Biometrika (2020), **103**, 1, *p.* 1
Printed in Great Britain

Nonparametric estimation of the intensity function of a spatial point process on a Riemannian manifold

BY S. WARD
Department of Mathematics, Imperial College London, London SW7 2AZ, U.K.
scott.ward12@imperial.ac.uk

H. S. BATTEY
Department of Mathematics, Imperial College London, London SW7 2AZ, U.K.
h.battey@imperial.ac.uk

AND E. A. K. COHEN
Department of Mathematics, Imperial College London, London SW7 2AZ, U.K.
e.cohen@imperial.ac.uk

SUMMARY

This paper is concerned with nonparametric estimation of the intensity function of a point process on a Riemannian manifold. It provides a first-order asymptotic analysis of the proposed kernel estimator for Poisson processes, supplemented by empirical work to probe the behaviour in finite samples and under other generative regimes. The investigation highlights the scope for finite-sample improvements by allowing the bandwidth to adapt to local curvature.

Some key words: Boundary-free manifold; Edge and shape correction; Kernel estimation; Point events.

1. INTRODUCTION

In the analysis of random collections of point events, a fundamental role is played by the intensity function, which determines the first-order properties of a spatial point process and is an essential component of second order analyses. It provides a complete characterization for the smaller class of Poisson processes.

Features of spatial point processes, as distinct from those along a time axis, are their inherent multidimensionality and the need to treat all directions equivalently, in contrast to the directionality of one-dimensional time. A further feature, sometimes ignored with little effect due to the scales involved, are the topological features of the space on which the point events occur.

In the present paper we are concerned with point processes on the surface of a Riemannian manifold, a situation of high relevance in cellular biology and microbiology, where super-resolution microscopy techniques can record the spatial arrangement of proteins and other molecules of interest on the cellular membranes of cells, bacteria and other microorganisms. At these scales the topology cannot be ignored, necessitating inferential procedures that adapt to local curvature. In this microbiological example, knowledge of the intensity functions of, say, two different molecular processes can guide scientific inference by suggesting possible dependencies between the processes, perhaps to be probed more formally. Alternatively, the intensity estimates might be used as outcomes, blocking factors or concomitant variables in an experimental context, concerned with assessing the efficacy of one or more treatments.

Intensity estimation under this framework is unexplored. Recent relevant work is due to Robeson et al. (2014), Lawrence et al. (2016) and Møller & Rubak (2016), who considered functional summary statistics for point processes on the surface of a d-dimensional unit sphere. These summarize the global properties of the point process. Ward et al. (2021) extended the construction of such statistics to convex manifolds using the Mapping Theorem (e.g. Kingman, 1993, p.18) to map the point events on the manifold to the surface of the unit sphere, performed statistical analysis there using the rotational invariance of the sphere, and mapped the conclusions back to the manifold of interest. Whilst the present paper is concerned with more general processes and manifolds, we must similarly assume the implicit equation $g(x_1, \dots, x_d) = 0$ of the manifold is known in analytic form, or can be well approximated as illustrated in Section 7.

The closest related work is that concerned with nonparametric density estimation from independent and identically distributed (i.i.d.) observations constrained to the surface of a manifold. Pelletier (2005) extended some of the theory of kernel density estimation to accommodate i.i.d. observations on a finite volume boundary-free Riemannian manifold, while Kerkyacharian et al. (2012) considered so-called needlet density estimation on compact homogeneous manifolds, motivated by applications in astrophysics. As with their Euclidean counterparts, the broad strategies appropriate for kernel density and kernel intensity estimation are rather similar, although the technical differences are considerable, most notably: the point process observations cannot be treated as i.i.d.; the number of event observations are, at least in the present context, treated as random; and the point process is frequently not observed over the entire manifold. The latter situation necessitates procedures that can seamlessly accommodate both boundary-free manifolds and manifolds with boundaries.

2. PRELIMINARIES

Consider a compact d-dimensional Riemannian manifold $(\mathcal{M}, g)$ with Riemannian metric tensor g. Our treatment here is coordinate-free, i.e., avoiding a fixed basis in which to express all calculations. This formulation comes at the expense of greater abstraction but leads to a more compact notation. Most of the details are deferred to the Supplementary Material along with the proofs of the main results.

Let X be a point process over $\mathcal{M}$, most naturally viewed as a random set formed of elements of $\mathcal{M}$. To distinguish between points in a realization of X and any point in the space $\mathcal{M}$, we shall refer to the former as *events* and the latter as *points*. We use x both to specify points in $\mathcal{M}$ and to index events in X, with the context ensuring no ambiguity.

For any Borel measurable subset $B \subseteq \mathcal{M}$, let $N_X(B)$ denote the number of events of X in B and let $d\mathrm{vol}$ denote the d-dimensional Riemannian volume form on $\mathcal{M}$ (see Supplementary Material). The intensity measure is defined as $\mu(B) = E\{N_X(B)\}$ and provided that μ is absolutely continuous with respect to $d\mathrm{vol}$, there exists a function $\rho : \mathcal{M} \to \mathbb{R}$ called the *intensity function* such that

$$\mu(B) = \int_B \rho(x) d\mathrm{vol}(x).$$

In other words, ρ is the Radon-Nikodyn derivative of the intensity measure with respect to the Riemannian volume form. A more precise formalization avoiding ambiguity in the asymptotic framework of Section 3 is

$$\rho(x) = \lim_{\delta_x \to 0} \mathrm{vol}(\delta_x)^{-1} E\{N_X(\delta_x)\}, \tag{1}$$

where $\delta_x \subset \mathcal{M}$ is a region centred on x, and the notation $\delta_x \to 0$ means that the geodesic distance $d_g : \mathcal{M} \times \mathcal{M} \to \mathbb{R}_+$ between any two elements of δ_x tends to zero. This is a natural adaptation of the Euclidean definition of Cressie (2015) and retains all the usual properties. Under the constraint that X is simple, that is $\mathrm{pr}\{N_X(\delta_x) > 1\} = o\{\mathrm{vol}(\delta_x)\}$, $\rho(x)d\mathrm{vol}(x)$ can be interpreted as the probability of a point occurrence in the infinitesimal volume $d\mathrm{vol}(x)$ at $x \in \mathcal{M}$.

