The Annals of Statistics
1997, Vol 25, No. 1, 263-292

KAPLAN-MEIER ESTIMATORS OF DISTANCE DISTRIBUTIONS FOR SPATIAL POINT PROCESSES

BY ADRIAN BADDELEY[1] AND RICHARD D. GILL

University of Western Australia and University of Leiden, and University of Utrecht

When a spatial point process is observed through a bounded window, edge effects hamper the estimation of characteristics such as the empty space function F, the nearest neighbor distance distribution G and the reduced second-order moment function K. Here we propose and study product-limit type estimators of F, G and K based on the analogy with censored survival data: the distance from a fixed point to the nearest point of the process is right-censored by its distance to the boundary of the window. The resulting estimators have a ratio-unbiasedness property that is standard in spatial statistics. We show that the empty space function F of *any* stationary point process is absolutely continuous, and so is the product-limit estimator of F. The estimators are strongly consistent when there are independent replications or when the sampling window becomes large. We sketch a CLT for independent replications within a fixed observation window and asymptotic theory for independent replications of sparse Poisson processes. In simulations the new estimators are generally more efficient than the "border method" estimator but (for estimators of K), somewhat less efficient than sophisticated edge corrections.

1. Introduction. The exploratory data analysis of observations of a spatial point process often starts with the estimation of certain distance distributions: $F(t)$, the distribution of the distance from an arbitrary point in space to the nearest point of the process; $G(t)$, the distribution of the distance from a typical point of the process to the nearest other point of the process; $K(t)$, the expected number of other points within distance t of a typical point of the process, divided by the intensity. For a homogeneous Poisson process F, G and K take known functional forms, and deviations of estimates of F, G, K from these forms are taken as indications of clustered or inhibited alternatives [11, 37, 38].

However, the estimation of F, G and K is hampered by edge effects arising because the point process is observed within a bounded window W.

Received July 1993; revised December 1995.

[1] Supported by CWI, Amsterdam.

AMS 1991 *subject classifications.* 62G05, 62H11, 60D05.

Key words and phrases. Border correction method, dilation, distance transform, edge corrections, edge effects, empty space statistic, erosion, functional delta-method, influence function, K-function, local knowledge principle, nearest-neighbor distance, product integration, reduced sample estimator, reduced second moment measure, sparse Poisson asymptotics, spatial statistics, survival data.

Essentially the distance from a given reference point to the nearest point of the process is *censored* by its distance to the boundary of W. Edge effects become rapidly more severe as the dimension of space increases, or as the distance t increases.

Traditionally in spatial statistics, one uses edge-corrected estimators which are weighted empirical distributions of the observed distances. The simplest approach is the "border method" [38] where we restrict attention (when estimating F, G or K at distance t) to those reference points lying more than t units away from the boundary of W. These are the points x for which distances up to t are observed without censoring. This approach is sometimes also justified by appealing to the "local knowledge principle" of mathematical morphology ([42], pages 49, 233). However, the border method discards much of the data; in three dimensions [5] it seems to be unacceptably wasteful, especially when estimating G.

In more sophisticated edge corrections, the weight $c(x, y)$ attached to the observed distance $\|x - y\|$ between two points x, y is the reciprocal of the probability that the distance will be observed under invariance assumptions (stationarity under translation and/or rotation). Corrections of this type were first suggested by Miles [34] and developed by Ripley, Lantuéjoul, Hanisch, Stoyan, Ohser and others ([11, 24, 35–37], [42], page 246). For surveys see [38], Chapter 3, [47], pages 122–131), ([9], Chapter 8) and [4].

The estimation problem for F, G and K from data in a bounded window W has a clear analogy, already implicitly drawn above, to the estimation of a survival function based on a sample of randomly censored survival times. This paper develops the analogy and proposes Kaplan–Meier [29] or product-limit estimators for F, G and K. Since the observed, censored distances are highly interdependent, classical theory from survival analysis has little to say about statistical properties of the new estimators. One may hope that the new estimators are better than the classical edge corrections, as in the survival analysis situation the Kaplan–Meier estimator has various large-sample optimality properties. In fact the border method for edge correction, described above, is analogous to the so-called reduced sample estimator, an inefficient competitor to the Kaplan–Meier estimator obtained using only those observations for which the censoring time is at least t when estimating the probability of survival to time t.

Surprisingly, the analogy between edge effects for point processes and random censoring of survival times has not been much explored. Laslett [30, 31] noted that when a spatial line segment process is clipped within a bounded window, the observed line segment lengths can be compared to censored survival times. However, a Kaplan–Meier type estimator for the segment length distribution is inconsistent and the NPMLE is a different, difficult estimator [50]. Zimmerman [51] proposed introducing artificial censoring in spatial sampling by restricting the maximum search distance from any reference point.

The estimation of F by a Kaplan–Meier type estimator poses a new (for survival analysis) problem, since one has a continuum of observations: for each point in the sampling window, a censored distance to the nearest point of the process. We tackle this using product integration [22, 23].

Together with estimators of F, G and K, one would like to evaluate their accuracy. We make a start on this by using linearization techniques (the functional delta-method; see [21]) and evaluate the asymptotic efficiency explicitly in a simple sparse Poisson limiting situation. This also leads to proposals for variance estimators.

The plan of the paper is as follows: Section 2 recalls some definitions from spatial statistics and from the analysis of survival data; Section 3 introduces our Kaplan–Meier style estimator of the empty space function F; Section 4 discusses asymptotic properties of this estimator; Sections 5 and 6 treat the estimation of G and K, respectively, in less detail. Critical comments are collected in Section 7.

2. Preliminaries.

2.1. *Spatial statistics.* Let Φ be a simple point process in $\mathbb{R}^k$, observed through a compact window $W \subset \mathbb{R}^k$. We consider Φ both as a random set in $\mathbb{R}^k$ and as a random measure. The problem is, based on the data $\Phi \cap W$ (and knowledge of W itself) to estimate the functions F, G and K defined as follows.

For $x \in \mathbb{R}^k$ and any closed $A \subset \mathbb{R}^k$, let

$$\rho(x, A) = \inf\{\|x - a\|_2 : a \in A\} \tag{1}$$

be the shortest Euclidean distance from x to A, and

$$\begin{aligned} A_{\oplus r} &= \{x \in \mathbb{R}^k : \rho(x, A) \le r\} \\ A_{\ominus r} &= \{x \in A : \rho(x, A^c) > r\}, \end{aligned}$$

where the superscript c denotes complement. Write $B(x, r)$ for the closed ball of radius r, center x in $\mathbb{R}^k$.

Assume now that Φ is a.s. stationary under translations, with intensity $0 < \alpha < \infty$. Thus $\mathbb{E}\Phi(A) = \alpha |A|_k$ for any bounded Borel $A \subset \mathbb{R}^k$, where $|\cdot|_k$ denotes k-dimensional Lebesgue volume. For $r \ge 0$, define

$$\begin{aligned} F(r) &= \mathbb{P}\{\rho(0, \Phi) \le r\} \\ &= \mathbb{P}\{\Phi(B(0, r)) > 0\}, \end{aligned} \tag{2}$$

$$\begin{aligned} G(r) &= \mathbb{P}\{\rho(0, \Phi \setminus \{0\}) \le r \mid 0 \in \Phi\} \\ &= \mathbb{P}\{\Phi(B(0, r)) > 1 \mid 0 \in \Phi\}, \end{aligned} \tag{3}$$

$$K(r) = \alpha^{-1}\mathbb{E}\{\Phi(B(0, r) \setminus \{0\}) \mid 0 \in \Phi\}. \tag{4}$$

The conditional expectations given $0 \in \Phi$ used above are expectations with respect to the Palm distribution of Φ at 0. By stationarity, the point 0 here may be replaced by any arbitrary point x. Using the Campbell–Mecke formula [47]

$$G(r) = \frac{\mathbb{E}\sum_{x \in \Phi \cap A} \mathbf{1}\{\rho(x, \Phi \setminus \{x\}) \le r\}}{\mathbb{E}\Phi(A)} \tag{5}$$

and

$$\alpha K(r) = \frac{\mathbb{E}\sum_{x \in \Phi \cap A} \Phi(B(x, r) \setminus \{x\})}{\mathbb{E}\Phi(A)} \tag{6}$$

for arbitrary Borel A with $0 < |A|_k < \infty$. The latter definition of K is the original one and it applies to any second-order stationary process [47].

Edge-corrected estimators for F, G and K based on observation of Φ in W are reviewed in [38], Chapter 3; [47], pages 122–131; and [9], Chapter 8. See [5, 6, 13–18, 20 and 43].

Many estimators in spatial statistics are not unbiased, but instead are ratios of two unbiased consistent estimators

$$\hat{\theta} = \frac{Y}{X} \quad \text{where } \theta = \frac{\mathbb{E}Y}{\mathbb{E}X}$$

with $X, Y \ge 0$, $\mathbb{P}\{X > 0\} > 0$, $X = 0 \Rightarrow Y = 0$ typically arising as the mean of a weighted empirical distribution where the weights are random variables [5, 38]. We call such estimators "ratio-unbiased" and accept this property as a substitute for the generally unobtainable unbiasedness.

2.2. *Survival data.* Next we recall some theory of the Kaplan–Meier and reduced sample estimators. Suppose $T_1, \dots, T_n$ are i.i.d. positive r.v.'s with distribution function F and survival function $S = 1 - F$. Let $C_1, \dots, C_n$ be independent of the T_i's and i.i.d. with d.f. H. Let $\tilde{T}_i = T_i \wedge C_i$, $D_i = \mathbf{1}\{T_i \le C_i\}$, where $a \wedge b$ denotes $\min\{a, b\}$. Then $(\tilde{T}_1, D_1), \dots, (\tilde{T}_n, D_n)$ is a sample of censored survival times $\tilde{T}_i$ with censoring indicators D_i. The *reduced-sample estimator* of F is

$$\hat{F}^{\mathrm{rs}}(t) = \frac{\#\{i: \tilde{T}_i \le t \le C_i\}}{\#\{i: C_i \ge t\}}. \tag{7}$$

This requires that we can observe the censoring times C_i themselves, or at least the event $\{C_i \ge t\}$ for all t for which $F(t)$ must be estimated. This estimator is clearly pointwise unbiased for F and has values in $[0, 1]$ but may not be a monotone function of t.

The *Kaplan–Meier estimator* of F is

$$\hat{F}(t) = 1 - \prod_{s \le t}\left(1 - \frac{\#\{i: \tilde{T}_i = s, D_i = 1\}}{\#\{i: \tilde{T}_i \ge s\}}\right). \tag{8}$$

Introduce

$$N_n(t) = \frac{1}{n}\#\{i: \tilde{T}_i \le t,\, D_i = 1\}, \tag{9}$$

$$Y_n(t) = \frac{1}{n}\#\{i: \tilde{T}_i \ge t\}, \tag{10}$$

$$\hat{\Lambda}_n(t) = \int_0^t \frac{dN_n(s)}{Y_n(s)} \tag{11}$$

$$\Lambda(t) = \int_0^t \frac{dF(s)}{1 - F(s-)}. \tag{12}$$

Then Λ is the cumulative hazard belonging to F, and $\hat{\Lambda}_n$ is the Nelson–Aalen estimator of it. One can write

$$\begin{aligned} 1 - F(t) &= \underset{0}{\overset{t}{\pi}}(1 - d\Lambda(s)), \\ 1 - \hat{F}_n(t) &= \underset{0}{\overset{t}{\pi}}\big(1 - d\hat{\Lambda}_n(s)\big), \end{aligned} \tag{13}$$

where π denotes *product integration*:

$$\underset{0}{\overset{t}{\pi}}(1 + dA(s)) = \lim_{\max|t_i - t_{i-1}| \to 0} \prod_{i=1}^{m} (1 + A(t_i) - A(t_{i-1})),$$

where $0 = t_0 < \cdots < t_m = t$ forms a partition of $(0, t]$.

