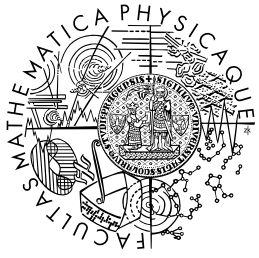

FACULTY
OF MATHEMATICS
AND PHYSICS
Charles University

HABILITATION THESIS

Jiří Dvořák

Nonparametric tests in spatial statistics

Department of Probability
and Mathematical Statistics

Mathematics – Probability
and Mathematical Statistics

Prague 2022

Contents

Preface

Spatial statistics has established its importance as a separate scientific field in the past few decades due to the increasing availability and the growing need to analyze spatially structured data.[1] In the classical statistics, the inference is drawn from a realization of a random sample (n-tuple of independent and identically distributed random variables) and makes strong use of the independence of observations. On the other hand, it is typical in spatial statistics to observe only a single realization of a random spatial object (random field, point process) exhibiting spatial dependencies. The lack of independent replications of the random object makes formal inference more challenging than in the context of classical statistics and justifies the need for development of new, tailor-made methods. This is similar to the analysis of time series data where often a single realization of a time series is observed, exhibiting temporal dependencies. However, the lack of natural ordering in the spatial setting prevents one from using the methods developed in the temporal context in a straightforward manner.

Parametric methods in statistics require specifying an appropriate model for the data and provide means of inference for the model parameters. However, they are prone to suffer from issues stemming from possible model misspecification or poor model fit. Also, verifying model assumptions is often difficult or impossible. The relevant choice of the model often needs to be based on expert knowledge of the underlying problem. On the other hand, if the model is capable of describing the data well, parametric tests usually exhibit higher powers than their nonparametric counterparts.

This thesis focuses on nonparametric tests of several important hypotheses in statistics for point processes and random fields. The tests are based on the Monte Carlo principle and use different ways of obtaining the required Monte Carlo replications. Chapter 1 provides an introduction to the necessary concepts of spatial statistics and Monte Carlo testing. In Chapter 2 permutation tests are discussed in different settings, ranging from random samples to spatio-temporal setting. Similarly, Chapter 3 focuses on random shift tests, now predominantly in the spatial setting. Finally, Chapter 4 provides a short discussion of tests based on stochastic reconstruction.

This thesis summarizes the results obtained in the following five papers and provides the relevant context:

[1] T. Mrkvička, J. Dvořák, J.A. González, J. Mateu (2021): Revisiting the random shift approach for testing in spatial statistics. *Spatial Statistics* 42, 100430.

[2] M. Ghorbani, N. Vafaei, J. Dvořák, M. Myllymäki (2021): Testing the first-order separability hypothesis for spatio-temporal point patterns. *Computational Statistics and Data Analysis* 161, 107245.

[3] J. Dvořák, T. Mrkvička (2022): Graphical tests of independence for general distributions. *Computational Statistics* 37, 671–699.

[1] In this thesis, the term *spatial statistics* is used in a slightly broader sense, also covering the inference for spatio-temporal data.

[4] J. Dvořák, T. Mrkvička, J. Mateu, J.A. González (2022): Nonparametric testing of the dependence structure among points-marks-covariates in spatial point patterns. *International Statistical Reviews* 90, 592–621.

[5] J. Dvořák, T. Mrkvička (2022): Nonparametric testing of the covariate significance for spatial point patterns under the presence of nuisance covariates. *Submitted.*

The motivation for this work came from the problem of testing the independence of marks and a covariate in the marked point process setting, eventually solved in paper [4] using the random shift approach. In the early stages of the project, permutation-based tests were considered instead, and the ideas turned out to be useful and relevant in the setting of classical statistics (test of independence in a bivariate random vector, paper [3]) and spatio-temporal point processes (test of the first-order separability hypothesis, paper [2]).

For the initial problem of testing independence between marks and a covariate, the permutation-based tests were not successful, and hence the random shift tests were investigated. Finding a relevant test turned out to be challenging due to possible preferential sampling issues, i.e., possible dependence between points of the process and the covariate. In the end, it is reasonable to investigate the complete dependence structure among the points, marks, and covariate simultaneously, even if the scientific interest lies specifically in the relationship between the marks and the covariate. This is covered in the paper [4].

During the course of the project, a new type of correction for the random shift tests was developed that remedies some important issues connected to the traditional torus correction (liberality of the tests and the necessity to work with rectangular observation windows). This new *variance correction* is presented in paper [1] for two classical problems: testing independence between a pair of random fields and testing independence between components in a bivariate point process.

Extending the ideas even further, the paper [5] develops a new methodology for testing the independence between a point process and a covariate of interest, taking into account the possible influence of nuisance covariates. To achieve this goal, several new concepts are defined: nonparametric residuals for point processes, covariate-weighted residual measure, and (partial) correlation coefficient between a point process and a covariate. These tools provide a fully nonparametric solution to the important problem of covariate selection.

Apart from testing, several other papers of the applicant address different problems in nonparametric spatial statistics. First of all, Koňasová and Dvořák (2021a) develop a stochastic reconstruction procedure for inhomogeneous point processes. It provides the means for obtaining the Monte Carlo replications used in [2] for inhomogeneous space-time cluster processes. The approach is based on minimization of the so-called energy functional. The energy functional provides a tool for quantifying the dissimilarity between a pair of point patterns based on their summary characteristics rather than precise point locations. Once such a dissimilarity measure is available, classification for replicated point patterns can be performed using the kernel regression approach (Pawlasová and Dvořák, 2022). Further directions where these ideas can be applied are discussed in Koňasová and Dvořák (2021b).

To facilitate the applicability of the developed methods, it is important to provide the potential users with a simple means of using them. For this reason, the computer codes that implement the methods proposed in all the papers contained in this thesis are publicly available on the website `https://www2.karlin.mff.cuni.cz/~dvorak/software.html`. Furthermore, parts of the codes are already available in the `R` package `GET` (Myllymäki et al., 2017; Myllymäki and Mrkvička, 2019).

1. Introduction

1.1 Basic definitions

Point processes play a central role in this thesis. Hence, we start with a short overview of some basic definitions from the theory of point processes. We focus on the point processes in $\mathbb{R}^d, d \geq 1$. The corresponding concepts can be developed for space-time point processes by considering $\mathbb{R}^d \times \mathbb{R}$ instead. We also consider simple point processes only, i.e., two points of the process cannot occur at the same location. This material is covered by the classical textbooks such as Daley and Vere-Jones (2008), Illian et al. (2008) or Møller and Waagepetersen (2004).

In the following, the symbol $\mathcal{B}^d$ denotes the set of Borel sets in $\mathbb{R}^d$ and $\mathcal{B}_0^d$ denotes the set of bounded Borel sets in $\mathbb{R}^d$. For a set $A \subset \mathbb{R}^d$, the cardinality of A is denoted $\#A$. The formal definition of a point process is given below. Informally, a point process is a collection of points randomly scattered in space.

Definition 1. *Let $(\Omega, \mathcal{A}, \mathbb{P})$ be an abstract probability space and $\mathcal{N}$ be the system of locally finite subsets of $\mathbb{R}^d$, i.e. $\mathcal{N} = \left\{N \subset \mathbb{R}^d : \#(N \cap B) < \infty \; \forall B \in \mathcal{B}_0^d\right\}$. Let $\mathcal{N}$ be equipped with the σ-algebra $\mathfrak{N} = \sigma\left\{U_{B,m} : m \in \mathbb{N}_0, B \in \mathcal{B}_0^d\right\}$ where $U_{B,m} = \{N \in \mathcal{N} : \#(N \cap B) = m\}$. The **point process** X **in** $\mathbb{R}^d$ is a measurable mapping $X : (\Omega, \mathcal{A}, \mathbb{P}) \to (\mathcal{N}, \mathfrak{N})$. Sample realization of X is called a **point pattern**.*

Equivalently, X can be viewed as a random locally finite counting measure satisfying $X(\{y\}) \leq 1 \; \forall y \in \mathbb{R}^d$ by identifying a locally finite set $A \subset \mathbb{R}^d$ with the measure $\sum_{x \in A} \delta_x$, where δ_x is the Dirac measure with the atom at the point $x \in \mathbb{R}^d$. As is often the case in the point process literature, in the following we will use both the set formalism (to write e.g. $x \in X$ to indicate that a point of the process X occurs at location x) and the measure formalism (to write e.g. $X(A) > 0$ to indicate that the process X has at least one point in the set A) to keep the notation as simple and comprehensible as possible.

Traditionally, point processes are defined on the entire $\mathbb{R}^d$. However, observations are made almost exclusively in a bounded observation window $W \subset \mathbb{R}^d$ with positive finite volume $|W|$, implying that $X(W)$ (the number of points of the process in W) is finite almost surely. The observed point pattern can then be represented as $\{x_1, \ldots, x_n\} \subset W$.