A point process X is said to be homogeneous if ρ is constant over $\mathcal{M}$ and otherwise inhomogeneous. By Campbell's theorem (Daley & Vere-Jones, 2010), for any measurable nonnegative function $f : W \subseteq \mathcal{M} \to \mathbb{R}_+$,

$$E\left\{ \sum_{x \in X \cap W} f(x) \right\} = \int_W f(x)\rho(x) d\mathrm{vol}(x). \tag{2}$$

Poisson processes can be characterized in the same way on $\mathcal{M}$ as in $\mathbb{R}^d$. Specifically, X is said to be a Poisson process with intensity function ρ if, for any Borel measurable subset $B \subseteq \mathcal{M}$, $N_X(B)$ is Poisson distributed with mean $\mu(B)$ and, for any disjoint Borel measurable subsets $A, B \subseteq \mathcal{M}$, $N_X(A)$ and $N_X(B)$ are independent random variables. This affords considerable simplification. In particular, for any measurable non-negative function $f : W \subseteq \mathcal{M} \to \mathbb{R}_+$,

$$\mathrm{Var}\left\{ \sum_{x \in X \cap W} f(x) \right\} = \int_W f^2(x)\rho(x) d\mathrm{vol}(x). \tag{3}$$

3. Intensity estimation on $\mathcal{M}$

Estimation of ρ is treated nonparametrically. As in simpler contexts (e.g. Bartlett, 1963; Cox, 1965, for events along a time axis) smoothing is required to achieve acceptable estimation variance. This entails some form of weighted averaging of nearby points, ideally with tapered weights for decreasing proximity. Intuitively, in regions of high curvature, neighbouring points appear closer in the Euclidean metric than the arc length of the shortest curve section between them, constrained to the surface of $\mathcal{M}$, namely the geodesic distance $d_g : \mathcal{M} \times \mathcal{M} \to \mathbb{R}_+$. This renders the standard Euclidean theory of kernel intensity estimation unusable. We pursue the natural approach of replacing the Euclidean metric in the kernel function by the geodesic distance, so that the kernel intensity estimator automatically adapts to local curvature.

A further complication in this setting is that a kernel function, typically non-compactly supported, centred at a particular point may not integrate to one over the manifold. This could be because the manifold has a boundary, or may only be observed over a convex compact subset of $\mathcal{M}$, a situation that is rather common in practice. A related problem arises when the manifold is of finite volume and boundary-free. Although this latter issue can be circumvented in certain special cases, for instance by using Fisher's (1953) density function as a kernel on the sphere or adopting finitely supported kernels (Pelletier, 2005), for more general manifolds and kernels a shape correction is needed, in effect to avoid double counting of points in the weighted average. Conveniently, the boundary correction required in the former situation is operationally the same as shape correction for finite-volume boundary-free manifolds. All cases can be encapsulated by defining a convex compact subset W of $\mathcal{M}$ over which the point process is observed. The corrections introduced in the forthcoming discussion are then either edge or shape corrections, the latter corresponding to $W = \mathcal{M}$ with $\mathcal{M}$ a finite-volume boundary-free manifold.

The intensity function estimator to be studied in the present paper is

$$\hat{\rho}_h(x) = \sum_{y \in X \cap W} \frac{c_h(x,y)^{-1}}{h^d} k\left\{ \frac{d_g(x,y)}{h} \right\}, \tag{4}$$

where W is as described above, $c_h(x,y)$ is the edge or shape correction and, for Euclidean norm $\|\cdot\|$, k is such that $k \circ \|\cdot\| : \mathbb{R}^d \to \mathbb{R}_+$ is a symmetric probability density function, specified for concreteness as Gaussian:

$$k\{d_g(\cdot, y)\} = (2\pi)^{-d/2} \exp\{-d_g^2(\cdot, y)/2\}.$$

In direct analogy to the corresponding corrections in $\mathbb{R}^d$ (Diggle, 1985; Berman & Diggle, 1989; van Lieshout, 2012), $c_h(x,y)$ is defined either globally or locally as

$$c_h^{\mathrm{glo}}(x,y) = c_h(x) = \frac{1}{h^d} \int_W k\left\{ \frac{d_g(x,z)}{h} \right\} d\mathrm{vol}(z), \tag{5}$$

$$c_h^{\mathrm{loc}}(x,y) = c_h(y) = \frac{1}{h^d} \int_W k\left\{ \frac{d_g(z,y)}{h} \right\} d\mathrm{vol}(z). \tag{6}$$

Specifically, the global correction depends only on the point at which the intensity is estimated, while the local correction adjusts for each event. The resulting estimator (4) is generally biased in finite samples regardless of which correction is used but, as shown in Proposition 1, the global version $\hat{\rho}_h^{\mathrm{glo}}$ is unbiased for homogeneous point processes. The local version $\hat{\rho}_h^{\mathrm{loc}}$ enjoys mass preservation for homogeneous and inhomogeneous processes, specifically,

$$\int_W \hat{\rho}_h^{\mathrm{loc}}(x) d\mathrm{vol}(x) \; = N_X(W), \tag{7}$$

as was demonstrated in the Euclidean case by van Lieshout (2012).

The estimators $\hat{\rho}_h^{\mathrm{glo}}(x)$ and $\hat{\rho}_h^{\mathrm{loc}}(x)$ are best justified by consideration of their first and second moment properties, stated as a series of Propositions of varying degrees of technical intricacy, and culminating in Proposition 3.