If F is absolutely continuous with density f then defining $\lambda(t) = f(t)/(1 - F(t))$, the hazard rate, one has $\Lambda(t) = \int_0^t \lambda(s)\,ds$ and

$$1 - F(t) = \underset{0}{\overset{t}{\pi}}(1 - d\Lambda(s)) = \exp(-\Lambda(t)).$$

However if F has a discrete component, the relation $\Lambda = -\log(1 - F)$ no longer holds. See [22] and [23] for further information on the product integral, including empirical process theory.

3. Kaplan–Meier estimator of the empty space function.

3.1. *Definition of estimator.* Return to the setup of Section 2.1. Every point x in the window W contributes one possibly censored observation of the distance from an arbitrary point in space to the point process Φ. The analogy with survival times is to regard $\rho(x, \Phi)$ as the distance (time) to failure and $\rho(x, \partial W)$ as the censoring distance, where ∂W denotes the boundary of W. The observation is censored if $\rho(x, \partial W) < \rho(x, \Phi)$.

From the data $\Phi \cap W$ we can compute $\rho(x, \Phi \cap W)$ and $\rho(x, \partial W)$ for each $x \in W$. Note that

$$\rho(x, \Phi) \wedge \rho(x, \partial W) = \rho(x, \Phi \cap W) \wedge \rho(x, \partial W) \tag{14}$$

(cf. [42], pages 49 and 233) so that we can indeed observe $\rho(x,\Phi) \wedge \rho(x,\partial W)$ and $\mathbf{1}\{\rho(x,\Phi) \le \rho(x,\partial W)\}$ for each $x \in W$. Then the set

$$\{x \in W: \rho(x,\Phi) \wedge \rho(x,\partial W) \ge r\}$$

can be thought of as the set of points "at risk of failure at distance r," and

$$\{x \in W: \rho(x,\Phi) = r,\ \rho(x,\Phi) \le \rho(x,\partial W)\}$$

are the "observed failures at distance r."

Geometrically the two sets are the closures of $W_{\ominus r} \setminus \Phi_{\oplus r}$ and $\partial(\Phi_{\oplus r}) \cap W_{\ominus r}$, respectively. See Figure 1.

DEFINITION 1. Let Φ be an a.s. stationary point process and $W \subset \mathbb{R}^k$ a fixed compact set. The Kaplan–Meier estimator $\hat{F}$ of the empty space function F of Φ, based on data $\Phi \cap W$, is defined by:

$$\hat{\Lambda}(r) = \int_0^r \frac{|\partial(\Phi_{\oplus s}) \cap W_{\ominus s}|_{k-1}}{|W_{\ominus s} \setminus \Phi_{\oplus s}|_k}\, ds \tag{15}$$

$$\hat{F}(r) = 1 - \prod_0^r \left(1 - d\hat{\Lambda}(s)\right) \tag{16}$$

$$= 1 - \exp\left(-\hat{\Lambda}(r)\right), \tag{17}$$

where $|\cdot|_{k-1}$ denotes $k-1$ dimensional Hausdorff measure (surface area). The reduced-sample estimator $\hat{F}^{\mathrm{rs}}$ of F is

$$\hat{F}^{\mathrm{rs}}(r) = \frac{|W_{\ominus r} \cap \Phi_{\oplus r}|_k}{|W_{\ominus r}|_k}, \tag{18}$$

that is, this is the border correction [38].

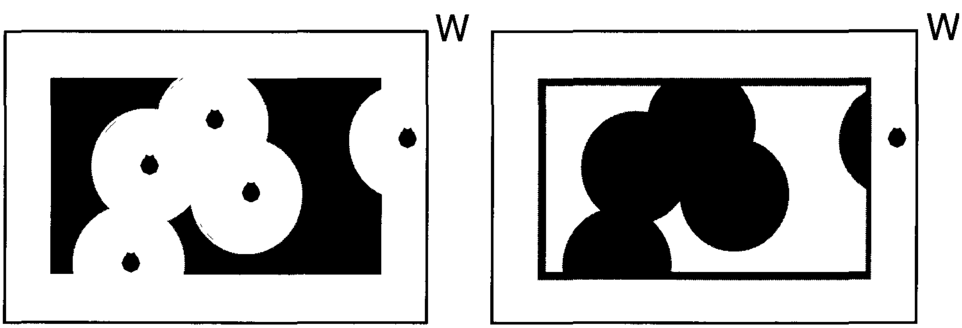

FIG. 1. *Geometry of the reduced sample (left) and Kaplan–Meier (right) estimators. Spatial process Φ indicated by filled dots. For Kaplan–Meier, points x at risk are shaded, and observed failures constitute the curved boundary of the shaded region.*

Here $\hat{F}$ is the Kaplan–Meier estimator based on the continuum of observations generated by all $x \in W$. Note that the estimator is a proper distribution function and is even absolutely continuous, with hazard rate

$$\hat{\lambda}(r) = \frac{|\partial(\Phi_{\oplus r}) \cap W_{\ominus r}|_{k-1}}{|W_{\ominus r} \setminus \Phi_{\oplus r}|_k}. \tag{19}$$

3.2. *Unbiasedness and continuity.*

THEOREM 1. *Let Φ be any stationary point process with intensity $0 < \alpha < \infty$. Then the following statements hold*:

(a) *The empty space function F is absolutely continuous.*
(b) *The hazard rate of F equals*

$$\lambda(r) = \frac{\mathbb{E}|W \cap \partial(\Phi_{\oplus r})|_{k-1}}{\mathbb{E}|W \setminus \Phi_{\oplus r}|_k}$$

for any compact window W such that the denominator is positive.
(c) *The Kaplan–Meier estimator* (19) *of λ is ratio-unbiased.*

Thus our estimator $\hat{F}(r)$ respects the smoothness of the true empty space function F. The reduced-sample estimator (18) is not even necessarily monotone.

To prove the theorem we need three regularity results. The first is an example of Crofton's perturbation or moving manifold formula [2, 10].

LEMMA 1. *Let $Z \subset \mathbb{R}^k$ be compact and $A \subset \mathbb{R}^k$ a finite union of compact convex sets. Then for $r \geq 0$,*

$$|Z \cap A_{\oplus r}|_k = |Z \cap A|_k + \int_0^r |Z \cap \partial(A_{\oplus s})|_{k-1}\, ds;$$

the integrand is Lebesgue measurable and integrable.

PROOF. The function $f(x) = \rho(x, A)$ is Lipschitz, $f(y) \leq f(x) + \|x - y\|$, and hence a.e. differentiable with approximate Jacobian ap $J_1 f \leq 1$. A geometric argument shows that up to a null set ap $J_1 f \equiv 1$ and $\{x\colon f(x) = s, J_1 f(x) > 0\} = \partial(A_{\oplus s})$. Apply the co-area formula ([19], page 258) to integration of 1_Z over $A_{\oplus r}$. □

The next lemma shows that the integrand $|Z \cap \partial(\Phi_{\oplus r})|_{k-1}$ is uniformly bounded (over possible realizations of Φ) in such a way that dominated convergence justifies interchanges of expectation and integration or differentiation with respect to r.

LEMMA 2. *For any compact set* Z,

$$|Z \cap \partial(\Phi_{\oplus r})|_{k-1} \leq \frac{k}{r}|Z_{\oplus r}|_k \wedge \omega_k r^{k-1}\Phi(Z_{\oplus r}) \quad a.s.,$$

where $\omega_k = |\partial B(0,1)|_{k-1} = 2\pi^{k/2}/\Gamma(k/2)$.

PROOF. The second term on the right is a trivial bound since $\omega_k r^{k-1} = |\partial B(0,r)|_{k-1}$. For the first term, fix a realization of Φ and let x_i, $i = 1,\dots,m$, be the a.s. distinct points in $\Phi \cap Z_{\oplus r}$. Then

$$Z \cap \partial(\Phi_{\oplus r}) = Z \cap \partial\left(\bigcup_{i=1}^{m} B(x_i, r)\right).$$

Construct the Dirichlet cells formed by the x_i,

$$C_i = C(x_i \mid x_1,\dots,x_m) = \left\{y \in \mathbb{R}^k : \|y - x_i\| = \min_j \|y - x_j\|\right\}.$$

Split the surface $\partial(\Phi_{\oplus r})$ into pieces of surface within each cell:

$$\begin{aligned}\partial\left(\bigcup_{i=1}^{m} B(x_i, r)\right) &= \bigcup_{i=1}^{m}\left\{y \in \mathbb{R}^k : \|y - x_i\| = r,\ \min_j \|y - x_j\| = r\right\}\\ &= \bigcup_{i=1}^{m}(C_i \cap \partial B(x_i, r))\\ &= \bigcup_{i=1}^{m} D_i, \quad \text{(say)}.\end{aligned}$$

The D_i are measure-disjoint since $D_i \cap D_j = \partial B(x_i, r) \cap \partial B(x_j, r)$ is $k-2$ dimensional (or empty) for $i \neq j$. Thus

$$\left|\partial\left(\bigcup_{i=1}^{m} B(x_i, r)\right) \cap Z\right|_{k-1} = \sum_{i=1}^{m} |D_i \cap Z|_{k-1}. \tag{20}$$

Any line segment joining x_i to a point on the corresponding surface piece $D_i \cap Z$ is contained entirely within the Dirichlet cell C_i, since this is convex. The union F_i of these segments is a solid angular cone of the sphere $B(x_i, r)$, and its curved surface area $|D_i \cap Z|_{k-1}$ equals k/r times its volume. The cones F_i are volume-disjoint since the C_i are, and $F_i \subseteq Z_{\oplus r}$, so the sum of the cone volumes is bounded by $|Z_{\oplus r}|_k$, yielding the result. □

LEMMA 3. *Let* Φ *be a simple point process in* $\mathbb{R}^k$ *and* $W \subset \mathbb{R}^k$ *compact. Then for fixed* r, $|W \cap \Phi_{\oplus r}|_k$ *and* $|W_{\ominus r} \cap \Phi_{\oplus r}|_k$ *are a.s. finite r.v.'s on the same probability space, and the following identities hold a.s.*:

$$|W \cap \Phi_{\oplus r}|_k = \int_0^r |W \cap \partial(\Phi_{\oplus s})|_{k-1}\, ds, \tag{21}$$

$$|\{x \in W : \rho(x, \Phi) \leq \rho(x, \partial W) \wedge r\}|_k = \int_0^r |W_{\ominus s} \cap \partial(\Phi_{\oplus s})|_{k-1}\, ds, \tag{22}$$

$$|W_{\ominus r} \setminus \Phi_{\oplus r}|_k = |W|_k - \int_0^r |\partial(W_{\ominus s} \setminus \Phi_{\oplus s})|_{k-1}\, ds, \tag{23}$$

where the integrands are well defined r.v.'s for each fixed s and are a.s. measurable and integrable functions of s.

PROOF. By [33], pages 9, 19, 47, $\Phi_{\oplus r}$ is a random closed set, so that $\partial(\Phi_{\oplus r})$ is a random closed set, the intersections with W are random compact sets and their measures are r.v.'s. Now apply Lemma 1 to each realization to get (21).