Definition 2. *A point process X on $\mathbb{R}^d$ is called **stationary** if its distribution is invariant with respect to the translations in $\mathbb{R}^d$, i.e. the distribution of the shifted process $X + y = \{x + y : x \in X\}$ is the same as the distribution of X for all $y \in \mathbb{R}^d$.*

Definition 3. *Let X be a point process on $\mathbb{R}^d$. Its k**th-order factorial moment measure** α_k is defined as*

$$\alpha_k(A) = \mathbb{E}\left(\sum_{x_1,\ldots,x_k \in X}^{\neq} \mathbf{1}\left[(x_1, \ldots, x_k) \in A\right]\right), \; A \in \left(\mathcal{B}^d\right)^k,$$

where $\mathbf{1}$ *is the indicator function and* $\neq$ *denotes that the summation is over* k*-tuples of distinct points of* X*. The measure* $\alpha_1(\cdot) = \mathbb{E}X(\cdot)$ *is called the* ***intensity measure****.*

Definition 4. *Consider a point process* X *on* $\mathbb{R}^d$*. If its* k*th-order factorial moment measure* α_k *has a density with respect to the Lebesgue measure on* $\left(\mathbb{R}^d\right)^k$*, it is denoted* λ_k *and called the* k***th-order product density of*** X *or the* k***th-order intensity function of*** X*. The first-order intensity function* λ_1 *is called the* ***intensity function*** *and denoted* λ*.*

Definition 5. *Let* X *be a point process on* $\mathbb{R}^d$*. If both* λ *and* λ_2 *exist, we define the* ***pair-correlation function*** g *by the formula*

$$g(x, y) = \frac{\lambda_2(x, y)}{\lambda(x)\lambda(y)},\ x, y \in \mathbb{R}^d : \lambda(x) > 0, \lambda(y) > 0.$$

It is often the case that some additional information about points is available, in addition to their location. The examples include tree species, tree height, area burned by a wildfire, magnitude of an earthquake, etc. Such information is called a mark, and such datasets can be modelled by so-called marked point processes.

Definition 6. *Let* $\mathbb{M}$ *be a complete separable metric space, called the mark space. A* ***marked point process*** X_m *is a point process on the product space* $\mathbb{R}^d \times \mathbb{M}$ *with the additional property that* $X_m(\cdot \times \mathbb{M})$ *is a point process on* $\mathbb{R}^d$*, meaning that only finitely many points of the process (disregarding the marks) occur in bounded Borel subsets of* $\mathbb{R}^d$*.*

A marked point process consists of random points (x, m_x), where $x \in \mathbb{R}^d$ is the location and $m_x \in \mathbb{M}$ is the mark. The additional property ensures that the underlying ground process (the process of point locations only) is a well-defined point process. Note that this definition of a marked point process requires the concept of point processes on more general spaces than $\mathbb{R}^d$, see e.g. Daley and Vere-Jones (2008). For simplicity, we consider only point processes on Euclidean spaces in Definition 1.

Definition 7. *Let* D *be a fixed subset of* $\mathbb{R}^d$*. A* ***random field*** $\{Z(u), u \in D\}$ *is a collection of real-valued random variables defined on the same abstract probability space* $(\Omega, \mathcal{A}, \mathbb{P})$*.*

Random fields are used in this work either to define clustered point process models such as log-Gaussian Cox processes, or to describe the covariate information accompanying a point process where available. Examples of covariate information include the terrain elevation, slope, mineral content in the soil, or distance to a particular landmark. For further reading on random fields, the books Cressie (1993) or Adler and Taylor (2007) provide a lot of interesting material.

1.2 Monte Carlo tests

When the distribution of a test statistic is too complicated to be derived analytically but there is a way of obtaining replications of the data under the null

hypothesis, it is possible to perform a formal test of the null hypothesis using the Monte Carlo approach (Davison and Hinkley, 1997). Monte Carlo tests compare the value of the test statistic T_0 computed from the observed data with a set of values $T_1, \ldots, T_N$ obtained from a set of replications. These are produced so that they follow the null hypothesis H_0, and hence the values $T_1, \ldots, T_N$ approximate the distribution of T_0 under H_0. The test outcome is then determined based on how typical or extreme the value of T_0 is among $T_0, T_1, \ldots, T_N$. The test statistic can be univariate, multivariate, or functional. The Monte Carlo p-value is

$$p = \frac{1}{N+1} \sum_{i=0}^{N} \mathbf{1}(T_i \prec T_0),$$

where $\mathbf{1}$ is the indicator function and where $T_i \prec T_0$ indicates that T_i is more extreme than (or the same as) T_0 in a certain ordering. For univariate test statistics, this means determining the rank of T_0. However, using multivariate or functional test statistics is also possible if a suitable ranking of the vectors or functions from the most typical to the most extreme is available, as e.g. in Myllymäki et al. (2017). Other means of ranking of vectors or functions can be employed, too, e.g. those based on (functional) data depth.

In order to be exact, Monte Carlo tests rely on the exchangeability (invariance of the distribution with respect to permutations of the components) of the vector $(T_0, T_1, \ldots, T_N)$. This is easily achieved if the replications are independent of the observed data and are independent, identically distributed under H_0. Whenever independent replications are not available, as is often the case when nonparametric methods are employed, a replication strategy must be chosen to achieve exchangeability. If this is not possible, the aim is to make the deviation from exchangeability as small as possible.

Since the exchangeability of $(T_0, T_1, \ldots, T_N)$ implies a uniform distribution of the p-value, potential deviations from the exchangeability can be detected in carefully designed simulation experiments studying the size of the test. Different tests of the same hypothesis can be compared in this way, with more severe deviations from the uniform distribution of p-values implying more severe deviations from exchangeability and hence inferiority of the test.

In classical statistics a typical way of obtaining the Monte Carlo replications is by simulation from a fitted parametric model. Alternatively, different permutation schemes can be applied when nonparametric approach is considered.

1.3 Monte Carlo tests in spatial statistics

Just as in the context of classical statistics, in spatial statistics it is possible to obtain Monte Carlo replications by simulation from a fitted parametric model. This provides means of goodness-of-fit testing. However, this thesis focuses on nonparametric methods, and there are several ways of producing the Monte Carlo replications in a nonparametric manner, depending on the particular context.

1.3.1 Permutation-based tests

One of the ways is to randomly permute the observed values. By generating N independent random permutations, the required number of replications is obtained,

in turn producing the test statistic values $T_1, \ldots, T_N$ to be compared with the value T_0 computed from the observed data in the Monte Carlo test.

One of the most prominent examples of this approach is the test of the independent marking hypothesis for marked point processes (Illian et al., 2008, Sec. 7.5). In this case, the null hypothesis states that the marks are independent, identically distributed random variables, independent of the ground process. The test statistic can be e.g. the mark-weighted K-function (Illian et al., 2008, Sec. 5.3.3). If the observed marked point pattern is $\{(x_1, m_{x_1}), \ldots, (x_n, m_{x_n})\}$ and π is a permutation on the set $\{1, \ldots, n\}$, the permuted sample is then obtained as $\{(x_1, m_{x_{\pi(1)}}), \ldots, (x_n, m_{x_{\pi(n)}})\}$. Generating N independent random permutations then produces the required Monte Carlo replications. Under the null hypothesis, the distribution of the data does not change by the permutation of marks.

Marked point processes are often accompanied by additional spatial covariates that may influence the distribution of points and/or marks. When investigating the problem of testing the independence of marks and a covariate in marked point process setting, the permutation strategy presented in the previous paragraph was considered as a starting point. Assuming that a realization of the covariate $C(x)$ is observed in the observation window W (where the marked point process is observed), the covariate values at the observed points can be added to the sample: $\{(x_1, m_{x_1}, C(x_1)), \ldots, (x_n, m_{x_n}, C(x_n))\}$.

Focusing on the relationship between the marks and the covariate only, one could focus on the sample $\{(m_{x_1}, C(x_1)), \ldots, (m_{x_n}, C(x_n))\}$ and perform a permutation-based Monte Carlo test in this bivariate sample. The test statistic might be e.g. the Pearson's correlation coefficient or the bivariate empirical cumulative distribution function. However, this strategy has only limited use since the permutations break the spatial structure of the observed phenomena and change the distribution of the test statistic, thus violating the requirement of exchangeability. It is only relevant to apply such a test for independent, identically distributed marks, which is very limiting in practice. Also, even in this case the possible dependence between the point process and the covariate might lead to preferential sampling issues (Diggle et al., 2010), invalidating the outcome of the test.

The problem of testing independence between marks and a covariate was eventually solved using the random shift approach in the paper Dvořák et al. (2022a), where the whole dependence structure between points, marks, and covariate is studied simultaneously, see Section 3.4. On the other hand, the idea of using the permutation-based test of independence in a bivariate sample, with the empirical cumulative distribution function as the test statistic and with the global envelope test indicating the significant regions (see Section 1.4 for details), turned out to be interesting on its own also in the context of classical statistics based on independent observations. These ideas were developed in the paper Dvořák and Mrkvička (2022a), see Section 2.1.

1.3.2 Random shift tests

Random shifts provide means of nonparametric testing of independence between a pair of spatial objects, such as a pair of random fields (Fortin and Payette, 2002) or a pair of point processes (Lotwick and Silverman, 1982). By randomly shifting one of the objects while keeping the other one fixed, any possible dependence

between them is broken. At least one of the spatial objects must be assumed to be stationary. By performing a certain amount of shifts along randomly generated vectors, one obtains the replications for performing a Monte Carlo test of independence.

