N(0,1).$$

But

$$\begin{aligned}\overset{\approx}{X}^* &= \frac{1}{\Lambda(K)} \sum_{i=1}^{N(K)} W_i X(\tau_i) \\ &= \frac{1}{\Lambda(K)} \sum_{j=1}^{L} \sum_{i=1}^{N(K)} Y_{ij} X(\tau_i).\end{aligned}$$

In addition, using the fact that $\sum_{j=1}^{L} p_{ij} = 1$ we have

$$\begin{aligned}\widetilde{X}_K &= \frac{1}{\Lambda(K)} \sum_{i=1}^{N(K)} X(\tau_i) \\ &= \frac{1}{\Lambda(K)} \sum_{i=1}^{N(K)} \left(\sum_{j=1}^{L} p_{ij} \right) X(\tau_i) \\ &= \frac{1}{\Lambda(K)} \sum_{j=1}^{L} \sum_{i=1}^{N(K)} p_{ij} X(\tau_i).\end{aligned}$$

Hence, $\sqrt{\Lambda(K)}(\overset{\approx}{X}^* - \widetilde{X}_K) = \sum_{j=1}^{L} T_j$ where $T_j = \frac{1}{\sqrt{\Lambda(K)}} \sum_{i=1}^{N(K)} (Y_{ij} - p_{ij}) X(\tau_i)$ and $p_{ij} = E^*[Y_{ij}]$ is the probability that $X(\tau_i)$ appears in block j in the resampled data. Also, notice that by our resampling scheme, the different blocks are independent.

The proof of the CLT would be complete by verifying the Liapunov Condition, that is, showing that

$$\frac{\sum_{j=1}^{L} E^*[|T_j|^6]}{\mathrm{Var}^*[\sqrt{\Lambda(K)}\overset{\approx}{X}^*]^3} \to 0 \qquad \text{as } L \to \infty \text{ (where } L = K/b).$$

But by Assumption A.4.1, there are $O_p(b)$ points in each of the L resampled blocks. Also, $X(\tau_i) = O_p(1)$. So, $\sum_{i=1}^{N(K)} (Y_{ij} - p_{ij}) X(\tau_i) = O_p(b)$ from which it follows that $T_j = O_p(\frac{b}{K^{1/2}})$.

Since $\mathrm{Var}^*[\sqrt{\Lambda(K)}\overset{\approx}{X}^*] \xrightarrow{p} \theta^2 + R(0) + \mu^2$ (by Lemma 4.14), we need only show $\sum_{j=1}^{L} E^*[|T_j|^6] \to 0$. Recalling that $L = K/b$, it follows that $\sum_{j=1}^{L} E^*[|T_j|^6] = O_p(\frac{Lb^6}{K^3}) = O_p(\frac{b^5}{K^2})$ which tends to 0 by Assumption A.4.6. □

Proof of Theorem 4.16. The proof is immediate using part (iii) of Theorem 4.15 and Lemma 2.6 together with a centering argument similar to the one given in the proof of the CLT of Garner and Politis [6]. □

Acknowledgements

Many thanks are due to the Editor, Eric Moulines, to the Associate Editor, and to two anonymous reviewers for their constructive comments that resulted into several improvements/corrections over multiple revisions of this manuscript. In addition, the authors would like to warmly thank Bruce Driver and Elias Masry for key insights, and to Jun Zhu for helpful suggestions. The research of the second author was partially supported by NSF Grants DMS-12-23137 and DMS-13-08319.

Supplementary Material

Supplement to "Local block bootstrap for inhomogeneous Poisson marked point processes" (DOI: 10.3150/16-BEJ889SUPP; .pdf). The supplementary material contains some technical proofs that were omitted from the paper to save space.

References

[1] Bolthausen, E. (1982). On the central limit theorem for stationary mixing random fields. *Ann. Probab.* **10** 1047–1050. MR0672305

[2] Brillinger, D.R. (1973). Estimation of the mean of a stationary time series by sampling. *J. Appl. Probab.* **10** 419–431. MR0373202

[3] Doukhan, P. (1994). *Mixing*: *Properties and Examples*. *Lecture Notes in Statistics* **85**. New York: Springer. MR1312160

[4] Garner, W. Resampling Inhomogeneous Marked Point Processes. Ph.D. Thesis, Dept. Mathematics, Univ. of California–San Diego.

[5] Garner, W. and Politis, D.N. Supplement to "Local block bootstrap for inhomogeneous Poisson marked point processes." DOI:10.3150/16-BEJ889SUPP. Available at http://www.math.ucsd.edu/~politis/PAPER/LBB-MPPsupplement.pdf.

