SLICED INVERSE REGRESSION AND INDEPENDENCE IN RANDOM MARKED SETS WITH COVARIATES

ONDŘEJ ŠEDIVÝ * ** AND

JAKUB STANĚK,*** *Charles University in Prague*

BLAŽENA KRATOCHVÍLOVÁ,**** *Palacký University Olomouc*

VIKTOR BENEŠ,* ***** *Charles University in Prague*

Abstract

Dimension reduction of multivariate data was developed by Y. Guan for point processes with Gaussian random fields as covariates. The generalization to fibre and surface processes is straightforward. In inverse regression methods, we suggest slicing based on geometrical marks. An investigation of the properties of this method is presented in simulation studies of random marked sets. In a refined model for dimension reduction, the second-order central subspace is analyzed in detail. A real data pattern is tested for independence of a covariate.

Keywords: Dimension reduction; central subspace; random marked set; Gaussian random field; covariate

2010 Mathematics Subject Classification: Primary 60D05
Secondary 62M30

1. Introduction

In this paper we deal with spatial point, fibre, and surface processes (see Stoyan *et al.* (1995), Beneš and Rataj (2004), and Schneider and Weil (2008)) and multivariate Gaussian random fields (GRFs) as covariates. The aim is to study the dimension reduction problem and the dependencies between these objects. Dimension reduction means finding a matrix B such that, for a random vector X of covariates $B^\top X$ is of lower dimension and it involves all information about the spatial process Y hidden in X.

A generalization of the sufficient dimension reduction paradigm for inhomogeneous spatial point processes developed in Guan and Wang (2010) is considered. Among inverse regression techniques, e.g. SAVE (see Cook and Weisberg (1991)) and the directional regression (see Li and Wang (2007)), we concentrate on the basic method called the sliced inverse regression (SIR; see Li (1991)). Guan (2008) claimed SIR to be hardly applicable in point processes because of nonexistence of natural slicing. We show that slicing can be realized in spatial processes, based on suitable geometrical marks. Thus, the concept of a random marked set from Ballani *et*

al. (2012) (an early reference is Molchanov (1983)) is used for spatial processes. Basic theorems on the structure of dimension reduction subspaces are derived for SIR. Guan and Wang (2010) refined the analysis defining the kth-order central intensity subspace and they studied the case $k = 1$. In the present paper, attention is paid to the case $k = 2$. A general result on the estimation of the second-order central intensity subspace is derived. Two special models of point processes are investigated to show the applicability of the method. An approach to slicing in this situation is suggested.

Furthermore, a model of a fibre process in a bounded planar domain is defined as a curve being a numerical solution of a stochastic differential equation (SDE), with coefficients conditioned by a GRF.

The statistical part of the paper involves the estimation of the directions in the central subspace, testing the hypotheses about its dimension and about the independence between the random set and covariates (cf. Schlather *et al.* (2004) and Illian *et al.* (2008) for point processes). The quality of estimators is quantified, and the power of the tests can be evaluated in repeated simulations.

Simulation studies demonstrate the methods for different models from stochastic geometry. In the Poisson–Voronoi tessellation, the fibre or surface process of its edges or faces, respectively, is considered, marked by the length of the edges or surface area of the faces. Finally, a test of independence based on the weighted L-function is applied to real data with a single covariate.

2. Theoretical background

2.1. Stochastic geometry

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space. Let $\mathcal{M}^d$ be the set of all locally finite measures on $(\mathbb{R}^d, \mathcal{B}^d)$ equipped with the smallest σ-algebra $\mathfrak{M}^d$ which makes the maps $\mu \mapsto \mu(A)$, $\mu \in \mathfrak{M}^d$, measurable for all $A \in \mathcal{B}^d$. Then $\Psi : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathcal{M}^d, \mathfrak{M}^d)$ is a random measure on $\mathbb{R}^d$ and $\Lambda(\cdot) = \mathbb{E}\Psi(\cdot)$ is the intensity measure of Ψ.

Let $\mathcal{H}^k$ be the Hausdorff measure of order k in $\mathbb{R}^d$. Zähle (1982) introduced the concept of random $\mathcal{H}^k$-sets in $\mathbb{R}^d$ as random closed sets which are $\mathcal{H}^k$-rectifiable. A random $\mathcal{H}^k$-set Y such that $\Psi_Y(\cdot) = \mathcal{H}^k(Y \cap \cdot)$ is a locally finite measure in $\mathbb{R}^d$ will be called a point, fibre, or surface process for $k = 0, 1, d - 1$, respectively. By $C \uparrow \mathbb{R}^d$ we mean that there exists a nondecreasing sequence of convex and compact sets C_n such that their union is $\mathbb{R}^d$. Ergodic theorems for Ψ_Y follow from the results in Nguyen and Zessin (1979); central limit theorems in the stationary case were considered in Pawlas (2003).

The lth moment measure of the random measure Ψ_Y is

$$\mu^{(l)}(A_1, \ldots, A_l) = \mathbb{E}[\Psi_Y(A_1) \cdots \Psi_Y(A_l)]$$

for bounded Borel sets $A_1, \ldots, A_l \subset \mathbb{R}^d$. The lth-order intensity function λ_l is defined by

$$\mu^{(l)}(\mathrm{d}s_1 \times \cdots \times \mathrm{d}s_l) = \lambda_l(s_1, \ldots, s_l) \, \mathrm{d}s_1 \cdots \mathrm{d}s_l \tag{1}$$

if it exists. In the case $k = 0$ (point processes), on the left-hand side of (1) there is the lth factorial moment measure $\mu^{(l)!}$ (see Stoyan *et al.* (1995, p. 111)) instead of $\mu^{(l)}$, which does not have a density with respect to the Lebesgue measure. The notation $\lambda = \lambda_1$ will be used.

Ballani *et al.* (2012) defined a random marked closed set as a pair (Y, Γ), which is a random element in a measurable space of hypographs of random upper-semicontinuous functions Γ (a mark) defined on a random closed set $Y \in \mathbb{R}^d$. Thus, for a random marked $\mathcal{H}^k$-set (Y, Γ), we

have the notion of a marked point, fibre, or surface process. Let the range of the real-valued mark Γ be divided into m disjoint intervals $J_1, \ldots, J_m$ called slices. This induces a partition of Y into m disjoint subsets $(Y^1, \ldots, Y^m)$. Alternatively, we may consider closures $\bar{Y}^i$ as processes with first-order intensities $\lambda^{(i)}$,

$$\int_A \lambda^{(i)}(s) \, \mathrm{d}s = \mathbb{E}\Psi_{\bar{Y}^i}(A), \qquad A \in \mathcal{B}^d, \, i = 1, \ldots, m,$$

respectively, corresponding to each slice, so that $\lambda = \sum_{i=1}^m \lambda^{(i)}$.

2.2. Dimension reduction

The dimension reduction problem in statistics concerns the situation where we have a response Y (random variable) dependent on a p-dimensional random column vector of covariates X and the aim is to reduce the number of covariates in order to use only the most significant ones. For vector data, Li (1991) suggested using the SIR method. The idea to regress X on Y inversely instead of the direct regression of Y on X stems from the fact that in such a way we obtain p univariate regressions instead of a multivariate one. Further methods of dimension reduction are described in Cook (1998), Li and Wang (2007), etc.

The dimension reduction problem for point processes was formulated in Guan (2008) and Guan and Wang (2010), from where Definitions 1 and 2 below were taken. In our setting, Y is a random point, fibre, or surface process in $\mathbb{R}^d$ and X is a p-dimensional stationary Gaussian random field in $\mathbb{R}^d$, $d = 2, 3$. Without loss of generality, assume that X is standardized, i.e. $\mathbb{E}X(s) = 0$ and $\mathrm{cov}\, X(s) = I_p$ for each s, where I_p is the unit matrix of order p. Let $A^\top$ be the transpose of a real-valued matrix A, and let $\mathcal{S}(A)$ be the linear subspace spanned by columns of A. All vectors are column vectors.

Definition 1. Let Y be conditionally independent of X given $B^\top X$ for a $p \times c$ matrix B, $c \leq p$. Then $\mathcal{S}(B)$ is called a sufficient dimension reduction subspace. Let $\mathcal{S}_{Y|X}$ be the intersection of all such subspaces. Assume that it is also a sufficient dimension reduction subspace. Then it is called the central subspace.

The definition says that $B^\top X = \{B^\top X(s), \, s \in \mathbb{R}^d\}$ contains all information of X about Y. Let c be the dimension of the central subspace.

A refined analysis of the dimension reduction is based on the following definition. Assume that the lth-order intensity functions λ_l, cf. (1), exist for all $l \geq 1$ (being dependent on X, they are random functions here).

Definition 2. Consider $l \in \mathbb{N}$, and assume that the relation

$$\lambda_l(s_1, \ldots, s_l) = f_l(B^\top X(s_1), \ldots, B^\top X(s_l)) \tag{2}$$

holds for some measurable function f_l and a $p \times c$ matrix B, $c \leq p$. Then $\mathcal{S}_l(B)$ is called the lth-order sufficient intensity dimension reduction subspace and the intersection of all such subspaces $\mathcal{S}_l = \cap \mathcal{S}_l(B)$ is called the lth-order central intensity subspace (if it is also an lth-order sufficient intensity dimension reduction subspace).