PROPOSITION 1. *Let $(\mathcal{M}, g)$ be a Riemannian manifold and let X be a homogeneous spatial point process over $\mathcal{M}$ with intensity function $\rho(x) = \rho$ for all $x \in \mathcal{M}$. Then for any h, $\hat{\rho}_h^{\mathrm{glo}}$ is unbiased for ρ while $E\{\hat{\rho}_h^{\mathrm{loc}}\} = \rho\eta$ with*

$$\eta = \int_W \frac{c_h^{\mathrm{loc}}(\,\cdot\,, y)^{-1}}{h^d} k\left\{ \frac{d_g(x,y)}{h} \right\} d\mathrm{vol}(y).$$

Proof. This is a special case of the more general result

$$E\{\hat{\rho}_h^{\bullet}(x)\} = \int_W \frac{c_h^{\bullet}(x,y)^{-1}}{h^d} k\left\{ \frac{d_g(x,y)}{h} \right\} \rho(y) \, d\mathrm{vol}(y), \quad \bullet \in \{\mathrm{glo}, \mathrm{loc}\},$$

which follows by Campbell's theorem. The result is immediate on noting the constancy of the intensity function. $\square$

Although $\hat{\rho}_h^{\mathrm{loc}}$ has multiplicative bias η for homogenous processes (and is expected to exhibit pointwise bias for inhomogeneous processes too) the functional $\int_W \hat{\rho}_h^{\mathrm{loc}}(x) d\mathrm{vol}(x)$ is always unbiased for $\mu(W)$ by taking expectations in (7).

PROPOSITION 2. *Let $(\mathcal{M}, g)$ be a Riemannian manifold and let X be a Poisson process on $\mathcal{M}$. Then*

$$\mathrm{Var}\{\hat{\rho}_h^{\mathrm{glo}}(x)\} = c_h^{\mathrm{glo}}(x,\,\cdot\,)^{-2} \int_W \left[\frac{1}{h^d} k\left\{\frac{d_g(x,y)}{h}\right\}\right]^2 \rho(y)\, d\mathrm{vol}(y),$$
$$\mathrm{Var}\{\hat{\rho}_h^{\mathrm{loc}}(x)\} = \int_W c_h^{\mathrm{loc}}(\,\cdot\,, y)^{-2} \left[\frac{1}{h^d} k\left\{\frac{d_g(x,y)}{h}\right\}\right]^2 \rho(y)\, d\mathrm{vol}(y).$$

Proof. The result is by direct calculation using (3). □

For homogeneous Poisson processes, it follows from Proposition 2 that the variance is not constant over $\mathcal{M}$ even though the intensity function is. This conclusion is equivalent to that of Rakshit et al. (2019) in the context of homogeneous point processes observed over linear networks.

Specification of the bandwidth h relies on a notional asymptotic regime in which the expected number of events $\mu(W)$ diverges. As in simpler contexts, the bias and variance are antagonistic as a function of h, and a suitable compromise between the two must be determined. In assessing the appropriate scaling of h with the expected number $\mu(W)$ of events, there are some subtleties that distinguish the present setting from the i.i.d. Euclidean case. In a Euclidean setting, one way to achieve $\mu(W) \to \infty$ is to consider an expanding W. This on its own is unsatisfactory, as the concentration of events around an arbitrary $x \in W$ could remain diffuse, as noted by Cucala (2008). The expanding W framework is also physically implausible in the context of boundary-free finite-volume manifolds where $W = \mathcal{M}$.

To ensure the target of inference is stable under the notional limiting operation $\mu(W) \to \infty$, the asymptotic properties of a suitably standardized version of (4) are considered, analogously to Cucala (2008). The standardized object of inference is $\rho_1(x) = \rho(x)/\mu(W)$, the relative concentration of events at each point of W ensuring that ρ_1 integrates to one over W. The corresponding estimators are

$$\hat{\rho}_{h,1}^{\bullet}(x) = \frac{1\{N_X(W) \neq 0\}}{N_X(W)} \sum_{y \in X \cap W} \frac{c_h^{\bullet}(x,y)^{-1}}{h^d} k\left\{\frac{-d_g(x,y)}{h}\right\}, \quad \bullet \in \{\mathrm{glo}, \mathrm{loc}\}, \tag{8}$$

where $1(A)$ denotes the indicator function of the event A, and the relationship to the estimator in (4) is $\hat{\rho}_h^{\bullet}(x) = N_X(W)\hat{\rho}_{h,1}^{\bullet}(x)$. For Poisson processes the following proposition gives the pointwise asymptotic properties of $\hat{\rho}_{h,1}^{\bullet}$ for $\bullet \in \{\mathrm{glo}, \mathrm{loc}\}$.

PROPOSITION 3. *Let $(\mathcal{M}, g)$ be a Riemannian manifold. Suppose X is a Poisson process parameterized by $\rho = \{\rho(x) : x \in \mathcal{M}\}$ and observed over the bounded window $W \subseteq \mathcal{M}$. Provided that ρ_1 is smooth, for any $x \in W \subseteq \mathcal{M}$ and $\bullet \in \{glo, loc\}$,*

$$E\{\hat{\rho}_{h,1}^{\bullet}(x)\} \to \rho_1(x),$$
$$\mathrm{Var}\{\hat{\rho}_{h,1}^{\bullet}(x)\} \to 0,$$

as $\mu(W) \to \infty$ provided that $h \to 0$ and $\mu(W)^{-1} = o(h^d)$.

This result supplies a degree of reassurance over the behaviour of the proposed estimators, as the conclusion coincides with that obtained in Euclidean space.