Assume that the spatial objects are denoted by Φ and Ψ and they are observed in the window W. The value of the test statistic computed directly from the observed data is denoted by $T_0 = T(\Phi, \Psi; W)$. After producing N random shift vectors $v_1, \ldots, v_N$, the value of the test statistic T_i is computed from Φ and Ψ shifted by v_i, i.e. $T_i = T(\Phi, \Psi + v_i; W)$, $i = 1, \ldots, N$. Clearly, some part of Ψ will be shifted outside of the observation window W and part of $\Psi + v_i$ will no longer overlap with Φ anymore. Hence, some form of correction is needed.

Torus correction

For a rectangular observation window W, one may identify its opposing edges, creating a toroidal geometry on W (Lotwick and Silverman, 1982; Upton and Fingleton, 1985). The version of Ψ shifted with respect to the toroidal geometry is denoted by $[\Psi + v_i]$ in the following, as opposed to $\Psi + v_i$ which denotes Ψ shifted with respect to the Euclidean geometry. The replications T_i are then obtained as $T_i = T(\Phi, [\Psi + v_i]; W)$, $i = 1, \ldots, N$.

As a result, all parts of the data are used for computing T_i. On the other hand, artificial cracks appear in the correlation structure of the data, as parts of the data originally far away are now "glued together". This means that exchangeability is violated, which in turn introduces liberality of the random shift tests (Fortin and Payette, 2002; Mrkvička et al., 2021). However, simulation studies show that when the spatial autocorrelations in the data are not very strong, the tests match the nominal significance level quite closely (Mrkvička et al., 2021; Dvořák et al., 2022a). Traditionally, the distribution of the random shift vectors is taken to be the uniform distribution on W, but other choices are also possible.

Variance correction

To remove the liberality of the torus correction, Mrkvička et al. (2021) propose the *variance correction*. It uses shifts respecting the Euclidean geometry and discards those parts of the data that are shifted outside of W. No artificial cracks are introduced to the correlation structure of the data, removing the liberality of the random shift tests. Also, irregular observation windows can be considered. On the other hand, different amounts of data are dropped for different shift vectors v_i and for typical choices of the test statistic the variance of T_i varies greatly, making it impossible to perform the Monte Carlo test directly. Therefore, the variance of T_i needs to be standardized before performing the test.

Formally, let W_i denote the smaller observation window where Φ and $\Psi + v_i$ overlap, i.e. $W_i = W \cap (W + v_i)$. The value T_i is computed from Φ and $\Psi + v_i$ restricted to W_i, denoted $\Phi|_{W_i}$ and $(\Psi + v_i)|_{W_i}$, respectively. Specifically, $T_i = T(\Phi|_{W_i}, (\Psi + v_i)|_{W_i}; W_i)$. The values $T_0, T_1, \ldots, T_N$ are then standardized to have zero mean and unit variance. This is achieved by subtracting the mean $\overline{T} = \frac{1}{N+1} \sum_{i=0}^{N} T_i$ and dividing by the square root of the variance: $S_i = \left(T_i - \overline{T}\right) / \sqrt{\mathrm{var}(T_i)}$. The standardized values $(S_0, S_1, \ldots, S_N)$ are closer to

exchangeability than $(T_0, T_1, \ldots, T_N)$ because their first two moments are the same. The standardized values are used to perform the Monte Carlo test. When a formula describing $\mathrm{var}(T_i)$ as a function of the size of W_i is known, at least asymptotically, it can be directly used in the standardization. If such a formula is not available, Mrkvička et al. (2021) suggest a kernel regression approach to estimating $\mathrm{var}(T_i)$.

Simulation studies in Mrkvička et al. (2021), Dvořák et al. (2022a), or Dvořák and Mrkvička (2022b) show that the random shift tests with variance correction match the nominal significance level even in the case of strong autocorrelation. In those papers, the shift vectors followed the uniform distribution on a disc with radius R centered at the origin. The choice of R is a compromise between two goals: longer shifts are more relevant for breaking the possible dependence between Φ and Ψ while shorter shifts mean that a larger amount of available data is used to compute T_i. Choosing R so that $|W_i|/|W| \geq 1/4$ for all i turned out to provide satisfactory results.

1.3.3 Tests based on stochastic reconstruction

Stochastic reconstruction provides a way to obtain replications with the same properties as the observed point pattern. While Tscheschel and Stoyan (2006) introduced the stochastic reconstruction procedure for stationary point processes, Wiegand et al. (2013) and Koňasová and Dvořák (2021a) adapted the procedure for inhomogeneous point processes. The properties to be preserved during the reconstruction are determined by the summary characteristics incorporated into the so-called *energy functional* and must be chosen by the user.

The energy functional quantifies the dissimilarity between the observed point pattern and another, arbitrary pattern. A reconstructed pattern (output pattern) is found by minimization of the energy functional, performed by iteratively proposing and accepting/rejecting small updates to the current candidate pattern. The procedure can be repeated N times to provide the required number of replications for the Monte Carlo test.

When using the stochastic reconstruction to provide replications for a Monte Carlo test of a certain hypothesis, the energy functional must be specified carefully so that the algorithm indeed produces outputs that fulfill the particular null hypothesis. Examples include the isotropy testing in Wong and Chiu (2016) or testing the separability hypothesis for the intensity function of space-time point processes in Ghorbani et al. (2021).

1.4 Global envelope tests

In recent years, the global envelope tests (Myllymäki et al., 2017) have become an industry standard in spatial statistics for performing Monte Carlo tests based on vector or functional test statistics. Their main benefits are in their graphical interpretation which helps guide further inference in case of rejection and the fact that they remedy a severe flaw of the pointwise envelopes, well-established in spatial statistics for decades, by properly treating the multiple testing problem.

Consider now a d-dimensional test statistics $T_i = (T_{i1}, \ldots, T_{id})$, $i = 0, \ldots, N$, where T_0 is computed from the observed data and $T_1, \ldots, T_N$ are computed from

the Monte Carlo replications. The ordering of the vectors $T_0, \ldots, T_N$ is based on the pointwise ranks R_{ik}, constructed below, of the element T_{ik} among the corresponding elements $T_{0k}, T_{1k}, \ldots, T_{Nk}$ of the $N+1$ vectors such that the lowest ranks correspond to the most extreme values of the statistics. Let $r_{0k}, r_{1k}, \ldots, r_{Nk}$ be the raw ranks of $T_{0k}, T_{1k}, \ldots, T_{Nk}$, such that the smallest value has rank 1 and the largest value is assigned rank $N+1$. In the case of ties, the raw ranks are averaged. The two-sided pointwise ranks are then calculated as

$$R_{ik} = \min(r_{ik}, N + 1 - r_{ik}).$$

The extreme rank length (ERL) measure defined in Myllymäki et al. (2017) induces the desired ordering. Roughly speaking, it ranks the vectors according to the number of the most extreme elements. Formally, the ERL measure of T_i is defined based on the vector of the pointwise ordered ranks $\mathbf{R}_i = (R_{i[1]}, R_{i[2]}, \ldots, R_{i[d]})$, where the ranks are arranged from smallest to largest, i.e., $R_{i[k]} \leq R_{i[k']}$ whenever $k \leq k'$. The ERL measure of T_i is defined as

$$E_i = \frac{1}{N+1} \sum_{i'=0}^{N} \mathbf{1}(\mathbf{R}_{i'} \prec \mathbf{R}_i) \tag{1.1}$$

where

$$\mathbf{R}_{i'} \prec \mathbf{R}_i \quad \Longleftrightarrow \quad \exists\, n \leq d : R_{i'[k]} = R_{i[k]} \forall\, k < n, \text{ and } R_{i'[n]} < R_{i[n]}.$$

The division by $N+1$ leads to normalized ranks that attain values between 0 and 1. Consequently, the ERL measure corresponds to the extremal depth of Narisetty and Nair (2016).

The probability of having a tie in the ERL measure is rather small in practical scenarios, thus the ERL effectively solves the ties problem which often appears in ordering of multivariate vectors using ranks. The final p-value of the Monte Carlo test is

$$p_{\text{erl}} = \frac{1}{N+1} \sum_{i=0}^{N} \mathbf{1}(E_i \leq E_0).$$

For a given $\alpha \in (0, 1)$, let $e_\alpha \in \mathbb{R}$ be the largest of the E_i such that the number of those i for which $E_i < e_{(\alpha)}$ is less than or equal to αN. Furthermore, let $I_\alpha = \{i \in 0, \ldots, N : E_i \geq e_{(\alpha)}\}$ be the index set of vectors less or as extreme as e_α. Then, the $100(1-\alpha)\%$ global ERL envelope induced by E_i is

$$T^{(\alpha)}_{\text{low}\,k} = \min_{i \in I_\alpha} T_{ik} \quad \text{and} \quad T^{(\alpha)}_{\text{upp}\,k} = \max_{i \in I_\alpha} T_{ik} \quad \text{for } k = 1, \ldots, d,$$

see Narisetty and Nair (2016) and Mrkvička et al. (2020).