For (22), we note that $\rho(x, \partial W)$ is continuous in x and $\rho(x, \Phi)$ is a random upper-semicontinuous (u.s.c.) function, so that $\rho(x, \Phi) - \rho(x, \partial W)$ is also a random u.s.c. function and $Z = \{x \in W\colon\ \rho(x, \Phi) \le \rho(x, \partial W)\}$ is a random closed set. Recognize the left-hand side of (22) as the volume of $Z \cap \Phi_{\oplus r}$ and the integrand as the surface area of $Z \cap \partial(\Phi_{\oplus s})$. Measurability arguments remain valid for the random closed set Z and we apply Lemma 1 to each realization.

For (23), observe that $W_{\ominus r} \setminus \Phi_{\oplus r} = W \setminus (\partial W \cup \Phi)_{\oplus r}$ and use the same technique as for (21). □

PROOF OF THEOREM 1. By Fubini (Robbins' theorem [33], page 47),

$$\begin{aligned} \mathbb{E}|W \cap \Phi_{\oplus r}|_k &= \mathbb{E}\int_W \mathbf{1}\{x \in \Phi_{\oplus r}\}\,dx \\ &= \int_W \mathbb{P}\{x \in \Phi_{\oplus r}\}\,dx \\ &= F(r)|W|_k. \end{aligned} \tag{24}$$

Since $r \mapsto |W \cap \Phi_{\oplus r}|_k$ is absolutely continuous with derivative given in Lemma 3 and bounded as in Lemma 2, its expectation is absolutely continuous too, with derivative

$$f(r)|W|_k = \mathbb{E}\big|W \cap \partial(\Phi_{\oplus r})\big|_{k-1}, \tag{25}$$

but complementarily to (24),

$$\mathbb{E}|W \setminus \Phi_{\oplus r}|_k = (1 - F(r))|W|_k. \tag{26}$$

Dividing (25) by (26) we obtain the first result of the theorem. The rest follows by replacing W with $W_{\ominus r}$. □

3.3. *Discretization and classical Kaplan–Meier estimator.* In practice one would not actually compute the surface areas and volumes for each $s \in [0, r]$ in order to estimate $F(r)$. Rather one would discretize W or $[0, r]$ or both. For standard estimators of F, one typically discretizes W on a regular lattice (see [12]), although Lotwick [32] showed the areas can be computed analytically.

A natural possibility here is to discretize W by superimposing a regular lattice L of points, calculating for each $x_i \in W \cap L$ the censored distance $\rho(x_i, \Phi) \wedge (x_i, \partial W)$ and the indicator $\mathbf{1}\{\rho(x_i, \Phi) \le \rho(x_i, \partial W)\}$. Then one would calculate the ordinary Kaplan–Meier estimator (8) based on this finite data set.

Our next result is that as the lattice becomes finer, the discrete Kaplan–Meier estimates converge to the theoretical continuous estimator $\hat{F}$.

THEOREM 2. *Let $\hat{F}_L$ be the Kaplan–Meier estimator* (8) *computed from the discrete observations at the points of $W \cap L$, where $L = \varepsilon M + b$ is a rescaled, translated copy of a fixed regular lattice M. Let*

$$R = \inf\{r \geq 0: W_{\ominus r} \setminus \Phi_{\oplus r} = \varnothing\}.$$

Then as the lattice mesh ε converges to zero, $\hat{F}_L(r) \to \hat{F}(r)$ for any $r < R$. The convergence is uniform on any compact interval in $[0, R)$. *Similarly the continuous reduced-sample estimator* (18) *is the uniform limit of the discrete reduced sample estimator* (7).

PROOF. For any compact set $A \subseteq \mathbb{R}^k$ with $|\partial A|_k = 0$, one can easily show that

$$\varepsilon^d \#(A \cap L) \to c|A|_k \quad \text{as } \varepsilon \to 0,$$

where c is a finite positive constant depending on M. The sets $W_{\ominus r}$, $\Phi_{\oplus r}$ and

$$V = \{x \in W: \rho(x, \Phi) \leq \rho(x, \partial W)\}$$

clearly have these properties for $r < R$. Hence the functions

$$N_L(r) = \frac{\#(L \cap \{x \in W: \rho(x, \Phi) \leq \rho(x, \partial W) \wedge r\})}{\#(L \cap W)}$$

and

$$Y_L(r) = \frac{\#(L \cap (W_{\ominus r} \setminus \Phi_{\oplus r}))}{\#(L \cap W)}$$

converge pointwise to

$$N(r) = \frac{|\{x \in W: \rho(x, \Phi) \leq \rho(x, \partial W) \wedge r\}|_k}{|W|_k} \tag{27}$$

and

$$Y(r) = \frac{|W_{\ominus r} \setminus \Phi_{\oplus r}|_k}{|W|_k}, \tag{28}$$

respectively. Since $N_L(r)$ is increasing in r and the limit is continuous, $N_L \to N$ uniformly in r. Recalling (23) and using the argument of Lemma 2 to bound the integrand, Y_L converges uniformly in r.

Given (22) and by continuity of the mapping from (N, Y) to $\hat{\Lambda}_n = \int dN/Y$ ([21], Lemma 3) the discrete Nelson–Aalen estimators

$$\hat{\Lambda}_L = \int \frac{dN_L}{Y_r}$$

converge to $\hat{\Lambda}$. By continuity of the product-integral mapping ([23], Theorem 7), $\hat{F}_L$ converges to $\hat{F}$. A similar, simpler argument establishes the result for the reduced-sample estimator. □

It does not seem to be widely known in spatial statistics (cf. [9], page 764; [12] and [15]) that computation of the distances $\rho(x, \Phi \cap W)$ and $\rho(x, \partial W)$ for all points x in a fine rectangular lattice can be performed very efficiently using the distance transform algorithm of image processing [7, 8, 39, 40] at the price of accepting a discrete approximation to the true Euclidean metric $\|\cdot\|_2$ in the definition of ρ at (1). Thus the reduced-sample and Kaplan–Meier estimators are equivalent in computational cost when a fine grid is used.

It is often of interest to replace Euclidean distance by another metric, either for computational convenience as above, or in order to obtain different information about the process Φ [26, 47], particularly in three dimensions [5]. It is possible to replace $\|\cdot\|_2$ by another vector space norm $\|\cdot\|$ in the above results, provided the unit ball of $\|\cdot\|$ is a polyhedron (in $\mathbb{R}^2$ a polygon) scaled so that

$$\sup\left\{\frac{\|x\|}{\|x\|_2} : \|x\|_2 \le 1\right\} = 1. \tag{29}$$

Examples are $\|\cdot\|_\infty$, $\|\cdot\|_1/\sqrt{2}$ and continuous versions of the standard chamfer metrics [7, 8] used in the distance transform algorithm. Redefine the ball of radius r as $B(x, r) = \{y: \|x - y\| \le r\}$ and the distance function ρ of (1) in terms of $\|\cdot\|$. Then it can be shown that

$$\frac{d}{dr}|B(x,r)|_k = |\partial B(x,r)|_{k-1}$$

and that Lemma 1 remains true when A is a finite set but not in general. Hence Theorems 1 and 2 continue to hold for the Kaplan–Meier estimator with respect to this more general metric.

Estimation of F for more general sets B, and for more general random sets instead of the point process Φ, is treated in [25] and [26].

3.4. *Simulations.* We have compared the performance of the Kaplan–Meier and reduced-sample estimators of F in Monte Carlo simulations of a Poisson process and of a randomly translated square grid.

Both processes were simulated as binary images on a 256×256 square grid. For the Poisson process of intensity α, the pixel values were i.i.d. Bernoulli variables with $p = \alpha/(256^2)$. We generated 100 realizations of each of Bernoulli $p = 0.001$, 0.0001, 0.00005, 0.00002 and randomly translated grids of side $s = 25$, 32, 50, 100 and 150. For each realization the distance transform was computed in the chamfer $(5,7)$ metric of Borgefors [7] and the two estimators were derived.

Figure 2 compares the sample standard deviation of the reduced-sample estimator with the sample root mean square error of the Kaplan–Meier estimator (since the reduced-sample estimator is unbiased pointwise for F). The Kaplan–Meier estimator appears to be uniformly more efficient.

Figure 3 is a similar comparison for a randomly translated grid. Here the comparison is not uniformly favorable to the Kaplan–Meier estimator, although it is generally better. One can attribute this to periodic effects. For certain values of r the reduced-sample estimator is exact; near these values it has small variance. The mse of the Kaplan–Meier estimator oscillates for similar reasons. An extreme case is Figure 4, where the grid dimension 32 is a divisor of the window dimension 256.

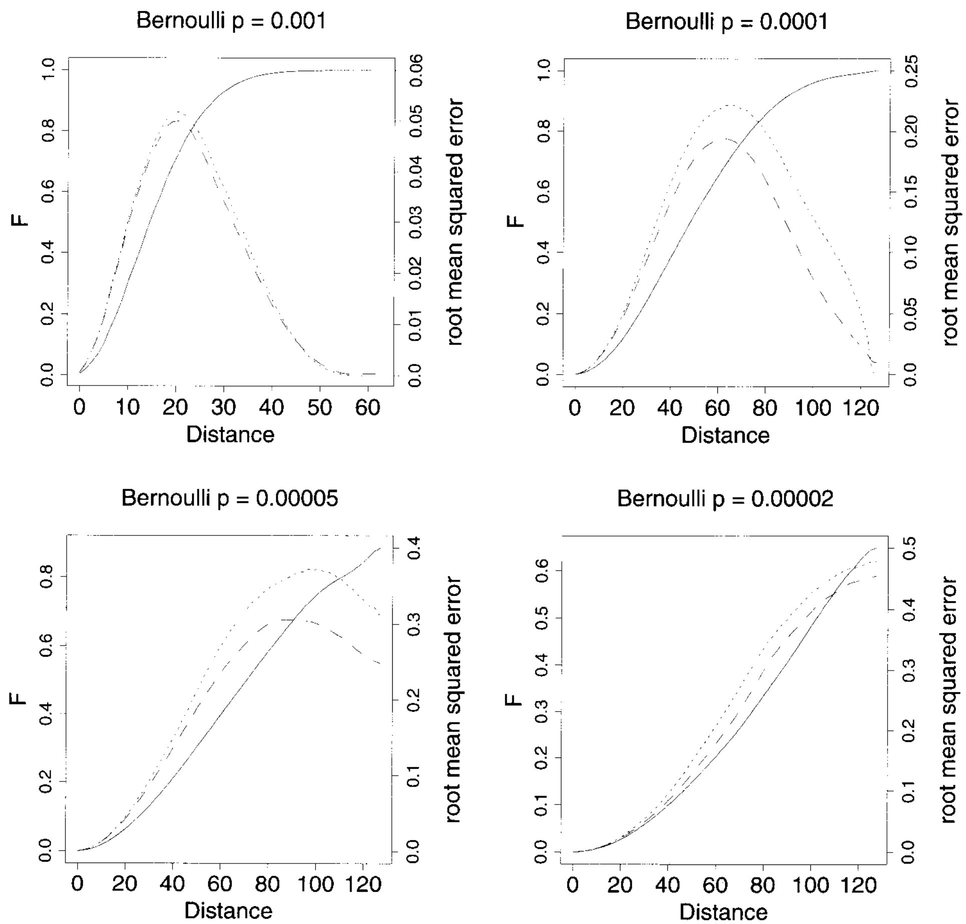

FIG. 2. *Root mean square error comparison for simulations of Poisson process. Dotted lines: reduced sample estimator; dashed lines: Kaplan–Meier estimator; solid lines: estimand F. Note different scales for rmse and F.*