The central subspace defined above can be expressed as

$$\mathcal{S}_{Y|X} = \bigcup_{l \geq 1} \mathcal{S}_l;$$

cf. Guan and Wang (2010). The aim is to investigate the structure of the kth-order sufficient intensity dimension reduction subspaces.

We concentrate mostly on the SIR method, which can be described in the following steps: (i) slicing the random set Y according to a suitable mark, (ii) finding the slice means of the random field X, (iii) applying the principal components analysis of slice means.

The first c directions from (iii) generate the central subspace of the corresponding order. Note that Y plays a role in step (i) only.

3. Results

We start with a lemma needed for the refined analysis of the dimension reduction.

Lemma 1. *Let $C \in \mathbb{R}^d$ be a compact convex set.*

(a) *Assume that (2) holds for $l = 1$. Then*

$$\int_C \mathbb{E}[X(s)\lambda(s)] \, \mathrm{d}s = \mathbb{E}\left[\int_{s \in Y \cap C} X(s) \mathcal{H}^k(\mathrm{d}s)\right].$$

(b) *Assume that (2) holds for $l = 2$. Then*

$$\int_C \int_C \mathbb{E}[X(s)X(t)^\top \lambda_2(s, t)] \, \mathrm{d}s \, \mathrm{d}t = \mathbb{E}\left[\int_{s,t \in Y \cap C} X(s)X(t)^\top \mathcal{H}^k(\mathrm{d}s) \mathcal{H}^k(\mathrm{d}t)\right].$$

Proof. Assertions (a) and (b) follow from the first- and second-order Campbell theorem, respectively, e.g. for (a),

$$\mathbb{E}\left[\int_{s \in Y \cap C} X(s) \mathcal{H}^k(\mathrm{d}s)\right] = \iint_{s \in Y \cap C} x(s) \mathcal{H}^k(\mathrm{d}s) P(\mathrm{d}\psi)$$

$$= \iiint_{s \in Y \cap C} x(s) \mathcal{H}^k(\mathrm{d}s) P_\Lambda(\mathrm{d}\psi) Q(\mathrm{d}\Lambda)$$

$$= \iint_C \int x(s) \Lambda(\mathrm{d}(x, s)) Q(\mathrm{d}\Lambda)$$

$$= \int_C \iint x(s) \lambda(s) M_s(\mathrm{d}x) Q(\mathrm{d}\Lambda) \, \mathrm{d}s$$

$$= \int_C \mathbb{E}[X(s)\lambda(s)] \, \mathrm{d}s,$$

where P is the distribution of the marked process $(Y, X(s))$ with random intensity measure

$$\Lambda(\mathrm{d}(x, s)) = \lambda(s) M_s(\mathrm{d}x) \, \mathrm{d}s = f(B^\top x(s)) M_s(\mathrm{d}x) \, \mathrm{d}s,$$

M_s is the distribution of the mark $X(s)$ at a point s, Q is the distribution of Λ, and P_Λ is the conditional distribution of the marked process given Λ.

For $k = 0$ and a point process Y, we have

$$\mathbb{E} \sum_{s \in Y \cap C} X(s) \quad \text{and} \quad \mathbb{E} \sum_{s,t \in Y \cap C}^{\neq} X(s)X(t)^\top$$

on the right-hand sides of the formulae of Lemma 1(a) and (b), respectively. In the second sum s and t are always different.

3.1. Investigation of the first-order central intensity subspace

In this subsection we assume that, for each $s \in \mathbb{R}^d$, it holds that

$$\lambda(s) = f(B^\top X(s)) \tag{3}$$

for a matrix B of size $p \times c$, $c \leq p$.

Lemma 2. *Let the mark $\Gamma(s) = g(B^\top X(s))$ for a measurable function g. Then*

$$\lambda^{(j)}(s) = f^j(B^\top X(s))$$

for some nonnegative measurable functions f^j, $j = 1, \ldots, m$.

Proof. For $|ds| \downarrow 0$, we have

$$\begin{aligned}
\lambda^{(j)}(s)|ds| &= \mathbb{E}[Y^j(ds) \mid X(s)] \\
&= \mathbb{E}[\mathcal{H}^k(\{y \in Y \cap ds; \Gamma(y) \in J_j\}) \mid X(s)] \\
&= \mathbb{E}[\mathcal{H}^k(\{y \in Y \cap ds; \Gamma(y) \in J_j\}) \mid B^\top X(s)] \\
&= f^j(B^\top X(s))|ds|.
\end{aligned}$$

This completes the proof.

Consider a convex compact window $C \subset \mathbb{R}^d$ and a statistic

$$\hat{V}_1 = \frac{1}{\Psi_Y(C)} \sum_{j=1}^{m} \frac{1}{\Psi_{Y^j}(C)} \int_{Y^j \cap C} X(s) \mathcal{H}^k(ds) \left[\int_{Y^j \cap C} X(s) \mathcal{H}^k(ds) \right]^\top. \tag{4}$$

Assume that Y and $\{X(s), s \in Y\}$ are ergodic. Then, from Lemma 1 for each j, it holds that

$$\frac{1}{\Psi_{Y^j}(C)} \int_{Y^j \cap C} X(s) \mathcal{H}^k(ds) \to \frac{\int_{\mathbb{R}^d} \mathbb{E}[\lambda^{(j)}(s)X(s)]\,ds}{\int_{\mathbb{R}^d} \mathbb{E}[\lambda^{(j)}(s)]\,ds} \tag{5}$$

in probability when $C \uparrow \mathbb{R}^d$. This limit is defined as a ratio of limits:

$$\lim_{C \uparrow \mathbb{R}^d} \frac{1}{|C|} \int_C \mathbb{E}[\lambda^{(j)}(s)X(s)]\,ds, \qquad \lim_{C \uparrow \mathbb{R}^d} \frac{1}{|C|} \int_C \mathbb{E}[\lambda^{(j)}(s)]\,ds.$$

Their finiteness can be verified easily, e.g. in our case when $X(\cdot)$ is a stationary random field, since then limit (5) is equal to

$$\frac{\mathbb{E}[\lambda^{(j)}(\cdot)X(\cdot)]}{\mathbb{E}[\lambda^{(j)}(\cdot)]}. \tag{6}$$

Let the theoretical counterpart of (4) be

$$V_1 = \frac{1}{\int_{\mathbb{R}^d} \mathbb{E}[\lambda(s)]\,ds} \sum_{j=1}^{m} \frac{\int_{\mathbb{R}^d} \mathbb{E}[\lambda^{(j)}(s)X(s)]\,ds \int_{\mathbb{R}^d} \mathbb{E}[\lambda^{(j)}(s)X(s)]^\top\,ds}{\int_{\mathbb{R}^d} \mathbb{E}[\lambda^{(j)}(s)]\,ds}.$$

Theorem 1. *Under the above assumptions, it holds that $\mathcal{S}(V_1) \subset \mathcal{S}_1$.*

Proof. Let B be a matrix with $\mathcal{S}(B) = \mathcal{S}_1$ and $j \in \{1, \ldots, m\}$. From Lemma 2 we have $\lambda^{(j)}(s) = f^j(B^\top X(s))$ for a measurable function f^j. It is enough to show that

$$\mathcal{S}\left(\int \mathbb{E}[f^j(B^\top X(s))X(s)]\,ds \int \mathbb{E}[f^j(B^\top X(s))X(s)]^\top\,ds \right) \subset \mathcal{S}_1.$$

Using the argument in Guan and Wang (2010, Equation (A.1), p. 385), for the projection matrix $P_B = B(B^\top B)^{-1} B^\top$, we obtain

$$\int \mathbb{E}[f^j(B^\top X(s))X(s)]\,\mathrm{d}s \int \mathbb{E}[f^j(B^\top X(s))X(s)]^\top\,\mathrm{d}s$$

$$= P_B \int \mathbb{E}[f^j(B^\top X(s))X(s)]\,\mathrm{d}s \int \mathbb{E}[f^j(B^\top X(s))X(s)]^\top\,\mathrm{d}s\, P_B.$$

Since this holds for each $j = 1, \ldots, m$, we have $\mathcal{S}(V_1) \subset \mathcal{S}(B) = \mathcal{S}_1$. This completes the proof.