From a technical point of view there are some limitations of this analysis. Firstly, Proposition 3 is proved only for Poisson processes. It is supplemented by empirical work in Section 6, which probes the behaviour in finite samples and for other generative processes. Secondly, the conclusions are first-order asymptotic in nature, and not optimized to exploit the interaction between the process and the manifold. Since the volume of a ball of radius r at two points x and y on a general manifold is not necessarily equal for $x \neq y$, the expected number of events in such a ball is, in general, not constant over the manifold. Intuitively then, an optimal estimator would have a bandwidth that adapted to the local curvature, thereby producing a better separation between clustering induced by the process and that induced by the geometry. Similar reasoning would lead one to allow asymmetric localization via elongated "balls" and so on.

4. PRACTICAL GUIDE TO BANDWIDTH SELECTION

While Proposition 3 specifies the properties of h under a notional asymptotic regime, and thereby provides some theoretical reassurance over the proposed intensity estimator, the practical problem of choosing the bandwidth for a given sample size is always present, as in almost all areas of nonparametric estimation.

One approach to selecting h is through a critical inspection of intensity plots in order to balance local and global features in the data (Møller & Waagepetersen, 2004). Other approaches involve optimization criteria. Baddeley et al. (2015, p. 176) suggest selecting the h that maximises the cross-validated Poisson log likelihood, which in the present setting is

$$\ell_{\text{cv}}(h|X) = \sum_{x \in X} \log\left\{\hat{\rho}_h^{-x}(x)\right\} - \int_{\mathcal{M}} \hat{\rho}_h(x)\, d\text{vol}(x), \tag{9}$$

where $\hat{\rho}_h^{-z}(x) = h^{-d} \sum_{y \in X \setminus \{z\}} k\{-d_g(x,y)/h\} c_h^{-1}(x,y)$ is an estimate of ρ constructed as in (4) but without the observation $z \in X$. Application of Campbell's Theorem shows ℓ_{cv} is unbiased for the log likelihood function

$$\ell(\rho; X) = \sum_{x \in X} \log\{\rho(x)\} - \int_{\mathcal{M}} \rho(x)\, d\text{vol}(x).$$

A nonparametric bandwidth selection procedure that can be readily extended to the Riemannian setting is given in Cronie & Van Lieshout (2018). On assuming that the intensity function is positive everywhere on $\mathcal{M}$ and applying Campbell's formula (2) to ρ^{-1},

$$E\Big\{\sum_{x \in X} \rho^{-1}(x)\Big\} = \text{Vol}(W).$$

Replacement of ρ by its estimate $\hat{\rho}_h$ points to a choice of h that minimizes

$$F(h) = \{T(\hat{\rho}_h) - \text{Vol}(W)\}^2\,, \tag{10}$$

where $T(\hat{\rho}_h) = \sum_{x \in X} \hat{\rho}_h^{-1}(x)$. In addition to being relatively free of modelling assumptions, (10) is less burdensome to compute than (9). The existence of a minima of F can be shown by consideration of its continuity and limiting properties. Proposition 4 extends Theorem 1 of Cronie & Van Lieshout (2018) to $\mathcal{M}$.

PROPOSITION 4. *Let $(\mathcal{M}, g)$ be a Riemannian manifold and let X be a point process observed through a bounded window $W \subseteq \mathcal{M}$. After disregarding the trivial case $X \cap W = \emptyset$, global and local corrections (5) or (6) both yield T continuous in $h \in (0, \infty)$. This conclusion also holds when no correction is used, i.e. $c_h^{\bullet}(x,y) = 1$. In all cases, $\lim_{h\to 0} T(\hat{\rho}_h) = 0$. For correction given by (5) and (6) $\lim_{h\to\infty} T(\hat{\rho}_h) = Vol(W)$ and if no correction is used $\lim_{h\to\infty} T(\hat{\rho}_h) = \infty$.*

The intermediate value theorem dictates that when $c_h(x,y) = 1$ there exists a minima for F, whilst if a correction is used a minimum occurs when $h \to \infty$. This is consistent with the Euclidean approach considered by Cronie & Van Lieshout (2018). The recommendation of the present paper is also to optimize F with no correction, including it instead after h has been selected.

5. Numerical confirmation in a test case

Numerical validity of the proposed procedure is checked empirically using an example in which standard methods are approximately valid in the limit as a key parameter becomes small, but more generally handled by the approach developed in Section 3.

We consider a Poisson process on the unit square with intensity function

$$\rho(x_1, x_2) = \frac{N}{2\pi\sigma^2 K} \exp\left\{-\frac{(x_1 - \frac{1}{2})^2 + (x_2 - \frac{1}{2})^2}{2\sigma^2}\right\},$$

where $N > 0$ and K is chosen to ensure the expected number of points in $[0,1]^2$ is N. With the unit square considered as the unit subset of the plane $x_3 = 0$ in $\mathbb{R}^3$, a Poisson process on a bounded Euclidean manifold, $\mathcal{M} = W$ say, is obtained by rotating the plane $x_3 = 0$ through an angle of θ about the x_2-axis, giving the intensity function

$$\rho_W(x_1, x_2, x_3) = \begin{cases} \rho\{(x_1^2 + x_3^2)^{1/2}, x_2\} & x_3 = x_1 \tan(\theta); \\ 0 & \text{otherwise.} \end{cases}$$

The first approach, as described in Sections 3 and 4, involves direct application of the recommended procedure on W to estimate ρ_W using a local correction. The second, for comparison, follows a cross-validation approach previously considered in Baddeley et al. (2015), whereby a standard bivariate Euclidean kernel intensity estimator is first applied to the orthogonal projection on the plane $x_3 = 0$, with the fitted intensity $\hat{\rho}_{\text{proj}}$ then mapped back onto W as

$$\hat{\rho}_W(x_1, x_2, x_3) = \begin{cases} \hat{\rho}_{\text{proj}}(x_1, x_2)\{1 + \tan^2(\theta)\}^{-1/2} & x_3 = x_1 \tan(\theta); \\ 0 & \text{otherwise.} \end{cases}$$

The two situations are depicted in Fig. 1(a).

