Methodology and Computing in Applied Probability (2020) 22:237–265
https://doi.org/10.1007/s11009-019-09704-w

Corrected Discrete Approximations for Multiple Window Scan Statistics of One-Dimensional Poisson Processes

Yi-Shen Lin[1] · **Xenos Chang-Shuo Lin**[2] · **Daniel Wei-Chung Miao**[3] · **Yi-Ching Yao**[4]

Received: 24 June 2018 / Revised: 15 October 2018 /
Accepted: 25 February 2019 / Published online: 5 March 2019

Abstract

In the literature on scan statistics, the distributions of continuous scan statistics for one-dimensional Poisson processes have been extensively studied, most of which deal with single window scan statistics under homogeneous Poisson processes. In this paper, we consider discrete approximations for the distributions of multiple window scan statistics of homogeneous/nonhomogeneous Poisson processes. We derive the first-order terms of the discrete approximations, which involve some functionals of the Poisson processes. We then apply Richardson's extrapolation to yield corrected (second-order) approximations. Numerical results are presented to show the accuracy of the approximations.

Keywords Scan statistics · Nonhomogeneous Poisson process · Richardson's extrapolation · Markov chain embedding

Mathematics Subject Classification (2010) 60E05 · 62E20 · 60J10 · 60D05

✉ Yi-Shen Lin
ychyslin@163.com

Xenos Chang-Shuo Lin
xenos.lin@gmail.com

Daniel Wei-Chung Miao
miao@mail.ntust.edu.tw

Yi-Ching Yao
yao@stat.sinica.edu.tw

[1] School of Mathematics and Statistics, Yancheng Teachers University, Yancheng 224002, Jiangsu, China

[2] Accounting Information Department, Aletheia University, New Taipei City, 25103, Taiwan, Republic of China

[3] Graduate Institute of Finance, National Taiwan University of Science and Technology, Taipei 106, Taiwan, Republic of China

[4] Institute of Statistical Science, Academia Sinica, Taipei 115, Taiwan, Republic of China

1 Introduction

Let $\Pi(t)$ be a Poisson process with intensity function $\lambda(t) > 0$ on $(0, 1]$. For a specified window size $0 < w < 1$, the continuous scan statistic is defined as

$$S_w = S_w(\Pi) := \max_{0 \leq t \leq 1-w} |\Pi \cap (t, t+w]|,$$

the maximum number of Poisson points within any window of size w, where $|A|$ is the cardinality of set A, which arises from the likelihood ratio test for the null hypothesis H_0: $\lambda(t) = \lambda$ (constant) versus H_a: $\lambda(t) = \lambda + \Delta$ or λ according to whether t is in or outside the interval $(a, a+w]$ for (unknown) $0 \leq a \leq 1-w$ and $\Delta > 0$. For a thorough review and comprehensive discussion of scan statistics, see Glaz and Naus (2010) and Glaz et al. (2001).

There have been numerous studies in the literature dealing with the conditional and unconditional probabilities $P(S_w \geq k \mid |\Pi| = N)$ and $P(S_w \geq k)$ for integers $N \geq k \geq 2$. In the conditional case, Huntington and Naus (1975) and Hwang (1977) gave exact expressions for $P(S_w \geq k \mid |\Pi| = N)$ under H_0 while Cressie (1977) derived exact formulas for $P(S_w \geq k \mid |\Pi| = N)$ under H_a. Unfortunately, these expressions involve summing up a large number of determinants of large matrices which are not easy to evaluate. Note that the unconditional probability $P(S_w \geq k)$ is a weighted average of the conditional probabilities $P(S_w \geq k \mid |\Pi| = N)$ over N with weights $P(|\Pi| = N)$, for which no computationally efficient (exact) formulas are available. Various bounds and approximations have been derived, see e.g. Naus (1982), Janson (1984), Glaz and Naus (1991), and Loader (1991). For more general large deviation approximation results, see Chan and Zhang (2007), Siegmund and Yakir (2000), and Fang and Siegmund (2016). The continuous scan statistic S_w may be approximated by a discrete analogue via time discretization. Specifically, assuming $w = p/q$ (p and q integers), partition the (time) interval $(0, 1]$ into n subintervals of length n^{-1}, n a multiple of q. The i-th subinterval $((i-1)/n, i/n]$ (independently) either contains exactly one point (with probability $\int_{(i-1)/n}^{i/n} \lambda(s)ds$) or no point (with probability $1 - \int_{(i-1)/n}^{i/n} \lambda(s)ds$). As a window of size w covers $nw = np/q$ subintervals, we define the discrete scan statistic $S_w^{(n)}$ to be the maximum number of points within any nw consecutive subintervals. Fu et al. (2012) showed that $P(S_w^{(n)} \geq k)$ can be computed via the Markov chain embedding method and that $P(S_w \geq k) - P(S_w^{(n)} \geq k) = \mathcal{O}(n^{-1})$. See also Fu (2001), Fu and Koutras (1994), and Koutras and Alexandrou (1995) for related results.

Some extensions of the (single window) scan statistic have been proposed in the literature. Nagarwalla (1996) introduced a scan statistic with a variable window whose size is not chosen *a priori*. Wu (2017) proposed a weighted scan statistic with the weight function inversely proportional to the intensity function $\lambda(t)$. Glaz and Zhang (2004) and Naus and Wallenstein (2004) proposed multiple window scan statistics involving a number of different window sizes w_r, $r = 1, \ldots, r^*$, and studied the joint distribution of S_{w_r}, $r = 1, \ldots, r^*$. Wu et al. (2013) considered approximating $P(S_{w_r} \geq k_r$ for some $1 \leq r \leq r^*)$ by $P(S_{w_r}^{(n)} \geq k_r$ for some $1 \leq r \leq r^*)$ where $S_{w_r}^{(n)}$ is the discrete analogue of S_{w_r} by partitioning the interval $(0, 1]$ into n subintervals.

In this paper, we consider approximating the (unconditional) distributions of multiple window scan statistics of Poisson processes with intensity function $\lambda(t)$ satisfying the following condition.

Condition (A) $\lambda(t)$ is bounded away from 0 and there exist an integer $\xi \geq 0$ and $0 = d_0 < d_1 < \cdots < d_\xi < d_{\xi+1} = 1$ such that $\lambda'(t)$ is bounded and continuous in $t \in (d_\ell, d_{\ell+1})$, $\ell = 0, 1, \ldots, \xi$.

Assuming the window sizes $w_1, \ldots, w_{r^*}$ are rational numbers, under condition (A), the corrected discrete approximations for single window scan statistics in Yao et al. (2017) are extended to improve the convergence rate of the discrete (first-order) approximation $P(S_{w_r}^{(n)} \geq k_r$ for some $1 \leq r \leq r^*)$ for $P(S_{w_r} \geq k_r$ for some $1 \leq r \leq r^*)$. Wu et al. (2013) showed that the discrete approximation $P(S_{w_r}^{(n)} \geq k_r$ for some $1 \leq r \leq r^*)$ can be computed via the Markov chain embedding method and that

$$P(S_{w_r} \geq k_r \text{ for some } 1 \leq r \leq r^*) - P(S_{w_r}^{(n)} \geq k_r \text{ for some } 1 \leq r \leq r^*) = \mathcal{O}(n^{-1}).$$

In Section 2, we make use of a coupling argument to derive the limit

$$\lim_{n\to\infty} n \left[P(S_{w_r} \geq k_r \text{ for some } 1 \leq r \leq r^*) - P(S_{w_r}^{(n)} \geq k_r \text{ for some } 1 \leq r \leq r^*) \right], \quad (1.1)$$

which involves some functionals of Π. As in Yao et al. (2017), to facilitate the proof, we introduce a slightly different discrete scan statistic (denoted by $S_{w_r}^{\prime(n)}$), which is stochastically smaller than S_{w_r} and $S_{w_r}^{(n)}$. We obtain the limits

$$\lim_{n\to\infty} n \left[P(S_{w_r} \geq k_r \text{ for some } 1 \leq r \leq r^*) - P(S_{w_r}^{\prime(n)} \geq k_r \text{ for some } 1 \leq r \leq r^*) \right]$$

and

$$\lim_{n\to\infty} n \left[P(S_{w_r}^{(n)} \geq k_r \text{ for some } 1 \leq r \leq r^*) - P(S_{w_r}^{\prime(n)} \geq k_r \text{ for some } 1 \leq r \leq r^*) \right],$$

which together yield the limit in Eq. 1.1. Based on these limit results, we apply in Section 3 Richardson's extrapolation to obtain corrected (second-order) approximations for $P(S_{w_r} \geq k_r$ for some $1 \leq r \leq r^*)$. Numerical results are presented to show the accuracy of the approximations. The proofs of two technical lemmas (stated in Section 2) are relegated to Section 4.

2 Main Results

Recall that for $r = 1, \ldots, r^*$,

$$S_{w_r} = S_{w_r}(\Pi) := \max_{0 \leq t \leq 1-w_r} |\Pi \cap (t, t + w_r]|,$$

where $w_r = p_r/q$ ($p_1, \ldots, p_{r^*}, q$ all integers) and the Poisson process Π has intensity function $\lambda(t)$ satisfying condition (A). To define the two discrete scan statistics $S_{w_r}^{(n)}$ and $S_{w_r}^{\prime(n)}$ mentioned in Section 1, for $n = mq$ ($m = 1, 2, \ldots$), let $I_1^n, I_2^n, \ldots, I_n^n$ be independent Bernoulli random variables with

$$P(I_i^n = 0) = \exp\left(-\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds \right) \text{ and } P(I_i^n = 1) = 1 - P(I_i^n = 0), \; i = 1, \ldots, n;$$

and let $H_1^n, H_2^n, \ldots, H_n^n$ be independent Bernoulli random variables with

$$P(H_i^n = 1) = \min\left\{ \int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds, 1 \right\} \text{ and } P(H_i^n = 0) = 1 - P(H_i^n = 1), \; i = 1, \ldots, n.$$

Then $(I_1^n, \ldots, I_n^n)$ approximates Π by matching the probability of no point in each subinterval, i.e.

$$P(I_i^n = 0) = P\left(\left|\Pi \cap \left(\frac{i-1}{n}, \frac{i}{n}\right]\right| = 0\right) = \exp\left(-\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right),$$

while $(H_1^n, \ldots, H_n^n)$ approximates Π by matching the expected number of points in each subinterval, i.e.

$$E(H_i^n) = E\left(\left|\Pi \cap \left(\frac{i-1}{n}, \frac{i}{n}\right]\right|\right) = \int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds < 1 \ \text{(for large } n).$$

For $r = 1, 2, \ldots, r^*$, the two discrete scan statistics $S_{w_r}^{\prime(n)}$ and $S_{w_r}^{(n)}$ are defined as

$$S_{w_r}^{\prime(n)} = S_{w_r,I}^{(n)} := \max_{i=1,\ldots,n-nw_r+1} \sum_{j=i}^{i+nw_r-1} I_j^n, \tag{2.1}$$

$$S_{w_r}^{(n)} = S_{w_r,H}^{(n)} := \max_{i=1,\ldots,n-nw_r+1} \sum_{j=i}^{i+nw_r-1} H_j^n. \tag{2.2}$$

Using a coupling argument, we now derive the limits

$$\lim_{n\to\infty} n\left[P(S_{w_r} \geq k_r \text{ for some } 1 \leq r \leq r^*) - P(S_{w_r,I}^{(n)} \geq k_r \text{ for some } 1 \leq r \leq r^*)\right]$$

$$\text{and} \quad \lim_{n\to\infty} n\left[P(S_{w_r} \geq k_r \text{ for some } 1 \leq r \leq r^*) - P(S_{w_r,H}^{(n)} \geq k_r \text{ for some } 1 \leq r \leq r^*)\right]$$

in Sections 2.1 and 2.2, respectively.