From the proof, it can be seen that, for each jth slice, $j = 1, \ldots, m$, vector (6) belongs to $\mathcal{S}_1$. Principal component analysis is applied to find just c vectors among them. Define the slice means as the conditional expectations

$$m_j = \mathbb{E}[X(s) \mid s \in Y^j],$$

and let $p_j = \mathbb{P}(x \in Y^j \mid x \in Y)$, $j = 1, \ldots, m$. The weighted covariance matrix

$$V = \sum_{j=1}^{m} p_j m_j m_j^\top \tag{7}$$

of size $p \times p$ has eigenvalues $\xi_1 \geq \cdots \geq \xi_p$. Then the eigenvectors η_l of V corresponding to the c largest eigenvalues form columns of the matrix B; cf. Li (1991). When dealing with data observed in a compact window C, the matrix V can be estimated as

$$\hat{V}_1 = \sum_{j=1}^{m} \frac{\Psi_{Y^j}(C)}{\Psi_Y(C)} \frac{\int_{Y^j \cap C} X(s)\mathcal{H}^k(\mathrm{d}s)}{\Psi_{Y^j}(C)} \frac{[\int_{Y^j \cap C} X(s)\mathcal{H}^k(\mathrm{d}s)]^\top}{\Psi_{Y^j}(C)};$$

cf. (4) and (7). For more information on estimation, see Section 4.

3.2. Investigation of the second-order central intensity subspace

Let $l = 2$ in Definition 2. Assume that

$$\lambda_2(s, t) = f_2(B^\top X(s), B^\top X(t)) \tag{8}$$

for a matrix B of size $p \times c$, $c \leq p$. The aim is to estimate the subspace $\mathcal{S}_2 = \mathcal{S}(B)$. Assume that Y and $\{X(s)X(t)^\top, s, t \in Y\}$ are ergodic, and that $C \subset \mathbb{R}^d$ is a convex compact window. Then, from Lemma 1,

$$\hat{M}_2 = \frac{\int_{s,t \in Y \cap C} X(s)X(t)^\top \mathcal{H}^k(\mathrm{d}s)\mathcal{H}^k(\mathrm{d}t)}{\Psi_Y(C)^2} \rightarrow \frac{\iint \mathbb{E}[\lambda_2(s,t)X(s)X(t)^\top]\,\mathrm{d}s\,\mathrm{d}t}{\iint \mathbb{E}[\lambda_2(s,t)]\,\mathrm{d}s\,\mathrm{d}t} = M_2 \tag{9}$$

in probability when $C \uparrow \mathbb{R}^d$, where limit (9) is defined as a ratio of two limits:

$$\lim_{C \uparrow \mathbb{R}^d} \frac{1}{|C|^2} \int_C \int_C \mathbb{E}[\lambda_2(s,t)X(s)X(t)^\top]\,\mathrm{d}s\,\mathrm{d}t, \qquad \lim_{C \uparrow \mathbb{R}^d} \frac{1}{|C|^2} \int_C \int_C \mathbb{E}[\lambda_2(s,t)]\,\mathrm{d}s\,\mathrm{d}t.$$

Their finiteness can be verified, e.g. in the case when $X(\cdot)$ is a stationary or stationary and isotropic random field, when the integrands depend only on the difference of variables or the modulus of the difference of variables, respectively.

For a point process ($k = 0$), we consider

$$\hat{M}_2 = \frac{\sum_{s,t \in Y \cap C}^{\neq} X(s) X(t)^\top}{\Psi_Y(C)(\Psi_Y(C) - 1)}.$$

Theorem 2. *Let $Q_B = I_p - P_B$. It holds that*

$$\begin{aligned}
M_2 &= M_2^P + M_2^Q \\
&= \frac{P_B \iint \mathbb{E}[f_2(B^\top X(s), B^\top X(t)) X(s) X(t)^\top] \, \mathrm{d}s \, \mathrm{d}t \, P_B}{\iint \mathbb{E}[f_2(B^\top X(s), B^\top X(t))] \, \mathrm{d}s \, \mathrm{d}t} \\
&\quad + \frac{\iint \mathbb{E}[f_2(B^\top X(s), B^\top X(t))] \mathbb{E}[Q_B X(s) X(t)^\top Q_B] \, \mathrm{d}s \, \mathrm{d}t}{\iint \mathbb{E}[f_2(B^\top X(s), B^\top X(t))] \, \mathrm{d}s \, \mathrm{d}t}.
\end{aligned} \tag{10}$$

Proof. Write $X(s) = P_B X(s) + Q_B X(s)$. Then

$$\begin{aligned}
&\mathbb{E}[\lambda_2(s, t) X(s) X(t)^\top] \\
&\quad = \mathbb{E}[\mathbb{E}[f_2(B^\top X(s), B^\top X(t)) \\
&\qquad\qquad \times [P_B X(s) X(t)^\top P_B + P_B X(s) X(t)^\top Q_B + Q_B X(s) X(t)^\top P_B \\
&\qquad\qquad\quad + Q_B X(s) X(t)^\top Q_B] \mid P_B X(s), P_B X(t)]] \\
&\quad = P_B \mathbb{E}[f_2(B^\top X(s), B^\top X(t)) X(s) X(t)^\top] P_B \\
&\qquad + \mathbb{E}[f_2(B^\top X(s), B^\top X(t)) P_B X(s) \mathbb{E}[X(t)^\top Q_B \mid P_B X(s), P_B X(t)]] \\
&\qquad + \mathbb{E}[f_2(B^\top X(s), B^\top X(t)) \mathbb{E}[Q_B X(s) \mid P_B X(s), P_B X(t)] X(t)^\top P_B] \\
&\qquad + \mathbb{E}[f_2(B^\top X(s), B^\top X(t)) \mathbb{E}[Q_B X(s) X(t)^\top Q_B \mid P_B X(s), P_B X(t)]].
\end{aligned}$$

The inner expectations in the second and third terms are equal to 0 from the assumptions, and

$$\mathbb{E}[Q_B X(s) X(t)^\top Q_B \mid P_B X(s), P_B X(t)] = \mathbb{E}[Q_B X(s) X(t)^\top Q_B];$$

thus, the assertion follows.

If the second term M_2^Q on the right-hand side of the formula for M_2 were 0 then

$$\mathcal{S}(M_2 M_2^\top) \subset \mathcal{S}(M_2) \subset \mathcal{S}_2.$$

We are interested in situations where M_2^Q is negligible with respect to M_2^P, i.e.

$$\|M_2^Q\| \ll \|M_2^P\|, \tag{11}$$

e.g. for the Euclidean matrix norm. This means that $\mathcal{S}_2$ can be approximately estimated by the SIR method applied to $M_2 M_2^\top$. Typically, $\mathbb{E}[Q_B X(s) X(t)^\top Q_B]$ is negligible for $\|s - t\| \to \infty$, while in some models (e.g. repulsive point processes) $\lambda_2(s, t)$ is around 0 for small $\|s - t\|$. Moreover, when there is a positive correlation between $P_B X(s) X(t)^\top P_B$ and $\lambda_2(s, t)$ then, intuitively, (11) may hold.

We make this reasoning precise in the following two examples, where we consider the centered Gaussian random field $X = (X_1, X_2)$ in $\mathbb{R}^2$ with independent components of correlation functions

$$\zeta(s, t) = \exp(-\|s - t\|^2),$$

$B = (1, 0)^\top$, so that $\lambda_2(s, t) = f_2(X_1(s), X_1(t))$. The 2×2 matrix functions

$$P_B \mathbb{E}[\lambda_2(s, t) X(s) X(t)^\top] P_B \tag{12}$$

and

$$\mathbb{E}[\lambda_2(s, t)] \mathbb{E}[Q_B X(s) X(t)^\top Q_B] \tag{13}$$

are the integrands in the numerators of (10). Nonzero elements of these matrices are in the upper-left and lower-right corners, respectively. These elements are evaluated as functions of the variable $z = \|s - t\|$. The expectations are evaluated with respect to the bivariate Gaussian probability density

$$g(x, y) = \frac{1}{2\pi \sqrt{1 - \zeta(s, t)}} \exp\left(-\frac{1}{2(1 - \zeta(s, t))}(x^2 + y^2 - 2xy\sqrt{\zeta(s, t)})\right), \qquad x, y \in \mathbb{R}.$$

Example 1. Consider a stationary Poisson point process Φ with intensity ρ. Let Y be a simple inhibition point process such that each pair of points $s, t \in \Phi$ satisfying

$$\max(Z(s), Z(t)) \geq \|s - t\|$$

is removed, where $Z(s) = g(X_1(s))$ is a nonnegative function of X_1. The process Y has second-order intensity (cf. Diggle (2003, p. 73))

$$\lambda_2(s, t) = \begin{cases} \rho^2 \exp(-\rho U(Z(s), Z(t))) & \text{if } \max(Z(s), Z(t)) \leq \|s - t\|, \\ 0 & \text{otherwise,} \end{cases} \tag{14}$$

where $U(Z(s), Z(t))$ is the area of the union of balls centered in s, t with radii $Z(s), Z(t)$, respectively. In this case assumption (8) holds with $B = (1, 0)^\top$. Put

$$Z(s) = a + b \, \mathbf{1}_{[X_1(s)<0]}, \qquad a, b > 0. \tag{15}$$

It holds that $\lambda_2(s, t) = 0$ if either $[X_1(s) < 0 \vee X_1(t) < 0] \wedge \|s - t\| < a + b$ or $[X_1(s) \geq 0 \wedge X_1(t) \geq 0] \wedge \|s - t\| < a$. In the opposite case there are variants

$$
\begin{array}{llll}
X_1(s) \geq 0, & X_1(t) \geq 0, & \|s - t\| \geq a, & \lambda_2(s, t) = \rho^2 e^{-\rho U_2(s,t,a,a)}, \\
X_1(s) \geq 0, & X_1(t) < 0, & \|s - t\| \geq a + b, & \lambda_2(s, t) = \rho^2 e^{-\rho U_2(s,t,a,a+b)}, \\
X_1(s) < 0, & X_1(t) \geq 0, & \|s - t\| \geq a + b, & \lambda_2(s, t) = \rho^2 e^{-\rho U_2(s,t,a+b,a)}, \\
X_1(s) < 0, & X_1(t) < 0, & \|s - t\| \geq a + b, & \lambda_2(s, t) = \rho^2 e^{-\rho U_2(s,t,a+b,a+b)}.
\end{array}
$$

The results are given in Figure 1(a).