The sample mean integrated squared error (MISE) of each approach is computed using 10,000 simulated replicates of the point pattern with parameters $\sigma^2 = 0.01$ and $N = 500$ for a range of values of θ. Fig. 1(b) shows that the MISE remains constant for increasing θ when estimation is performed directly on the manifold, whereas the projection approach differs considerably for large θ, but coincides, as expected, for small θ.

6. Simulations

Point patterns are simulated on the surface of three ellipsoids of increasing eccentricity: manifolds $\mathcal{E}_1$, $\mathcal{E}_2$ and $\mathcal{E}_3$, respectively. A common local chart used to describe an ellipsoid $\mathcal{E}$ is $x \equiv (x_1, x_2, x_3) = \{a\sin(\theta)\cos(\phi), b\sin(\theta)\cos(\phi), c\cos(\theta)\}$ where $\theta \in [0, \pi)$ and $\phi \in [0, 2\pi)$. Manifold $\mathcal{E}_1$ is a sphere of radius $a = b = c = (4\pi)^{-1/2}$, $\mathcal{E}_2$ has $a = b = 0.8(4\pi)^{-1/2}$, and $\mathcal{E}_3$ has $a = b = 0.6(4\pi)^{-1/2}$. To enable comparison, the value of c for $\mathcal{E}_2$ and $\mathcal{E}_3$ is set to ensure that they each have unit Riemannian volume measure (surface area).

The intensity function is estimated using point patterns sampled from three Poisson process models. Details and results for log Gaussian Cox processes and Strauss processes are presented in Supplementary Materials Section 3, alongside a detailed explanation of how the processes were simulated in Supplementary Materials Section 4. The three Poisson process models considered are: (PP1) homogeneous Poisson process, i.e. with intensity function $\rho_1(x) = \rho_1$; (PP2) inhomogeneous Poisson process with log-linear intensity function $\rho_2(x) =$

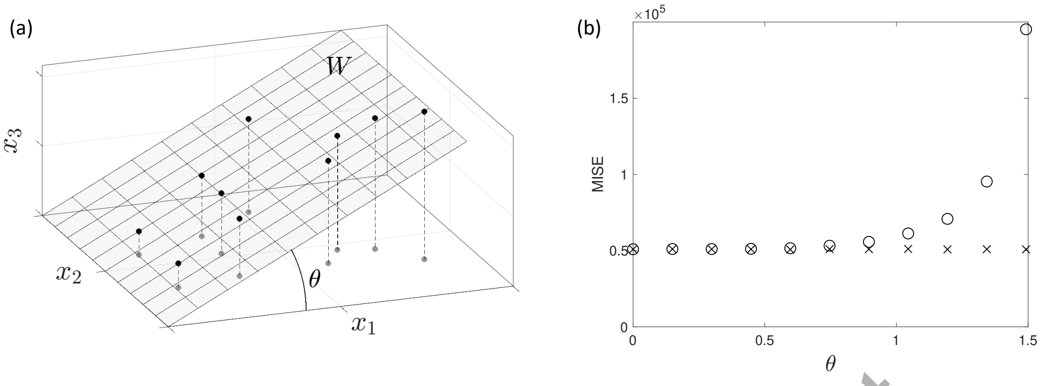

Fig. 1. (a) Example point pattern on W (black) with its orthogonal projection on the plane $x_3 = 0$ (grey). (b) Sample mean integrated squared error (MISE) of the intensity estimate across 10,000 simulations for the projected point pattern ($\circ$) and working directly on the manifold ($\times$)

$\exp(3 + \alpha_2 x_1)$; (PP3) inhomogeneous Poisson process with log-modulation intensity function $\rho_3(x) = \exp\{2 + \alpha_3 \cos(8x_2)\}$. Parameters ρ_1, α_2 and α_3 each take three values to give a low, medium and high number of expected events.

The results from this 3^3 factorial experiment are presented in Table 1, where the last two columns display the integrated squared error of the estimate, $\|\hat{\rho} - \rho\|^2$, averaged over Monte Carlo replications and standardized by the square of the expected number of events $\|\rho\|^2$ to make the rows comparable. The norm is the $L_2(\mathcal{M})$ norm, i.e. with $\mathcal{M} \in \{\mathcal{E}_1, \mathcal{E}_2, \mathcal{E}_3\}$,

$$\begin{aligned}\|\hat{\rho} - \rho\|^2 &= \int_{\mathcal{M}} \{\hat{\rho}(x) - \rho(x)\}^2 d\mathrm{vol}(x) \\ &= \int_0^{\pi} \int_0^{2\pi} \left[\hat{\rho}\left\{\psi^{-1}(\theta, \phi)\right\} - \rho\left\{\psi^{-1}(\theta, \phi)\right\}\right]^2 \{\det(g_{ij})\}^{1/2}\, d\theta d\phi, \qquad (11)\end{aligned}$$

where ψ is the local chart for $\mathcal{M}$ and where (g_{ij}) is the matrix representation of the metric under the corresponding local coordinate system. For the chosen chart it can be shown that

$$\det(g_{ij}) = \sin^2(\theta) a^2 b^2 \{1 - (1 - c^2/a^2)\sin^2(\theta)\cos^2(\phi) - (1 - c^2/b^2)\sin^2(\theta)\sin^2(\phi)\}.$$

The bandwidth is selected using the two methods outlined in Section 4, referred to here as *cross validation* (CV) (Baddeley et al., 2015) and *nonparametric* (NP) (Cronie & Van Lieshout, 2018). Intensity function estimates are then computed using the local correction. The integral in (11) is computed using a numerical approximation.