2.1 Matching the Probability of no Point

Since $(I_1^n, \ldots, I_n^n)$ and Π match in the probability of no point in each subinterval, we may define $I_i^n = \mathbf{1}\{\Pi \cap (\frac{i-1}{n}, \frac{i}{n}] \neq \emptyset\}$, $i = 1, \ldots, n$, so that $(I_1^n, \ldots, I_n^n)$ and Π are defined on the same probability space. (Here $\mathbf{1}A = \mathbf{1}_A$ denotes the indicator function of set A.) For $w_r = p_r/q$ with $p_1 < p_2 < \cdots < p_{r^*}$, and for integers $2 \leq k_1 < k_2 < \cdots < k_{r^*}$, let

$$\alpha = P(\mathcal{A}), \quad \text{where} \ \mathcal{A} = \bigcup_{r=1}^{r^*}\{S_{w_r} \geq k_r\}, \tag{2.3}$$

$$\text{and} \quad \alpha_n = P(\mathcal{A}_n), \quad \text{where} \ \mathcal{A}_n = \bigcup_{r=1}^{r^*}\{S_{w_r,I}^{(n)} \geq k_r\}. \tag{2.4}$$

Let $M := |\Pi|$, which is Poisson distributed with mean $\int_0^1 \lambda(s)ds$. Writing $\Pi = \{Q_1, Q_2, \ldots, Q_M\}$, we assume (with probability 1) that $0 < Q_1 < \cdots < Q_M < 1$. Note that, with probability 1, for $r = 1, 2, \ldots, r^*$ and $\ell = 1, 2, \ldots, M$,

$$w_r \notin \Pi, \ 1 - w_r \notin \Pi, \ Q_\ell \pm w_r \notin \Pi, \ Q_\ell \notin \mathcal{D}, \ Q_\ell \pm w_r \notin \mathcal{D}, \tag{2.5}$$

where $\mathcal{D} := \{d_1, d_2, \ldots, d_\xi\}$, the set of discontinuities of $\lambda(t)$. (Note that $\mathcal{D} = \emptyset$ if $\xi = 0$.) Define

$$\nu(\Pi) := \sum_{r=1}^{r^*} \sum_{\{\ell:\ Q_\ell<1-w_r\}} \lambda(Q_\ell + w_r)\mathbf{1}\big\{S_{w_m} < k_m \text{ for all } m=1,\ldots,r^*, |\Pi \cap (Q_\ell, Q_\ell+w_r]| = k_r - 2,$$

$$|\Pi \cap (t, t+w_r]| \le k_r - 2 \text{ for all } t \text{ with } Q_\ell \le t \le Q_\ell + w_r\big\}, \quad (2.6)$$

$$\tilde{\nu}(\Pi) := \sum_{\ell=1}^{M} \lambda(Q_\ell)\mathbf{1}\Big\{S_{w_r} < k_r \text{ for all } r = 1,\ldots,r^*,$$

$$\max_{0\le t\le 1-w_r} |(\Pi \cup \{Q_\ell\}) \cap (t, t+w_r]| = k_r \text{ for some } 1 \le r \le r^*\Big\}, \quad (2.7)$$

where $\Pi \cup \{Q_\ell\}$ is interpreted as a multiset with Q_ℓ having multiplicity 2.

Theorem 2.1 *For* $n = mq$ $(m = 1, 2, \ldots)$*, define* α*,* α_n*,* $\nu(\Pi)$ *and* $\tilde{\nu}(\Pi)$ *as in* Eqs. 2.3, 2.4, 2.6 and 2.7*, respectively. Then*

$$\lim_{n\to\infty} n(\alpha - \alpha_n) = \frac{1}{2} E\left[\nu(\Pi) + \tilde{\nu}(\Pi)\right].$$

To prove Theorem 2.1, we need the following lemma, whose proof is relegated to Section 4. We denote the complement of $\mathcal{A}_n$ by $\mathcal{A}_n^c$, and for $i = 1, \ldots, n$, let

$$\tilde{I}_i^n = \left|\Pi \cap \left(\frac{i-1}{n}, \frac{i}{n}\right]\right|, \text{ the number of Poisson points in the } i\text{-th subinterval.} \quad (2.8)$$

Then $\tilde{I}_i^n = 0$ implies $I_i^n = 0$ and $\tilde{I}_i^n \ge 1$ implies $I_i^n = 1$.

Lemma 2.1 *For* $n = mq$ $(m = 1, 2, \ldots)$*, define* $\tilde{I}_i^n$*,* $i = 1, \ldots, n$*, as in* Eq. 2.8*, and let*

$$\mathcal{P}_{i,r}^{(n)} = P\left(\mathcal{A}_n^c, \sum_{j=i+1}^{i+nw_r-1} I_j^n = k_r - 2, I_i^n = I_{i+nw_r}^n = 1\right), \quad r = 1, \ldots, r^*, \quad i = 1, \ldots, n - nw_r, \quad (2.9)$$

and

$$\tilde{\mathcal{P}}_i^{(n)} = P\Bigg(\mathcal{A}_n^c, \tilde{I}_i^n = 2, \sum_{j=i'}^{i'+nw_r-1} I_j^n = k_r - 1 \textit{ for some } (i', r)$$

$$\textit{with } 1 \le i' \le i \le i' + nw_r - 1 \le n \textit{ and } 1 \le r \le r^*\Bigg), \quad i = 1, \ldots, n.$$

Then

$$\alpha - \alpha_n = \frac{1}{2}\sum_{r=1}^{r^*}\sum_{i=1}^{n-nw_r} \mathcal{P}_{i,r}^{(n)} + \sum_{i=1}^{n} \tilde{\mathcal{P}}_i^{(n)} + \mathcal{O}(n^{-2}).$$

Proof of Theorem 2.1 By Lemma 2.1, we have

$$\alpha - \alpha_n = \frac{1}{2}\sum_{r=1}^{r^*}\sum_{i=1}^{n-nw_r} \mathcal{P}_{i,r}^{(n)} + \sum_{i=1}^{n} \tilde{\mathcal{P}}_i^{(n)} + \mathcal{O}(n^{-2}). \quad (2.10)$$

For $r = 1, \ldots, r^*$ and $i = 1, \ldots, n - nw_r$, let $\mathcal{P}'^{(n)}_{i,r} = P(\mathcal{F}_{i,r})$ where

$$\mathcal{F}_{i,r} := \mathcal{A}_n^c \cap \left\{ \sum_{j=i+1}^{i+nw_r-1} I_j^n = k_r - 2,\ I_i^n = 1,\ I_{i+nw_r}^n = 0, \right.$$
$$\left. \text{sum of any } nw_r \text{ consecutive } I_j^n \text{ including } j = i + nw_r \text{ is at most } k_r - 2 \right\}.$$

We claim, for $1 \le r \le r^*$, that

$$\mathcal{P}^{(n)}_{i,r} / \mathcal{P}'^{(n)}_{i,r} = \rho^{(n)}_{i,r}, \quad i = 1, \ldots, n - nw_r, \tag{2.11}$$

where

$$\rho^{(n)}_{i,r} := \frac{P(I^n_{i+nw_r} = 1)}{P(I^n_{i+nw_r} = 0)} = \exp\left(\int_{w_r + \frac{i-1}{n}}^{w_r + \frac{i}{n}} \lambda(s) ds \right) - 1.$$

To establish the claim, note that the event inside the parentheses on the right-hand side of Eq. 2.9, $\mathcal{A}_n^c \cap \{\sum_{j=i+1}^{i+nw_r-1} I_j^n = k_r - 2,\ I_i^n = I^n_{i+nw_r} = 1\}$, depends only on $(I_1^n, \ldots, I_n^n)$ and can be interpreted as a collection of configurations $(I_1^n, \ldots, I_n^n) = (h_1, \ldots, h_n)$ where $(h_1, \ldots, h_n)$ satisfies

$$h_\ell = 0 \text{ or } 1 \text{ for all } \ell, \quad \max_{j=1,\ldots,n-nw_m+1} \sum_{\ell=j}^{j+nw_m-1} h_\ell < k_m \text{ for all } m = 1, \ldots, r^*,$$
$$\sum_{\ell=i+1}^{i+nw_r-1} h_\ell = k_r - 2 \text{ and } h_i = h_{i+nw_r} = 1. \tag{2.12}$$

Likewise, the event $\mathcal{F}_{i,r}$ is a collection of configurations $(I_1^n, \ldots, I_n^n) = (h'_1, \ldots, h'_n)$ where $(h'_1, \ldots, h'_n)$ satisfies

$$h'_\ell = 0 \text{ or } 1 \text{ for all } \ell, \quad \max_{j=1,\ldots,n-nw_m+1} \sum_{\ell=j}^{j+nw_m-1} h'_\ell < k_m \text{ for all } m = 1, \ldots, r^*, \quad \sum_{\ell=i+1}^{i+nw_r-1} h'_\ell = k_r - 2,$$
$$h'_i = 1, h'_{i+nw_r} = 0, \text{ and sum of any } nw_r \text{ consecutive } h'_\ell \text{ including } \ell = i + nw_r \text{ is at most } k_r - 2. \tag{2.13}$$

It is readily seen that a configuration $(I_1^n, \ldots, I_n^n) = (h_1, \ldots, h_n)$ satisfies (2.12) if and only if $(I_1^n, \ldots, I_n^n) = (h'_1, \ldots, h'_n)$ satisfies (2.13) where $(h'_1, \ldots, h'_n) = (h_1, \ldots, h_n) - \mathbf{e}_{i+nw_r}$ with $\mathbf{e}_{i+nw_r}$ being the vector of zeroes except for the $(i + nw_r)$-th entry being 1. The claim (2.11) now follows from the independence property of $I_1^n, \ldots, I_n^n$.

Let $\mathcal{D}_{r,n} := \{1 \le j \le n - nw_r : (\frac{j-1}{n}, \frac{j}{n}) \cap \mathcal{D} \ne \emptyset \text{ or } (w_r + \frac{j-1}{n}, w_r + \frac{j}{n}) \cap \mathcal{D} \ne \emptyset\}$. By condition (A), we have for $i \in \{1, \ldots, n - nw_r\} \backslash \mathcal{D}_{r,n}$,

$$\rho^{(n)}_{i,r} = \frac{1}{n} \lambda \left(w_r + \frac{i-1}{n} + \frac{1}{2n} \right) + \mathcal{O}(n^{-2}), \tag{2.14}$$

where the $\mathcal{O}(n^{-2})$ term is uniform in $i \in \{1, \ldots, n - nw_r\} \backslash \mathcal{D}_{r,n}$. Since $|\mathcal{D}_{r,n}| \leq 2\xi < \infty$, it follows from Eqs. 2.11 and 2.14 that

$$\begin{aligned}\sum_{r=1}^{r^*} \sum_{i=1}^{n-nw_r} \mathcal{P}_{i,r}^{(n)} &= \sum_{r=1}^{r^*} \sum_{i=1}^{n-nw_r} \rho_{i,r}^{(n)} \mathcal{P}_{i,r}^{\prime(n)} \\ &= \sum_{r=1}^{r^*} \sum_{i=1}^{n-nw_r} \rho_{i,r}^{(n)} P(\mathcal{F}_{i,r}) \\ &= \sum_{r=1}^{r^*} \sum_{i=1}^{n-nw_r} \frac{1}{n} \lambda \left(w_r + \frac{i-1}{n} + \frac{1}{2n} \right) P(\mathcal{F}_{i,r}) + \mathcal{O}(n^{-2}) \\ &= \frac{1}{n} E[\nu^{(n)}(\Pi)] + \mathcal{O}(n^{-2}), \end{aligned} \tag{2.15}$$

where

$$\nu^{(n)}(\Pi) = \sum_{r=1}^{r^*} \sum_{i=1}^{n-nw_r} \lambda \left(w_r + \tfrac{i-1}{n} + \tfrac{1}{2n} \right) \mathbf{1} \Bigg\{ \mathcal{A}_n^c, \sum_{\ell=i+1}^{i+nw_r-1} I_\ell^n = k_r - 2, I_i^n = 1, I_{i+nw_r}^n = 0, \\ \text{sum of any } nw_r \text{ consecutive } I_\ell^n \text{ including } \ell = i + nw_r \text{ is at most } k_r - 2 \Bigg\}.$$

To deal with $\tilde{\mathcal{P}}_i^{(n)}$, $i = 1, 2, \ldots, n$, let

$$\tilde{\mathcal{P}}_i^{\prime(n)} = P\Big(\mathcal{A}_n^c, I_i^n = 1, \sum_{\ell=i'}^{i'+nw_r-1} I_\ell^n = k_r - 1 \text{ for some } (i', r) \\ \text{with } 1 \leq i' \leq i \leq i' + nw_r - 1 \leq n \text{ and } 1 \leq r \leq r^* \Big).$$

Following an argument similar to the proof of Eq. 2.11, we have $\tilde{\mathcal{P}}_i^{(n)} / \tilde{\mathcal{P}}_i^{\prime(n)} = \tilde{\rho}_i^{(n)}$ for $i = 1, \ldots, n$, where

$$\tilde{\rho}_i^{(n)} = \frac{P(\tilde{I}_i^n = 2)}{P(I_i^n = 1)} = \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s) ds \right)^2 \Big/ 2 \left[\exp\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s) ds \right) - 1 \right].$$