Example 2. The stationary determinantal point process Y has second-order intensity equal to the determinant

$$\lambda_2(s, t) = \begin{vmatrix} C_0(0) & C_0(s - t) \\ C_0(t - s) & C_0(0) \end{vmatrix},$$

where C_0 is a covariance function. For parameters $\alpha, \rho > 0$, we use the covariance function with finite range α:

$$C_0(x) = \frac{2\rho}{\pi}\left(\arccos\frac{\|x\|}{\alpha} - \frac{\|x\|}{\alpha}\sqrt{1 - \left(\frac{\|x\|}{\alpha}\right)^2}\right) \mathbf{1}_{[\|x\|<\alpha]}.$$
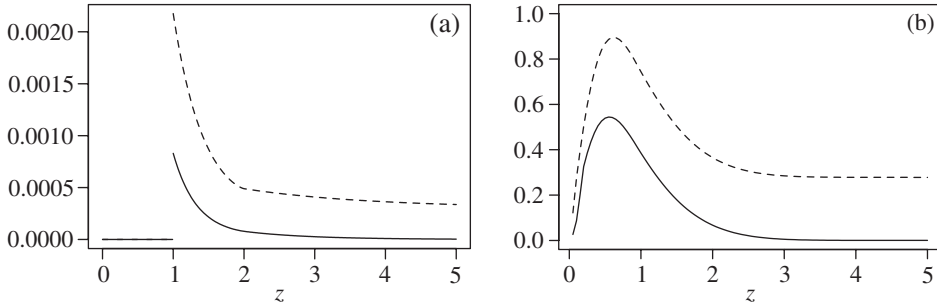
FIGURE 1: Comparison of functions of the variable $z = \|s - t\|$ and the nonzero elements of matrices (12) (*dashed lines*) and (13) (*solid lines*). (a) Example 1 (simple inhibition) for $a = b = \rho = 1$. (b) Example 2 (determinantal process) for $\alpha = 1$. The dashed graph values are much larger than the solid graph values in both cases.

The parameter ρ is randomized, and it depends on the first component of X:

$$\rho = \frac{4}{\pi^2 \alpha^2} \left(\arctan(X_1(s) X_1(t)) + \frac{\pi}{2} \right).$$

Under this scaling, given X, the determinantal process always exists; cf. Lavancier *et al.* (2012). The results are given in Figure 1(b).

Theorem 2 also enables us to understand the availability of a slicing procedure in the analysis of $\mathcal{S}_2$. In this case, the cartesian product $Y \times Y$ should be marked. Let

$$
\begin{aligned}
\mathcal{Y} &= Y \times Y, &\Psi_\mathcal{Y}(C) &= \Psi_Y(C)^2, &k &> 0, \\
\mathcal{Y} &= \{(s, t); s \in Y, t \in Y, s \neq t\}, &\Psi_\mathcal{Y}(C) &= \Psi_Y(C)(\Psi_Y(C) - 1), &k &= 0.
\end{aligned}
\tag{16}
$$

Consider a mark $\Gamma \colon \mathcal{Y} \to \mathbb{R}$ which is a measurable symmetric function, $\Gamma(s, t) = \Gamma(t, s)$ for each $(s, t) \in \mathcal{Y}$. Let the range of Γ be divided into m disjoint intervals called slices. This induces a (random) partition of $\mathcal{Y}$ into m disjoint subsets $(\mathcal{Y}^1, \ldots, \mathcal{Y}^m)$. Let

$$\Psi_{\mathcal{Y}^j}(C) = \iint_{\mathcal{Y}^j \cap C^2} \mathcal{H}^k(\mathrm{d}s) \mathcal{H}^k(\mathrm{d}t).$$

Define conditional expectation matrices (slice means)

$$o_j = \mathbb{E}[X(s)X(t)^\top \mid (s, t) \in \mathcal{Y}^j]$$

and $q_j = \mathbb{P}((s, t) \in \mathcal{Y}^j \mid (s, t) \in \mathcal{Y})$, $j = 1, \ldots, m$. The matrix $U_2 = \sum_{j=1}^m q_j o_j o_j^\top$ is then subject to the principal component analysis. The empirical version of the matrix U_2 calculated from data is

$$
\begin{aligned}
\hat{U}_2 &= \frac{1}{\Psi_\mathcal{Y}(C)} \sum_{j=1}^m \Psi_{\mathcal{Y}^j}(C) \hat{M}_2^j [\hat{M}_2^j]^\top \\
&= \frac{1}{\Psi_\mathcal{Y}(C)} \sum_{j=1}^m \frac{1}{\Psi_{\mathcal{Y}^j}(C)} \iint_{\mathcal{Y}^j \cap C^2} X(s) X(t)^\top \mathcal{H}^k(\mathrm{d}s) \mathcal{H}^k(\mathrm{d}t) \\
&\quad \times \left[\iint_{\mathcal{Y}^j \cap C^2} X(s) X(t)^\top \mathcal{H}^k(\mathrm{d}s) \mathcal{H}^k(\mathrm{d}t) \right]^\top.
\end{aligned}
\tag{17}
$$

In the special case $k = 0$ we can express $\hat{M}_2^j$ as

$$\hat{M}_2^j = \frac{1}{\Psi_{\mathcal{Y}^j}(W)} \sum_{\mathcal{Y}^j \cap C^2} X(s) X(t)^\top.$$

3.3. Fibre process based on diffusion

As a model of an $\mathcal{H}^1$-set in $\mathbb{R}^2$, we will consider a numerical solution of an SDE for $Y_t = (Y_t^{(1)}, Y_t^{(2)}) \in \mathbb{R}^2$, $t \geq 0$:

$$
\begin{aligned}
\mathrm{d}Y_t^{(1)} &= -\frac{a}{2} Y_t^{(1)} \, \mathrm{d}t + b(Y_t^{(1)}, Y_t^{(2)}) h(Y_t^{(1)}, Y_t^{(2)}) \, \mathrm{d}W_t^{(1)}, \\
\mathrm{d}Y_t^{(2)} &= -\frac{a}{2} Y_t^{(2)} \, \mathrm{d}t + b(Y_t^{(1)}, Y_t^{(2)}) h(Y_t^{(1)}, Y_t^{(2)}) \, \mathrm{d}W_t^{(2)}.
\end{aligned}
\tag{18}
$$

Here

$$
b(y) = b_1 \exp\{b_2 g(y)\}, \quad b_1, b_2 \in \mathbb{R}, \qquad g \colon \mathbb{R}^2 \to \mathbb{R},
\tag{19}
$$
$$
h(y) = (1 - \|y\|^k), \qquad y \in \mathbb{R}^2,
$$

and $W_t = (W_t^{(1)}, W_t^{(2)})$ is a Brownian motion. In this subsection, let C be the unit circle in $\mathbb{R}^2$.

Theorem 3. *Let a Borel function $g(y)$ be bounded on C. Then, for an arbitrary initial condition $y_0 \in C$, there exists a solution $\{Y_t, t \geq 0\}$ of (18) such that $Y_0 = y_0$ and $\{Y_t, t \geq 0\} \subset C$ almost surely.*

Moreover, if $g(y)$ is Lipschitz continuous on $C^\varepsilon := \{y \colon \|y\| < 1 + \varepsilon\}$ for some $\varepsilon > 0$ then $\{Y_t, t \geq 0\} \subset C$ almost surely for an arbitrary solution of (18) with initial condition $Y_0 = y_0 \in C$.