The $100(1-\alpha)\%$ global ERL envelope $[T^{(\alpha)}_{\text{low}\,k}, T^{(\alpha)}_{\text{upp}\,k}], k = 1, \ldots, d$, has an important *intrinsic graphical interpretation* property, i.e. for every $i = 0, \ldots, N$:

1. $T_{ik} < T^{(\alpha)}_{\text{low}\,k}$ or $T_{ik} > T^{(\alpha)}_{\text{upp}\,k}$ for some $k = 1, \ldots, d \iff E_i < e_{(\alpha)}$,
2. $T^{(\alpha)}_{\text{low}\,k} \leq T_{ik} \leq T^{(\alpha)}_{\text{upp}\,k}$ for all $k = 1, \ldots, d \iff E_i \geq e_{(\alpha)}$.

This means that the outcome of the test may be equivalently determined by the data curve T_0 leaving/not leaving the envelope or by comparing p_{erl} with the nominal significance level α.

2. Permutation tests

In this chapter, we provide a brief overview of two papers, included in this thesis, which use permutation-based Monte Carlo tests:

- Dvořák and Mrkvička (2022a) which proposes new independence tests in the context of classical statistics, see Section 2.1,
- Ghorbani et al. (2021) which proposes tests of the first-order separability hypothesis for space-time point processes, see Section 2.2.

2.1 Tests for bivariate random samples

The problem of testing the null hypothesis of independence of two random variables is essential in statistics. There are many methods to discover the possible dependence. Still, all of them concentrate on summarizing the information over the whole distribution and as such cannot detect the combinations of quantiles where a deviation from independence occurs. Our aim is to introduce general tests of independence (applicable to arbitrary bivariate distributions without any assumptions, i.e. for continuous and discrete distributions, mixtures of those, distributions with heavy tails, etc.) which are graphical in nature (providing a two-dimensional visual output where the combinations of quantiles causing the possible rejection are indicated). All details are given in the paper Dvořák and Mrkvička (2022a), and a summary is given below.

Consider a bivariate random vector (X, Y) and the null hypothesis of independence of X and Y which needs to be tested. Assume that a sample $\{(x_1, y_1), \dots, (x_n, y_n)\}$ is observed. We propose two multivariate test statistics to be used in connection with the global envelope test described in Section 1.4 of this thesis.

The first test statistic is the sample cumulative distribution function (CDF), computed on a fine grid of points. This test statistic is of cumulative nature, and as such it is sensitive to small departures from the null model which accumulate into a possibly significant departure.

The second test statistic provides a local test with easier interpretation. It is the kernel estimate of the intensity of point occurrence in the two-dimensional QQ-plot, constructed as follows. Let $\mathbb{Q}$ denote the two-dimensional point pattern $\mathbb{Q} = \{(q_1^1, q_1^2), \dots, (q_n^1, q_n^2)\}$, where q_i^1 is the sample quantile corresponding to the value x_i in the sample $x_1, \dots, x_n$, and similarly q_i^2 is the sample quantile corresponding to the value y_i in the sample $y_1, \dots, y_n$. $\mathbb{Q}$ is a point process in the observation window $[0, 1]^2$.

The intensity of point occurrence in the process $\mathbb{Q}$, including its local variations, can be estimated by kernel smoothing, similarly to estimating the intensity function of a point process or a probability density function of a random vector. The constant intensity of point occurrence in $\mathbb{Q}$ indicates independence. In contrast, the accumulation of points in a certain area indicates a steeper increase of the CDF values at the given combination of quantiles, meaning observations with such a combination of values are more likely to appear than what is expected under independence. On the other hand, the absence of points in a certain area

indicates a less steep increase of the CDF values at the given combination of quantiles, meaning observations with such a combination of values are less likely to appear than what is expected under independence.

The second proposed test statistic $\hat{g}(x, y)$ is motivated by the nonparametric kernel estimator of the intensity function of a point process (Illian et al., 2008), and is defined as

$$\hat{g}(x, y) = \frac{1}{e(x, y)} \sum_{i=1}^{n} k_\sigma((x, y) - (q_i^1, q_i^2)),$$

where k_σ is a probability density function on $\mathbb{R}^2$, usually called the kernel function with bandwidth σ, and $e(x, y) = \int_{[0,1]^2} k_\sigma((x, y) - (x', y'))\,\mathrm{d}x'\,\mathrm{d}y'$ is the correction for bias due to edge effects. The edge effects, often appearing in statistics for point processes, are in this context closely related to the problem of bounded support in kernel estimation of a probability density function in classical statistics.

Note that the proposed test statistic is defined as a two-dimensional kernel estimate, and hence it is appropriate for continuous distributions for which the probability of ties appearing in any coordinate of the observed data points is 0, implying that $\mathbb{Q}$ is a simple point process. If an atom appears in the distribution of one or both of the marginals or the data are purely categorical, appropriate adjustments need to be made, see Section 3 of Dvořák and Mrkvička (2022a).

The Monte Carlo replications are obtained by permuting the order of $y_1, \ldots y_n$. Specifically, the test statistic value $T_i, i = 1, \ldots, N$, is computed from the sample $\{(x_1, \pi_i(y_1)), \ldots, (x_n, \pi_i(y_n))\}$, where $\pi_1, \ldots, \pi_N$ are independent random permutations. This permutation scheme secures exchangeability under the null hypothesis, therefore, any such permutation test will achieve the prescribed significance level α, as given in the following theorem.

Theorem 1. *Let T_0 be the test statistic computed from the observed data and $T_i, i = 1, \ldots, N$, be the test statistics computed from randomly permuted observations. Furthermore, let '$\prec$' be an ordering for which ties occur with probability 0 and let*

$$p = \frac{1}{N + 1} \sum_{i=0}^{N} \mathbf{1}(T_i \prec T_0),$$

where we stress that in the notation established in Section 1.2 of this thesis, it holds that $T_i \prec T_i$. Let $\alpha \in (0, 1)$ be the nominal significance level and let the test reject the null hypothesis of independence if and only if $p \leq \alpha$. Assuming that $\alpha(N + 1)$ is an integer, the test rejects the null hypothesis at the prescribed significance level α.

We remark that in the ordering discussed in Section 1.4 of this thesis ties can occur with positive probability. This can be remedied, e.g., by breaking the ties at random.

The choice of the number of permutations N is important for the practical application of the test. Under the alternative hypothesis, the power of the test is expected to increase with an increasing number of permutations N. Based on the approach of Oden (1991) and Marozzi (2016, Sec. 3), we investigate in the following theorem how the power of the test depends on N.

Theorem 2. *Assume that a particular hypothesis H_1 holds. Let T_0 be the test statistic computed from the observed data and $T_i, i = 1, \ldots, N$, be the test statistics computed from randomly permuted observations. Conditionally on the observed data, let P_d denote the probability that $T_1 \prec T_0$. Let the (unconditional) distribution of P_d be given by the distribution function $W_{H_1}(t) = \mathbb{P}(P_d \leq t), t \in [0, 1]$. The true power P_N of the test is then given by*

$$P_N = \int_0^1 \sum_{k=0}^{\lfloor \alpha(N+1)-1 \rfloor} \binom{N}{k} t^k (1-t)^{N-k} \, \mathrm{d}W_{H_1}(t),$$

where $\lfloor u \rfloor$ denotes the largest integer which is smaller than or equal to u.

The simulation study in Section 4 of Dvořák and Mrkvička (2022a) provides insight into the performance of the tests and indicates in which settings the CDF test performs well (departures from independence appearing over the whole range of data values) and in which the QQ test performs well (localized departures from independence). Hence, the two tests are complementary in a sense, and the user can choose which test to use based on prior assumptions about the dataset at hand.

The simulation study also provides a comparison with several benchmark independence tests. As generally accepted, no test of independence is uniformly better than others. The main benefit of our tests lies in the graphical interpretation of the outcome of the tests and their generality, not in outperforming all others.

2.2 Tests for space-time point processes

Analysis of space-time point patterns presents a challenge due to higher dimensionality and due to the temporal coordinate playing a distinct role. The first-order separability (product structure of the space-time intensity function) is often a convenient working assumption which greatly simplifies the inference (Gabriel and Diggle, 2009; Møller and Ghorbani, 2012; Møller et al., 2019). In such a case, the inference can be often based on the lower-dimensional component processes, i.e. the spatial projection process and the temporal projection process (Møller and Ghorbani, 2012; Prokešová and Dvořák, 2014; Dvořák and Prokešová, 2016).

In practice, the first-order separability hypothesis is often not fulfilled. Therefore, a formal test of this hypothesis is needed to decide whether the inference for a particular dataset can be based on the first-order separability assumption or not. It is rather straightforward to construct a permutation-based test of this hypothesis where the time coordinates of the observed points are randomly permuted. All the details are given in the paper Ghorbani et al. (2021), and a summary is given below. Such a test is exact for inhomogeneous space-time Poisson processes but does not perform well for processes with interactions. For such processes, Ghorbani et al. (2021) propose a test based on the stochastic reconstruction procedure which is discussed in Chapter 4 of this thesis. Also, various visualisation techniques for informal assessment of the first-order separability hypothesis are proposed in the online Supplementary material to the paper

Ghorbani et al. (2021).[1]

Formally speaking, a space-time point process X is a point process on the space $\mathbb{R}^d \times \mathbb{R}$. To be consistent with the paper Ghorbani et al. (2021) we consider the case $d = 2$ only. For each point $(u, t) \in X$ the spatial location is given by $u \in \mathbb{R}^2$ and the temporal coordinate is given by $t \in \mathbb{R}$. The points of a space-time point process are traditionally called *events*. In practice, X is observed within a spatio-temporal window $W \times T$, where $W \subset \mathbb{R}^2$ is a bounded region of area $|W| > 0$, $T \subset \mathbb{R}$ is a bounded time interval of length $|T| > 0$. The observed space-time point pattern is $\{(u_i, t_i),\ i = 1 \dots n\}$.