By condition (A), we have uniformly in $i \in \{1, \ldots, n\} \backslash \{1 \leq j \leq n : \left(\frac{j-1}{n}, \frac{j}{n} \right) \cap \mathcal{D} \neq \emptyset\}$,

$$\tilde{\rho}_i^{(n)} = \frac{1}{2n} \lambda \left(\frac{i-1}{n} + \frac{1}{2n} \right) + \mathcal{O}(n^{-2}).$$

So,

$$\begin{aligned}\sum_{i=1}^{n} \tilde{\mathcal{P}}_i^{(n)} &= \sum_{i=1}^{n} \tilde{\rho}_i^{(n)} \tilde{\mathcal{P}}_i^{\prime(n)} \\ &= \frac{1}{2n} \sum_{i=1}^{n} \lambda \left(\frac{i-1}{n} + \frac{1}{2n} \right) \tilde{\mathcal{P}}_i^{\prime(n)} + \mathcal{O}(n^{-2}) \\ &= \frac{1}{2n} E[\tilde{\nu}^{(n)}(\Pi)] + \mathcal{O}(n^{-2}), \end{aligned} \tag{2.16}$$

where

$$\tilde{\nu}^{(n)}(\Pi) := \sum_{i=1}^{n} \lambda\left(\frac{i-1}{n}+\frac{1}{2n}\right) \mathbf{1}\left\{\mathcal{A}_n^c, I_i^n = 1, \sum_{\ell=i'}^{i'+nw_r-1} I_\ell^n = k_r - 1 \right.$$

$$\left. \text{for some } (i', r) \text{ with } 1 \leq i' \leq i \leq i' + nw_r - 1 \leq n \text{ and } 1 \leq r \leq r^* \right\}.$$

It follows from Eqs. 2.10, 2.15 and 2.16 that

$$n(\alpha-\alpha_n) = \frac{1}{2} E[\nu^{(n)}(\Pi) + \tilde{\nu}^{(n)}(\Pi)] + \mathcal{O}(n^{-1}). \tag{2.17}$$

By Eq. 2.5, $\nu^{(n)}(\Pi)$ and $\tilde{\nu}^{(n)}(\Pi)$ converge a.s. to $\nu(\Pi)$ and $\tilde{\nu}(\Pi)$ respectively. Since λ is bounded, we have

$$\max\left\{\nu^{(n)}(\Pi), \tilde{\nu}^{(n)}(\Pi)\right\} \leq \left(\sup_{0<t<1} \lambda(t)\right) \sum_{r=1}^{r^*} \sum_{i=1}^{n} \mathbf{1}\{I_i^n = 1\} \leq \left(\sup_{0<t<1} \lambda(t)\right) r^* |\Pi|.$$

By the dominated convergence theorem,

$$E\left[\nu^{(n)}(\Pi) + \tilde{\nu}^{(n)}(\Pi)\right] \to E\left[\nu(\Pi) + \tilde{\nu}(\Pi)\right] \quad \text{as} \quad n \to \infty,$$

which together with Eq. 2.17 completes the proof. □

Remark 2.1 Since λ is bounded and $P(\tilde{I}_i^n \geq 2) = \mathcal{O}(n^{-2})$ uniformly in i, it can be shown that

$$E\left[\nu^{(n)}(\Pi) + \tilde{\nu}^{(n)}(\Pi)\right] - E\left[\nu(\Pi) + \tilde{\nu}(\Pi)\right] = \mathcal{O}(n^{-1}),$$

which together with Eq. 2.17 implies

$$\alpha - \alpha_n = \frac{E\left[\nu(\Pi) + \tilde{\nu}(\Pi)\right]}{2n} + \mathcal{O}(n^{-2}). \tag{2.18}$$

2.2 Matching the Expected Number of Points

Recall that $H_1^n, H_2^n, \ldots, H_n^n$ are independent with

$$P(H_i^n = 1) = \min\left\{\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds, 1\right\} = 1 - P(H_i^n = 0), \ i = 1, 2, \ldots, n.$$

Let

$$\mathcal{B}_n = \bigcup_{r=1}^{r^*} \{S_{w_r,H}^{(n)} \geq k_r\} = \bigcup_{r=1}^{r^*} \left\{ \max_{i=1,\ldots,n-nw_r+1} \sum_{j=i}^{i+nw_r-1} H_j^n \geq k_r \right\}$$

and $\beta_n = P(\mathcal{B}_n)$. For $u \in (0,1)$, let $f(u) = \lambda^2(u) \Big/ \int_0^1 \lambda^2(s)ds$ and

$$\mathcal{E}(u) = \bigcup_{r=1}^{r^*} \left\{ \max_{0 \leq t \leq 1-w_r} |(\Pi \cup \{u\}) \cap (t, t+w_r]| \geq k_r \right\}.$$

Theorem 2.2 *Define* α, $\nu(\Pi)$ *and* $\tilde{\nu}(\Pi)$ *as in* Eqs. 2.3, 2.6 and 2.7, *respectively. For* $n = mq$ $(m = 1, 2, \ldots)$,

$$\lim_{n\to\infty} \frac{2n(\alpha - \beta_n)}{\int_0^1 \lambda^2(s)ds} = \frac{E[\nu(\Pi) + \tilde{\nu}(\Pi)]}{\int_0^1 \lambda^2(s)ds} + \alpha - \int_0^1 f(u)P\left(\mathcal{E}(u)\right)du.$$

To prove Theorem 2.2, we need the following lemma, whose proof is also relegated to Section 4.

Lemma 2.2 *Define* α_n *as in* Eq. 2.4. *Then, for* $n = mq$ $(m = 1, 2, \ldots)$,

$$\lim_{n\to\infty} \frac{2n\left(\beta_n - \alpha_n\right)}{\int_0^1 \lambda^2(s)ds} = -\alpha + \int_0^1 f(u)P(\mathcal{E}(u))du.$$

Proof of Theorem 2.2 Since

$$\frac{2n(\alpha - \beta_n)}{\int_0^1 \lambda^2(s)ds} = \frac{2n(\alpha - \alpha_n)}{\int_0^1 \lambda^2(s)ds} - \frac{2n(\beta_n - \alpha_n)}{\int_0^1 \lambda^2(s)ds},$$

Theorem 2.2 follows from Theorem 2.1 and Lemma 2.2. The proof is complete. □

Remark 2.2 Similarly to Eq. 2.18, it can be shown that

$$\alpha - \beta_n = \frac{1}{2n}\left\{E\left[\nu(\Pi) + \tilde{\nu}(\Pi)\right] + \int_0^1 \lambda^2(s)ds\left[\alpha - \int_0^1 f(u)P\left(\mathcal{E}(u)\right)du\right]\right\} + \mathcal{O}(n^{-2}). \tag{2.19}$$

Remark 2.3 Theorems 2.1, 2.2 and Lemma 2.1 of Yao et al. (2017) are, respectively, special cases of our Theorems 2.1, 2.2 and Lemma 2.2 when $r^* = 1$ and $\lambda(t) = \lambda$ (constant). Our proofs of the main results follow those in Yao et al. (2017) closely.

3 Numerical Results and Discussion

In this section, we compare several approximations numerically for $r^* = 1, 2, 3$ and various combinations of $(w_1, \ldots, w_{r^*})$ and $(k_1, \ldots, k_{r^*})$. Four intensity functions are considered: $\lambda(t) = 5$ (constant),

$$\lambda_1(t) = 2 + (12 - 16t)\mathbf{1}_{(1/4,3/4]}(t), \ t \in (0, 1],$$
$$\lambda_2(t) = 2 + \left(8 - 32\left|t - \frac{1}{2}\right|\right)\mathbf{1}_{(1/4,3/4]}(t), \ t \in (0, 1],$$
$$\lambda_3(t) = 2 + 8 \cdot \mathbf{1}_{(1/4,3/4]}(t), \ t \in (0, 1],$$

where the last three intensity functions are plotted in Fig. 1. While $\alpha - \alpha_n = \mathcal{O}(n^{-1})$ and $\alpha - \beta_n = \mathcal{O}(n^{-1})$ (implying that α_n and β_n are first-order approximations of α), we have by Eqs. 2.18 and 2.19 that

$$\alpha = \alpha_n + C_\alpha n^{-1} + \mathcal{O}(n^{-2}) \ \text{ and } \ \alpha = \beta_n + C_\beta n^{-1} + \mathcal{O}(n^{-2}), \tag{3.1}$$

where

$$C_\alpha = \frac{1}{2}E[\nu(\Pi) + \tilde{\nu}(\Pi)]$$

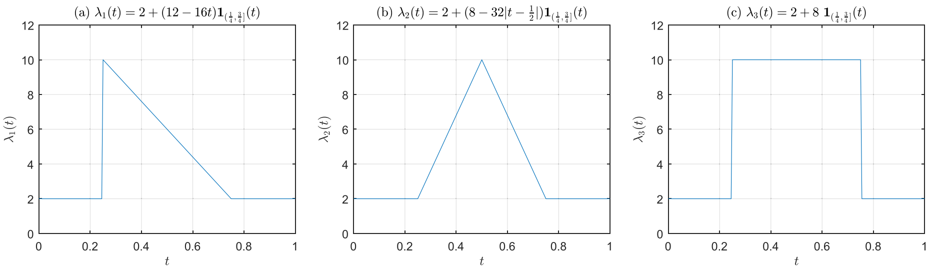

Fig. 1 The three intensity functions $\lambda_i(t)$, $i = 1, 2, 3$

and

$$C_\beta = \frac{1}{2}\left\{E[\nu(\Pi) + \tilde{\nu}(\Pi)] + \int_0^1 \lambda^2(s)ds\left[\alpha - \int_0^1 f(u)P(\mathcal{E}(u))du\right]\right\}.$$

For an even integer n with $n/2$ a multiple of q, define the corrected discrete approximations

$$\tilde{\alpha}_n = 2\alpha_n - \alpha_{n/2} \quad \text{and} \quad \tilde{\beta}_n = 2\beta_n - \beta_{n/2},$$

which are Richardson's extrapolations. It follows from Eq. 3.1 that $\alpha - \tilde{\alpha}_n = \mathcal{O}(n^{-2})$ and $\alpha - \tilde{\beta}_n = \mathcal{O}(n^{-2})$, i.e. $\tilde{\alpha}_n$ and $\tilde{\beta}_n$ are second-order approximations of α.

Table 1 presents the values of $n(\alpha - \alpha_n)$ and $n(\alpha - \beta_n)$ for $n = 100, 200, 400, 800$. If the exact value of α is available, $n(\alpha - \alpha_n)$ and $n(\alpha - \beta_n)$ should converge as $n \to \infty$ to C_α and C_β, respectively. Except for the case $r^* = 1$ and $\lambda(t) = 5$ for which the exact α is available from Neff and Naus (1980), we carried out Monte Carlo simulations to estimate α as well as C_α and C_β. For the case $r^* = 1$ and $\lambda(t) = 5$, $n(\alpha - \alpha_n)$ and $n(\alpha - \beta_n)$ appear to be monotonically increasing in n. For some other cases (e.g. $(w_1, w_2, w_3) = (0.05, 0.1, 0.2)$), the convergence of $n(\alpha - \alpha_n)$ and $n(\alpha - \beta_n)$ appears to be slow, which is due partly to the simulation error in α and partly to the small value of $w_1 = 0.05$. Note that in order for the discrete approximations to be accurate, nw_r, $r = 1, \ldots, r^*$, need to be large (i.e. each window is required to cover a large number of subintervals). Thus, if some w_r is small, n needs to be large in order for the approximations to be accurate. Additional numerical results on the convergence of $n(\alpha - \alpha_n)$ and $n(\alpha - \beta_n)$ are given in Table 2 for different combinations of $(w_1, \ldots, w_{r^*})$ and $(k_1, \ldots, k_{r^*})$.

Tables 3, 4 and 5 present approximation errors $\alpha - \alpha_n$, $\alpha - \beta_n$, $\alpha - \tilde{\alpha}_n$ and $\alpha - \tilde{\beta}_n$. According to the tables, when n doubles, the errors of α_n and β_n decrease by roughly a factor of 2 as expected. On the other hand, as n doubles, the errors of $\tilde{\alpha}_n$ and $\tilde{\beta}_n$ are expected to decrease by roughly a factor of 4, which, however, is not so in some cases. Again, it is due, to some extent, to the simulation error in α. As an example, in Table 3, for the four intensity functions and for $(w_1, w_2) = (0.1, 0.2)$ and $(k_1, k_2) = (3, 5)$, $\alpha - \tilde{\alpha}_{400}$, $\alpha - \tilde{\alpha}_{800}$, $\alpha - \tilde{\beta}_{400}$ and $\alpha - \tilde{\beta}_{800}$ are either comparable (in magnitude) to or smaller than the s.e. of the estimated value of α.