Proof. Let

$$\beta(y) = \beta(y_1, y_2) = \left(-\frac{ay_1}{2}, -\frac{ay_2}{2}\right)^\top$$

and

$$\sigma(y) = \begin{pmatrix} b(y)h(y) & 0 \\ 0 & b(y)h(y) \end{pmatrix}.$$

Then (18) with initial condition $Y_0 = y_0 \in C$ can be rewritten as

$$\mathrm{d}Y_t = \beta(Y_t) \, \mathrm{d}t + \sigma(Y_t) \, \mathrm{d}W_t, \qquad Y_0 = y_0. \tag{20}$$

Define $\hat{\sigma}(y) = \sigma(y)$, $y \in C$, and $\hat{\sigma}(y) = 0$, $y \notin C$, and consider the equation

$$\mathrm{d}\hat{Y}_t = \beta(\hat{Y}_t) \, \mathrm{d}t + \hat{\sigma}(\hat{Y}_t) \, \mathrm{d}W_t, \qquad \hat{Y}_0 = y_0. \tag{21}$$

According to Lemma 1 of Staněk and Štěpán (2010), $\hat{Y} \in C$ almost surely for an arbitrary solution $\hat{Y}$ to (21). The existence of a solution $\hat{Y}$ to (21) follows from local boundedness of the coefficients β and $\hat{\sigma}$. Since any solution $\hat{Y}$ to (21) is also a solution to (20), the first part of the theorem is proved.

If p is Lipschitz continuous on C^ε, then $\hat{Y}$ and Y have a strong uniqueness property (according to Theorem 12.1 of Rogers and Williams (1994, p. 132)) in the time interval $(0, \tau_{C^\varepsilon})$, where $\tau_{C^\varepsilon} := \inf\{t > 0 \colon Y_t \notin C\}$ is the time of the first exit of the process Y from the set C^ε. Since the solution $\hat{Y}$ coincides with Y in C, $\hat{Y}$ is a unique solution to (20) as well.

In the following $g(s)$, $s \in C$, will be a linear combination, $\mathcal{L}_X$, of realizations of components of a vector GRF $\{X(s),\ s \in C\}$. In order to simulate a realization of the random set $Y \subset C$, we first simulate $g(s)$ and then a trajectory $Y = \{Y_t,\ 0 \le t \le T\}$ of the SDE solution by means of the Euler method (conditionally on $g = \mathcal{L}_X$). From Theorem 3, the theoretical solution of the SDE remains in the circular region. The numerical solution may cross the boundary; therefore, a condition is added that in this case the trajectory is projected on the boundary.

4. Statistical methods

4.1. Estimation

When we deal with simulated or real data, the theoretical quantities from Section 3 are estimated by their sampling analogues. Generally, for a stationary Gaussian p-dimensional random field $\tilde{X}$ observed in a bounded window $C \subset \mathbb{R}^d$ with Lebesgue measure $|C|$, put

$$\bar{X} = \frac{1}{|C|} \int_C \tilde{X}(s)\, ds, \quad \Sigma = \frac{1}{|C|} \int_C [\tilde{X}(s) - \bar{X}][\tilde{X}(s) - \bar{X}]^\top\, ds.$$

Then the standardized p-dimensional random field is $X(s) = \Sigma^{-1/2}[\tilde{X}(s) - \bar{X}]$. Based on observation of the random field on a set of grid points G, the sampling analogues of $\bar{X}$ and Σ are

$$\hat{\bar{X}} = \frac{1}{\operatorname{card} G} \sum_{s \in G} \tilde{X}(s)\, ds, \qquad \hat{\Sigma} = \frac{1}{\operatorname{card} G} \sum_{s \in G} [\tilde{X}(s) - \hat{\bar{X}}][\tilde{X}(s) - \hat{\bar{X}}]^\top\, ds,$$

and the empirical standardized random field at the arbitrary point $s \in C$ is given by

$$\hat{X}(s) = \hat{\Sigma}^{-1/2}[\tilde{X}(s_G) - \hat{\bar{X}}],$$

where $s_G \in G$ is the nearest grid point to s.

Characteristics of $\mathcal{H}^k$-sets can be estimated by choosing a finite set

$$T = \{t_i\}_{i=1}^n \subset Y \cap C \tag{22}$$

of random test points. Generally, probes T_p of complementary dimension $d - k$ are used to get test points as intersections $T = T_p \cap Y \cap C$; cf. Stoyan *et al.* (1995, p. 293). Let

$$T_j = T \cap Y^j, \quad j = 1, \ldots, m, \qquad n_j = \operatorname{card} T_j,$$

and (cf. (16))

$$\mathcal{T} = \{(s, t) \in T \times T,\ s \ne t\}, \qquad \mathcal{T}_j = \mathcal{T} \cap \mathcal{Y}^j,$$

$l_j = \operatorname{card} \mathcal{T}_j$, $\sum_{j=1}^m l_j = n(n-1)$. We have the estimators

$$\hat{p}_j = \frac{n_j}{n}, \qquad \hat{q}_j = \frac{l_j}{n(n-1)},$$

$$\hat{m}_j = \frac{1}{n_j} \sum_{t \in T_j} X(t), \qquad \hat{o}_j = \frac{1}{l_j} \sum_{(s,t) \in \mathcal{T}_j} X(s)X(t)^\top.$$

Using these estimates, we put

$$\hat{V} = \sum_{j=1}^m \hat{p}_j \hat{m}_j \hat{m}_j^\top, \qquad \hat{U} = \sum_{j=1}^m \hat{q}_j \hat{o}_j \hat{o}_j^\top. \tag{23}$$

The eigenvectors $\hat{\eta}_l$ of $\hat{V}$ and $\hat{U}$ corresponding to the c largest eigenvalues are evaluated and transformed to

$$\hat{\beta}_l = \hat{\Sigma}^{-1/2}\hat{\eta}_l, \qquad l = 1, \ldots, c. \tag{24}$$

Vectors $\hat{\beta}_l$ form columns of an estimator $\hat{B}$ of matrix B for the dimension reduction problem of $(\tilde{X}, Y)$ under assumptions (3) and (8), respectively.

Guan and Wang (2010) suggested the estimation error

$$\triangle(B, \hat{B}) = \|B(B^\top B)^{-1}B^\top - \hat{B}(\hat{B}^\top\hat{B})^{-1}\hat{B}^\top\|_{\max} \tag{25}$$

to compare the estimated and true matrix of the central subspace. Here $\|A\|_{\max}$ denotes the maximum of the absolute singular value of a matrix A.

Having n data sets and getting $\hat{\beta}_l^i$, $i = 1, \ldots, n$, from (24) we want to obtain an estimator of β_l, $l = 1, \ldots, c$, from this information. In fact, the directions of the vectors are crucial. In $\mathbb{R}^d$ consider an arbitrary unit vector w, let $\hat{\beta}^1, \ldots, \hat{\beta}^n$ be independent and identically distributed (i.i.d.) unit random vectors, and let $\langle\cdot,\cdot\rangle$ be the inner product. Then

$$\hat{\beta} = \underset{\{v\colon \|v\|=1, \langle w,v\rangle\geq 0\}}{\arg\max} \sum_{i=1}^n |\langle\hat{\beta}^i, v\rangle|. \tag{26}$$

Obviously, different choices of the vector w do not effect the direction of the estimator $\hat{\beta}$, but can affect its orientation, i.e. if $\hat{\beta}^1$ and $\hat{\beta}^2$ are estimators with choices $w = w^1$ and $w = w^2$, respectively, where w^1 and w^2 are arbitrary units vectors in $\mathbb{R}^d$, then $\hat{\beta}^1 = \hat{\beta}^2$ or $\hat{\beta}^1 = -\hat{\beta}^2$.

This estimator is unbiased in the following sense.

Proposition 1. *Let the distribution of $\hat{\beta}^i$ be symmetric with respect to the axis given by $\tilde{\beta}$, i.e. $\hat{\beta}^i$ has the same distribution as $-\hat{\beta}^i + 2\tilde{\beta}\langle\tilde{\beta}, \hat{\beta}^i\rangle$ and choose $w = \tilde{\beta}$. Then there exists $a \in [0, 1]$ such that*

$$\mathbb{E}\hat{\beta} = a\tilde{\beta}.$$

Proof. Without loss of generality, consider $\tilde{\beta} = (1, 0, \ldots, 0)$. Furthermore, let $^sv = (v_1, -v_2, \ldots, -v_d)$, $^s\hat{\beta} = (\hat{\beta}_1, -\hat{\beta}_2, \ldots, -\hat{\beta}_d)$, and $^s\hat{\beta}^i = (\hat{\beta}_1^i, -\hat{\beta}_2^i, \ldots, -\hat{\beta}_d^i)$ for all $i = 1, \ldots, n$. Then $\hat{\beta}^i$ and $^s\hat{\beta}^i$ have the same distribution. Obviously, $|^sv| = |v| = 1$, $\langle^sv, \tilde{\beta}\rangle = \langle v, \tilde{\beta}\rangle$, and $\langle^sv, ^s\hat{\beta}^i\rangle = \langle v, \hat{\beta}^i\rangle$. Therefore, $\hat{\beta}$ and $^s\hat{\beta}$ have the same distribution, and, hence, $\mathbb{E}\hat{\beta} = a\tilde{\beta}$ for some $a \in [0, 1]$.