[6] Garner, W. and Politis, D.N. (2013). The correct asymptotic variance for the sample mean of a homogeneous Poisson marked point process. *J. Appl. Probab.* **50** 889–892. MR3102522

[7] Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. *Springer Series in Statistics*. New York: Springer. MR1145237

[8] Hall, P., Horowitz, J.L. and Jing, B.-Y. (1995). On blocking rules for the bootstrap with dependent data. *Biometrika* **82** 561–574. MR1366282

[9] Ibragimov, I.A. and Linnik, Y.V. (1971). *Independent and Stationary Sequences of Random Variables*. Groningen: Wolters-Noordhoff Publishing. MR0322926

[10] Ivanov, A.V. and Leonenko, N.N. (1989). *Statistical Analysis of Random Fields*. *Mathematics and Its Applications* (*Soviet Series*) **28**. Dordrecht: Kluwer Academic. MR1009786

[11] Karr, A.F. (1991). *Point Processes and Their Statistical Inference*, 2nd ed. *Probability*: *Pure and Applied* **7**. New York: Dekker. MR1113698

[12] Kingman, J.F.C. (1993). *Poisson Processes*. *Oxford Studies in Probability* **3**. New York: The Clarendon Press. MR1207584
[13] Künsch, H.R. (1989). The jackknife and the bootstrap for general stationary observations. *Ann. Statist.* **17** 1217–1241. MR1015147
[14] Lahiri, S.N. (2003). Central limit theorems for weighted sums of a spatial process under a class of stochastic and fixed designs. *Sankhyā* **65** 356–388. MR2028905
[15] Lahiri, S.N. (2003). *Resampling Methods for Dependent Data*. *Springer Series in Statistics*. New York: Springer. MR2001447
[16] Lahiri, S.N. and Zhu, J. (2006). Resampling methods for spatial regression models under a class of stochastic designs. *Ann. Statist.* **34** 1774–1813. MR2283717
[17] Law, A.M. and Kelton, W.D. (1982). *Simulation Modeling and Analysis*. New York: McGraw-Hill Book. MR0630193
[18] Leemis, L.M. (1991). Nonparametric estimation of the cumulative intensity function for a nonhomogeneous Poisson process. *Manage. Sci.* **37** 886–900.
[19] Lewis, P.A.W. and Shedler, G.S. (1976). Statistical analysis of non-stationary series of events in a data base system. *IBM J. Res. Develop.* **20** 465–482. MR0443292
[20] Lewis, P.A.W. and Shedler, G.S. (1979). Simulation of nonhomogeneous Poisson processes by thinning. *Naval Res. Logist. Quart.* **26** 403–413. MR0546120
[21] Paparoditis, E. and Politis, D.N. (2002). Local block bootstrap. *C. R. Math. Acad. Sci. Paris* **335** 959–962. MR1952557
[22] Pasupathy, R. (2010). Generating nonhomogeneous Poisson processes. In *Wiley Encyclopedia of Operations Research and Management Science.* Wiley, New York.
[23] Politis, D.N., Paparoditis, E. and Romano, J.P. (1998). Large sample inference for irregularly spaced dependent observations based on subsampling. *Sankhyā Ser. A* **60** 274–292. MR1711685
[24] Politis, D.N., Paparoditis, E. and Romano, J.P. (1999). Resampling marked point processes. In *Multivariate Analysis, Design of Experiments, and Survey Sampling*. *Statist. Textbooks Monogr.* **159** 163–185. New York: Dekker. MR1719074
[25] Politis, D.N., Romano, J.P. and Wolf, M. (1999). *Subsampling*. *Springer Series in Statistics*. New York: Springer. MR1707286
[26] Politis, D.N. and Sherman, M. (2001). Moment estimation for statistics from marked point processes. *J. R. Stat. Soc. Ser. B. Stat. Methodol.* **63** 261–275. MR1841414
[27] Rosenblatt, M. (1985). *Stationary Sequences and Random Fields*. Boston, MA: Birkhäuser. MR0885090
[28] Roussas, G.G. and Ioannides, D. (1987). Moment inequalities for mixing sequences of random variables. *Stoch. Anal. Appl.* **5** 61–120. MR0882698
[29] Rozanov, J.A. (1965). On the density of Gaussian distributions and Wiener–Hopf integral equations. *Dokl. Akad. Nauk SSSR* **165** 1000–1002. MR0189131
[30] Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. New York: Wiley. MR0595165

Received November 2015 and revised May 2016