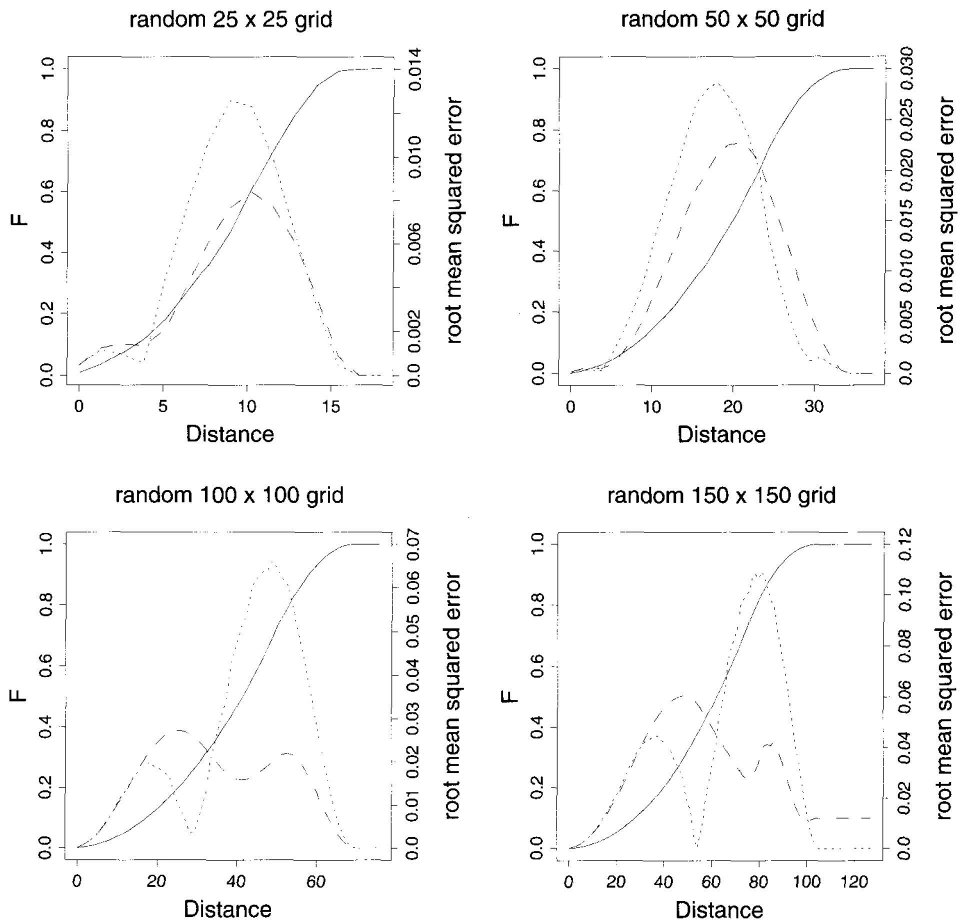

FIG. 3. *RMSE comparison for simulations of randomly translated grids. Dotted lines: reduced sample estimator; dashed lines: Kaplan–Meier estimator; solid lines: estimand F.*

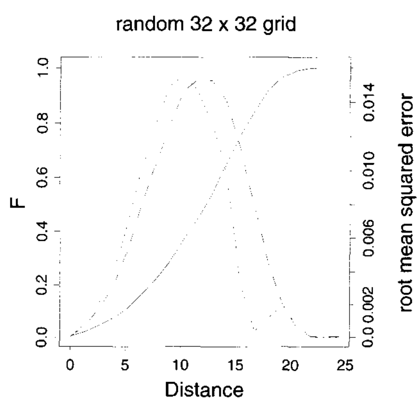

FIG. 4. *Extreme resonance case of Figure 3.*

4. Asymptotic properties of estimators of F.

4.1. *Large sample theory.* In spatial statistics, many large sample limiting regimes are possible (see [9], page 480; [38]; [45], page 224). It is common to consider the limit in which the window W expands to fill $\mathbb{R}^k$ [3, 27, 28, 44]. Under the additional assumption of ergodicity, it is clear that the reduced-sample and Kaplan–Meier estimators of F are pointwise consistent as $W \nearrow \mathbb{R}^k$. Edge effects are asymptotically negligible in this limit.

However, edge effects are appreciable in practical applications, so it would be more relevant to study asymptotic regimes in which the edge effect remains equally severe for all sample sizes. One such limit is considered by Stein [45].

We shall consider the situation where there are n independent replicated observations Φ_i of a process Φ within a fixed window W. This is becoming increasingly common in applications: the data consist of 10–100 binary images which may be treated as independent replications of the same process (e.g., [5, 26]). Equivalently, if Φ satisfies a mixing assumption, we may consider observation of the same point process through n distantly separated windows W_i of fixed size and shape (cf. [5]). Apart from its practical relevance, study of this limiting regime ($n \to \infty$ replicates) enables qualitative comparison of different estimators and may provide suggestions for variance estimation.

Given n replicated observations Φ_i in W, the pooled statistics $\hat{F}$ and $\hat{F}^{\mathrm{rs}}$ are obtained, not as the mean of the separate estimators for each window, but by analogues of (16) and (18) in which the numerators and denominators of (15) and (18) are replaced by the sums of these expressions over all replicates Φ_i. Asymptotics as $n \to \infty$ are now straightforward using empirical process theory.

THEOREM 3. *Let $\Phi_1, \Phi_2, \dots$ be i.i.d. copies of an a.s. stationary point process Φ with finite positive intensity α. Fix a compact set $W \subset \mathbb{R}^k$ and let $\hat{F}_n$ be the Kaplan–Meier estimator of F obtained from $\Phi_1, \dots, \Phi_n$ in W by pooling as above. Let $\tau > 0$ satisfy $F(\tau) < 1$. Then $\hat{F}_n$ is consistent and $\sqrt{n}(\hat{F}_n - F)$ converges weakly in $C[0, \tau]$ to a Gaussian process with linear approximation*

$$\hat{F}_n(r) - F(r) = \frac{1}{n} \sum_{i=1}^{n} I\big(\hat{F}, \Phi_i, r\big) + o_p(n^{-1/2}) \tag{30}$$

uniformly in $0 \le r \le \tau$, where I is the influence function,

$$I(\hat{F}, \Phi, r) = (1 - F(r)) \int_0^r \frac{|W_{\ominus s} \cap \partial(\Phi_{\oplus s})|_{k-1} - |W_{\ominus s} \setminus \Phi_{\oplus s}|_k \lambda(s)}{y(s)}\, ds \tag{31}$$

and $y(s) = \mathbb{E}|W_{\ominus s} \setminus \Phi_{\oplus s}|_k = (1 - F(s))|W_{\ominus s}|_k$. A similar statement holds for the reduced-sample estimator $\hat{F}_n^{\mathrm{rs}}$ with the influence function replaced by $I(\hat{F}^{\mathrm{rs}}, \Phi, r) = \hat{F}^{\mathrm{rs}}(r) - F(r)$.

We use the following lemma.

LEMMA 4. *Any nonnegative, monotone nondecreasing, cadlag process X on $[0,1]$ with finite second moment satisfies a uniform LLN and CLT in $D[0,1]$.*

SKETCH OF PROOF. Since $\mathbb{E}\sup_t X(t)^2 = \mathbb{E}X(1)^2 < \infty$, for the bracketing CLT of [49], Theorem 2.11.9 (cf. [1]), it suffices to find for each $\varepsilon > 0$ a partition of $[0,1]$ into N_ε sets $I_{j,\varepsilon}$ such that $\mathbb{E}\sup_{s,t\in I_{j,\varepsilon}}(X(t)-X(s))^2 < \varepsilon^2$ for each j and $\int_0^1\sqrt{\log N_\varepsilon}\,d\varepsilon < \infty$. For $s < t$, $(X(t)-X(s))^2 \le 2X(1)(X(t)-X(s))$ a.s. The function $h(t) := 2\mathbb{E}[X(1)X(t)]$ is finite, monotone nondecreasing and right-continuous. Given $\varepsilon > 0$, there are at most $h(1)/\varepsilon^2$ points where h jumps by more than ε^2. By adding further points we can partition $[0,1]$ into $N_\varepsilon \le 2h(1)/\varepsilon^2$ intervals $[t_i, t_{i+1})$ with $t_i < t_{i+1}$ such that $h(t_{i+1}-) - h(t_i) \le \varepsilon^2$. The result follows. □

PROOF OF THEOREM 3 (Sketch). Let $N_i(r), Y_i(r)$ for $i = 1,2,\dots$ be the fraction of failures and fraction at risk processes (27) and (28) for Φ_i in W. They are monotone and uniformly bounded by 1. By Lemma 4, they satisfy a LLN and CTL uniformly on $[0,\tau]$. A joint CLT follows immediately. Apply the functional delta-method ([21], Theorem 3) to the sequence of mappings from (N_n, Y_n) to $(N_n, 1/Y_n)$, then to $\hat{\Lambda}_n = \int (dN_n)/Y_n$, then to $1-\hat{F}_n = \pi(1-d\hat{\Lambda}_n)$. Each mapping is Hadamard differentiable or compactly differentiable ([21], [22], and [23], Theorem 8). Hence $\sqrt{n}(\hat{F}_n - F)$ is asymptotically equivalent (in the sense that the supremum of the difference over any bounded interval converges in probability to zero) to the linear functional of the empirical processes

$$\sqrt{n}\,(1-F(t))\int_0^t \frac{dN_n(s) - Y_n(s)\,d\Lambda(s)}{(1-\Delta\Lambda(s))\,y(s)}, \qquad 0 \le t < \tau,$$

where $y(s) = \mathbb{E}Y_n(s)$. □

4.2. *Calculations for the sparse Poisson limit.* From Theorem 3 we can obtain the asymptotic variance of the Kaplan–Meier estimator as the variance of the influence function (31). However, this expression is unwieldy, and further simplifying assumptions are needed to obtain explicit results.

In this section we calculate variances of (31) for the extreme case of a Poisson process whose intensity α is sent to zero. Edge effects become increasingly severe for small α.

This sparse Poisson limit is chosen because it is mathematically tractable, yet is stringent enough to reveal qualitative differences between the competing estimators. The differences emerge in the first-order approximation and not (as is usual) at higher orders. The limit also facilitates comparisons with results in survival analysis. It is, of course, an extreme situation which may not have direct practical impact. It may be relevant to applications where data are observed in a large number of windows, each window containing relatively little information.

There are just two situations to consider as $\alpha \to 0$: (i) no random point in W, with probability $\exp(-\alpha|W|_k) = 1 + \mathscr{O}(\alpha)$, and (ii) one random point in W at a position x uniformly distributed over W, occurring with probability $\alpha|W|_k \exp(-\alpha|W|_k) = \alpha|W|_k + \mathscr{O}(\alpha^2)$; the remaining possibilities have probability $\mathscr{O}(\alpha^2)$.

The influence function (31) for Kaplan–Meier is the difference of two terms: a part depending on surface areas at some distances from a point of Φ and a part depending on volumes at risk and involving the hazard rate of F. In case (i) only the second part is present and is of order α; in case (ii) the first part is also present and is of constant order.