4.2. Statistical testing

Generally, the dimension c of the central subspace is not known. A starting point would be the test of the null hypothesis

$$H_0\colon c = 0 \quad \text{against} \quad H_A\colon c > 0, \tag{27}$$

where by $c = 0$ we mean the independence of X and Y. Consider the coefficient (see Li (1991))

$$R^2 = R^2(\hat{\beta}_1) = \max_{\beta\in\mathscr{S}_{Y|X}} \frac{(\hat{\beta}_1^\top\beta)^2}{\hat{\beta}_1^\top\hat{\beta}_1\beta^\top\beta}, $$

where $\hat{\beta}_1$ is from (24). From the independence of X and Y, it follows that $R^2 = 0$. Therefore, if we reject the null hypothesis of orthogonality

$$H_0\colon R^2 = 0 \quad \text{against} \quad H_A\colon R^2 > 0, \tag{28}$$

then (27) also has to be rejected. The test of orthogonality for the known distribution of X can proceed via the following steps.

- Calculate R^2 from the observed data Y and X.

- Calculate R^2 from observed Y and each of the n independently simulated realizations of X; thus, we have R_j^2, $j = 1, \ldots, n$.

- The p-value of the test is (card $\{R_j^2 \geq R^2\} + 1)/(n + 1)$.

In practice, we cannot simulate independent realizations of X since its distribution is unknown. On a planar rectangular or circular region W we can use this testing algorithm with the assumption that, under H_0, the joint distribution of Y and X is invariant with respect to translation or rotation of Y, respectively. Then instead of independent realizations of X we use n systematic translations or rotations, respectively, of X with respect to fixed Y. For translations, the window is wrapped on a cylinder.

An alternative test of independence of X and Y, when X is one dimensional (not necessarily Gaussian), can be based on the work of Schlather *et al.* (2004). They developed tests for stationary marked point processes using the mark-weighted K-function or L-function. Under the second-order intensity reweighted stationarity, cf. Illian *et al.* (2008), this test can be generalized to the use for an $\mathcal{H}^k$-set Y, $0 \leq k < d$.

Assume that we observe X and Y in a window $C \subset \mathbb{R}^d$. Consider a partition of $Y \cap C$ into disjoint subsets B_j, points $z_j \in B_j$, $j = 1, 2, \ldots, e$, and let $\mathcal{H}^k(B_j) = \triangle_{z_j}^k$. The algorithm of the test is as follows.

(a) Choose random test points $\{t_1, \ldots, t_n\}$ as in (22), typically $e \gg n$.

(b) Estimate the intensity function λ at points t_i, $i = 1, \ldots, n$, and z_j, $j = 1, \ldots, e$.

(c) Put

$$\hat{K}_X(r) = \frac{1}{n} \sum_{t_i} f(X(t_i)) \sum_j \triangle_{z_j}^k \frac{\mathbf{1}_{[\|t_i - z_j\| < r]}}{\lambda(t_i)\lambda(z_j)}, \tag{29}$$

where f is a suitable nonnegative increasing function.

(d) *Random reallocation.* In (c) perform q permutations of $\{X(t_i),\ i = 1, \ldots, n\}$, evaluate $\hat{K}_X$ for each permutation, and evaluate bounds $\hat{K}_{\max}$ and $\hat{K}_{\min}$.

(e) Transform the K-function to the L-function

$$\hat{L}(r) = \left(\frac{\hat{K}(r)}{\omega_d} \right)^{1/d},$$

where ω_d is the volume of unit d-balls, and draw envelopes

$$\hat{L}_{\max}(r) - \hat{L}_X(r), \qquad \hat{L}_{\min}(r) - \hat{L}_X(r). \tag{30}$$

When the horizontal axis lies between the envelopes, the null hypothesis of independence is not rejected on the significance level $2/(q + 1)$.

For complete estimation of the dimension c, the statistics

$$\hat{\Lambda}_c = n \sum_{i=c+1}^{p} \xi_i \tag{31}$$

might be of use, where $\xi_1 \geq \xi_2 \geq \cdots \geq \xi_p$ are eigenvalues of the weighted covariance matrix $\hat{V}$ or $\hat{U}$ given by (23). The number of slices m must be chosen greater than $c+1$. To estimate c, we start with $c_0 = 0$. If hypothesis (27) is rejected, we increase $c_0 = c_0 + 1$ and repeat the same procedure sequentially until

$$H_0 : c = c_0 \quad \text{against} \quad H_A : c > c_0$$

is not rejected or $c_0 = p$. Under the validity of H_0, Li (1991) and Cook (1998) proved that $\hat{\Lambda}_c$ has an asymptotic chi-square distribution with $(p - c_0)(m - c_0 - 1)$ degrees of freedom when $X(t_i)$, $i = 1, \ldots, n$, are i.i.d. This is not the case in the GRF, so the test is approximate here and can be tried only when the t_i are rather sparse. An analogous reasoning is necessary when thinking of other sampling properties of SIR (consistency, etc.) as summarized in Li (2000) for i.i.d. observations of random vectors X.

5. Simulation studies

5.1. Description of the simulation

In simulations $X = (X_1, X_2, X_3)$ is a three-dimensional Gaussian random field in $\mathbb{R}^d$ with independent components which have zero mean and covariance function

$$\zeta(s, t) = \exp(-\gamma \|s - t\|^\alpha), \qquad s, t \in \mathbb{R}^d, 1 \leq \alpha \leq 2, \gamma > 0. \tag{32}$$

The method GaussRF from the RandomFields library in software R was used for generating realizations $X = \{X(s), s \in J_d = [0, 1]^d\}$ for either $d = 2$ or $d = 3$. A random set Y is simulated so that it depends on a linear combination, $\mathcal{L}_X$, of components of the GRF X. Obviously, we have, for $\mathcal{L}_X = \sum_{i=1}^3 e_i X_i$, $e_i \in \mathbb{R}$,

$$\lambda(s) = f(B^\top X(s)), \qquad B = (e_1, e_2, e_3)^\top,$$

and the dimension of the central subspace $c = 1$. Slicing of Y is based on its geometrical properties.

In simulation study I, given the GRF X, consider an inhomogeneous Poisson point process Y_p with intensity

$$\lambda(s) = a \exp(\mathcal{L}_X(s)), \qquad s \in J_d, a > 0, d = 2 \text{ or } 3.$$

In fact, Y_p is the log-Gaussian Cox process (see Møller and Waagepetersen (2004)), the slicing is based on the nearest neighbour distance as the mark Γ_p. Furthermore, a Poisson–Voronoi tessellation is simulated in J_d with germs corresponding to events of Y_p; see Figure 2(a) for $d = 2$. In $\mathbb{R}^2$, the system of edges forms a random fibre process Y. A piecewise constant mark Γ at a point of Y is the length of the corresponding edge ($\mathcal{H}^1$-almost surely unique). In $\mathbb{R}^3$, the system of faces forms a random surface process Y. A piecewise constant mark Γ at a point of Y is the area of the corresponding face ($\mathcal{H}^2$-almost surely unique). For the estimation, test points T are centroids of edges or faces randomly chosen with probability proportional to the length of the edge or surface area of the face, respectively. Besides the basic dependent case as in Figure 2(a) we consider $n - 1$ independent cases where the same realization of Y is independent of each component of X. The whole procedure is repeated in order to get a number q of simulated sets of data.

In simulation study II, given the GRF X, we evaluate a fibre process Y based on diffusion from Subsection 3.3. The linear combination $g(s) = \mathcal{L}_X(s)$ enters in (19). In the numerical
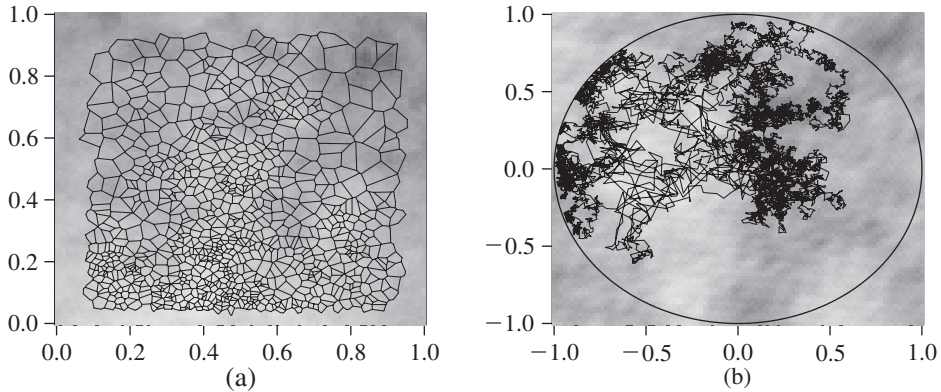
FIGURE 2: (a) A Poisson–Voronoi tessellation generated by a log-Gaussian Cox process with intensity $\lambda(s) = a \exp(X_1(s))$, where X_1 is a Gaussian random field with its surface plotted in a grey scale (white is the largest value), $\gamma = 1$ and $\alpha = 1$. High values of X_1 (*white*) imply a higher intensity of cells, which implies shorter lengths. (b) Simulated fibre process Y from (18) and a Gaussian random field X_1 with parameters $\gamma = \frac{5}{3}$ and $\alpha = 1$. The high values of X_1 (*white*) imply a higher speed of the motion and longer segments in the numerical solution.

solution of (18) using the Euler method with a fixed temporal step the curve Y (a fibre process) is formed by segments whose length is proportional to the speed of motion; see Figure 2(b). In any point of Y the length of the corresponding segment is the mark Γ ($\mathcal{H}_1$-almost surely unique). In each of the q simulations of X and Y a number of n systematic rotations of X are taken in angular steps $2\pi j/n$, $j = 0, 1, \ldots, n - 1$. The number of test points along Y is s (taken equidistantly systematically in time).