In the Poisson setting outlined here, the CV method for bandwidth selection outperforms the NP method, while inspection of the results for log Gaussian Cox and Strauss processes shows the opposite is true. This is unsurprising since the CV method is based on a Poisson likelihood. The simulation results are consistent with the Euclidean analysis considered in Cronie & Van Lieshout (2018).

7. APPLICATION TO THE BEILSCHMIEDIA PENDULA DATASET

In applying the proposed estimator to the Beilschmiedia Pendula data set (Hubbell, 1983; Condit et al., 1996; Condit, 1998), a number of important practical considerations are isolated.

Table 1. *Performance of kernel intensity estimators*

Manifold	Poisson model	Process parameters	Expected number of events	$\{\hat{E}(\lVert\hat{\rho}-\rho\rVert^2/\lVert\rho\rVert^2)\}^{1/2}$ CV	NP
$\mathcal{E}_1$	PP1	$\rho_1 = 50$	50.00	0.285	0.307
$\mathcal{E}_1$	PP1	$\rho_1 = 150$	150.0	0.182	0.220
$\mathcal{E}_1$	PP1	$\rho_1 = 300$	300.0	0.128	0.185
$\mathcal{E}_1$	PP2	$\alpha_2 = 10$	59.57	0.306	0.598
$\mathcal{E}_1$	PP2	$\alpha_2 = 18$	317.2	0.179	0.765
$\mathcal{E}_1$	PP2	$\alpha_2 = 22$	802.2	0.141	0.806
$\mathcal{E}_1$	PP3	$\alpha_3 = 3$	49.98	0.462	0.624
$\mathcal{E}_1$	PP3	$\alpha_3 = 4$	116.1	0.378	0.671
$\mathcal{E}_1$	PP3	$\alpha_3 = 5$	280.2	0.305	0.706
$\mathcal{E}_2$	PP1	$\rho_1 = 50$	50.00	0.285	0.303
$\mathcal{E}_2$	PP1	$\rho_1 = 150$	150.0	0.178	0.224
$\mathcal{E}_2$	PP1	$\rho_1 = 300$	300.0	0.137	0.186
$\mathcal{E}_2$	PP2	$\alpha_2 = 10$	43.60	0.343	0.538
$\mathcal{E}_2$	PP2	$\alpha_2 = 18$	153.8	0.236	0.717
$\mathcal{E}_2$	PP2	$\alpha_2 = 22$	313.7	0.183	0.763
$\mathcal{E}_2$	PP3	$\alpha_3 = 3$	58.60	0.449	0.562
$\mathcal{E}_2$	PP3	$\alpha_3 = 4$	135.9	0.380	0.617
$\mathcal{E}_2$	PP3	$\alpha_3 = 5$	326.7	0.303	0.659
$\mathcal{E}_3$	PP1	$\rho_1 = 50$	50.00	0.292	0.297
$\mathcal{E}_3$	PP1	$\rho_1 = 150$	150.0	0.197	0.229
$\mathcal{E}_3$	PP1	$\rho_1 = 300$	300.0	0.147	0.187
$\mathcal{E}_3$	PP2	$\alpha_2 = 10$	32.34	0.432	0.511
$\mathcal{E}_3$	PP2	$\alpha_2 = 18$	75.05	0.313	0.662
$\mathcal{E}_3$	PP2	$\alpha_2 = 22$	123.1	0.279	0.718
$\mathcal{E}_3$	PP3	$\alpha_3 = 3$	74.88	0.447	0.478
$\mathcal{E}_3$	PP3	$\alpha_3 = 4$	175.7	0.368	0.514
$\mathcal{E}_3$	PP3	$\alpha_3 = 5$	423.1	0.290	0.562

Square root of the mean integrated squared error from a 3^3 factorial experiment. The mean is taken over 100 Monte Carlo replicates.

These data provide locations of 3605 trees in a 1000m by 500m rectangular sampling region of a tropical rainforest. Each event is characterized by its longitude, latitude and elevation and is assumed to lie on a two-dimensional Riemannian manifold within $\mathbb{R}^3$, inheriting the canonical metric tensor by the embedding. As with most practical settings, there is no analytical formula for the manifold. Instead, it was approximated using a 201×101 regular grid of longitude and latitude coordinates each with an elevation. Let U be the union of the recorded events and the manifold grid points. The manifold was approximated as a triangular mesh $\mathcal{M} = (U, T)$, where $T = \{F_m; m = 1, \ldots, M\}$ is the set of triangular faces of the mesh with each element represented by the three elements of U that form its vertices. This was constructed by implementing a 2D Delaunay triangular mesh on the longitude and latitude, with each vertex then raised by its respective elevation. The full mesh and an illustrative subsection are shown in Fig. 3(a) and Fig. 3(b), respectively. Including the events as vertices of the triangular mesh aids the computation of event-to-point geodesic distances, here conducted with the Fast Marching Algorithm (Peyre, 2021).