Remark 3.1 In Tables 1, 2, 3, 4 and 5, the Monte Carlo simulation estimates of α, C_α and C_β were based on 10^7 replications for each case. As a result, the standard error (s.e.) for the estimate of α in each case is about 10^{-4}. The first-order approximations α_n and β_n were both computed via the Markov chain embedding method. More precisely, for $\lambda(t) = \lambda$ (constant), $1 - \alpha_n$ and $1 - \beta_n$ were obtained from formula (2.2) in Lemma 2.1 of Wu et al. (2013) where the parameter p in the (essential) transition matrix $\Lambda^E(k; p)$ equals $p = P(I_i^n = 1) = 1 - e^{-\lambda/n}$ for α_n and $p = P(H_i^n = 1) = \lambda/n$ for β_n. For time-varying

Table 1 Convergence of $n(\alpha - \alpha_n)$, $n(\alpha - \beta_n)$

			Constant intensity $\lambda = 5$			Varying intensity $\lambda_1(t)$		
w	k	n	α (Sim.)	$n(\alpha - \alpha_n)$	$n(\alpha - \beta_n)$	α (Sim.)	$n(\alpha - \alpha_n)$	$n(\alpha - \beta_n)$
0.2	5	100		1.783875	1.421022		2.166863	1.501655
		200	0.052686*	1.871392	1.443581	0.067985	2.253884	1.495892
		400		1.914215	1.453364	s.e. 0.000080	2.302179	1.498821
		800		1.935372	1.457903		2.339399	1.513656
		(C_α, C_β)		1.955056	1.462400		2.316428	1.471481
				s.e. 0.001547	s.e. 0.001520		s.e. 0.001942	s.e. 0.001597
(0.1, 0.2)	(3, 5)	100		5.579473	4.343425		5.116425	3.469959
		200	0.270642	5.547467	4.220263	0.258479	5.063975	3.326967
		400	s.e. 0.000140	5.542565	4.174429	s.e. 0.000138	5.041127	3.265011
		800		5.567864	4.180371		5.040317	3.246068
		(C_α, C_β)		5.543283	4.136165		5.063219	3.251636
				s.e. 0.002125	s.e. 0.002902		s.e. 0.002159	s.e. 0.002503
(0.05, 0.1, 0.2)	(2, 3, 5)	100		6.899022	5.287630		6.402231	4.395108
		200	0.595317	6.547673	4.934020	0.528701	6.067560	4.056523
		400	s.e. 0.000155	6.430567	4.820527	s.e. 0.000158	5.897306	3.889827
		800		6.466191	4.858929		5.787813	3.783242
		(C_α, C_β)		6.616034	5.015142		6.262502	4.259704
				s.e. 0.002923	s.e. 0.004019		s.e. 0.002586	s.e. 0.003378
			Varying intensity $\lambda_2(t)$			Varying intensity $\lambda_3(t)$		
0.2	5	100		2.352675	1.566094		6.350349	3.507697
		200	0.077496	2.427743	1.544600	0.242347	6.379634	3.338109
		400	s.e. 0.000085	2.465611	1.536070	s.e. 0.000136	6.357193	3.229034

Table 1 (continued)

			Varying intensity $\lambda_2(t)$			Varying intensity $\lambda_3(t)$		
w	k	n	α (Sim.)	$n(\alpha - \alpha_n)$	$n(\alpha - \beta_n)$	α (Sim.)	$n(\alpha - \alpha_n)$	$n(\alpha - \beta_n)$
		800		2.489524	1.537280		6.288143	3.119705
		(C_α, C_β)		2.490623	1.514652		6.444651	3.227893
				s.e. 0.002016	s.e. 0.001639		s.e. 0.003707	s.e. 0.004476
(0.1, 0.2)	(3, 5)	100		5.163134	3.413016		8.021877	4.260943
		200	0.269728	5.090393	3.256269	0.538275	7.711972	3.961684
		400	s.e. 0.000140	5.052224	3.182563	s.e. 0.000158	7.602448	3.868851
		800		5.032508	3.146552		7.630110	3.907191
		(C_α, C_β)		5.097693	3.197983		7.516576	3.806683
				s.e. 0.002174	s.e. 0.002513		s.e. 0.003638	s.e. 0.005565
(0.05, 0.1, 0.2)	(2, 3, 5)	100		6.351726	4.312354		6.014357	3.140592
		200	0.533713	6.008806	3.970403	0.785341	5.473962	2.762562
		400	s.e. 0.000158	5.829622	3.797060	s.e. 0.000130	5.199738	2.567171
		800		5.705238	3.676684		5.006989	2.413049
		(C_α, C_β)		6.241222	4.214804		5.610689	3.049688
				s.e. 0.002594	s.e. 0.003373		s.e. 0.003744	s.e. 0.005303

Note: 1. All α's are from Monte Carlo simulation (together with s.e.) except the one with * which is exact taken from Neff and Naus (1980)

2. The intensity functions $\lambda_i(t)$, $i = 1, 2, 3$ are shown in Fig. 1

3. (C_α, C_β) are evaluated by simulation with s.e. given

Table 2 Convergence of $n(\alpha - \alpha_n)$, $n(\alpha - \beta_n)$

			Constant intensity $\lambda = 5$			Varying intensity $\lambda_1(t)$		
w	k	n	α (Sim.)	$n(\alpha - \alpha_n)$	$n(\alpha - \beta_n)$	α (Sim.)	$n(\alpha - \alpha_n)$	$n(\alpha - \beta_n)$
0.3	6	50		1.210475	0.970163		1.590956	1.200471
		100	0.042262*	1.348097	1.007559	0.053299	1.628382	1.042952
		200		1.416403	1.022659	s.e. 0.000071	1.693663	1.033281
		400		1.450285	1.029387		1.717407	1.019878
		(C_α, C_β)		1.485985	1.037587		1.771787	1.036963
				s.e. 0.001479	s.e. 0.001345		s.e. 0.001825	s.e. 0.001389
(0.2, 0.3)	(4, 6)	50		3.701318	2.713337		3.965497	2.650760
		100	0.197787	3.765826	2.611111	0.200351	3.631711	2.049978
		200	s.e. 0.000126	3.782585	2.553607	s.e. 0.000127	3.645939	1.992292
		400		3.777033	2.513183		3.661708	1.975424
		(C_α, C_β)		3.854309	2.554010		3.662855	1.946738
				s.e. 0.001913	s.e. 0.002318		s.e. 0.002043	s.e. 0.002034
(0.1, 0.2, 0.3)	(3, 5, 6)	50		5.601585	4.578668		5.554214	4.217994
		100	0.270926	5.545310	4.302990	0.258919	5.097212	3.442906
		200	s.e. 0.000141	5.495034	4.162755	s.e. 0.000139	5.044659	3.301585
		400		5.452492	4.079797		5.019683	3.238230
		(C_α, C_β)		5.579558	4.165819		5.094206	3.277857
				s.e. 0.002157	s.e. 0.002922		s.e. 0.002189	s.e. 0.002520
			Varying intensity $\lambda_2(t)$			Varying intensity $\lambda_3(t)$		
0.3	6	50		1.603078	1.087410		5.381152	2.765323
		100	0.059688	1.740869	1.068338	0.232019	5.551470	2.536197
		200	s.e. 0.000075	1.808842	1.060345	s.e. 0.000133	5.606385	2.431738

Table 2 (continued)

			Varying intensity $\lambda_2(t)$			Varying intensity $\lambda_3(t)$		
w	k	n	α (Sim.)	$n(\alpha - \alpha_n)$	$n(\alpha - \beta_n)$	α (Sim.)	$n(\alpha - \alpha_n)$	$n(\alpha - \beta_n)$
		400		1.846485	1.060811		5.624760	2.379737
		(C_α, C_β)		1.862787	1.041713		5.653973	2.338708
				s.e. 0.001858	s.e. 0.001396		s.e. 0.003750	s.e. 0.004171
(0.2, 0.3)	(4, 6)	50		3.685295	2.103135		6.695968	2.652311
		100	0.215789	3.680820	1.941068	0.488074	6.376600	2.305578
		200	s.e. 0.000130	3.676232	1.873913	s.e. 0.000158	6.208021	2.156671
		400		3.678060	1.847993		6.115070	2.079816
		(C_α, C_β)		3.684996	1.832350		6.155314	2.135779
				s.e. 0.002038	s.e. 0.002026		s.e. 0.003326	s.e. 0.004886
(0.1, 0.2, 0.3)	(3, 5, 6)	50		5.267547	3.733335		8.578640	4.870774
		100	0.270142	5.135646	3.377127	0.539592	7.899934	4.111400
		200	s.e. 0.000140	5.058063	3.217566	s.e. 0.000158	7.575927	3.806940
		400		5.007592	3.132378		7.421202	3.672221
		(C_α, C_β)		5.118366	3.211826		7.610207	3.885833
				s.e. 0.002204	s.e. 0.002532		s.e. 0.003738	s.e. 0.005612

Table 3 Convergence of $\alpha_n, \beta_n, \tilde{\alpha}_n, \tilde{\beta}_n$

			Constant intensity $\lambda = 5$					Varying intensity $\lambda_1(t)$				
w	k	n	α (Sim.)	α_n	β_n	$\tilde{\alpha}_n$	$\tilde{\beta}_n$	α (Sim.)	α_n	β_n	$\tilde{\alpha}_n$	$\tilde{\beta}_n$
0.2	5	100		0.034848	0.038476				0.046317	0.052969		
				(0.017839)	(0.014210)				(0.021669)	(0.015017)		
		200		0.043329	0.045469	0.051811	0.052461		0.056716	0.060506	0.067115	0.068043
			0.052686*	(0.009357)	(0.007218)	(0.000875)	(0.000226)	0.067985	(0.011269)	(0.007479)	(0.000870)	(-0.000058)
		400		0.047901	0.049053	0.052472	0.052638	s.e. 0.000080	0.062230	0.064238	0.067744	0.067971
				(0.004786)	(0.003633)	(0.000214)	(0.000049)		(0.005755)	(0.003747)	(0.000241)	(0.000015)
		800		0.050267	0.050864	0.052634	0.052675		0.065061	0.066093	0.067892	0.067948
				(0.002419)	(0.001822)	(0.000053)	(0.000011)		(0.002924)	(0.001892)	(0.000093)	(0.000037)
(0.1, 0.2)	(3, 5)	100		0.214848	0.227208				0.207315	0.223779		
				(0.055795)	(0.043434)				(0.051164)	(0.034700)		
		200		0.242905	0.249541	0.270962	0.271874		0.233159	0.241844	0.259003	0.259909
			0.270642	(0.027737)	(0.021101)	(-0.000320)	(-0.001232)	0.258479	(0.025320)	(0.016635)	(-0.000525)	(-0.001430)
		400	s.e. 0.000140	0.256786	0.260206	0.270667	0.270871	s.e. 0.000138	0.245876	0.250316	0.258593	0.258789
				(0.013856)	(0.010436)	(-0.000025)	(-0.000229)		(0.012603)	(0.008163)	(-0.000114)	(-0.000310)
		800		0.263682	0.265417	0.270579	0.270627		0.252179	0.254421	0.258481	0.258526
				(0.006960)	(0.005225)	(0.000063)	(0.000015)		(0.006300)	(0.004058)	(-0.000002)	(-0.000047)
(0.05, 0.1, 0.2)	(2, 3, 5)	100		0.526327	0.542441				0.464678	0.484749		
				(0.068990)	(0.052876)				(0.064022)	(0.043951)		
		200		0.562579	0.570647	0.598831	0.598853		0.498363	0.508418	0.532047	0.532086
			0.595317	(0.032738)	(0.024670)	(-0.003513)	(-0.003536)	0.528701	(0.030338)	(0.020283)	(-0.003347)	(-0.003386)
		400	s.e. 0.000155	0.579241	0.583266	0.595903	0.595885	s.e. 0.000158	0.513957	0.518976	0.529552	0.529534
				(0.016076)	(0.012051)	(-0.000586)	(-0.000567)		(0.014743)	(0.009725)	(-0.000851)	(-0.000833)
		800		0.587234	0.589244	0.595228	0.595221		0.521466	0.523971	0.528974	0.528967
				(0.008083)	(0.006074)	(0.000089)	(0.000096)		(0.007235)	(0.004729)	(-0.000274)	(-0.000266)

Table 3 (continued)