The empty space function for the Poisson process is

$$F(r) = 1 - \exp(-\alpha|B_r|_k)$$

and its hazard rate is

$$\lambda(r) = \frac{d}{dr}[-\log(1 - F(r))] = \alpha|\partial B_r|_{k-1},$$

where $B_r = B(0, r)$ is a ball of radius r in the Euclidean metric, so that $|B_r|_k = r^d\omega_d/d$ and $|\partial B_r|_{k-1} = r^{d-1}\omega_d$. The expected volume at risk is

$$y(r) = (1 - F(r))|W_{\ominus r}|_k.$$

In case (i), no random points in W, (31) is therefore

$$\begin{aligned}
I(\hat{F}, \varnothing, r) &= (1 - F(r))\left\{-\int_0^r \frac{\alpha|\partial B_s|_{k-1}|W_{\ominus s}|_k}{|W_{\ominus s}|_k \exp(-\alpha|B_s|_k)}\, ds\right\} \\
&= \exp(-\alpha|B_r|_k)\left\{-\int_0^r \alpha|\partial B_s|_{k-1} \exp(\alpha|B_s|_k)\, ds\right\} \\
&= -(1 - \exp(-\alpha|B_r|_k)) \\
&= -\alpha|B_r|_k + \mathscr{O}(\alpha^2).
\end{aligned}$$

In case (ii) the influence function is

$$\begin{aligned}
&I(\hat{F}, \{x\}, r) \\
&= (1 - F(r))\left\{\int_0^r \frac{|\partial B(x,s) \cap W_{\ominus s}|_{k-1} - \alpha|\partial B_s|_{k-1}|W_{\ominus s} \setminus B(x,s)|_k}{|W_{\ominus s}|_k \exp(-\alpha|B_s|_k)}\, ds\right\} \\
&= \exp(-\alpha|B_r|_k)\int_0^r \frac{|\partial B(x,s) \cap W_{\ominus s}|_{k-1}}{|W_{\ominus s}|_k \exp(-\alpha|B_s|_k)}\, ds + \mathscr{O}(\alpha) \\
&= \int_0^r \frac{|\partial B(x,s) \cap W_{\ominus s}|_{k-1}}{|W_{\ominus s}|_k}\, ds + \mathscr{O}(\alpha).
\end{aligned}$$

It is an interesting exercise to check this by verifying that the expected influence function is zero to first order, using integral geometry ([41], page 97).

Hence the variance of the influence function is to first order,

$$\operatorname{var} I(\hat{F}, \Phi, r) \sim \alpha |W|_k \mathbb{E}\left(\int_0^r \frac{|\partial B(x,s) \cap W_{\ominus s}|_{k-1}}{|W_{\ominus s}|_k} \, ds \right)^2 \tag{32}$$

since case (i) is now $\mathscr{O}(\alpha^2)$. For the reduced sample estimator, in case (i) the estimator is identically zero; in case (ii) it is

$$\hat{F}^{\mathrm{rs}}(r) = |B(x,r) \cap W_{\ominus r}|_k / |W_{\ominus r}|_k.$$

Since $F(r) = 1 - \exp(-\alpha |B_r|_k) = \alpha |B_r|_k + \mathscr{O}(\alpha^2)$, the influence function (= estimator − estimand in this linear case) is in case (i),

$$I(\hat{F}^{\mathrm{rs}}, 0, r) = -\alpha |B_r|_k + \mathscr{O}(\alpha^2);$$

in case (ii),

$$I(\hat{F}^{\mathrm{rs}}, \{x\}, r) = \frac{|B(x,r) \cap W_{\ominus r}|_k}{|W_{\ominus r}|_k} + \mathscr{O}(\alpha).$$

Again, it can be verified using integral geometry that the expectation of the influence function is zero to first order. The variance is

$$\operatorname{var} I(\hat{F}^{\mathrm{rs}}, \Phi, r) \sim \alpha |W|_k \mathbb{E}\left(\left(\frac{|B(x,r) \cap W_{\ominus r}|_k}{|W_{\ominus r}|_k} \right)^2 \right). \tag{33}$$

For convenience in calculation of (32) and (33), we will take W to be the d-dimensional unit cube centered at $(\frac{1}{2}, \dots, \frac{1}{2})$, and replace the Euclidean metric $\|\cdot\|_2$ by the L_∞ metric in the definition (1) of ρ and $A_{\oplus r}$, $A_{\ominus r}$ (see comments at the end of Section 3.3). Thus F becomes the empty square space function obtained by replacing $B(x,r)$ by a cube $B_\infty(x,r)$ of center x and side length $2r$.

In this case it becomes feasible to enumerate all possible ways the cubes $B_\infty(x,r)$ and $W_{\ominus r}$ intersect. Expressing the volume and surface area contributions in terms of x in each case, we integrate over r (for Kaplan–Meier only) and then over x.

In one dimension with $W = [-1/2, 1/2]$ the variance of $n^{1/2}(\hat{F}_W(r) - F(r))$, ignoring terms of order $O(\alpha^2)$, equals α times the following expression:

$$\begin{cases} 2r + (1-4r)\log(1-2r) - \frac{1}{2}(\log(1-2r))^2, & \text{for } 0 \le r \le \frac{1}{4}, \\ 2r + \displaystyle\int_{1/2}^{2r} \log u \log(1-u)\, du - 2r \log 2r \log(1-2r), & \text{for } \frac{1}{4} \le r < \frac{1}{2}. \end{cases}$$

For the reduced sample estimator $|\Phi_{\oplus r} \cap W_{\ominus r}|_k / |W_{\ominus r}|_k$, the correspondin formula is

$$\begin{cases} 4r^2(1 - \frac{8r}{3})/(1 - 2r)^2, & \text{for } 0 \le r \le \frac{1}{4}, \\ (8r - 1)/3, & \text{for } \frac{1}{4} \le r < \frac{1}{2}. \end{cases}$$

These functions are plotted in Figure 5 together with the correspondin curves for two and three dimensions; the latter have been calculated (b Mathematica) with a mixture of computer algebra, numerical integration (fo

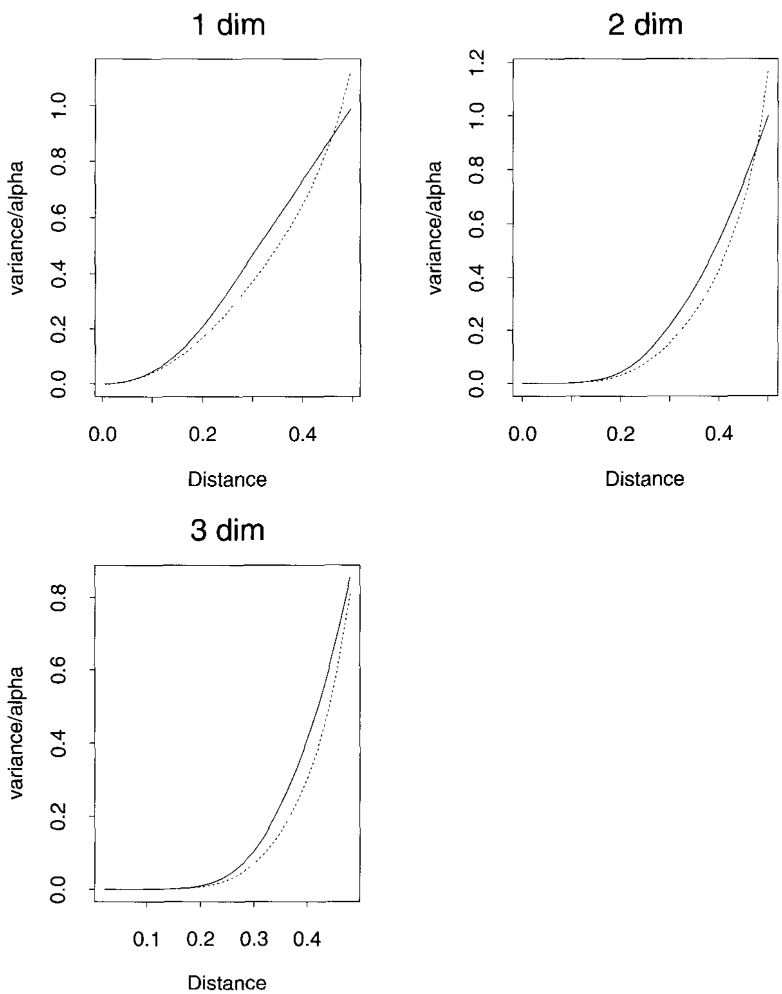

FIG. 5. *Sparse Poisson limit asymptotic variances (divided by α) in dimensions* 1, 2 *and* 3 *using the L_∞ metric. Solid lines: reduced sample estimator; dotted lines: Kaplan–Meier estimator.*

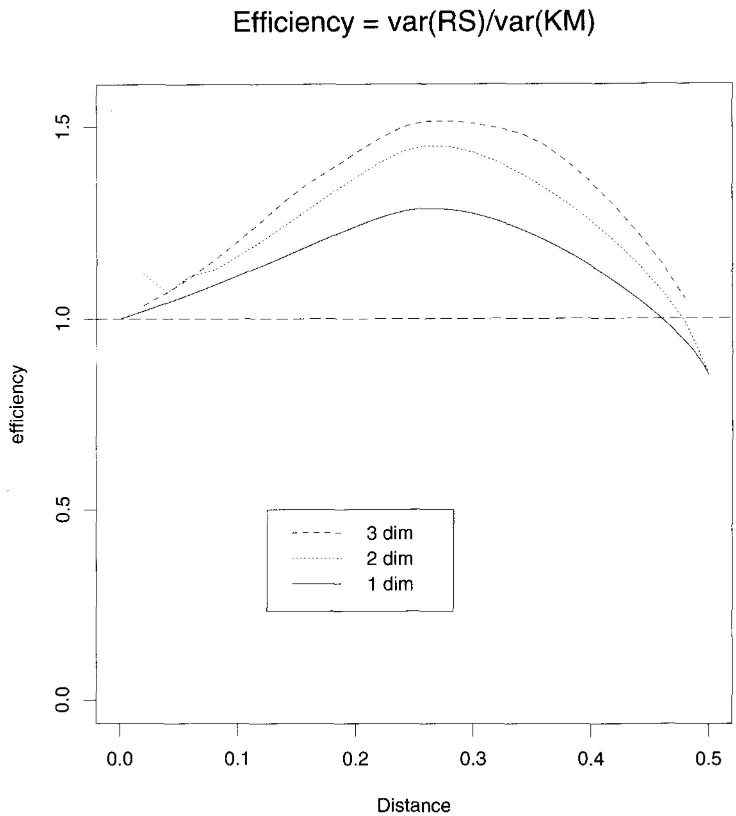

FIG. 6. *Asymptotic relative efficiency (ratio of variances) in* 1, 2 *and* 3 *dimensions, calculated for the* L_∞ *metric.*

integrals over s) and Monte Carlo integration (for integrals over x). The new estimator is superior over a broad range of distances r, but surprisingly deteriorates at very large distances. Apparently the dependence here has destroyed the uniform optimality enjoyed by the Kaplan–Meier estimator in the i.i.d. case.

Figure 6 shows the asymptotic relative efficiency in dimensions 1–3. The greatest gain is achieved at intermediate distances (near $\frac{1}{4}$); only for very large distances (near $\frac{1}{2}$) is there a loss in efficiency. As the dimension d increases and hence as edge effects become more severe, Kaplan–Meier represents an ever more convincing improvement on the reduced sample estimator.

5. The nearest neighbor function G. The nearest neighbor distance distribution function G was defined in (3) and (5). Note that G need not have any special continuity properties, in contrast to F; in fact G may be degenerate, as in the case of a randomly translated lattice.