Simulation study III corresponds to the theoretical Example 1 in Subsection 3.2. Given the GRF X, a simple inhibition point process Y with second-order intensity λ_2 as in (14) was simulated on J_2. We consider two choices of Z. Let (Z1) denote the process $Z(s) = a(\arctan g(X_1(s)) + \pi/2)$, $a > 0$, and let (Z2) denote the process given in (15). Slicing is performed in the same way as explained in (16), where the criterion for slicing (the mark) is the theoretical second-order intensity $\lambda_2(s, t)$. Its value is calculated for each pair of points and the range is divided into several slices with approximately equal cardinality. Finally, the matrix $\hat{U}_2$ described in (17) is a subject for the principal components method.

5.2. Numerical results

In Figure 3, the histograms of p-values for the test of orthogonality (28) are presented for both simulations I and II. Each row represents one type of random marked set, namely a point, fibre (twice), or surface process. We observe how the power of the test increases with the number of slices (from left to right). This is more apparent in the upper two rows.

A question of an optimal number of slices arises. It was investigated by means of the estimation errors criterion (25). In Table 1 we summarize the results for five models of random marked sets. We denote by (B1) and (B2) the choices $B = (1, 0, 0)^\top$ and $B = (1, 1, 0)^\top$, respectively. Both alternatives, (Z1) with $a = 0.02$ and (Z2) with $a = b = 0.04$, $\rho = 100$, are considered for simulation III. In each case, $q = 100$ simulations were realized and the sample means of $\triangle(B, \hat{B})$ with different numbers of slices were computed. The results vary: while in simulation I optimal numbers are below 10, in simulation III the opposite is true, as can be seen in Table 1.
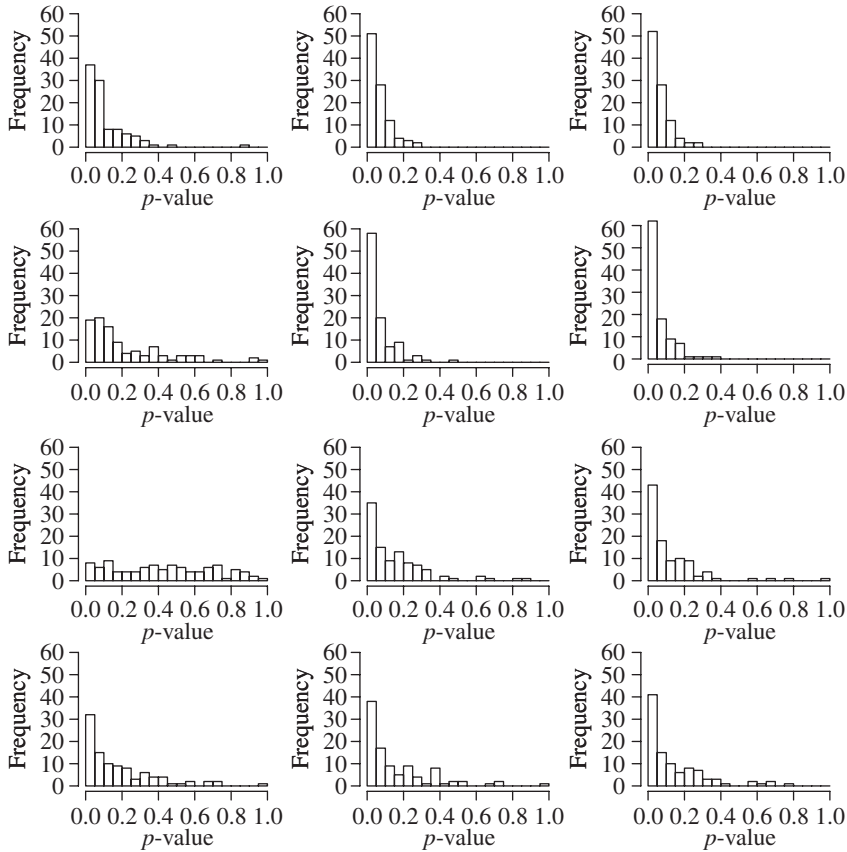
FIGURE 3: Histograms of *p*-values for the orthogonality test for $m = 1$ (*first column*), $m = 2$ (*second column*), and $m = 4$ (*third column*) slices obtained using the SIR method, based on $q = 100$, $n = 20$, $p = 3$, and $B = (1, 0, 0)$. From simulation I we present a marked point process of tessellation generators in $\mathbb{R}^2$ (*first row*), a fibre process of tessellation edges in $\mathbb{R}^2$ (*second row*), and a surface process of tessellation faces in $\mathbb{R}^3$ (*third row*), and from simulation II we present a fibre process based on diffusion (*fourth row*). The mean number of generators is 1000 in $\mathbb{R}^2$ and 10 000 in $\mathbb{R}^3$.

TABLE 1.

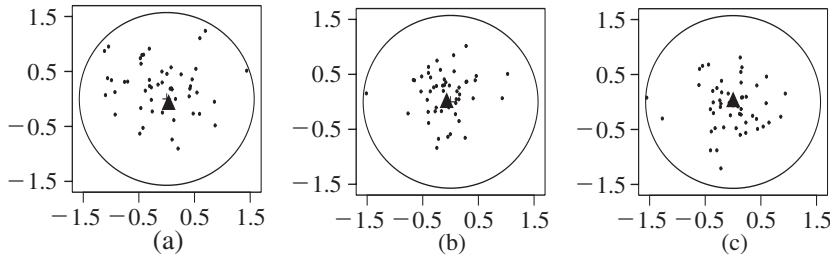| Slices | Simulation I, tessellations | | | | | | Simulation III | |
| | Points in $\mathbb{R}^2$ | | Edges in $\mathbb{R}^2$ | | Faces in $\mathbb{R}^3$ | | Z_1 | Z_2 |
	B_1	B_2	B_1	B_2	B_1	B_2	B_1	B_1
1	0.296	0.243	0.418	0.329	0.640	0.538	0.582	0.463
2	0.236	0.216	0.214	0.220	0.341	0.346	0.343	0.407
4	0.231	0.212	0.211	0.207	0.309	0.315	0.195	0.405
8	0.227	0.212	0.212	0.208	0.309	0.307	0.167	0.385
16	0.230	0.214	0.220	0.210	0.318	0.309	0.159	0.378
32	0.244	0.220	0.231	0.215	0.340	0.318	0.160	0.361

FIGURE 4: Estimators $\hat{\beta}_1$ from (26) based on simulation II. Panels (a) and (b) show estimators of $B = (1, 0, 0)$, where we have weaker ($b_2 = 0.3$ in (a)) and stronger ($b_2 = 0.9$ in (b)) dependence between X_1 and Y. Panel (c) presents $b_2 = 0.9$ and estimators of $B = (\frac{1}{2}, -1, \frac{1}{2})$. Points correspond to 50 simulations (estimators given by (24)), and the triangles represent the final estimator (26).

TABLE 2.

c	$\gamma = 1.666$		$\gamma = 10$	
	$s = 50$	$s = 100$	$s = 50$	$s = 100$
1	39	31	41	37
2	11	19	9	13

Estimators $\hat{\beta}_1$ from (26) based on simulation II are presented graphically in Figure 4. The centre of the circle represents the true vector B and each vector $\hat{\beta}_1^i$ is represented by a point x, whose distance from the centre corresponds to the angle between $\hat{\beta}_1^i$ and B. We observe that the spread of the xs is smaller when the dependence between X_1 and Y is bigger; cf. Figure 4(a) and (b). The method also works for the general vector B in Figure 4(c). The triangle represents $\hat{\beta}_1$, which, in each case, lies close to the true vector.

In order to demonstrate the estimation of the dimension c of the central subspace using statistic $\hat{\Lambda}_c$ in (31), in simulation II we made $q = 50$ simulations of X with $\alpha = 1$ and Y as above (without rotations). The number of slices $m = 3$ was chosen. Since $p = 3$, the maximal value of c_0 we can use for testing $H_0 : c = c_0$ against $H_A : c > c_0$ is $c_0 = 2$, true $c = 1$. The numbers of accepted hypotheses for different cases are given in Table 2.

It arises that, for a larger number, s, of the test points as well as for a smaller coefficient γ in the covariance function (32), the test points involve more dependence and the conclusion is false. In the opposite case, their dependence decreases and the approximative test yields better results.