			Varying intensity $\lambda_2(t)$					Varying intensity $\lambda_3(t)$				
w	k	n	α (Sim.)	α_n	β_n	$\tilde{\alpha}_n$	$\tilde{\beta}_n$	α (Sim.)	α_n	β_n	$\tilde{\alpha}_n$	$\tilde{\beta}_n$
0.2	5	100		0.053969	0.061835				0.178844	0.207270		
				(0.023527)	(0.015661)				(0.063503)	(0.035077)		
		200		0.065357	0.069773	0.076745	0.077711		0.210449	0.225657	0.242055	0.244043
			0.077496	(0.012139)	(0.007723)	(0.000751)	(-0.000215)	0.242347	(0.031898)	(0.016691)	(0.000293)	(-0.001696)
		400	s.e. 0.000085	0.071332	0.073656	0.077306	0.077538	s.e. 0.000136	0.226454	0.234275	0.242460	0.242893
				(0.006164)	(0.003840)	(0.000189)	(-0.000043)		(0.015893)	(0.008073)	(-0.000112)	(-0.000545)
		800		0.074384	0.075574	0.077436	0.077493		0.234487	0.238448	0.242520	0.242621
				(0.003112)	(0.001922)	(0.000060)	(0.000003)		(0.007860)	(0.003900)	(-0.000173)	(-0.000273)
(0.1, 0.2)	(3, 5)	100		0.218097	0.235598				0.458056	0.495665		
				(0.051631)	(0.034130)				(0.080219)	(0.042609)		
		200		0.244276	0.253447	0.270455	0.271295		0.499715	0.518466	0.541374	0.541267
			0.269728	(0.025452)	(0.016281)	(-0.000727)	(-0.001567)	0.538275	(0.038560)	(0.019808)	(-0.003099)	(-0.002993)
		400	s.e. 0.000140	0.257097	0.261772	0.269919	0.270097	s.e. 0.000158	0.519269	0.528603	0.538822	0.538739
				(0.012631)	(0.007956)	(-0.000191)	(-0.000369)		(0.019006)	(0.009672)	(-0.000548)	(-0.000464)
		800		0.263437	0.265795	0.269777	0.269818		0.528737	0.533391	0.538206	0.538179
				(0.006291)	(0.003933)	(-0.000049)	(-0.000090)		(0.009538)	(0.004884)	(0.000069)	(0.000096)
(0.05, 0.1, 0.2)	(2, 3, 5)	100		0.470196	0.490589				0.725198	0.753935		
				(0.063517)	(0.043124)				(0.060144)	(0.031406)		
		200		0.503669	0.513861	0.537142	0.537132		0.757971	0.771528	0.790745	0.789122
			0.533713	(0.030044)	(0.019852)	(-0.003429)	(-0.003420)	0.785341	(0.027370)	(0.013813)	(-0.005404)	(-0.003780)
		400	s.e. 0.000158	0.519139	0.524220	0.534609	0.534580	s.e. 0.000130	0.772342	0.778923	0.786712	0.786318
				(0.014574)	(0.009493)	(-0.000896)	(-0.000867)		(0.012999)	(0.006418)	(-0.001371)	(-0.000977)
		800		0.526581	0.529117	0.534024	0.534014		0.779083	0.782325	0.785823	0.785727
				(0.007132)	(0.004596)	(-0.000311)	(-0.000301)		(0.006259)	(0.003016)	(-0.000482)	(-0.000385)

Note 1. Given in parentheses are approximations errors ($\alpha - \alpha_n, \alpha - \beta_n, \alpha - \tilde{\alpha}_n, \alpha - \tilde{\beta}_n$) where the value of α is found by simulation

Table 4 Convergence of $\alpha_n, \beta_n, \tilde{\alpha}_n, \tilde{\beta}_n$

			Constant intensity $\lambda = 5$					Varying intensity $\lambda_1(t)$				
w	k	n	α (Sim.)	α_n	β_n	$\tilde{\alpha}_n$	$\tilde{\beta}_n$	α (Sim.)	α_n	β_n	$\tilde{\alpha}_n$	$\tilde{\beta}_n$
0.3	6	50		0.018052	0.022858				0.021480	0.029290		
				(0.024209)	(0.019403)				(0.031819)	(0.024009)		
		100		0.028781	0.032186	0.039509	0.041514		0.037015	0.042870	0.052551	0.056449
			0.042262*	(0.013481)	(0.010076)	(0.002752)	(0.000748)	0.053299	(0.016284)	(0.010430)	(0.000749)	(-0.003150)
		200		0.035180	0.037148	0.041579	0.042111	s.e. 0.000071	0.044831	0.048133	0.052646	0.053396
				(0.007082)	(0.005113)	(0.000683)	(0.000151)		(0.008468)	(0.005166)	(0.000653)	(-0.000097)
		400		0.038636	0.039688	0.042092	0.042228		0.049006	0.050749	0.053180	0.053366
				(0.003626)	(0.002573)	(0.000169)	(0.000034)		(0.004294)	(0.002550)	(0.000119)	(-0.000067)
(0.2, 0.3)	(4, 6)	50		0.123761	0.143521				0.121041	0.147336		
				(0.074026)	(0.054267)				(0.079310)	(0.053015)		
		100		0.160129	0.171676	0.196497	0.199832		0.164034	0.179851	0.207027	0.212367
			0.197787	(0.037658)	(0.026111)	(0.001290)	(-0.002045)	0.200351	(0.036317)	(0.020500)	(-0.006676)	(-0.012016)
		200	s.e. 0.000126	0.178874	0.185019	0.197620	0.198362	s.e. 0.000127	0.182121	0.190390	0.200209	0.200928
				(0.018913)	(0.012768)	(0.000168)	(-0.000575)		(0.018230)	(0.009961)	(0.000142)	(-0.000577)
		400		0.188345	0.191504	0.197815	0.197989		0.191197	0.195413	0.200272	0.200435
				(0.009443)	(0.006283)	(-0.000028)	(-0.000202)		(0.009154)	(0.004939)	(0.000079)	(-0.000084)
(0.1, 0.2, 0.3)	(3, 5, 6)	50		0.158894	0.179352				0.147835	0.174559		
				(0.112032)	(0.091573)				(0.111084)	(0.084360)		
		100		0.215472	0.227896	0.272051	0.276439		0.207947	0.224490	0.268059	0.274421
			0.270926	(0.055453)	(0.043030)	(-0.001125)	(-0.005514)	0.258919	(0.050972)	(0.034429)	(-0.009140)	(-0.015502)
		200	s.e. 0.000141	0.243450	0.250112	0.271428	0.272328	s.e. 0.000139	0.233696	0.242411	0.259445	0.260332
				(0.027475)	(0.020814)	(-0.000503)	(-0.001402)		(0.025223)	(0.016508)	(-0.000526)	(-0.001413)
		400		0.257294	0.260726	0.271138	0.271340		0.246370	0.250824	0.259044	0.259236
				(0.013631)	(0.010199)	(-0.000213)	(-0.000415)		(0.012549)	(0.008096)	(-0.000125)	(-0.000317)

Table 4 (continued)

w	k	n	Varying intensity $\lambda_2(t)$					Varying intensity $\lambda_3(t)$				
			α (Sim.)	α_n	β_n	$\tilde{\alpha}_n$	$\tilde{\beta}_n$	α (Sim.)	α_n	β_n	$\tilde{\alpha}_n$	$\tilde{\beta}_n$
0.3	6	50		0.027627	0.037940				0.124396	0.176713		
				(0.032062)	(0.021748)				(0.107623)	(0.055306)		
		100		0.042280	0.049005	0.056932	0.060070		0.176505	0.206657	0.228613	0.236602
			0.059688	(0.017409)	(0.010683)	(0.002756)	(-0.000381)	0.232019	(0.055515)	(0.025362)	(0.003406)	(-0.004583)
		200	s.e. 0.000075	0.050644	0.054386	0.059008	0.059768	s.e. 0.000133	0.203987	0.219861	0.231470	0.233064
				(0.009044)	(0.005302)	(0.000680)	(-0.000080)		(0.028032)	(0.012159)	(0.000549)	(-0.001045)
		400		0.055072	0.057036	0.059500	0.059686		0.217958	0.226070	0.231928	0.232279
				(0.004616)	(0.002652)	(0.000188)	(0.000002)		(0.014062)	(0.005949)	(0.000092)	(-0.000260)
(0.2, 0.3)	(4, 6)	50		0.142083	0.173726				0.354155	0.435028		
				(0.073706)	(0.042063)				(0.133919)	(0.053046)		
		100		0.178981	0.196378	0.215879	0.219030		0.424308	0.465019	0.494462	0.495009
			0.215789	(0.036808)	(0.019411)	(-0.000090)	(-0.003241)	0.488074	(0.063766)	(0.023056)	(-0.006387)	(-0.006935)
		200	s.e. 0.000130	0.197408	0.206420	0.215835	0.216461	s.e. 0.000158	0.457034	0.477291	0.489760	0.489563
				(0.018381)	(0.009370)	(-0.000046)	(-0.000672)		(0.031040)	(0.010783)	(-0.001686)	(-0.001489)
		400		0.206594	0.211169	0.215780	0.215919		0.472787	0.482875	0.488539	0.488459
				(0.009195)	(0.004620)	(0.000009)	(-0.000130)		(0.015288)	(0.005200)	(-0.000465)	(-0.000384)
(0.1, 0.2, 0.3)	(3, 5, 6)	50		0.164791	0.195475				0.368019	0.442176		
				(0.105351)	(0.074667)				(0.171573)	(0.097415)		
		100		0.218786	0.236371	0.272780	0.277266		0.460592	0.498478	0.553166	0.554779
			0.270142	(0.051356)	(0.033771)	(-0.002638)	(-0.007124)	0.539592	(0.078999)	(0.041114)	(-0.013574)	(-0.015187)
		200	s.e. 0.000140	0.244852	0.254054	0.270918	0.271738	s.e. 0.000158	0.501712	0.520557	0.542832	0.542636
				(0.025290)	(0.016088)	(-0.000776)	(-0.001596)		(0.037880)	(0.019035)	(-0.003240)	(-0.003045)
		400		0.257623	0.262311	0.270395	0.270568		0.521039	0.530411	0.540365	0.540265
				(0.012519)	(0.007831)	(-0.000252)	(-0.000426)		(0.018553)	(0.009181)	(-0.000774)	(-0.000674)

Table 5 Results with parameters from Table 6 of Wu et al. (2013)

			Constant intensity $\lambda = 5$					Varying intensity $\lambda_1(t)$				
w	k	n	α (Sim.)	α_n	β_n	$\tilde{\alpha}_n$	$\tilde{\beta}_n$	α (Sim.)	α_n	β_n	$\tilde{\alpha}_n$	$\tilde{\beta}_n$
(0.01, 0.02, 0.03)	(2, 3, 4)	200		0.111753	0.114378				0.099866	0.103162		
				(0.095000)	(0.092375)				(0.084674)	(0.081378)		
		400		0.161156	0.162964	0.210559	0.211549		0.143955	0.146210	0.188044	0.189258
			0.206753	(0.045597)	(0.043789)	(-0.003806)	(-0.004796)	0.184540	(0.040585)	(0.038330)	(-0.003505)	(-0.004718)
		800	s.e. 0.000128	0.184421	0.185433	0.207686	0.207901	s.e. 0.000123	0.164715	0.165973	0.185475	0.185735
				(0.022332)	(0.021320)	(-0.000933)	(-0.001149)		(0.019825)	(0.018567)	(-0.000935)	(-0.001195)
		1600		0.195715	0.196246	0.207009	0.207059		0.174793	0.175452	0.184871	0.184931
				(0.011038)	(0.010507)	(-0.000256)	(-0.000306)		(0.009747)	(0.009088)	(-0.000331)	(-0.000391)
(0.02, 0.04)	(3, 4)	200		0.009208	0.009560				0.011338	0.012011		
				(0.013451)	(0.013099)				(0.015398)	(0.014725)		
		400		0.015332	0.015616	0.021456	0.021671		0.018492	0.019014	0.025647	0.026017
			0.022660	(0.007328)	(0.007044)	(0.001204)	(0.000989)	0.026736	(0.008243)	(0.007722)	(0.001088)	(0.000719)
		800	s.e. 0.000047	0.018799	0.018970	0.022265	0.022324	s.e. 0.000051	0.022490	0.022800	0.026487	0.026585
				(0.003861)	(0.003690)	(0.000395)	(0.000335)		(0.004246)	(0.003936)	(0.000249)	(0.000150)
		1600		0.020628	0.020722	0.022458	0.022473		0.024586	0.024753	0.026682	0.026707
				(0.002032)	(0.001938)	(0.000202)	(0.000187)		(0.002150)	(0.001982)	(0.000054)	(0.000029)
(0.03, 0.06)	(3, 6)	200		0.025733	0.026660				0.029436	0.031025		
				(0.018426)	(0.017499)				(0.020181)	(0.018592)		
		400		0.034579	0.035186	0.043424	0.043712		0.039202	0.040220	0.048968	0.049415
			0.044159	(0.009581)	(0.008973)	(0.000735)	(0.000447)	0.049617	(0.010415)	(0.009397)	(0.000649)	(0.000202)
		800	s.e. 0.000065	0.039306	0.039647	0.044033	0.044108	s.e. 0.000069	0.044366	0.044931	0.049530	0.049643
				(0.004854)	(0.004513)	(0.000126)	(0.000052)		(0.005251)	(0.004686)	(0.000087)	(-0.000026)
		1600		0.041741	0.041921	0.044177	0.044196		0.047013	0.047309	0.049659	0.049687
				(0.002418)	(0.002238)	(-0.000017)	(-0.000036)		(0.002604)	(0.002308)	(-0.000042)	(-0.000070)