5.1. *Kaplan–Meier estimator.* Let $\Phi \cap W = \{x_1, \dots, x_m\}$ be the observed point pattern. A Kaplan–Meier estimator for G is more immediate than for F; for each point x_i of the process Φ observed in the window W, one has a censored distance from x_i to the nearest other point of Φ,

$$s_i = \rho(x_i, \Phi \setminus \{x_i\}),$$

censored by its distance to ∂W,

$$b_i = \rho(x_i, \partial W).$$

Counting observed failures and numbers at risk as for censored data,

$$\begin{aligned} N^G(r) &= \#\{x \in \Phi \cap W: \rho(x, \Phi \setminus \{x\}) \le \rho(x, \partial W) \wedge r\} \\ &= \#\{i: s_i \le b_i \wedge r\} \end{aligned}$$

and

$$\begin{aligned} Y^G(r) &= \#\{x \in \Phi \cap W: r \le \rho(x, \Phi \setminus \{x\}) \wedge \rho(x, \partial W)\} \\ &= \#\{i: s_i \wedge b_i \ge r\}, \end{aligned}$$

define the Nelson–Aalen estimator

$$\hat{\Lambda}^G(r) = \int_0^r \frac{dN^G(s)}{Y^G(s)} \tag{34}$$

and the Kaplan–Meier estimator of G,

$$\begin{aligned} \hat{G}(r) &= 1 - \prod_0^r (1 - d\hat{\Lambda}^G(s)) \\ &= 1 - \prod_s \left(1 - \frac{\#\{i: s_i = s, s_i \le b_i\}}{\#\{i: s_i \ge s, b_i \ge s\}}\right), \end{aligned} \tag{35}$$

where s in the product ranges over the finite set $\{s_i\}$.

It follows from the Campbell–Mecke formula [see (5)] that the numerator and denominator of (34) satisfy the same mean-value relation as for ordinary randomly censored data,

$$\mathbb{E}N^G(r) = \int_0^r \mathbb{E}Y^G(s)\, d\Lambda^G(s), \tag{36}$$

where $d\Lambda^G(s) = dG(s)/(1 - G(s-))$.

Compare this to the reduced-sample estimator

$$\hat{G}_1(r) = \frac{\sum_{x \in \Phi \cap W_{\ominus r}} \mathbf{1}\{\rho(x, \Phi \setminus \{x\}) \le r\}}{\Phi(W_{\ominus r})} = \frac{\#\{i: s_i \le r, b_i \ge r\}}{\#\{i: b_i \ge r\}} \tag{37}$$

and the modification

$$\hat{G}_2(r) = \frac{|W|_k}{n} \frac{\#\{i: s_i \le r, b_i \ge r\}}{|W_{\ominus r}|_k} \tag{38}$$

obtained by replacing $\Phi(W_{\ominus r})$ by an estimate of its expectation. Other estimators are described in [47], page 128; [9], pages 614 and 637, 638; [13] and [18].

5.2. *Large sample theory for $\hat{G}$.* Linearization can be applied to $\hat{G} - G$ just as well as for $\hat{F} - F$ and the results used to study variances.

Analogously to Theorem 3, since $N^G(r)$, $Y^G(r) \le \Phi(W)$, it follows from Lemma 4 that provided $\mathbb{E}\Phi(W)^2 < \infty$, each of N^G, Y^G satisfies a LLN and CLT uniformly on an interval $[0, \tau]$, where $\mathbb{E}Y(\tau) > 0$. A joint LLN and CLT for (N^G, Y^G) follow immediately. Then differentiability of the product-integral mapping implies weak convergence of $\hat{G}$ to a Gaussian process at rate $\sqrt{n}$ and the asymptotic variance is equivalent to that of the influence function.

The Kaplan–Meier influence function for the nearest neighbor distances equals the sum over points $x \in \Phi \cap W$ of the usual influence function based on a censored observation $(\rho(x, \Phi \setminus \{x\}) \wedge \rho(x, \partial W),\ \mathbf{1}\{\rho(x, \Phi \setminus \{x\}) \le \rho(x, \partial W)\})$. The effective censoring distribution is that of the distance to ∂W from a uniformly distributed random point in W.

Fix the window W, an arbitrary compact set with Lebesgue measure 1. The information we need about W and the metric $\|\cdot\|$ is contained in the functions $b(r) = |B_d(0, r)|_k$, $c(r) = |\partial B_d(0, r)|_{k-1}$ and $e(r) = |W_{\ominus r}|_k$, where the erosion $W_{\ominus r}$ is defined in terms of $\|\cdot\|$. For the L_∞ metric and $W = [0,1]^k$, we have $e(r) = (1 - 2r)^k$, $b(r) = (2r)^k$ and $c(r) = 2k(2r)^{k-1}$.

The influence function for the Kaplan–Meier estimator of G is thus

$$\sum_{x \in \Phi \cap W} (1 - G(r)) \left[\frac{\mathbf{1}\{\rho(x, \Phi \setminus \{x\}) \le r \wedge \rho(x, \partial W)\}}{y(\rho(x, \Phi \setminus \{x\}))} - \int_0^{r \wedge \rho(x, \Phi\setminus\{x\}) \wedge \rho(x, \partial W)} \frac{\Lambda(ds)}{y(s)} \right],$$

where Λ is the cumulative hazard function associated with G and

$$\begin{aligned} y(r) &= \mathbb{E}\left\{ \sum_{x \in \Phi \cap W} \mathbf{1}\{\rho(x, \Phi \setminus \{x\}) \ge r,\ \rho(x, \partial W) \ge r\} \right\} \\ &= \alpha(1 - G(r))e(r). \end{aligned}$$

The factor α is the expected number of points in W since $|W|_k = 1$.

5.3. *Sparse Poisson asymptotics for $\hat{G}$.* Suppose the process is homogeneous Poisson with intensity α; then $G(r) = \exp(-\alpha b(r))$ and $\Lambda(ds) = \alpha c(s)\,ds$. For α small, $1 - G(r) \sim 1$ and $y(r) \sim \alpha e(s)$. The cases $\Phi(W) = 0, 1, 2$ have probabilities ~ 1, α and $\frac{1}{2}\alpha^2$ and result in influence functions

$$I(\hat{G}, \varnothing, r) = 0,$$

$$I(\hat{G}, \{x\}, r) = -(1 - G(r)) \int_0^{r \wedge \rho(x, \partial W)} \frac{\Lambda(ds)}{y(s)} \sim -\int_0^{r \wedge \rho(x, \partial W)} \frac{c(s)}{e(s)}\, ds,$$

$$I(\hat{G},\{x,z\},r)$$
$$= (1-G(r))\Bigg\{\mathbf{1}\{d(x,z)\le r\} \times \frac{\mathbf{1}\{d(x,z)\le \rho(x,\partial W)\} + \mathbf{1}\{d(x,z)\le \rho(z,\partial W)\}}{y(d(x,z))} - \int_0^{r\wedge d(x,z)\wedge \rho(x,\partial W)} \frac{\Lambda(ds)}{y(s)} - \int_0^{r\wedge d(x,z)\wedge \rho(z,\partial W)} \frac{\Lambda(ds)}{y(s)}\Bigg\}$$
$$\sim \mathbf{1}\{d(x,z)\le r\} \times\left(\frac{\mathbf{1}\{d(x,z)\le \rho(x,\partial W)\} + \mathbf{1}\{d(x,z)\le \rho(z,\partial W)\}}{\alpha e(d(x,z))}\right) - \int_0^{r\wedge d(x,z)\wedge \rho(x,\partial W)} \frac{c(s)}{e(s)}\,ds - \int_0^{r\wedge d(x,z)\wedge \rho(z,\partial W)} \frac{c(s)}{e(s)}\,ds.$$

Larger values of $\Phi(W)$ have probability of order α^3 and influence functions of order α^{-1}.

The required asymptotic variance is the expectation of the square of the influence function. The leading term comes from the first part of the case $\Phi(W)=2$ and is (of constant order)

$$\operatorname{var} I(\hat{G},\Phi,r) \to \frac{1}{2}\mathbb{E}\left\{\mathbf{1}\{d(U,V)\le r\} \times\left(\frac{(\mathbf{1}\{\rho(U,\partial W)\ge d(U,V)\} + \mathbf{1}\{\rho(V,\partial W)\ge d(U,V)\})^2}{e(d(U,V))^2}\right)\right\} \tag{39}$$

where U, V are independent uniformly distributed random points in W.

We now look at the reduced-sample estimator (37) in the same way. The expectations of numerator and denominator are $\alpha G(r)|W_{\ominus r}|_k$ and $\alpha|W_{\ominus r}|_k$ respectively, so that the linearized estimator minus estimand is

$$I(\hat{G}_1,\Phi,r) = \frac{\sum_{x\in\Phi\cap W_{\ominus r}} \mathbf{1}\{\rho(x,\Phi\setminus\{x\})\le r\} - G(r)\Phi(W_{\ominus r})}{\alpha|W_{\ominus r}|_k}$$
$$= \sum_{x\in\Phi\cap W} \frac{\mathbf{1}\{\rho(x,\partial W)\ge r\}(\mathbf{1}\{\rho(x,\Phi\setminus\{x\})\le r\} - G(r))}{\alpha|W_{\ominus r}|_k}.$$

The cases $\Phi(W)=0,1,2$ give influence functions [up to higher order terms and putting $G(r)\sim\alpha b(r)$],

$$I(\hat{G}_1,\varnothing,r)=0,$$

$$I(\hat{G}_1,\{x\},r) = -\frac{b(r)}{c(r)}\mathbf{1}\{\rho(x,\partial W)\ge r\}$$

$$I(\hat{G}_1, \{x, z\}, r) = \frac{\mathbf{1}\{d(x,z) \le r\}(\mathbf{1}\{\rho(x,\partial W) \ge r\} + \mathbf{1}\{\rho(z,\partial W) \ge r\})}{\alpha e(r)} - \frac{b(r)}{e(r)}(\mathbf{1}\{\rho(x,\partial W) \ge r\} + \mathbf{1}\{\rho(z,\partial W) \ge r\}).$$

For the asymptotic variance's leading term, again only the first part of the case $\Phi(W) = 2$ contributes, giving a term (of constant order)

$$\begin{aligned} &\operatorname{var} I(\hat{G}_1, \Phi, r) \\ &\quad \to \frac{1}{2}\mathbb{E}\left\{\mathbf{1}\{d(U,V) \le r\}\frac{(\mathbf{1}\{\rho(U,\partial W) \ge r\} + \mathbf{1}\{\rho(V,\partial W) \ge r\})^2}{e(r)^2}\right\}, \end{aligned} \tag{40}$$

where again U, V are independent uniformly distributed random points in W.

It is also easy to calculate the influence function of the estimator $\hat{G}_2$ defined in (38). Its asymptotic variance turns out to be asymptotically equivalent to that of $\hat{G}_1$ given above.

Compare (40) with the result (39) for Kaplan–Meier. These have leading terms of constant order, because only a fraction α of the realizations provide any data at all; this amplifies an asymptotic variance of order α by the factor $1/\alpha$ to constant order. In the case of F, asymptotic variances are of order α as we would expect.