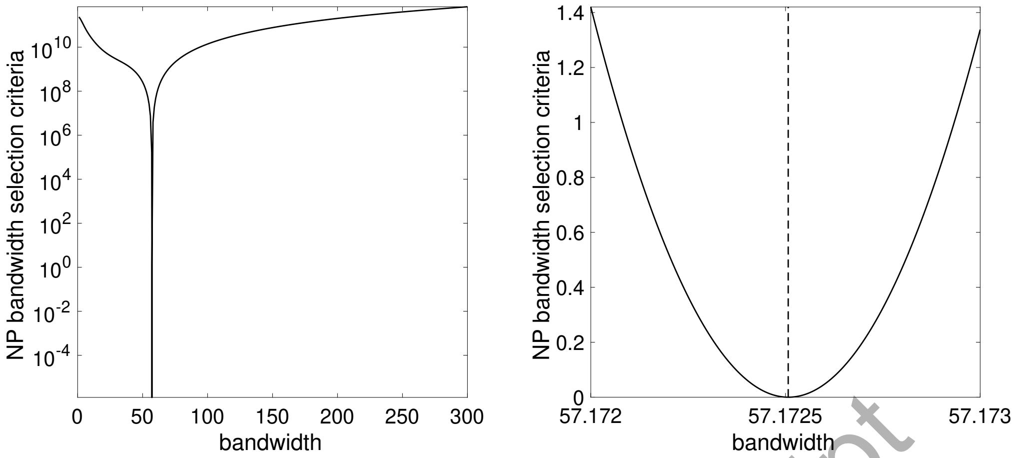

Fig. 2. (a) Nonparametric (NP) bandwidth selection criteria $F(h)$ for the Beilschmiedia Pendula Dataset, plotted with a log scale y-axis. (b) $F(h)$ in the locality of its minimum, plotted with a linear scale y-axis.

The kernel intensity estimator of (4) was constructed using the local correction with (6) becoming

$$c_h(y) = \frac{1}{h^d} \sum_{m=1}^{M} \int_{F_m} k\left\{\frac{d_g(z,y)}{h}\right\} d\lambda_m(z)$$

on replacement of $d\mathrm{vol}(\cdot)$ by $d\lambda_m(\cdot)$ for the surface area element over the mth face of the mesh. This is approximated by

$$\frac{1}{h^d} \sum_{m=1}^{M} k\left\{\frac{d_g(z_m,y)}{h}\right\} \lambda_m(F_m),$$

where z_m is a representative point of F_m, here computed as the arithmetic average of its three vertices. At this junction, a second triangular mesh was constructed in an identical manner to the first but including $\{z_1,\ldots,z_M\}$ as additional vertices such that all required geodesic distances could be computed with the Fast Marching Algorithm.

To avoid modelling the data generating process, only the nonparametric (NP) approach to bandwidth selection was used. If the CV method was to be applied, the integral in (9) would instead be approximated as

$$\int_{\mathcal{M}} \hat{\rho}_h(x) d\mathrm{vol}(x) \approx \sum_{m=1}^{M} \hat{\rho}_h(z_m)\lambda_m(F_m).$$

As recommended, the bandwidth was selected without correction, which was only applied subsequently in the construction of the kernel intensity estimate. The selection criterion was evaluated at bandwidths $h \in \{1,2,\ldots,300\}$ in the units of metres. Additional refinement around the minimizing value of h gave a final bandwidth choice of 57.17 m (2 d.p.). Fig. 2 shows a well behaved convex function with a pronounced minimum. The resulting NP intensity estimate is shown in Fig. 3(c). Fig. 3(d) displays the relative difference between this and a simple alternative. The latter, written $\hat{\rho}_{\text{flat}}$ constructs the intensity estimate on the plane and projects it onto the landscape using local gradients, as in (Baddeley et al., 2015, p. 176). The elevation scale has

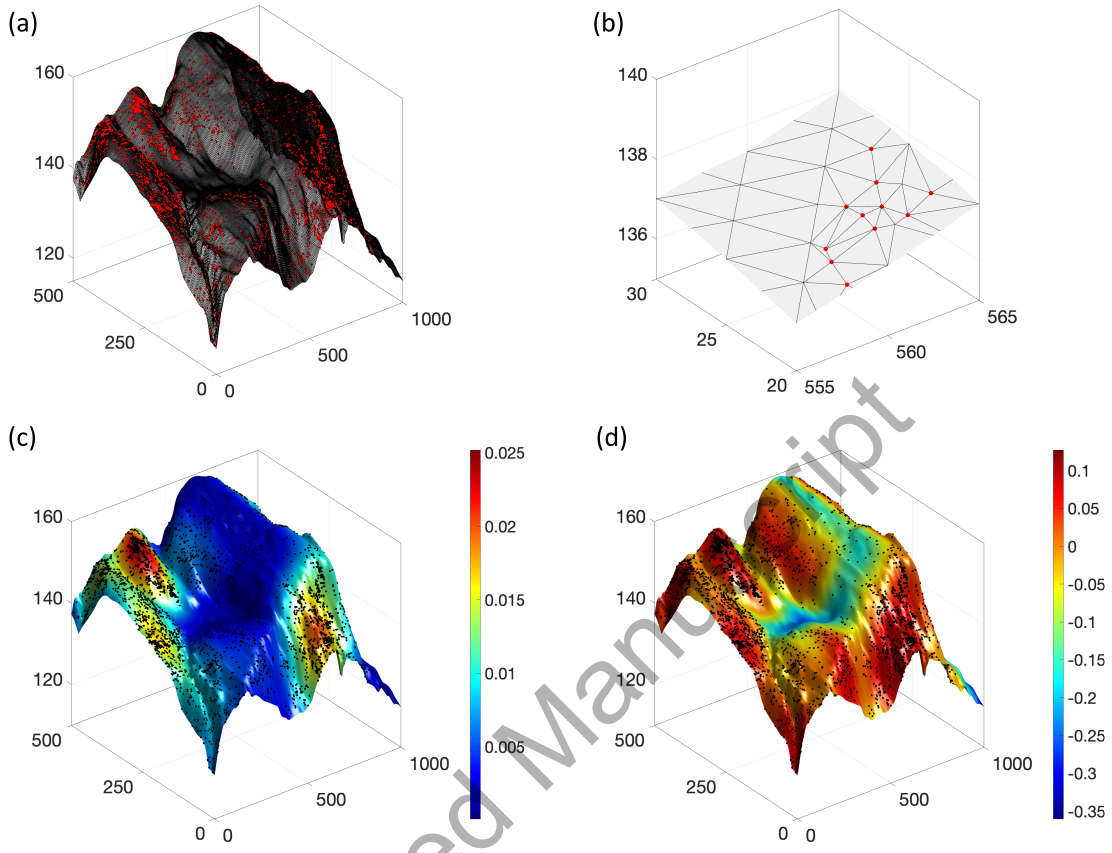

Fig. 3. (a) Beilschmiedia Pendula events (red dots) on the triangular mesh representation of the geographical surface. (b) Small portion of (a), depicting the events (red dots) and constructed triangular mesh. (c) NP intensity estimate in units m^{-1} depicted using the presented colormap, shown with the events (black dots). (d) Relative difference between the estimated intensity directly on the manifold and estimated intensity assuming a flat surface, $\{\hat{\rho}_{\text{manifold}}(x)/\hat{\rho}_{\text{flat}}(x)\} - 1$.