Table 5 (continued)

			Varying intensity $\lambda_2(t)$					Varying intensity $\lambda_3(t)$				
w	k	n	α (Sim.)	α_n	β_n	$\tilde{\alpha}_n$	$\tilde{\beta}_n$	α (Sim.)	α_n	β_n	$\tilde{\alpha}_n$	$\tilde{\beta}_n$
(0.01, 0.02, 0.03)	(2, 3, 4)	200		0.100449	0.103778				0.210142	0.219089		
				(0.085310)	(0.081981)				(0.154060)	(0.145113)		
		400		0.144736	0.147010	0.189022	0.190243		0.292950	0.298653	0.375758	0.378217
			0.185759	(0.041023)	(0.038748)	(-0.003263)	(-0.004484)	0.364202	(0.071252)	(0.065549)	(-0.011556)	(-0.014015)
		800	s.e. 0.000123	0.165575	0.166843	0.186414	0.186676	s.e. 0.000152	0.329924	0.333001	0.366899	0.367349
				(0.020184)	(0.018915)	(-0.000656)	(-0.000917)		(0.034278)	(0.031201)	(-0.002697)	(-0.003147)
		1600		0.175688	0.176353	0.185802	0.185863		0.347418	0.349005	0.364913	0.365009
				(0.010070)	(0.009406)	(-0.000043)	(-0.000104)		(0.016784)	(0.015197)	(-0.000711)	(-0.000807)
(0.02, 0.04)	(3, 4)	200		0.011718	0.012423				0.035314	0.037972		
				(0.015745)	(0.015040)				(0.044157)	(0.041499)		
		400		0.019058	0.019601	0.026397	0.026779		0.056184	0.058164	0.077054	0.078356
			0.027463	(0.008405)	(0.007862)	(0.001066)	(0.000684)	0.079471	(0.023287)	(0.021307)	(0.002417)	(0.001115)
		800	s.e. 0.000052	0.023151	0.023474	0.027245	0.027346	s.e. 0.000086	0.067585	0.068740	0.078985	0.079316
				(0.004312)	(0.003989)	(0.000218)	(0.000117)		(0.011886)	(0.010731)	(0.000486)	(0.000155)
		1600		0.025296	0.025470	0.027441	0.027467		0.073494	0.074113	0.079404	0.079487
				(0.002167)	(0.001993)	(0.000022)	(-0.000004)		(0.005977)	(0.005358)	(0.000067)	(-0.000016)
(0.03, 0.06)	(3, 6)	200		0.030470	0.032135				0.085390	0.091114		
				(0.020659)	(0.018995)				(0.053870)	(0.048145)		
		400		0.040519	0.041583	0.050568	0.051031		0.111760	0.115311	0.138131	0.139508
			0.051129	(0.010610)	(0.009546)	(0.000561)	(0.000098)	0.139260	(0.027499)	(0.023949)	(0.001129)	(-0.000248)
		800	s.e. 0.000070	0.045823	0.046413	0.051127	0.051243	s.e. 0.000109	0.125372	0.127312	0.138984	0.139313
				(0.005306)	(0.004716)	(0.000002)	(-0.000114)		(0.013887)	(0.011948)	(0.000275)	(-0.000054)
		1600		0.048539	0.048849	0.051255	0.051284		0.132264	0.133273	0.139155	0.139235
				(0.002590)	(0.002281)	(-0.000126)	(-0.000155)		(0.006996)	(0.005986)	(0.000105)	(0.000025)

Note: 1. The discrete approximation results for $\lambda = 5$ from Wu et al. (2013) are: 0.2076, 0.0206, 0.0404 (upper panel on the left, from top to bottom)

$\lambda(t)$, α_n and β_n were found by the formula in Remark 4.1 of Wu et al. (2013) where the value of p varies from subinterval to subinterval. Specifically, for the i-th subinterval,

$$p_i = P(I_i^n = 1) = 1 - \exp\left(-\int_{(i-1)/n}^{i/n} \lambda(s)ds\right) \text{ for } \alpha_n$$

$$\text{and} \qquad p_i = P(H_i^n = 1) = \int_{(i-1)/n}^{i/n} \lambda(s)ds \text{ for } \beta_n.$$

Remark 3.2 Scan statistics arise naturally in search of clusters of events. In the one-dimensional setting, the sequence of events (denoted by Π) is often modeled as a Poisson process with (unknown) intensity function $\lambda(t)$, which is typically assumed to be constant under the null hypothesis. The referee of this paper raised the question of how to apply a classic scan statistic when the null hypothesis is that $\lambda(t) = \lambda_0(t)$ for some specified (nonconstant) function $\lambda_0(t)$. While to address this issue thoroughly is beyond the scope of this paper, it seems reasonable to consider the following approach via a change of time. Assuming $\lambda_0(t) > 0$ for $t \in (0, 1]$, define

$$T(t) = \int_0^t \lambda_0(s)ds \Big/ \overline{\lambda},\ 0 < t \le 1 \ \left(\overline{\lambda} := \int_0^1 \lambda_0(s)ds\right),$$

which makes a one-to-one transformation from $(0, 1]$ onto $(0, 1]$. Denoting $\Pi = \{Q_1, \ldots, Q_M\}$ with $M := |\Pi|$, let $T(\Pi) := \{T(Q_1), \ldots, T(Q_M)\}$. Since Π is a Poisson process with intensity function $\lambda_0(t)$, it is readily seen that $T(\Pi)$ is a (homogeneous) Poisson process with constant intensity $\overline{\lambda}$. For given $w > 0$, consider the scan statistic S_w applied to $T(\Pi)$, i.e.

$$S_w(T(\Pi)) = \max_{0 \le t \le 1-w} |T(\Pi) \cap (t, t+w]|.$$

The (corrected) discrete approximations for $P(S_w(T(\Pi)) \ge k)$ can be readily computed via the Markov chain embedding method. Note that

$$\begin{aligned} S_w(T(\Pi)) &= \max_{0 \le s \le 1-w} |\{T(Q_1), \ldots, T(Q_M)\} \cap (s, s+w]| \\ &= \max_{0 \le s \le 1-w} |\{Q_1, \ldots, Q_M\} \cap (T^{-1}(s), T^{-1}(s+w)]| \\ &= \max_{0 \le t \le T^{-1}(1-w)} |\Pi \cap (t, T^{-1}(T(t)+w)]| \\ &= \max_{0 \le t \le T^{-1}(1-w)} |\Pi \cap (t, t+w_T(t)]|, \end{aligned}$$

where $w_T(t) = T^{-1}(T(t)+w) - t$, $0 \le t \le T^{-1}(1-w)$. This shows that $S_w(T(\Pi))$ is equivalent to the classic scan statistic applied to Π with (time-varying) window size $w_T(t)$. Furthermore, the window size $w_T(t)$ at t is small (large, *resp.*) if λ_0 is large (small, *resp.*) in the vicinity of t. It is also worth noting the relationship between $S_w(T(\Pi))$ and the weighted scan statistic proposed in Wu (2017).

Remark 3.3 In this paper, we have only considered approximating the unconditional probability $P(S_{w_r} \ge k_r$ for some $1 \le r \le r^*)$ for multiple window scan statistics of general one-dimensional Poisson processes. While, in principle, a discrete approximation to the conditional probability $P(S_{w_r} \ge k_r$ for some $1 \le r \le r^* \mid |\Pi| = N)$ can be made in a similar fashion, it is unclear how to compute the resulting discrete approximation efficiently for nonhomogeneous Poisson processes. In particular, the Markov chain embedding method is not readily applicable to the setting of nonhomogeneous Poisson processes.

4 Proofs of Lemmas 2.1 and 2.2

Proof of Lemma 2.1 Noting that $\mathcal{A}_n \subset \mathcal{A}$, we have $\alpha - \alpha_n = P(\mathcal{A}) - P(\mathcal{A}_n) = P(\mathcal{A} \cap \mathcal{A}_n^c)$. Consider the following disjoint events

$$\begin{aligned}\mathcal{G}_1 &= \{\tilde{I}_j^n \le 1,\ j = 1, \ldots, n\},\\ \mathcal{G}_{2,i} &= \{\tilde{I}_i^n = 2, \tilde{I}_j^n \le 1 \text{ for all } j \ne i\},\ i = 1, \ldots, n,\\ \mathcal{G}_3 &= \{\tilde{I}_j^n = \tilde{I}_{j'}^n = 2 \text{ for some } j \ne j'\} \cup \{\tilde{I}_j^n \ge 3 \text{ for some } j\}.\end{aligned}$$

We have

$$\begin{aligned}\alpha - \alpha_n &= P(\mathcal{A} \cap \mathcal{A}_n^c)\\ &= P(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_1) + \sum_{i=1}^{n} P(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_{2,i}) + P(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_3). \qquad (4.1)\end{aligned}$$

By Eq. 4.1, the lemma follows if we can show that

$$P(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_1) = \frac{1}{2} \sum_{r=1}^{r^*} \sum_{i=1}^{n-nw_r} \mathcal{P}_{i,r}^{(n)} + \mathcal{O}(n^{-2}), \qquad (4.2)$$

$$\sum_{i=1}^{n} P(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_{2,i}) = \sum_{i=1}^{n} \tilde{\mathcal{P}}_i^{(n)} + \mathcal{O}(n^{-2}), \qquad (4.3)$$

$$P(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_3) = \mathcal{O}(n^{-2}). \qquad (4.4)$$

By condition (A), we have

$$\begin{aligned} P(\tilde{I}_i^n = 2) &= \tfrac{1}{2!}\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right)^2 \exp\left(-\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right) = \mathcal{O}(n^{-2}),\\ \text{and } P(\tilde{I}_i^n = 3) &= \tfrac{1}{3!}\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right)^3 \exp\left(-\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right) = \mathcal{O}(n^{-3}),\end{aligned}$$

where the $\mathcal{O}(n^{-2})$ and $\mathcal{O}(n^{-3})$ terms are uniform in i, which implies that $P(\mathcal{G}_3) = \mathcal{O}(n^{-2})$. This establishes (4.4).