Assuming that X has the intensity function $\rho(\cdot)$, the spatial component process X_{space} consisting of the locations of events with times in T and the temporal component process X_{time} consisting of the times of events with locations in W, i.e. $X_{\text{space}} = \{u : (u, t) \in X,\ t \in T\}$ and $X_{\text{time}} = \{t : (u, t) \in X,\ u \in W\}$, are well-defined point processes on $\mathbb{R}^2$ and $\mathbb{R}$, respectively, with well-defined intensity functions. The intensity functions of these component processes are given by

$$\rho_{\text{space}}(u) = \int_T \rho(u, t)\,\mathrm{d}t,\ u \in W, \quad \rho_{\text{time}}(t) = \int_W \rho(u, t)\,\mathrm{d}u,\ t \in T.$$

Nonparametric kernel estimates of ρ_{space} and ρ_{time} are, respectively, given by

$$\hat{\rho}_{\text{space}}(u) = \sum_{i=1}^{n} k_\epsilon^2(u - u_i)/C_{W,\epsilon}(u_i), \quad u \in W,$$

and

$$\hat{\rho}_{\text{time}}(t) = \sum_{i=1}^{n} k_\delta^1(t - t_i)/C_{T,\delta}(t_i), \quad t \in T,$$

where k_b^m is an m-dimensional kernel with bandwidth $b > 0$, $m \in \mathbb{N}$, and $C_{W,\epsilon}(u_i) = \int_W k_\epsilon^2(u - u_i)\,\mathrm{d}u$ and $C_{T,\delta}(t_i) = \int_T k_\delta^1(t - t_i)\,\mathrm{d}t$ are edge correction factors in space and time, respectively. A nonseparable kernel estimator of the space-time intensity function is

$$\hat{\rho}(u, t) = \sum_{i=1}^{n} \frac{k_\epsilon^2(u - u_i)}{C_{W,\epsilon}(u_i)} \frac{k_\delta^1(t - t_i)}{C_{T,\delta}(t_i)}, \quad (u, t) \in W \times T. \tag{2.1}$$

The first-order separability hypothesis assumes that the space-time intensity function of the process has a product form:

$$\rho(u, t) = \rho_1(u)\rho_2(t), \quad (u, t) \in \mathbb{R}^2 \times \mathbb{R},$$

where ρ_1 and ρ_2 are non-negative measurable functions. Under this hypothesis, we have:

$$\rho_{\text{space}}(u) = \rho_1(u) \int_T \rho_2(t)\,\mathrm{d}t,\ u \in W, \quad \rho_{\text{time}}(t) = \rho_2(t) \int_W \rho_1(u)\,\mathrm{d}u,\ t \in T,$$

and the intensity function of X fulfills

$$\rho(u, t) = \frac{\rho_{\text{space}}(u)\rho_{\text{time}}(t)}{\int_{W \times T} \rho(v, s)\,\mathrm{d}(v, s)}, \quad (u, t) \in \mathbb{R}^2 \times \mathbb{R}.$$

[1] See also the website `https://www2.karlin.mff.cuni.cz/~dvorak/software/testing_space-time_first_order_separability/ST_separability_suppl.html`

This leads to a separable kernel estimator of the space-time intensity function

$$\hat{\rho}_{\text{sep}}(u,t) = \hat{\rho}_{\text{space}}(u)\hat{\rho}_{\text{time}}(t)/n, \quad (u,t) \in W \times T. \tag{2.2}$$

In Ghorbani et al. (2021), three test statistics are proposed for testing the first-order separability hypothesis, based on the comparison between the nonseparable estimator (2.1) and the separable estimator (2.2):

$$\begin{aligned} S(u,t) &= \frac{\hat{\rho}(u,t)}{\hat{\rho}_{\text{sep}}(u,t)},\ (u,t) \in W \times T, \\ S_{\text{time}}(t) &= \int_W S(u,t)\,\mathrm{d}u,\ t \in T, \\ S_{\text{space}}(u) &= \int_T S(u,t)\,\mathrm{d}t,\ u \in W, \end{aligned}$$

for $\hat{\rho}_{\text{sep}}(u,t) > 0$. Under the null hypothesis, these functions are expected to be approximately constant. The function $S(u,t)$ provides detailed information about possible deviations from separability, including which locations and which times are responsible for such deviations. On the other hand, the functions $S_{\text{time}}(t)$ and $S_{\text{space}}(u)$ aggregate such information, but provide more straightforward ways of visualisation due to their lower-dimensional domains.

The Monte Carlo replications are obtained by permuting the order of the observed times $t_1, \dots t_n$. Specifically, the test statistic value $T_i, i = 1, \dots, N$, is computed from the sample $\{(u_1, \pi_i(t_1)), \dots, (u_n, \pi_i(t_n))\}$, where $\pi_1, \dots, \pi_N$ are independent random permutations. Under the assumption that X is an inhomogeneous Poisson process on $W \times T$ with the intensity function $\rho(s,t)$, the permutation of times does not change the distribution of the process, as seen from the form of the density of X with respect to the unit-rate space-time Poisson process on the same domain, see Ghorbani et al. (2021, Section 4.2).

According to the simulations in Ghorbani et al. (2021), the permutation-based tests match the nominal significance level as expected. The test based on the $S(u,t)$ test statistic achieves the highest power. For processes with weak space-time clustering, the permutation-based tests matched the nominal significance level rather well, but for stronger space-time clustering, the tests exhibited a high degree of liberality. This motivates the use of the test based on the stochastic reconstruction procedure suggested in Ghorbani et al. (2021, Section 7) and discussed in Chapter 4 of this thesis.

3. Random shift tests

In this chapter, we provide a brief overview of three papers, included in this thesis, which use random shift Monte Carlo tests:

- Mrkvička et al. (2021) which introduces the variance correction for random shift tests described in Section 1.3.2, leading to a new independence test for a pair of random fields and a pair of point processes, see Sections 3.1 and 3.2,
- Dvořák et al. (2022a) which proposes a new test of independence between a point process and a covariate and a new test of independence between a marked point process and a covariate, see Sections 3.3 and 3.4,
- Dvořák and Mrkvička (2022b) which introduces the partial correlation analysis for a point process and a set of covariates, and proposes a test of independence between a point process and a covariate of interest, taking into account possible influence of nuisance covariates, see Section 3.6.

We also briefly describe two related problems which are solved by random shift tests in the works of González et al. (2022) and Veselý (2021), not included in this thesis, see Sections 3.5 and 3.7. We structure the content of this chapter according to the different problems addressed rather than according to the three papers. This provides the reader with a better overview of the work.

3.1 Testing independence of two random fields

As noted already in Fortin and Payette (2002), the torus correction may lead to liberality of the test and hence to incorrect conclusions. To remedy this issue, Mrkvička et al. (2021) introduce the variance correction to replace the torus correction. The variance correction successfully removes the liberality and can also be applied for irregular observation windows. The details are given in Section 1.3.2 of this thesis.

Assume now that Φ and Ψ are random fields observed in W. Assume also that Ψ is observed in any location, at least on a fine pixel grid. If this is not the case, kriging can be used to provide an estimate of the unobserved values. Let X be the set of sampling locations where the values of Φ are observed. X can be either random or nonrandom, depending on the design of the experiment. We denote by $\Phi(X)$ the vector of values of Φ observed at the sampling points X, and similarly for $\Psi(X)$.

For testing the independence hypothesis for the two random fields Φ and Ψ, a natural choice of the test statistic is the sample covariance

$$T = \mathrm{cov}(\Phi(X), \Psi(X)).$$

It is known in classical statistics that the variance of the sample covariance, computed under the null hypothesis of independence from a random sample of size n, is of the order $1/n$. Below, we argue that the same holds also in the case with spatial autocorrelation, see Theorem 3. Therefore, $\mathrm{var}(T_i) \approx C/n_i$ where n_i

is the number of sampling locations in $W_i, i = 0, 1, \ldots, N, W_0 = W$, and C is a constant. Thus, setting $\mathrm{var}(T_i) \approx 1/n_i$ allows one to use the variance correction for this test statistic.