Integration techniques of geometrical probability applied to (39) and (40) give, for the L_∞ metric and $W = [0,1]^k$,

$$\lim \operatorname{var} \hat{G}_1(r) = \frac{(2r)^k}{(1-2r)^k} + \frac{v(r)^k}{(1-2r)^{2k}},$$

where

$$\begin{aligned} v(r) &= 2(1-2r)(r \wedge (1-2r)) - (r \wedge (1-2r))^2 \\ &= \begin{cases} 2r - 5r^2, & \text{for } r \le 1/3, \\ (1-2r)^2, & \text{for } 1/3 \le r \le 1/2, \end{cases} \end{aligned}$$

and

$$\lim \operatorname{var} \hat{G}(r) = \int_0^r \frac{2k(2s)^{k-1}}{(1-2s)^k}\,ds + \int_0^{r \wedge 1/3} \frac{2k(1-3s)(2s-5s^2)^{k-1}}{(1-2s)^{2k}}\,ds.$$

The results are plotted in Figure 7 for dimensions 1, 2 and 3. They show a superiority of Kaplan–Meier over the reduced-sample estimator more marked than in the case of the empty space function. Moreover, the deterioration of the Kaplan–Meier estimator at large distances is no longer observed.

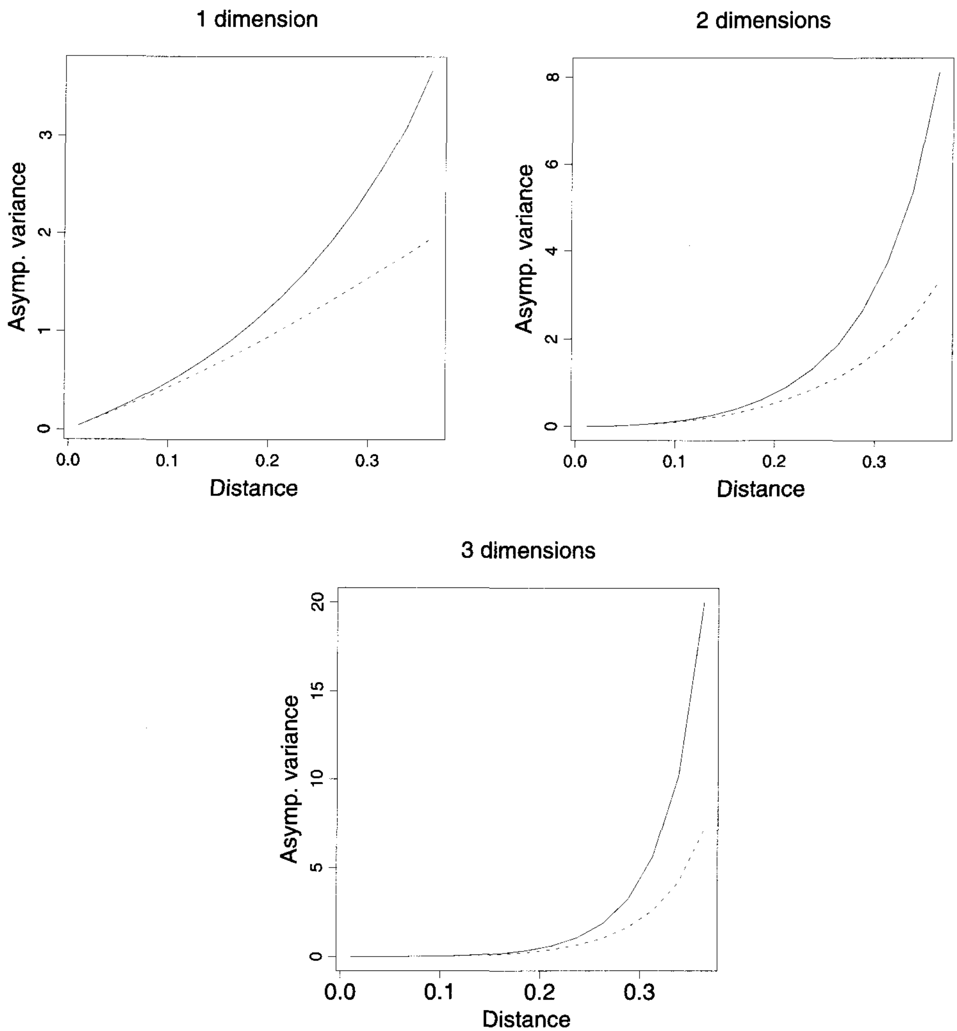

FIG. 7. *Asymptotic variance of estimators of G, sparse Poisson limit. Solid lines: reduced sample; dotted lines: Kaplan–Meier.*

6. The *K* function. $K(r)$ was defined in (4). Equivalently

$$\alpha K(r) = \sum_{n=0}^{\infty} G_n(r), \tag{41}$$

where $G_n(r) = \mathbb{P}\{\Phi(B(0,r)) > n \mid 0 \in \Phi\}$ is the distribution function of the distance from a typical point of Φ to the nth nearest point. For each G_n one can form a Kaplan–Meier estimator, since the distance from a point $x \in \Phi$ to its nth nearest neighbor is censored just as before by its distance to the boundary. The sequence of Kaplan–Meier estimators always satisfies the natural stochastic ordering of the distance distributions.

The large-sample theory we sketched for F and G can also be developed for K. Again we require $\mathbb{E}\Phi(W)^2 < \infty$. For the estimator of G_n the influence function has a similar form to that given for G. Since, for a Poisson process,

$$G_n(r) = e^{-\alpha b(r)} \sum_{k=n}^{\infty} \frac{(\alpha b(r))^k}{k!},$$

terms in the influence function for $\Phi(W) = 3, 4, \ldots$ remain of the same (lower) order, while those for $\Phi(W) = 0, 1, 2$ are unchanged. That is, *sparse Poisson asymptotics for $\hat{K}$ coincide with those for $\hat{G}$.* Hence our conclusions are similar to those of the previous section.

For estimating K, a number of sophisticated edge corrections exist; see [11]; [35–37]; [38], Chapter 3; [47], pages 122–131; [9], pages 616–619, 639–644; and recent investigations in [14], [17] and [43]. The asymptotic variances of these estimators are the variances of weighted analogues of the influence function given in the previous analysis. Figure 8 shows asymptotic variances for the rigid motion correction, translation correction and isotropic correction (estimated by Monte Carlo simulation of the influence function) together with the asymptotic variances of reduced sample and Kaplan–Meier estimators carried over from Figure 7. It turns out that under sparse Poisson asymptotics, the sophisticated edge corrections are equally as good and better than Kaplan–Meier, which in turn is better than the classical border method (reduced-sample) estimator.

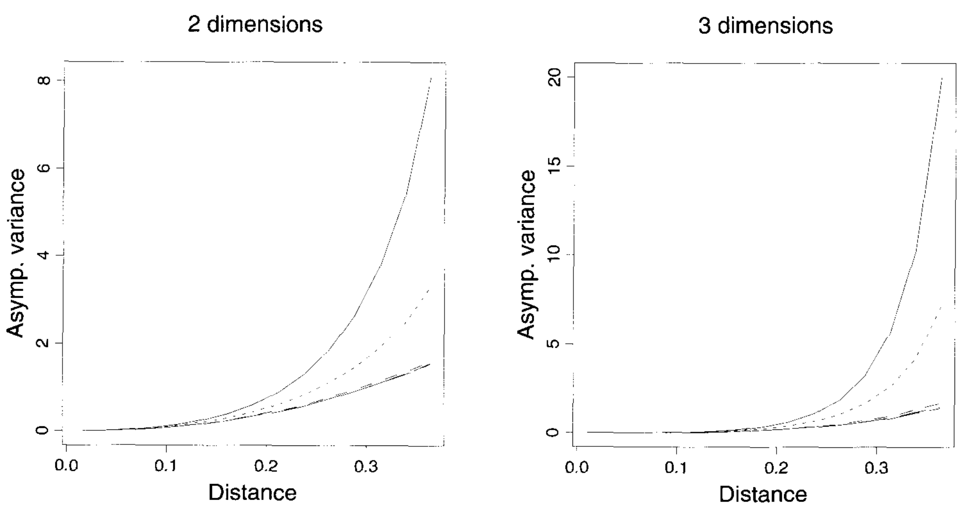

FIG. 8. *Asymptotic variance of estimators of K, sparse Poisson limit. Solid lines: reduced sample; dotted lines: Kaplan–Meier; dashed lines: weighted edge corrections; see text.*

In two-dimensional spatial statistics, it is common to transform K into $L(r) = \sqrt{K(r)/\pi}$. Our efficiency comparisons remain the same and all asymptotic variances are multiplied by a constant factor $(2\pi\sqrt{K(r)}\,)^{-1}$.

7. General discussion. The Kaplan–Meier technique has been shown to provide good estimators of all three distributions F, G and K. It appears to be substantially more efficient than the simple border correction (reduced-sample) estimators in most situations. However, in the case of K, the Kaplan–Meier estimator is less efficient (asymptotically in the sparse Poisson limit) than the more sophisticated edge corrections currently in favor. This loss of efficiency is offset by the ease of implementing the Kaplan–Meier estimator for arbitrary windows W, while the popular edge corrections are only easy to apply in rectangular windows.

Experimentation is needed to compare the worth of the various estimators in practical situations (see, e.g., [13]). Heinrich [27] proved large-domain limit theorems concerning the estimation of K in Poisson cluster processes, and Stoyan, Bertram and Wendrock [46] derived approximations to the variance of kernel estimators of the pair-correlation function.

The Kaplan–Meier estimator casts new light on the local knowledge principle ([42], pages 49, 233). This states, for example, that for all X, $W \subset \mathbb{R}^k$,

$$X_{\oplus r} \cap W_{\ominus r} = (X \cap W)_{\oplus r} \cap W_{\ominus r}$$

and that $W_{\ominus r}$ is the largest set Z satisfying

$$X_{\oplus r} \cap Z = (X \cap W)_{\oplus r} \cap Z \quad \text{for all } X.$$

In words, given observation of a set X within a window W, the dilation of X is known only within the mask $W_{\ominus r}$. While this principle has been used to justify the border method (reduced-sample) estimators, it is not in conflict with the construction of the Kaplan–Meier estimator since $\hat{F}(r)$ is based on hazard estimates for distances $s \le r$.

The Kaplan–Meier estimators use more information than the corresponding reduced-sample estimators, but not all information, in the following sense. Write $C(x)$ for the censoring distance $\rho(x, \partial W)$ at a point x and $T(x)$ for the observed failure distance $\rho(x, \Phi)$ or $\rho(x, \Phi \setminus \{x\})$ as appropriate. Then the reduced-sample estimate at distance r depends only on those points x where $C(x) \ge r$, while the Kaplan–Meier estimate also involves cases where $T(x) \le C(x)$ but $C(x) < r$. However, neither estimator makes use of cases where $C(x) < T(x)$ and it seems plausible that these may contain usable information. The sophisticated edge-correction estimators for K use information from the case $C(x) < T(x) \le r$. Doguwa [15] argues that information should be used from all six possible orderings of $C(x)$, $T(x)$, r.

A bootstrap result for the estimators of F, G and K in the independent replications case is available from the Giné–Zinn equivalence theorem that the bootstrap works if and only if the CLT holds; see, for example, [22], Section 11.

One might wonder whether it is possible to improve the Kaplan–Meier estimators of F, G and K by considering the observed distances as interval censored rather than just right censored. This seems possible since, for a point $x \in W$, which is closer to ∂W than to other points in $\Phi \cap W$, one does know that its distance to $\Phi \setminus \{x\}$ is not greater than its distance to $(\Phi \setminus \{x\}) \cap W$, so

$$\rho(x, \partial W) \leq \rho(x, \Phi \setminus \{x\}) \leq \rho(x, (\Phi \setminus \{x\}) \cap W).$$

Similar statements can be made for the distance to the kth nearest neighbor. However, treating this data as randomly interval-censored data would produce asymptotically biased estimators, since the upper limit $\rho(x, (\Phi \setminus \{x\}) \cap W)$ is strongly dependent on $\rho(x, \Phi \setminus \{x\})$, unlike the lower limit $\rho(x, \partial W)$.