To prove (4.2), note that when $\tilde{I}_i^n \le 1$ for all i (i.e. on the event $\mathcal{G}_1$), each subinterval $((i-1)/n, i/n]$ contains at most one Poisson point. If $\tilde{I}_i^n = 1$, denote the only Poisson point in $((i-1)/n, i/n]$ by $Q_{(i)}$ whose location is distributed with the probability density function

$$\lambda(x) \Big/ \int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds, \quad x \in \left(\frac{i-1}{n}, \frac{i}{n}\right].$$

When $\tilde{I}_i^n \le 1$ for all i, in order for $\mathcal{A} \cap \mathcal{A}_n^c$ to occur, there must exist some pair (i, r) such that $i + nw_r \le n$ and

$$\sum_{j=i+1}^{i+nw_r-1} \tilde{I}_j^n = k_r - 2,\ \tilde{I}_i^n = \tilde{I}_{i+nw_r}^n = 1 \text{ and } Q_{(i+nw_r)} - Q_{(i)} < w_r.$$

So we have $\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_1 = \bigcup_{r=1}^{r^*} \bigcup_{i=1}^{n-nw_r} \mathcal{G}_{1,i}^{(r)}$ where for $r = 1, \dots, r^*$ and $i = 1, \dots, n - nw_r$,

$$\mathcal{G}_{1,i}^{(r)} = \mathcal{A}_n^c \cap \left\{ \tilde{I}_j^n \le 1 \text{ for all } j, \sum_{j=i+1}^{i+nw_r-1} \tilde{I}_j^n = k_r - 2, \tilde{I}_i^n = \tilde{I}_{i+nw_r}^n = 1 \text{ and } Q_{(i+nw_r)} - Q_{(i)} < w_r \right\}.$$

Letting $\tilde{I}_i^n = 0$ for $i > n$, we have

$$\sum_{\substack{1 \le i, j \le n \\ 1 \le r, s \le r^* \\ (i,r) \neq (j,s)}} P\left(\tilde{I}_i^n = \tilde{I}_{i+nw_r}^n = \tilde{I}_j^n = \tilde{I}_{j+nw_s}^n = 1\right) = \mathcal{O}(n^{-2}).$$

It follows that

$$P(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_1) = \sum_{r=1}^{r^*} \sum_{i=1}^{n-nw_r} P(\mathcal{G}_{1,i}^{(r)}) + \mathcal{O}(n^{-2}). \tag{4.5}$$

For $r = 1, \dots, r^*$ and $i = 1, \dots, n - nw_r$, let

$$\mathcal{G}_{1,i}^{\prime(r)} = \mathcal{A}_n^c \cap \left\{ \tilde{I}_j^n \le 1 \text{ for all } j, \sum_{j=i+1}^{i+nw_r-1} \tilde{I}_j^n = k_r - 2, \tilde{I}_i^n = \tilde{I}_{i+nw_r}^n = 1 \right\}.$$

Note that $\tilde{I}_1^n, \dots, \tilde{I}_n^n$ are independent and that $\mathcal{G}_{1,i}^{\prime(r)} \in \sigma(\tilde{I}_1^n, \dots, \tilde{I}_n^n)$. Given $\tilde{I}_i^n = \tilde{I}_{i+nw_r}^n = 1$, $Q_{(i)}$ and $Q_{(i+nw_r)}$ are (conditionally) independent with joint probability density function

$$f_{Q_{(i)}, Q_{(i+nw_r)}}(x, y) = \frac{\lambda(x)\lambda(y)}{\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right)\left(\int_{w_r+\frac{i-1}{n}}^{w_r+\frac{i}{n}} \lambda(s)ds\right)}, \quad \frac{i-1}{n} < x < \frac{i}{n}, \quad w_r + \frac{i-1}{n} < y < w_r + \frac{i}{n}.$$

So,

$$\begin{aligned} P(\mathcal{G}_{1,i}^{(r)} \mid \mathcal{G}_{1,i}^{\prime(r)}) &= P(Q_{(i+nw_r)} - Q_{(i)} < w_r \mid \mathcal{G}_{1,i}^{\prime(r)}) \\ &= \frac{\int_{\frac{i-1}{n}}^{\frac{i}{n}} \int_{w_r+\frac{i-1}{n}}^{w_r+x} \lambda(x)\lambda(y)dydx}{\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right)\left(\int_{w_r+\frac{i-1}{n}}^{w_r+\frac{i}{n}} \lambda(s)ds\right)} \\ &= \frac{1}{2} + \mathcal{O}(n^{-1}), \end{aligned} \tag{4.6}$$

for $i \notin \mathcal{D}_{r,n} := \{1 \le j \le n - nw_r : (\frac{j-1}{n}, \frac{j}{n}) \cap \mathcal{D} \neq \emptyset \text{ or } (w_r + \frac{j-1}{n}, w_r + \frac{j}{n}) \cap \mathcal{D} \neq \emptyset\}$. The $\mathcal{O}(n^{-1})$ term in Eq. 4.6 is uniform in $i \in \{1, \dots, n - nw_r\} \backslash \mathcal{D}_{r,n}$. Combining (4.5) and (4.6) yields that

$$\begin{aligned} P(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_1) &= \sum_{r=1}^{r^*} \sum_{i=1}^{n-nw_r} P(\mathcal{G}_{1,i}^{(r)} \mid \mathcal{G}_{1,i}^{\prime(r)}) P(\mathcal{G}_{1,i}^{\prime(r)}) + \mathcal{O}(n^{-2}) \\ &= \frac{1}{2} \sum_{r=1}^{r^*} \sum_{i=1}^{n-nw_r} P(\mathcal{G}_{1,i}^{\prime(r)}) + \mathcal{O}(n^{-2}). \end{aligned} \tag{4.7}$$

For $r = 1, \ldots, r^*$ and $i = 1, \ldots, n - nw_r$, define

$$\mathcal{G}_{1,i}^{\prime\prime(r)} = \mathcal{A}_n^c \cap \left\{ \sum_{j=i+1}^{i+nw_r-1} I_j^n = k_r - 2,\, I_i^n = I_{i+nw_r}^n = 1 \right\},$$

so that $\mathcal{P}_{i,r}^{(n)} = P(\mathcal{G}_{1,i}^{\prime\prime(r)})$. Note that $\mathcal{G}_{1,i}^{\prime\prime(r)} \subset \mathcal{G}_{1,i}^{\prime\prime(r)}$ and $\mathcal{G}_{1,i}^{\prime\prime(r)} \backslash \mathcal{G}_{1,i}^{\prime(r)}$ is contained in $\{I_i^n = I_{i+nw_r}^n = 1, \tilde{I}_j^n \geq 2$ for some $j\}$, which has a probability of order n^{-3} uniformly in i. It follows from Eq. 4.7 that

$$\begin{aligned} P(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_1) &= \frac{1}{2} \sum_{r=1}^{r^*} \sum_{i=1}^{n-nw_r} P(\mathcal{G}_{1,i}^{\prime\prime(r)}) + \mathcal{O}(n^{-2}) \\ &= \frac{1}{2} \sum_{r=1}^{r^*} \sum_{i=1}^{n-nw_r} \mathcal{P}_{i,r}^{(n)} + \mathcal{O}(n^{-2}), \end{aligned}$$

proving (4.2).

It remains to prove (4.3). Let $\mathcal{H} := \{I_j^n = I_{j+nw_r}^n = 1$ for some (j, r) with $1 \leq j \leq n - nw_r$ and $1 \leq r \leq r^*\}$. On $\mathcal{G}_{2,i} \cap \mathcal{H}^c$, in order for $\mathcal{A} \cap \mathcal{A}_n^c$ to occur, there must exist some pair (i', r) with $1 \leq i' \leq i \leq i' + nw_r - 1 \leq n$ and $1 \leq r \leq r^*$ such that

$$\sum_{j=i'}^{i'+nw_r-1} I_j^n = k_r - 1 \quad \left(\text{implying that} \quad \sum_{j=i'}^{i'+nw_r-1} \tilde{I}_j^n = k_r \right).$$

It follows that $\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_{2,i} \cap \mathcal{H}^c \subset \mathcal{G}_{2,i}' \subset \mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_{2,i}$, where

$$\begin{aligned} \mathcal{G}_{2,i}' = \mathcal{A}_n^c \cap \Bigg\{ &\tilde{I}_i^n = 2, \tilde{I}_j^n \leq 1 \text{ for all } j \neq i, \\ &\sum_{j=i'}^{i'+nw_r-1} I_j^n = k_r - 1 \text{ for some } (i', r) \text{ with } 1 \leq i' \leq i \leq i' + nw_r - 1 \leq n \text{ and } 1 \leq r \leq r^* \Bigg\}. \end{aligned}$$

Since $P(\mathcal{G}_{2,i} \cap \mathcal{H}) = \mathcal{O}(n^{-3})$, we have $\sum_{i=1}^n P(\mathcal{G}_{2,i} \cap \mathcal{H}) = \mathcal{O}(n^{-2})$, implying that

$$\sum_{i=1}^n P(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_{2,i}) = \sum_{i=1}^n P\left(\mathcal{G}_{2,i}'\right) + \mathcal{O}(n^{-2}). \tag{4.8}$$

Letting

$$\begin{aligned} \mathcal{G}_{2,i}'' = \mathcal{A}_n^c \cap \Bigg\{ &\tilde{I}_i^n = 2, \sum_{j=i'}^{i'+nw_r-1} I_j^n = k_r - 1 \text{ for some } (i', r) \\ &\text{with } 1 \leq i' \leq i \leq i' + nw_r - 1 \leq n \text{ and } 1 \leq r \leq r^* \Bigg\}, \end{aligned}$$

we have $\tilde{\mathcal{P}}_i^{(n)} = P(\mathcal{G}''_{2,i})$. Note that $\mathcal{G}'_{2,i} \subset \mathcal{G}''_{2,i}$ and $\mathcal{G}''_{2,i} \backslash \mathcal{G}'_{2,i}$ is contained in $\{\tilde{I}_i^n = 2, \tilde{I}_j^n \geq 2$ for some $j \neq i\}$, which has a probability of order n^{-3} uniformly in i. Hence, $P\left(\mathcal{G}'_{2,i}\right) = P(\mathcal{G}''_{2,i}) + \mathcal{O}(n^{-3}) = \tilde{\mathcal{P}}_i^{(n)} + \mathcal{O}(n^{-3})$, which together with Eq. 4.8 implies that

$$\sum_{i=1}^n P(\mathcal{A} \cap \mathcal{A}_n^c \cap \mathcal{G}_{2,i}) = \sum_{i=1}^n \tilde{\mathcal{P}}_i^{(n)} + \mathcal{O}(n^{-2}),$$

proving (4.3). The proof is complete. □

Proof of Lemma 2.2 Let $L_1^n, L_2^n, \ldots, L_n^n$ be independent Bernoulli random variables such that $L_i^n, I_i^n, i = 1, \ldots, n$, are all independent and

$$P(L_i^n = 0) = \left(1 - \int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right) \exp\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right) := \ell_i$$

and $P(L_i^n = 1) = 1 - P(L_i^n = 0) = 1 - \ell_i$. (Note that all ℓ_i are between 0 and 1 for large n.) Set $\tilde{L}_i^n = \max\{I_i^n, L_i^n\}$, $i = 1, \ldots, n$, and let

$$\tilde{\mathcal{B}}_n = \bigcup_{r=1}^{r^*} \left\{ \max_{i=1,\ldots,n-nw_r+1} \sum_{j=i}^{i+nw_r-1} \tilde{L}_j^n \geq k_r \right\}.$$

Since

$$P(\tilde{L}_i^n = 0) = P(I_i^n = 0 \text{ and } L_i^n = 0) = 1 - \int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds = P(H_i^n = 0),$$

we have $\mathcal{L}(\tilde{L}_1^n, \tilde{L}_2^n, \ldots, \tilde{L}_n^n) = \mathcal{L}(H_1^n, H_2^n, \ldots, H_n^n)$ where $\mathcal{L}(Z)$ denotes the law of a random vector Z. So,

$$\beta_n = P(\mathcal{B}_n) = P(\tilde{\mathcal{B}}_n). \tag{4.9}$$

Let $S_0^n = 0$ and $S_j^n = \sum_{i=1}^j L_i^n$, $j = 1, \ldots, n$. Claim that

$$P(S_n^n = 0) = 1 - \frac{1}{2n}\int_0^1 \lambda^2(s)ds + \mathcal{O}(n^{-2}), \tag{4.10}$$

$$P(S_n^n = 1) = \frac{1}{2n}\int_0^1 \lambda^2(s)ds + \mathcal{O}(n^{-2}), \tag{4.11}$$

$$\text{and} \quad P(S_n^n \geq 2) = \mathcal{O}(n^{-2}). \tag{4.12}$$

To see this, note that

$$\begin{aligned}
\ell_i &= \left(1 - \int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right) \exp\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right) \\
&= \left(1 - \int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right) \left[1 + \int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds + \frac{1}{2}\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right)^2 + \mathcal{O}(n^{-3})\right] \\
&= 1 - \frac{1}{2}\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right)^2 + \mathcal{O}(n^{-3}),
\end{aligned}$$

where the $\mathcal{O}(n^{-3})$ term is uniform in i, implying that

$$\begin{aligned} P(S_n^n = 0) &= \prod_{i=1}^n \ell_i \\ &= \exp\left[\sum_{i=1}^n \log\left(1 - \frac{1}{2}\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right)^2 + \mathcal{O}(n^{-3})\right)\right] \\ &= \exp\left[\sum_{i=1}^n \left(-\frac{1}{2}\left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right)^2 + \mathcal{O}(n^{-3})\right)\right] \\ &= \exp\left[-\frac{1}{2}\sum_{i=1}^n \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right)^2 + \mathcal{O}(n^{-2})\right] \\ &= 1 - \frac{1}{2}\sum_{i=1}^n \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right)^2 + \mathcal{O}(n^{-2}). \end{aligned} \tag{4.13}$$

By condition (A),

$$\sum_{i=1}^n \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \lambda(s)ds\right)^2 = \frac{1}{n}\int_0^1 \lambda^2(s)ds + \mathcal{O}(n^{-2}). \tag{4.14}$$

Combining (4.13) and (4.14) yields (4.10).