Theorem 3. *Let Φ and Ψ be two independent stationary random fields on $\mathbb{R}^d$ with finite second moments and non-negative autocovariance functions C_Φ, C_Ψ. Assume that there is a constant $R > 0$ such that $C_\Phi(u - v) = C_\Psi(u - v) = 0$ for $\|u - v\| > R$. Let $X = \{x_i, i \in \mathbb{N}\}$ be a sequence of observation points such that for each point there are at most K other points within distance R. Let $s_n, n = 2, 3, \ldots,$ be the sample covariance defined as*

$$s_n = \frac{1}{n-1} \sum_{i=1}^{n} \left(\Phi(x_i) - \bar{\Phi}_n\right) \left(\Psi(x_i) - \bar{\Psi}_n\right)$$

where $\bar{\Phi}_n = \frac{1}{n}\sum_{i=1}^{n} \Phi(x_i)$ and $\bar{\Psi}_n = \frac{1}{n}\sum_{i=1}^{n} \Psi(x_i)$ are the sample means. Then $\mathbb{E}s_n = 0$ for each $n \in \mathbb{N}$ and

$$\mathrm{var}\, s_n = \frac{1}{(n-1)^2} \sum_{i=1}^{n} \sum_{j=1}^{n} C_\Phi(x_i - x_j) C_\Psi(x_i - x_j) + o(1/n)$$

and there are constants $0 < C_1 \leq C_2 < \infty$ such that

$$C_1 \leq \liminf_{n\to\infty} (n\, \mathrm{var}\, s_n) \leq \limsup_{n\to\infty} (n\, \mathrm{var}\, s_n) \leq C_2.$$

Without more specific assumptions on the positions of the observation points $X = \{x_i, i \in \mathbb{N}\}$, such as $X = \mathbb{Z}^d$, it is not possible to establish a limit for $n\, \mathrm{var}\, s_n$. On the other hand, the theorem ensures that the asymptotic order of the variance is $1/n$. Furthermore, it is not possible to drop the assumption of the bounded support of C_Φ, C_Ψ without additional assumptions on the properties of Φ, Ψ such as α-mixing.

The simulation experiments in Section 3 of Mrkvička et al. (2021) indicate that the random shift test with torus correction is in fact valid for very rough random fields with nearly independent observations. However, it becomes increasingly liberal with increasing smoothness of the random fields. For very smooth random fields, the actual significance level can be even more than twice as high as the nominal significance level. In the power study, the torus correction approach achieved the highest empirical rejection rates, but this is only a consequence of the fact that the test rejects inadequately often even under the null hypothesis (liberality of the test). Concerning the random shift tests with variance correction, very little liberality was observed but only for very smooth random fields. At the same time, the test had a high power, compared to several benchmark methods.

3.2 Testing independence of two point processes

The paper Mrkvička et al. (2021) also studied the random shift tests of independence between a pair of point processes and the possibility to apply the variance correction in place of the torus correction. Assume now that Φ and Ψ are stationary point processes observed in W. For testing the independence between

two point processes, the usual choice of the test statistic is the sample cross K-function (Illian et al., 2008) computed for a number of different ranges. Roughly speaking, the cross K-function carries information about the mean number of points of Ψ up to distance r from an arbitrary point of Φ. By considering several different values of r, we obtain a vector test statistic, and the global envelope test can be used to perform the Monte Carlo test.

The globally corrected "Ohser-type" estimator of the cross K-function is of the form (Illian et al., 2008, p.230)

$$\widehat{K}_{12}(r) = \frac{c(r)}{\widehat{\lambda}_\Phi \widehat{\lambda}_\Psi} \sum_{x \in \Phi \cap W} \sum_{y \in \Psi \cap W} \mathbb{I}(\|x - y\| \leq r), \ r > 0,$$

where $\widehat{\lambda}_\Phi$ and $\widehat{\lambda}_\Psi$ are estimated intensities of Φ and Ψ, i.e. $\widehat{\lambda}_\Phi = \sum_{x \in \Phi \cap W} \frac{1}{|W|}$ and similarly for $\widehat{\lambda}_\Psi$, and $c(r)$ is an edge correction factor given by

$$c(r) = \frac{\pi r^2}{\Gamma_W(r)},$$

where

$$\Gamma_W(r) = \int_W \int_W \mathbb{I}(\|x - y\| \leq r) \, \mathrm{d}x \, \mathrm{d}y = 2\pi \int_0^r t \bar{\gamma}_W(t) \, \mathrm{d}t,$$

with $\bar{\gamma}_W(t)$ being the isotropized set covariance function of W.

The variance of $\widehat{K}_{12}(r)$ was studied for Poisson processes in Rajala et al. (2019) under the assumption of a fixed number of observed points of Φ and Ψ in W. The paper also presents a formula for general non-Poisson processes, but without a proof. The following theorem provides the first- and second-order moment properties of the quantities used to estimate the variance of $\widehat{K}_{12}(r)$ in the case of a random number of points in W. The theorem can be proved by using the appropriate versions of the Campbell theorem, which are given in the appendix of Mrkvička et al. (2021). This allows one to use the variance correction for the test statistic $\widehat{K}_{12}(r)$, with the standardization performed separately for each considered value of r.

Theorem 4. *Let Φ and Ψ be two independent stationary point processes on $\mathbb{R}^d$ with intensities λ_1 and λ_2 and pair-correlation functions g_1 and g_2, respectively. Let W be the observation window where both processes are observed. For a given $r > 0$, let $f_r : \mathbb{R}^d \to [0, \infty)$ be a Borel function and denote*

$$R = \sum_{x \in \Phi \cap W} \sum_{y \in \Psi \cap W} f_r(x - y).$$

Let

$$\widehat{\lambda}_1 = \sum_{x \in \Phi \cap W} \frac{1}{|W|}, \qquad \widehat{\lambda}_2 = \sum_{y \in \Psi \cap W} \frac{1}{|W|}$$

be the estimated intensities and $S = \widehat{\lambda}_1 \widehat{\lambda}_2$. Then, if $f_r(x - y) = \mathbb{I}(\|x - y\| \leq r)$,

$$\frac{R}{S} = \frac{1}{\widehat{\lambda}_1 \widehat{\lambda}_2} \sum_{x \in \Phi \cap W} \sum_{y \in \Psi \cap W} f_r(x - y)$$

is, up to a deterministic multiplicative constant, equal to $\widehat{K}_{12}(r)$*. It holds that*

$$\begin{aligned}
\mathbb{E}R = \mu_R =& \lambda_1\lambda_2 \int_{W^2} f_r(u-v)\,\mathrm{d}u\,\mathrm{d}v,\\
\mathbb{E}S = \mu_S =& \lambda_1\lambda_2,\\
\mathrm{var}R = \sigma_R^2 =& \lambda_1^2\lambda_2^2 \int_{W^4} [g_1(u-u')g_2(v-v')-1] f_r(u-v) f_r(u'-v')\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}u'\,\mathrm{d}v'\\
&+ \lambda_1^2\lambda_2 \int_{W^3} g_1(u-u') f_r(u-v) f_r(u'-v)\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}u'\\
&+ \lambda_1\lambda_2^2 \int_{W^3} g_2(v-v') f_r(u-v) f_r(u-v')\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}v'\\
&+ \lambda_1\lambda_2 \int_{W^2} f_r(u-v)\,\mathrm{d}u\,\mathrm{d}v,\\
\mathrm{var}S = \sigma_S^2 =& \frac{1}{|W|^4}\left(\lambda_1^2 \int_{W^2} g_1(u-v)\,\mathrm{d}u\,\mathrm{d}v + \lambda_1|W|\right)\cdot\\
&\cdot\left(\lambda_2^2 \int_{W^2} g_2(u-v)\,\mathrm{d}u\,\mathrm{d}v + \lambda_2|W|\right) - \lambda_1^2\lambda_2^2,\\
\mathrm{cov}(R,S) =& \frac{\lambda_1^2\lambda_2^2}{|W|^2} \int_{W^4} [g_1(u-u')g_2(v-v')-1] f_r(u-v)\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}u'\,\mathrm{d}v'\\
&+ \frac{\lambda_1^2\lambda_2}{|W|^2} \int_{W^3} g_1(u-u') f_r(u-v)\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}u'\\
&+ \frac{\lambda_1\lambda_2^2}{|W|^2} \int_{W^3} g_2(v-v') f_r(u-v)\,\mathrm{d}u\,\mathrm{d}v\,\mathrm{d}v'\\
&+ \frac{\lambda_1\lambda_2}{|W|^2} \int_{W^2} f_r(u-v)\,\mathrm{d}u\,\mathrm{d}v.
\end{aligned}$$

Using the notation of Theorem 4, the approach of Stuart and Ord (1994, p.351), based on the Taylor expansion of the function $f(R,S) = R/S$, provides an approximation of the variance of the ratio R/S:

$$\mathrm{var}\left(\frac{R}{S}\right) \approx \left(\frac{\mu_R}{\mu_S}\right)^2 \left[\frac{\sigma_R^2}{\mu_R^2} - 2\frac{\mathrm{cov}(R,S)}{\mu_R\mu_S} + \frac{\sigma_S^2}{\mu_S^2}\right].$$

All required quantities are given in Theorem 4 and therefore the variance correction for $\widehat{K}_{12}(r)$ can be practically used with this approximate variance and plugged-in estimates of the pair-correlation functions g_1, g_2 and intensities λ_1, λ_2. However, the computational demands of such approach are very high.

We remark that the sample cross K-function is estimated for a given number of arguments, which results in the same number of simultaneous Monte Carlo tests. The multiple correction was resolved in our study using the global envelope test (Myllymäki et al., 2017). Due to this multiple testing correction and the fact that the cross K-function summarizes the information from some neighborhood of the observed points (which was not the case for the sample covariance of two random fields in Section 3.1), the variance correction tests are conservative and less powerful than the torus correction approach, as shown in the simulation study in Mrkvička et al. (2021).