6. Real data example

A real data specimen from Frcalová *et al.* (2010) which involves a planar fibre system is analyzed. A spatiotemporal point pattern of action potentials (called spikes) of a place cell of hippocampus of a rat looking for food in a circular arena $C \subset \mathbb{R}^2$ is shown in Figure 5(a). The experiment lasted $T = 614$ seconds and the location of the rat was recorded each $\frac{1}{60}$ of a second, so that we observe a track, which is modeled as a random $\mathcal{H}^1$-set.

The aim is to find out whether there is a significant dependence between the spiking activity of the neuron and the movement of the rat. The method based on the weighted K-function (29) is presented. Even if there is a temporal element in the experiment, we consider a purely spatial problem within our context. Thus, the spiking activity is modeled as a planar intensity $\{X_1(s), s \in C\}$ of the point process of spikes and Y is a planar fibre process (the track).

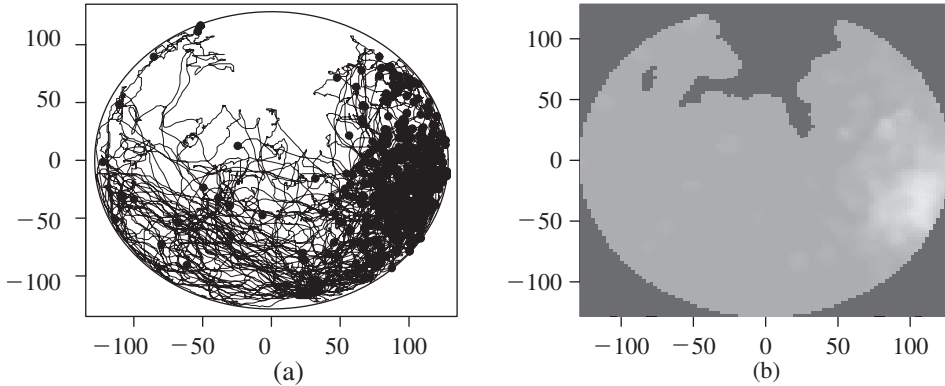(a)   (b)

FIGURE 5: (a) The line describes the track Y of a rat moving in a circular arena and the points are the locations where the neuron fired. Altogether 1096 spikes were recorded. (b) Estimate of the spiking intensity in a grey scale (white is the largest value). The black area indicates the region where we were not able to estimate the intensity because of a lack of data.
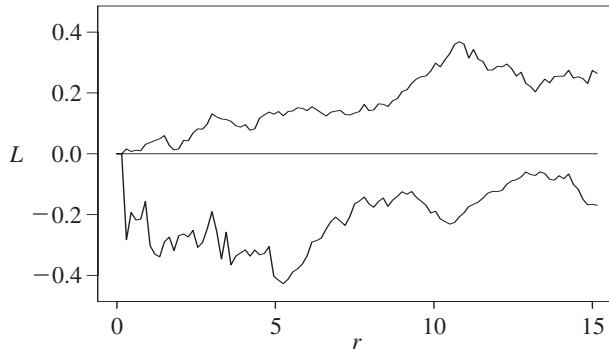

FIGURE 6: Test of independence of real data from neurophysiological experiment. It is based on the estimated X-weighted K-function (29) transformed to the L-function, $L(r) = \sqrt{K(r)/\pi}$. After $k = 39$ permutations the horizontal axis lies inside the envelopes (30); we do not reject the hypothesis of independence of the spiking activity on the track (on significance level 0.05).

Let $Y(t)$ be the location of the rat at time t. Since we observe the spikes along the track (point process on a curve; see Frcalová and Beneš (2009)), we use a kernel estimator of the intensity X_1 of the form

$$\hat{X}_1(s) = \frac{\sum_{s_i \in C} k(\|s - s_i\|)}{\sum_{t_i \in H} k(\|s - Y(t_i)\|)\triangle_t},$$

where s_i is the location of a spike, H is a set of equidistant times, $\triangle_t = t_{i+1} - t_i$, and k is a kernel function, here given by

$$k(r) = \begin{cases} (b^2 - r^2)\dfrac{2}{\pi b^4}, & r \le b, \\ 0, & r > b. \end{cases}$$

Figure 5(b) shows the corresponding intensity estimate $\hat{X}_1$, where we chose the bandwidth $b = 10$. The result of the test is given in Figure 6. Independence hypothesis of the spiking activity and the random track cannot be rejected.

7. Conclusions

The dimension reduction problem for random $\mathcal{H}^k$-sets brings interesting mathematical problems on the frontier of stochastic geometry and linear algebra; see Subsection 3.2. Further development in this area is desirable. Our improvement of the sliced inverse regression method was demonstrated in simulations in Section 5, based on both standard and new statistical methods from Section 4. The simulation of an $\mathcal{H}^1$-set (a fibre process) based on the diffusion in Section 5 was motivated by the presented data of a track of an experimental rat (Section 6); cf. Brillinger (2010) for further discussion on track modeling.

Acknowledgements

References

BALLANI, F., KABLUCHKO, Z. AND SCHLATHER, M. (2012). Random marked sets. *Adv. Appl. Prob.* **44,** 603–616.

BENEŠ, V. AND RATAJ, J. (2004). *Stochastic Geometry: Selected Topics*. Kluwer Academic, Boston, MA.

BRILLINGER, D. R. (2010). Modeling spatial trajectories. In *Handbook of Spatial Statistics*, eds A. E. Gelfand *et al.*, Chapman & Hall/CRC, Boca Raton, FL, pp. 463–476.

COOK, R. D. (1998). *Regression Graphics*. John Wiley, New York.

COOK, R. D. AND WEISBERG, S. (1991). Sliced inverse regression for dimension reduction: comment. *J. Amer. Statist. Assoc.* **86,** 328–332.

DIGGLE, P. J. (2003). *Statistical Analysis of Spatial Point Patterns*, 2nd edn. Arnold, London.

FRCALOVÁ, B. AND BENEŠ, V. (2009). Spatio-temporal modelling of a Cox point process sampled by a curve, filtering and inference. *Kybernetika* **45,** 912–930.

FRCALOVÁ, B., BENEŠ, V. AND KLEMENT, D. (2010). Spatio-temporal point process filtering methods with an application. *Environmetrics* 21, 240–252.

GUAN, Y. (2008). On consistent nonparametric intensity estimation for inhomogeneous spatial point processes. *J. Amer. Statist. Assoc.* **103,** 1238–1247.

GUAN, Y. AND WANG, H. (2010). Sufficient dimension reduction for spatial point processes directed by Gaussian random fields. *J. R. Statist. Soc. B* **72,** 367–387.

ILLIAN, J., PENTTINEN, A., STOYAN, H. AND STOYAN, D. (2008). *Statistical Analysis and Modelling of Spatial Point Patterns*. John Wiley, Chichester.

LAVANCIER, F., MØLLER, J. AND RUBAK, E. (2012). Statistical aspects of determinantal point processes. Preprint. Available at http://arxiv.org/abs/.1205.4818v2.

LI, K.-C. (1991). Sliced inverse regression for dimension reduction. *J. Amer. Statist. Assoc.* **86,** 316–327.

LI, K.-C. (2000). High dimensional data analysis via the SIR/PHD approach. Lecture Notes, UCLA.

LI, B. AND WANG, S. (2007). On directional regression for dimension reduction. *J. Amer. Statist. Assoc.* **102,** 997–1008.

MOLCHANOV, I. (1983). Labelled random sets (Russian). *Teor. Veroyat. Matem. Statist.* **29,** 93–98 (in Russian). English translation: *Theory Prob. Math. Statist.* **29** (1984), 113–119.

MØLLER, J. AND WAAGEPETERSEN, R. (2004). *Statistics and Simulations of Spatial Point Processes*. World Scientific, Singapore.

NGUYEN, X. AND ZESSIN, H. (1979). Ergodic theorems for spatial processes. *Z. Wahrscheinlichkeitsth.* **48,** 133–158.

PAWLAS, Z. (2003). Central limit theorem for random measures generated by stationary processes of compact sets. *Kybernetika* **39,** 719–729.

ROGERS, L. C. G. AND WILLIAMS, D. (1994). *Diffusions, Markov Processes and Martingales*, Vol. 2. Cambridge University Press.

SCHLATHER, M., RIBEIRO, P. J. AND DIGGLE, P. J. (2004). Detecting dependence between marks and locations of marked point processes. *J. R. Statist. Soc. B* **66,** 79–93.

SCHNEIDER, R. AND WEIL, W. (2008). *Stochastic and Integral Geometry*. Springer, Berlin.

STANĚK, J. AND ŠTĚPÁN, J. (2010). Diffusion with a reflecting and absorbing level set boundary—a simulation study. *Acta Univ. Carol. Math. Phys.* **51,** 73–86.

STOYAN, D., KENDALL, W. S. AND MECKE, J. (1995). *Stochastic Geometry and Its Applications*. John Wiley, New York.

ZÄHLE, M. (1982). Random processes of Hausdorff rectifiable closed sets. *Math. Nachr.* **108,** 49–72.