Furthermore, we have $P(L_i^n = 1) = 1 - \ell_i = \mathcal{O}(n^{-2})$ uniformly in i, so that

$$P(S_n^n \geq 2) \leq \sum_{1 \leq i < j \leq n} P(L_i^n = L_j^n = 1) = \mathcal{O}(n^{-2}),$$

and

$$\begin{aligned} P(S_n^n = 1) &= 1 - P(S_n^n = 0) - P(S_n^n \geq 2) \\ &= 1 - \left(1 - \frac{1}{2n}\int_0^1 \lambda^2(s)ds + \mathcal{O}(n^{-2})\right) + \mathcal{O}(n^{-2}) \\ &= \frac{1}{2n}\int_0^1 \lambda^2(s)ds + \mathcal{O}(n^{-2}), \end{aligned}$$

establishing (4.11) and (4.12).

Noting that $\tilde{\mathcal{B}}_n \cap \{S_n^n = 0\} = \mathcal{A}_n \cap \{S_n^n = 0\}$, we have by Eqs. 4.10–4.12,

$$\begin{aligned} P(\tilde{\mathcal{B}}_n) &= P(\tilde{\mathcal{B}}_n \mid S_n^n = 0)P(S_n^n = 0) + P(\tilde{\mathcal{B}}_n \mid S_n^n = 1)P(S_n^n = 1) + P(\tilde{\mathcal{B}}_n \mid S_n^n \geq 2)P(S_n^n \geq 2) \\ &= P(\mathcal{A}_n \mid S_n^n = 0)P(S_n^n = 0) + \frac{1}{2n}\left(\int_0^1 \lambda^2(s)ds\right) P(\tilde{\mathcal{B}}_n \mid S_n^n = 1) + \mathcal{O}(n^{-2}) \\ &= \alpha_n\left(1 - \frac{1}{2n}\int_0^1 \lambda^2(s)ds\right) + \frac{1}{2n}\left(\int_0^1 \lambda^2(s)ds\right) P(\tilde{\mathcal{B}}_n \mid S_n^n = 1) + \mathcal{O}(n^{-2}), \end{aligned} \tag{4.15}$$

where the last equality follows from the fact $P(\mathcal{A}_n \mid S_n^n = 0) = P(\mathcal{A}_n) = \alpha_n$ (since the L_i^n's and I_i^n's are independent). By Eqs. 4.9, and 4.15 and the fact that $\lim_{n\to\infty} \alpha_n = \alpha$, to complete the proof, it remains to show

$$\lim_{n\to\infty} P(\tilde{\mathcal{B}}_n \mid S_n^n = 1) = \int_0^1 f(u) P\left(\mathcal{E}(u)\right) du. \tag{4.16}$$

To prove (4.16), let Q be a random point in (0, 1], which is independent of Π and has the probability density function f. For $i = 1, 2, \ldots, n-1$, let $q_{i,n}$ satisfy

$$P(Q \le q_{i,n}) = P(S_i^n = 1 \mid S_n^n = 1).$$

Clearly, $q_{0,n} = 0 < q_{1,n} < \cdots < q_{n-1,n} < 1 = q_{n,n}$, and

$$\begin{aligned} P(q_{i-1,n} < Q \le q_{i,n}) &= P(S_i^n = 1 \mid S_n^n = 1) - P(S_{i-1}^n = 1 \mid S_n^n = 1) \\ &= P(L_i^n = 1 \mid S_n^n = 1). \end{aligned}$$

We claim that

$$\max_{1\le i\le n} \left| q_{i,n} - \frac{i}{n} \right| = \max_{1\le i\le n-1} \left| q_{i,n} - \frac{i}{n} \right| \to 0 \ \text{ as } \ n \to \infty. \tag{4.17}$$

To establish the claim, note that Eqs. 4.10–4.12 can be readily extended to

$$\begin{aligned} P(S_{\lfloor nv\rfloor}^n - S_{\lfloor nu\rfloor}^n = 0) &= 1 - \frac{1}{2n}\int_u^v \lambda^2(s)ds + \mathcal{O}(n^{-2}), \\ P(S_{\lfloor nv\rfloor}^n - S_{\lfloor nu\rfloor}^n = 1) &= \frac{1}{2n}\int_u^v \lambda^2(s)ds + \mathcal{O}(n^{-2}), \\ P(S_{\lfloor nv\rfloor}^n - S_{\lfloor nu\rfloor}^n \ge 2) &= \mathcal{O}(n^{-2}), \end{aligned}$$

where $0 \le u < v \le 1$ and $\lfloor x \rfloor$ denotes the largest integer not exceeding x. As a consequence,

$$\lim_{n\to\infty} P(S_{\lfloor nv\rfloor}^n = 1 \mid S_n^n = 1) = \frac{\int_0^v \lambda^2(s)ds}{\int_0^1 \lambda^2(s)ds} = P(Q \le v)$$

for $0 < v \le 1$. For a fixed (large) N,

$$\lim_{n\to\infty} P(S_{\lfloor nj/N\rfloor}^n = 1 \mid S_n^n = 1) = P(Q \le j/N) \ \text{ for } \ j = 1, \ldots, N-1. \tag{4.18}$$

By Eq. 4.18, there exists an $n_0 > 0$ such that

$$P(Q \le \max\{(j-1)/N, 0\}) \le P(S_{\lfloor nj/N\rfloor}^n = 1 \mid S_n^n = 1) \le P(Q \le \min\{(j+1)/N, 1\}) \tag{4.19}$$

for $j = 0, 1, \ldots, N$ and $n \ge n_0$. For $n \ge n_0$, for each $i = 1, 2, \ldots, n-1$, there is a $j \in \{1, 2, \ldots, N\}$ such that $\frac{j-1}{N} < \frac{i}{n} \le \frac{j}{N}$, so that by Eq. 4.19

$$\begin{aligned} P(Q \le q_{i,n}) &= \mathbb{P}(S_i^n = 1 \mid S_n^n = 1) \\ &\le P(S_{\lfloor nj/N\rfloor}^n = 1 \mid S_n^n = 1) \\ &\le P(Q \le \min\{(j+1)/N, 1\}) \end{aligned} \tag{4.20}$$

and

$$\begin{aligned} P(Q \le q_{i,n}) &= \mathbb{P}(S_i^n = 1 \mid S_n^n = 1) \\ &\ge P(S_{\lfloor n(j-1)/N\rfloor}^n = 1 \mid S_n^n = 1) \\ &\ge P(Q \le \max\{(j-2)/N, 0\}). \end{aligned} \tag{4.21}$$

By Eqs. 4.20 and 4.21, we have

$$\max\left\{\frac{j-2}{N},0\right\} \le q_{i,n} \le \min\left\{\frac{j+1}{N},1\right\},$$

which together with $\frac{j-1}{N} < \frac{i}{n} \le \frac{j}{N}$ implies that $\left|q_{i,n} - \frac{i}{n}\right| \le \frac{2}{N}$. So,

$$\max_{1\le i\le n-1}\left|q_{i,n} - \frac{i}{n}\right| \le \frac{2}{N} \quad \text{for } \ n \ge n_0.$$

Since N can be chosen arbitrarily large, Eq. 4.17 follows.

Let $Q_n = i/n$ if $q_{i-1,n} < Q \le q_{i,n}$. Then $Q_n \to Q$ a.s. as $n \to \infty$. Let $I_i'^n = \mathbf{1}\{(\Pi \cup \{Q_n\}) \cap (\frac{i-1}{n}, \frac{i}{n}] \ne \emptyset\}$, $i = 1, 2, \ldots, n$. Since Q_n is independent of Π and since $P(Q_n = i/n) = P(q_{i-1,n} < Q \le q_{i,n}) = P(L_i^n = 1 \mid S_n^n = 1)$, it follows that

$$\mathcal{L}(I_1'^n, \ldots, I_n'^n) = \mathcal{L}(\tilde{L}_1^n, \ldots, \tilde{L}_n^n \mid S_n^n = 1).$$

Letting

$$\tilde{\mathcal{B}}_n' = \bigcup_{r=1}^{r^*}\left\{\max_{i=1,\ldots,n-nw_r+1}\sum_{j=i}^{i+nw_r-1} I_j'^n \ge k_r\right\},$$

we have

$$P(\tilde{\mathcal{B}}_n') = P(\tilde{\mathcal{B}}_n \mid S_n^n = 1).$$

Since $\mathbf{1}_{\tilde{\mathcal{B}}_n'}$ converges a.s. to $\mathbf{1}\{\max_{0\le t\le 1-w_r} |(\Pi \cup \{Q\}) \cap (t, t+w_r]| \ge k_r$ for some $1 \le r \le r^*\}$,

$$\begin{aligned}\lim_{n\to\infty} P(\tilde{\mathcal{B}}_n \mid S_n^n = 1) &= \lim_{n\to\infty} P(\tilde{\mathcal{B}}_n') \\ &= P\left(\bigcup_{r=1}^{r^*}\left\{\max_{0\le t\le 1-w_r} |(\Pi \cup \{Q\}) \cap (t, t+w_r]| \ge k_r\right\}\right) \\ &= \int_0^1 f(u)P(\mathcal{E}(u))du,\end{aligned}$$

proving (4.16). The proof is complete. □

Acknowledgements The authors gratefully acknowledge support from the Ministry of Science and Technology of Taiwan, R.O.C.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

References

Chan HP, Zhang NR (2007) Scan statistics with weighted observations. J Amer Statist Assoc 102:595–602

Cressie N (1977) On some properties of the scan statistic on the circle and the line. J Appl Probab 14:272–283

Fang X, Siegmund D (2016) Poisson approximation for two scan statistics with rates of convergence. Ann Appl Probab 26:2384–2418

Fu JC (2001) Distribution of the scan statistic for a sequence of bistate trials. J Appl Probab 38:908–916

Fu JC, Koutras MV (1994) Distribution theory of runs: a Markov chain approach. J Amer Statist Assoc 89:1050–1058

Fu JC, Wu TL, Lou WYW (2012) Continuous, discrete, and conditional scan statistics. J Appl Probab 49:199–209

Glaz J, Naus JI (1991) Tight bounds and approximations for scan statistic probabilities for discrete data. Ann Appl Probab 1:306–318

Glaz J, Naus JI (2010) Scan statistics. In: Balakrishnan N (ed) Methods and applications of statistics in the life and health sciences. Wiley, New Jersey, pp 733-747

Glaz J, Naus J, Wallenstein S (2001) Scan statistics. Springer, New York

Glaz J, Zhang Z (2004) Multiple window discrete scan statistics. J Appl Statist 31:967–980

Huntington RJ, Naus JI (1975) A simpler expression for Kth nearest neighbor coincidence probabilities. Ann Probab 3:894–896

Hwang FK (1977) A generalization of the Karlin-McGregor theorem on coincidence probabilities and an application to clustering. Ann Probab 5:814–817

Janson S (1984) Bounds on the distributions of extremal values of a scanning process. Stoch Process Appl 18:313–328

Koutras MV, Alexandrou VA (1995) Runs, scans, and urn model distributions: A unified Markov chain approach. Ann Inst Stat Math 47:743–776

Loader C (1991) Large deviation approximations to the distribution of scan statistics. Adv Appl Probab 23:751–771

Nagarwalla N (1996) A scan statistic with a variable window. Statist Med 15:845–850

Naus JI (1982) Approximations for distributions of scan statistics. J Amer Statist Assoc 77:177–183

Naus JI, Wallenstein S (2004) Multiple window and cluster size scan procedures. Methodology Comput Appl Probab 6:380–400

Neff ND, Naus JI (1980) Selected tables in mathematical statistics, vol VI. American Mathematical Society, Providence

Siegmund D, Yakir B (2000) Tail probabilities for the null distribution of scanning statistics. Bernoulli 6:191–213

Wu T-L (2017) Approximations of distributions of scan statistics of inhomogeneous Poisson processes. J Statist Plann Inference 190:52–59

Wu T-L, Glaz J, Fu JC (2013) Discrete, continuous and conditional multiple window scan statistics. J Appl Probab 50:1089–1101

Yao Y-C, Miao DW-C, Lin XC-S (2017) Corrected discrete approximations for the conditional and unconditional distributions of the continuous scan statistic. J Appl Probab 54:304–319