Therefore, it is desirable to look for a test statistic which would be scalar (avoiding the multiple correction problem, which amplifies the deviation of the

random shift strategy from exchangeability) and which would be less affected by summarizing the information from a neighborhood of the observed points (which amplifies the effect of cracks in the autocorrelation structure). Mrkvička et al. (2021) propose to use the expectation of the cross nearest-neighbor distance D_{12}, which is the (random) distance from an arbitrary point of Φ to the nearest point of Ψ. To estimate the expectation $\mathbb{E}D_{12}$, we use the Lebesgue-Stieltjes integral

$$\int r \, \widehat{G}_{12}(\mathrm{d}r)$$

where $\widehat{G}_{12}$ is the Kaplan-Meier estimator of the cross nearest-neighbor distance distribution function G_{12} (Baddeley and Gill, 1997; Illian et al., 2008), which is the distribution function of the random variable D_{12}.

The point process case is more complex than the random field case since the test statistic accumulates information from a certain neighborhood of the observed points and hence the variance correction methods do not perform very well in the simulation experiments in Mrkvička et al. (2021): the effect of dropping a part of the information after the shifts is more severe here. They are too conservative and have smaller power than tests with torus correction. The effect is more pronounced for the cross K-function than for $\mathbb{E}D_{12}$ which is caused by the multiple testing problem. The torus correction with the cross K-function shows the same liberality for clustered processes as in the random field case. On the other hand, it does not show liberality for the Poisson process and repulsive processes. The torus correction with $\mathbb{E}D_{12}$ seems to show very little to no liberality at all.

3.3 Testing independence between a point process and a covariate

One of the problems investigated in Dvořák et al. (2022a) concerns testing the null hypothesis of independence between a point process and a covariate. While the previous tests of this hypothesis assume the covariate values are observed only at the points of the process, see Sections 3.1 and 3.2 of Dvořák et al. (2022a), the paper aims at investigating the situation where the covariate values are available everywhere in the observation window and proposing nonparametric tests that fully exploit the available covariate information. This type of test is denoted P-C in Dvořák et al. (2022a) to indicate that the null hypothesis is the independence between the point process (P) and the covariate (C).

The proposed tests are based on the random shift approach with either the torus correction or the variance correction described in Section 1.3.2 of this thesis. Let Φ be the point process in question and $\Psi = Z$ be the covariate. Assuming that the covariate has numeric values, the test statistic is the sample mean of the covariate values observed at the point pattern locations:

$$T = \frac{1}{\Phi(W)} \sum_{x \in \Phi \cap W} Z(x). \tag{3.1}$$

The choice of the test statistic is motivated by the most natural scenario, where the points of the process are more likely to appear in locations with high (or low)

covariate values. The test statistic T is able to capture this type of dependence, attaining a higher (or lower) value for the observed data than for the shifted data, resulting in a test with high power in this setting.

For application of the variance correction, the variance of the test statistic must be known, at least asymptotically. The asymptotic order of the variance of the sample mean is $1/n$ when computed from a sample of n independent, identically distributed observations. In our case, the following theorem gives, for a stationary point process Φ with intensity λ, that $\mathrm{var}(T) \approx 1/(\lambda|W|)$. Since the true intensity λ is unknown, in practice we plug-in its estimator $\widehat{\lambda} = \Phi(W)/|W|$ and for variance correction we use the (estimated) correction term $\mathrm{var}(T) \approx 1/\Phi(W)$.

Theorem 5. *Let Φ be a stationary point process in $\mathbb{R}^2$ with intensity λ and pair-correlation function g, observed in the observation window W. Let $Z(u), u \in W$, be a centered stationary random field with finite second moments, independent of Φ, having a non-negative covariance function C. Assume that there is a constant $R > 0$ such that $C(u-v) = 0$ for $\|u-v\| > R$. Define the random variables S, U in the following way:*

$$S = \sum_{x \in \Phi \cap W} Z(x), \qquad U = \Phi(W).$$

Then there exist constants $0 < C_1 \leq C_2 < \infty$, depending on the properties of Φ and Z but not on $|W|$, such that

$$C_1 \leq \frac{\mathrm{var}(S)}{\lambda|W|} \leq C_2.$$

Moreover, the variance of S/U can be approximated by

$$\mathrm{var}\left(\frac{S}{U}\right) \approx \frac{\mathrm{var}(S)}{\lambda^2|W|^2}$$

and hence $\mathrm{var}(S/U)$ is of order $1/(\lambda|W|)$.

We note that the fraction $S/U = T$ gives the test statistic from (3.1). Appropriate mixing conditions can replace the assumption of bounded support of the covariance function C.

Other test statistics can, of course, be used if there is a particular indication that other properties of the covariate Z might influence the occurrence of points in Φ. For example, the histogram (vector of counts of observations with values in disjoint intervals) can be used so that the whole distribution of the covariate values at points of the process is captured. In this case, the global envelope tests of Myllymäki et al. (2017) can be used to perform the Monte Carlo test with the multivariate test statistic.

Simulation experiments in Section 4.1 of Dvořák et al. (2022a) showed that, as expected, the random shift test with torus correction is liberal under the null hypothesis and has very high power for the studied alternatives. The random shift test with variance correction is slightly conservative under the null hypothesis and has slightly lower power than the test with torus correction, but always higher power than the considered benchmark tests. Comparison with a wider range of

tests available in the literature is given in the paper Dvořák et al. (2022b), which is not included in this thesis.

Finally, we remark that the methods proposed in the paper Dvořák and Mrkvička (2022b), described in Section 3.6 of this thesis, can also be used to test the null hypothesis of independence between a point process and a covariate, considering the special case with no nuisance covariates.

3.4 Testing independence between a marked point process and a covariate

The main goal of the paper Dvořák et al. (2022a) was to propose a nonparametric test of independence between marks and a covariate in a marked point process setting. Such a test would allow one to decide whether the covariate should be included in the analysis of the marks or not.

However, when the nonparametric test is based on random shifts, the marks cannot be shifted without simultaneously shifting the points (in fact, the marks cannot exist separately from the points). This means that possible dependence between the (unmarked) point process and the covariate may affect the inference about the relationship between the marks and the covariate, through the preferential sampling effects (Diggle et al., 2010). This term is used in the geostatistical literature to describe the general situation where the set of sampling points is not independent of the studied random field, e.g., when more samples are taken at the locations where high-grade ore is thought likely to be found. It has been reported that preferential sampling introduces bias into the estimation of the covariance structure of the random field (Diggle et al., 2010).

Hence, Dvořák et al. (2022a) propose a test of independence between a marked point process and a covariate (instead of a test of independence between marks and a covariate), and suggest performing several tests to inquire the complete dependence structure in the triangle points-marks-covariate. This type of test is denoted PM-C in Dvořák et al. (2022a) to indicate that the null hypothesis is the independence between the marked point process (PM) and the covariate (C).

In the following, we assume that both the marks and the covariate have numeric (continuous) values. The case of categorical marks and/or categorical covariate is briefly discussed in Section 3.3 of Dvořák et al. (2022a).

For a PM-C test, our starting point is choosing the sample covariance as the test statistic. This is motivated by the use of sample covariance for testing the independence between two random fields in Mrkvička et al. (2021). Let $\{(x_1, m_1), \ldots, (x_n, m_n)\}$ be the observed realization of the marked point process and let $z(u), u \in W$, denote the observed realization of the covariate Z. We define

$$T_C = \frac{1}{n-1} \sum_{i=1}^{n} (m_i - \overline{m})(z(x_i) - \overline{z}),$$

where

$$\overline{m} = \frac{1}{n} m_i, \qquad \overline{z} = \frac{1}{n} z(x_i)$$

are the observed sample means. The asymptotic order of variance of T_C can be determined under the assumption of the geostatistical marking model (Illian et al., 2008) using Theorem 5 above.

Although this choice of a test statistic is perfectly appropriate in situations where the unmarked point process of sampling locations is independent of the marks and the covariate, it may perform poorly in cases where there is a dependence between them. In the case of dependence between the point process and the covariate, the problem of preferential sampling occurs. Preferential sampling may introduce bias when estimating the covariance between marks and a covariate. Random shifts then violate preferential sampling, changing the distribution of the test statistic computed from the shifted distribution. This, in turn, damages exchangeability, and the resulting test is far from exact. Whether the test would be conservative or liberal depends on the particular type of preferential sampling. The same issues can be caused by the dependence between points and marks. However, we do not use the term "preferential sampling" in this case to avoid confusion.

This leads us to define different test statistics that will be less affected by the bias in the estimated covariance structure of the covariate and the marks, more specifically, less affected by the sample variance. We choose Pearson's correlation coefficient and, assuming no ties are present in the data, Kendall's rank correlation coefficient:

$$T_P = \frac{\sum_{i=1}^{n}(m_i - \overline{m})(z(x_i) - \overline{z})}{\sqrt{\sum_{i=1}^{n}(m_i - \overline{m})^2}\sqrt{\sum_{i=1}^{n}(z(x_i) - \overline{z})^2}},$$

$$T_K = \frac{1}{n(n-1)} \sum_{i \neq j} \operatorname{sgn}(m_i - m_j)\operatorname{sgn}(z(x_i) - z(x_j)).$$

The asymptotic order of the variance of T_P, when computed from a sample of n i.i.d. observations, is $1/n$ (van der Vaart, 1998, p. 30). The same holds for T_K (van der Vaart, 1998, pp. 164–165). Therefore, we use $\operatorname{var}(T_P), \operatorname{var}(T_K) \approx 1/n$ for variance correction. This is justified for T_K by the following theorem which states that, for a stationary point process Φ with intensity λ, the variance of T_K is of the order $1/(\lambda|W|)$. Since the true intensity λ is unknown, in practice we plug-in its estimator $\widehat{\lambda} = \Phi(W)/|W|$ and for the variance correction the (estimated) correction term is $\operatorname{var}(T) \approx 1/\Phi(W)$.