The asymptotic theory also suggests variance estimators. In Theorem 3, the variance of $\hat{F}(r)$ can be approximated by the sum of the squares of the summands in (30), with $\lambda(\cdot)$ and F replaced by their Kaplan–Meier estimates. The expression (31) for $I(\hat{F}, \Phi, r)$ can be rewritten as an integral over $x \in W$ of the one-point influence function

$$(1 - F(r))\left(\frac{\mathbf{1}\{\rho(x,\Phi) \leq r, \rho(x,\Phi) \leq \rho(x, \partial W)\}}{y(\rho(x,\Phi))} - \int_0^{r \wedge \rho(x,\Phi) \wedge \rho(x,\partial W)} \frac{\lambda(s)}{y(s)}\, ds\right). \tag{42}$$

The integral over x can be approximated by a sum over lattice points as above. In order to implement this proposal one only has to numerically tabulate an estimate of the function $\int_0^r (\lambda(s)/y(s))\, ds$ together with the functions y and $1 - F$.

Alternatively (and this is applicable for $n = 1$ replicate) one can write down the variance of $I(\hat{F}, \Phi, r)$ in terms of the covariance structures of the random function $r(x) = \rho(x, \Phi)$ and of the window W:

$$\begin{aligned} &\operatorname{cov}(\hat{F}(r), \hat{F}(r')) \\ &\quad \approx (1 - F(r))(1 - F(r')) \int_{\mathbb{R}^k} \int_0^r \int_0^r C_{W_{\ominus s}, W_{\ominus s'}}(x) \frac{h(ds, ds', x)}{y(s) y(s')}\, dx, \end{aligned} \tag{43}$$

where $C_{A,B}(x)$ is the set cross-covariance function of $A, B \subset \mathbb{R}^k$,

$$C_{A,B}(x) = |A \cap (B + x)|_k, \qquad x \in \mathbb{R}^k,$$

with $B + x$ being the translate of B by x, and

$$h(t, t', x) = \operatorname{cov}(M(0,t), M(x,t')),$$

where

$$M(x,t) = \mathbf{1}\{\rho(x,\Phi) \leq t\} - \int_0^t \mathbf{1}\{\rho(x,\Phi) > s\} \lambda(s)\, ds,$$

a martingale in t for each $x \in \mathbb{R}^k$. Further work is needed to find good estimators of (43).

Finally, Kaplan–Meier estimators can also be developed for contact distributions, the analogues of F for random closed sets [47]. This is investigated in [25] and [26].

Acknowledgments. The simulations in Section 3.4 were performed using the image processing package `scilimage` [48] and a library written by Adri Steenbeek. We thank Martin Hansen, Marie-Colette van Lieshout, David Pollard, Aad van der Vaart and Jon Wellner for many helpful suggestions.

REFERENCES

[1] ANDERSEN, N. T., GINÉ, E., OSSIANDER, M. and ZINN, J. (1988). The central limit theorem and the law of the iterated logarithm for empirical processes under local conditions. *Probab. Theory Related Fields* **77** 271–305.

[2] BADDELEY, A. J. (1977). Integrals on a moving manifold and geometrical probability. *Adv. in Appl. Probab.* **9** 588–603.

[3] BADDELEY, A. J. (1980). A limit theorem for statistics of spatial data. *Adv. in Appl. Probab.* **12** 447–461.

[4] BADDELEY, A. J. (1993). Stereology and survey sampling theory. *Bulletin of the International Statistical Institute* **50** 435–449.

[5] BADDELEY, A. J., MOYEED, R. A., HOWARD, C. V. and BOYDE, A. (1993). Analysis of a three-dimensional point pattern with replication. *J. Roy. Statist. Soc. Ser. C* **42** 641–668.

[6] BARENDREGT, L. G. and ROTTSCHÄFER, M. J. (1991). A statistical analysis of spatial point patterns. A case study. *Statist. Neerlandica* **45** 345–363.

[7] BORGEFORS, G. (1984). Distance transformations in arbitrary dimensions. *Computer Vision, Graphics and Image Processing* **27** 321–345.

[8] BORGEFORS, G. (1986). Distance transformations in digital images. *Computer Vision, Graphics and Image Processing* **34** 344–371.

[9] CRESSIE, N. A. C. (1991). *Statistics for Spatial Data.* Wiley, New York.

[10] CROFTON, M. W. (1869). Sur quelques théorèmes du calcul intégral. *C. R. Acad. Sci. Paris* **68** 1469–1470.

[11] DIGGLE, P. J. (1983). *Statistical Analysis of Spatial Point Patterns.* Academic Press, London.

[12] DIGGLE, P. J. and MATÉRN, B. (1980). On sampling designs for the estimation of point-event nearest neighbor distributions. *Scand. J. Statist.* **7** 80–84.

[13] DOGUWA, S. I. (1989). A comparative study of the edge-corrected kernel-based nearest neighbor density estimators for point processes. *J. Statist. Comp. Simulation* **33** 83–100.

[14] DOGUWA, S. I. (1990). On edge-corrected kernel-based pair correlation function estimators for point processes. *Biometrical J.* **32** 95–106.

[15] DOGUWA, S. I. (1992). On the estimation of the point-object nearest neighbor distribution $F(y)$ for point processes. *J. Statist. Comp. Simulation* **41** 95–107.

[16] DOGUWA, S. I. and CHOJI, D. N. (1991). On edge-corrected probability density function estimators for point processes. *Biometrical J.* **33** 623–637.

[17] DOGUWA, S. I. and UPTON, G. J. G. (1989). Edge-corrected estimators for the reduced second moment measure of point processes. *Biometrical J.* **31** 563–575.

[18] DOGUWA, S. I. and UPTON, G. J. G. (1990). On the estimation of the nearest neighbour distribution, $G(t)$, for point processes. *Biometrical J.* **32** 863–876.

[19] FEDERER, H. (1969). *Geometric Measure Theory.* Springer, Heidelberg.

[20] FIKSEL, T. (1988). Edge-corrected density estimators for point processes. *Statistics* **19** 67–75.

[21] GILL, R. D. (1989). Non- and semiparametric maximum likelihood estimators and the von Mises method, I (with discussion). *Scand. J. Statist.* **16** 97–128.

[22] GILL, R. D. (1994). Lectures on survival analysis. *Ecole d'Eté de Probabilités de Saint-Flour 1992. Lecture Notes in Math.* **1581**. Springer, Berlin.

[23] GILL, R. D. and JOHANSEN, S. (1990). A survey of product-integration with a view toward application in survival analysis. *Ann. Statist.* **18** 1501–1555.

[24] HANISCH, K.-H. (1984). Some remarks on estimators of the distribution function of nearest neighbor distance in stationary spatial point patterns. *Mathematische Operationsforschung und Statistik—Statistics* **15** 409–412.

[25] HANSEN, M. B., BADDELEY, A. J. and GILL, R. D. (1994). Some regularity properties for first contact distributions. Preprint 890, Mathematics Institute, Univ. Utrecht.

[26] HANSEN, M. B., GILL, R. D. and BADDELEY, A. J. (1996). Kaplan–Meier type estimators for linear contact distributions. *Scand. J. Statist.* **23** 129–155.

[27] HEINRICH, L. (1988). Asymptotic Gaussianity of some estimators for reduced factorial moment measures and product densities of stationary Poisson cluster processes. *Statistics* **19** 87–106.

[28] JOLIVET, E. (1980). Central limit theorem and convergence of empirical processes for stationary point processes. In *Point Processes and Queueing Problems* (P. Bastfai and J. Tomko, eds.) 117–161. North-Holland, Amsterdam.

[29] KAPLAN, E. L. and MEIER, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc.* **53** 457–481.

[30] LASLETT, G. M. (1982). Censoring and edge effects in areal and line transect sampling of rock joint traces. *Math. Geol.* **14** 125–140.

[31] LASLETT, G. M. (1982). The survival curve under monotone density constraints with applications to two-dimensional line segment processes. *Biometrika* **69** 153–160.

[32] LOTWICK, H. W. (1981). Spatial stochastic point processes. Ph.D. thesis, Univ. Bath.

[33] MATHERON, G. (1975). *Random Sets and Integral Geometry.* Wiley, New York.

[34] MILES, R. E. (1974). On the elimination of edge-effects in planar sampling. In *Stochastic Geometry: A Tribute to the Memory of Rollo Davidson* (E. F. Harding and D. G. Kendall, eds.) 228–247. Wiley, New York.

[35] OHSER, J. (1983). On estimators for the reduced second moment measure of point processes. *Mathematische Operationsforschung und Statistik—Statistics* **14** 63–71.

[36] RIPLEY, B. D. (1977). Modelling spatial patterns (with discussion). *J. Roy. Statist. Soc. Ser. B* **39** 172–212.

[37] RIPLEY, B. D. (1981). *Spatial Statistics.* Wiley, New York.

[38] RIPLEY, B. D. (1988). *Statistical Inference for Spatial Processes.* Cambridge Univ. Press.

[39] ROSENFELD, A. and PFALZ, J. L. (1966). Sequential operations in digital picture processing. *J. Assoc. Comput. Mach.* **13** 471.

[40] ROSENFELD, A. and PFALZ, J. L. (1968). Distance functions on digital pictures. *Pattern Recognition* **1** 33–61.

[41] SANTALÓ, L. A. (1976). *Integral Geometry and Geometric Probability. Encyclopedia of Mathematics and Its Applications* **1**. Addison-Wesley, Reading, MA.

[42] SERRA, J. (1982). *Image Analysis and Mathematical Morphology.* Academic Press, London.

[43] STEIN, M. L. (1991). A new class of estimators for the reduced second moment measure of point processes. *Biometrika* **78** 281–286.

[44] STEIN, M. L. (1993). Asymptotically optimal estimation for the reduced second moment measure of point processes. *Biometrika* **80** 443–449.

[45] STEIN, M. L. (1995). An approach to asymptotic inference for spatial point processes. *Statistica Sinica* **5** 221–234.

[46] STOYAN, D., BERTRAM, U. and WENDROCK, H. (1993). Estimation variances for estimators of product densities and pair correlation functions of planar point processes. *Ann. Inst. Statist. Math.* **45** 211–221.

[47] STOYAN, D., KENDALL, W. S. and MECKE, J. (1987). *Stochastic Geometry and Its Applications.* Wiley, Chichester.

[48] TEN KATE, T. K., VAN BALEN, R., SMEULDERS, A. W. M., GROEN, F. C. A. and DEN BOER, G. A. (1990). SCILAIM: a multi-level interactive image processing environment. *Pattern Recognition Letters* **11** 429–441.

[49] VAN DER VAART, A. and WELLNER, J. A. (1993). *Weak Convergence and Empirical Processes.* IMS, Hayward, CA.

[50] WIJERS, B. J. (1991). Consistent non-parametric estimation for a one-dimensional line segment process observed in an interval. Preprint 683, Dept. Mathematics, Univ. Utrecht.

[51] ZIMMERMAN, D. (1991). Censored distance-based intensity estimation of spatial point processes. *Biometrika* **78** 287–294.

DEPARTMENT OF MATHEMATICS
UNIVERSITY OF WESTERN AUSTRALIA
NEDLANDS WA 6907
AUSTRALIA
E-MAIL: adrian@maths.uwa.edu.au

MATHEMATICAL INSTITUTE
UNIVERSITY OF UTRECHT
BUDAPESTLAAN 6
3584 CD UTRECHT
THE NETHERLANDS
E-MAIL: gill@math.ruu.nl