been magnified to aid visualization. For a manifold without a boundary, such as a sphere or an ellipsoid, it is unclear how $\hat{\rho}_{\text{flat}}$ could be implemented.

8. Discussion and open problems

The constructions presented in the present work have first-order asymptotic guarantees for the estimation of intensity functions of Poisson processes observed over general Riemannian manifolds, with or without boundaries. Their properties under other generative point processes have been assessed by simulation. As discussed in Section 3, intuitive reasoning suggests that the proposed estimator is not optimal in finite samples. A finite-sample theoretical analysis seems challenging and may involve extension of the classical probability inequalities.

ACKNOWLEDGEMENTS

We are grateful to Niall Adams for joint supervision of S.W.'s Ph.D. at Imperial College London, and Karthik Bharath for a helpful discussion. Financial support was provided by the Wellcome Trust (to S.W.) and the UK Engineering and Physical Sciences Research Council (to H.S.B.).

REFERENCES

BADDELEY, A., RUBAK, E. & TURNER, R. (2015). *Spatial Point Patterns: Methodology and Applications with R*. CRC press.

BARTLETT, M. S. (1963). The spectral analysis of point processes. *J. Roy. Statist. Soc. Ser. B* **25**, 264–296.

BERMAN, M. & DIGGLE, P. (1989). Estimating weighted integrals of the second-order intensity of a spatial point process. *J. Roy. Statist. Soc. Ser. B* **51**, 81–92.

CONDIT, R. (1998). *Tropical Forest Census Plots*. Springer Berlin, Heidelberg.

CONDIT, R., HUBBELL, S. P. & FOSTER, R. B. (1996). Changes in tree species abundance in a neotropical forest: impact of climate change. *J. Trop. Ecol* , 231–256.

COX, D. R. (1965). On the estimation of the intensity function of a stationary point process. *J. Roy. Statist. Soc. Ser. B* **27**, 332–337.

CRESSIE, N. (2015). *Statistics for Spatial Data*. John Wiley & Sons.

CRONIE, O. & VAN LIESHOUT, M. N. (2018). A non-model-based approach to bandwidth selection for kernel estimators of spatial intensity functions. *Biometrika* **105**, 455–462.

CUCALA, L. (2008). Intensity estimation for spatial point processes observed with noise. *Scand. J. Stat.* **35**, 322–334.

DALEY, D. & VERE-JONES, D. (2010). *An Introduction to the Theory of Point Processes*, vol. I. Springer New York, 2nd ed.

DIGGLE, P. (1985). A kernel method for smoothing point process data. *J. Roy. Statist. Soc. Ser. C* **34**, 138–147.

HUBBELL, S. P. (1983). Diversity of canopy trees in a neotropical forest and implications for conservation. In *Tropical Rain Forest: Ecology and Management*, T. Whitmore, A. Chadwick & A. Sutton, eds. The British Ecological Society, pp. 25–41.

KERKYACHARIAN, G., NICKL, R. & PICARD, D. (2012). Concentration inequalities and confidence bands for needlet density estimators on compact homogeneous manifolds. *Probab. Theory Relat. Fields* **153**, 363–404.

KINGMAN, J. F. C. (1993). *Poisson Processes*. Oxford University Press.

LAWRENCE, T., BADDELEY, A., MILNE, R. K. & NAIR, G. (2016). Point pattern analysis on a region of a sphere. *Stat* **5**, 144–157.

MØLLER, J. & RUBAK, E. (2016). Functional summary statistics for point processes on the sphere with an application to determinantal point processes. *Spat. Stat.* **18**, 4–23.

MØLLER, J. & WAAGEPETERSEN, R. P. (2004). *Statistical Inference and Simulation for Spatial Point Processes*. Florida: CRC Press.

PELLETIER, B. (2005). Kernel density estimation on Riemannian manifolds. *Statist. Prob. Lett.* **73**, 297–304.

PEYRE, G. (2021). Toolbox Fast Marching. `https://www.mathworks.com/matlabcentral/fileexchange/6110-toolbox-fast-marching`, [Accessed: 13 September 2022].

RAKSHIT, S., DAVIES, T., MORADI, M. M., MCSWIGGAN, G., NAIR, G., MATEU, J. & BADDELEY, A. (2019). Fast kernel smoothing of point patterns on a large network using two-dimensional convolution. *Int. Stat. Rev.* **87**, 531–556.

ROBESON, S. M., LI, A. & HUANG, C. (2014). Point-pattern analysis on the sphere. *Spat. Stat.* **10**, 76–86.

VAN LIESHOUT, M.-C. N. (2012). On estimation of the intensity function of a point process. *Methodol. Comput. Appl. Probab.* **14**, 567–578.

WARD, S., COHEN, E. A. K. & ADAMS, N. (2021). Testing for complete spatial randomness on three dimensional bounded convex shapes. *Spat. Stat.* **41**.

[*Received on 2 January 2017. Editorial decision on 1 April 2017*]