Theorem 6. *Let Ψ be a stationary marked point process in $\mathbb{R}^2$, observed in the observation window W. Let Ψ follow the geostatistical marking model, i.e. it is obtained by sampling the (random) mark field Z_1 at points of the unmarked point process Φ. Assume that the product densities of Φ up to the fourth order exist and are bounded by finite positive constants both from above and from below. They will be denoted by $\lambda, \lambda_2, \lambda_3$ and λ_4 in the following.*

Let the covariate be given by the random field Z_2 and let the random fields Z_1, Z_2 be independent, identically distributed, centered stationary Gaussian random fields with a non-negative covariance function C. Assume that there is a constant R such that $C(u - v) = 0$ for $\|u - v\| > R$. Furthermore, assume that there are constants $\delta > 0$ and $r > 0$ such that $C(u - v) \geq \delta$ for $\|u - v\| \leq r$.

Define the random variables S, U in the following way:

$$S = \sum_{x,y\in\Phi\cap W}^{\neq} \operatorname{sgn}(Z_1(x) - Z_1(y)) \operatorname{sgn}(Z_2(x) - Z_2(y)), \quad U = \Phi(W)(\Phi(W) - 1).$$

Then there exist constants $0 < c_1 \leq c_2 < \infty$ and $0 < d_1 \leq d_2 < \infty$, depending on the properties of Φ and Z_1, Z_2 but not on $|W|$, such that $\operatorname{var}(S) = A + B$ with

$$c_1 \leq \frac{A}{\lambda^3|W|^3} \leq c_2, \qquad d_1 \leq \frac{B}{\lambda^2|W|^2} \leq d_2.$$

Moreover, the variance of S/U can be approximated by

$$\operatorname{var}\left(\frac{S}{U}\right) \approx \frac{\operatorname{var}(S)}{(\mathbb{E}U)^2}$$

with $u_1 \leq \mathbb{E}U/(\lambda|W|)^2 \leq u_2$ for some finite positive constants u_1, u_2. Therefore, $\operatorname{var}(S/U)$ is of the order $1/(\lambda|W|)$.

Note that the fraction $S/U = T$ gives the test statistic T_K. The assumption of bounded support of the covariance function C can be replaced by appropriate mixing conditions.

In the simulation experiments in Dvořák et al. (2022a) the PM-C tests with torus correction are all slightly liberal under the null hypothesis, regardless of the test statistic used. Using the variance correction reduces the liberality but does not make the tests conservative. The performance of all PM-C tests under different alternatives is comparable. The variance correction performs only marginally better than the torus correction, and the Pearson's correlation coefficient results in a slightly higher power than the other two test statistics.

3.5 Testing independence between a covariate and functional marks

The approach from Section 3.4 can be extended to the setting with functional marks $m_i(r)$ attached to the observed points x_i. For testing the independence between a covariate Z and the functional marks, it is natural to use a functional test statistic such as

$$T_P(r) = \frac{\sum_{i=1}^n (m_i(r) - \overline{m}(r))(z(x_i) - \overline{z})}{\sqrt{\sum_{i=1}^n (m_i(r) - \overline{m}(r))^2}\sqrt{\sum_{i=1}^n (z(x_i) - \overline{z})^2}},$$
$$T_K(r) = \frac{1}{n(n-1)} \sum_{i\neq j} \operatorname{sgn}(m_i(r) - m_j(r))\operatorname{sgn}(z(x_i) - z(x_j)),$$

where

$$\overline{m}(r) = \frac{1}{n} m_i(r), \qquad \overline{z} = \frac{1}{n} z(x_i).$$

In the random shift test of independence between the covariate and the functional marks, the points with their corresponding marks are shifted against the covariate.

The global envelope test of Myllymäki et al. (2017) is then used to perform the Monte Carlo test.

An interesting application of such test is in testing the assumption of the second-order intensity-reweighted stationarity of the point process in question (Baddeley et al., 2000). In this case, the functional marks can be constructed as the local K-functions or local L-functions corresponding to the individual points of the process (Anselin, 1995; Cressie and Collins, 2001a,b). This idea is investigated in the planned paper González et al. (2022).

3.6 Partial correlation analysis for a point process and a set of covariates

When a point process is accompanied by one or more covariates, determining which covariates influence the positions of points is crucial for the upcoming inferential steps. The paper Dvořák and Mrkvička (2022b) studies this problem. Also, the paper proposes a correlation coefficient and a partial correlation coefficient between a point process and a covariate, which allow quantification of the spatial dependence between a point process and a covariate, both without and with the presence of nuisance covariates.

The approach of Dvořák and Mrkvička (2022b) is based on the newly proposed notion of nonparametric residuals for point processes. They are a counterpart of the parametric residuals from Baddeley et al. (2005) which can be used to check whether a fitted model for the intensity function is appropriate. We remark that the version of the residuals considered here is based on the intensity function, as suggested by R. Waagepetersen in the discussion to the paper Baddeley et al. (2005), rather than based on the conditional intensity function as discussed in the paper itself.

To recall the parametric residuals, let $\hat{\beta}$ be the vector of the estimated regression parameters in a parametric model for the intensity function λ depending on some covariates. The *residual measure* is defined as

$$\mathcal{R}(B) = n(X \cap B) - \int_B \lambda(u; \hat{\beta}) \, \mathrm{d}u, \tag{3.2}$$

where $B \subseteq W$ is a Borel set and W is the observation window. The *smoothed residual field* is obtained as

$$s(u) = \frac{1}{e(u)} \left[\sum_{x_i \in X \cap W} k(u - x_i) - \int_W k(u - v) \lambda(v; \hat{\beta}) \, \mathrm{d}v \right], \tag{3.3}$$

where $e(u) = \int_W k(u - v) \, \mathrm{d}v$ is the edge-correction factor and k is a probability density function in $\mathbb{R}^2$. In fact, the first term in (3.3) gives the nonparametric kernel estimate of the intensity function, the covariates not being taken into account, while the second term gives the smoothed parametric estimate which incorporates the covariates. If the estimated model $\lambda(v; \hat{\beta})$ describes the point process X well, the smoothed residual field $s(u)$ is expected to fluctuate around 0. Its deviations from 0 indicate a disagreement between $\lambda(v; \hat{\beta})$ and the true intensity function in the corresponding parts of the observation window.

Nonparametric residuals

As opposed to fitting a parametric model, the dependence of the intensity function on a set of covariates $C_1, \ldots, C_m$ can be captured nonparametrically. Baddeley et al. (2012) assume that there is an unknown function $\rho : \mathbb{R}^m \to [0, \infty)$ such that $\lambda(u) = \rho(C_1(u), \ldots, C_m(u))$. Assuming absolute continuity of the distribution of the vector of covariates $(C_1(u), \ldots, C_m(u))$ on $\mathbb{R}^m$, the function ρ can be estimated using kernel smoothing in the space of covariate values, see Baddeley et al. (2012). This opens up the possibility to define nonparametric residuals. Using the nonparametric estimate of the intensity function $\hat{\lambda}(u) = \hat{\rho}(C_1(u), \ldots, C_m(u))$, the nonparametric version of the residual measure (3.2) can be defined as

$$\tilde{\mathcal{R}}(B) = n(X \cap B) - \int_B \hat{\rho}(C_1(u), \ldots, C_m(u)) \, \mathrm{d}u.$$

The corresponding nonparametric smoothed residual field is then

$$\tilde{s}(u) = \frac{1}{e(u)} \left[\sum_{x_i \in X \cap W} k(u - x_i) - \int_W k(u - v) \hat{\rho}(C_1(u), \ldots, C_m(u)) \, \mathrm{d}v \right]. \quad (3.4)$$

If $\hat{\rho}(C_1(u), \ldots, C_m(u))$ describes the intensity function of X well, meaning e.g. that no relevant covariate was left out, $\tilde{s}(u)$ is expected to fluctuate around 0. Deviations from 0 indicate disagreement between the estimated model and the true intensity function of X.

Partial correlation coefficient

When several possibly correlated covariates are available, one might be interested in assessing the strength of dependence between the point process X and the covariate of interest C_{m+1} after removing the possible influence of the remaining (nuisance) covariates $C_1, \ldots, C_m$, in the spirit of the partial correlation coefficient.

The strength of dependence can be quantified by some measure of dependence between the covariate of interest C_{m+1} and the smoothed residual field $\tilde{s}$ from (3.4) where the possible influence of the nuisance covariates $C_1, \ldots, C_m$ on X has been removed. When a parametric model for the intensity function of X is available, parametric residuals (3.3) may be used instead.

We suggest using Kendall's correlation coefficient to quantify the dependence. We consider a set of sampling points $\{y_1, \ldots, y_n\}$, independently and uniformly distributed in W, independent of X and $C_1, \ldots, C_{m+1}$, and define the sample version of the partial correlation coefficient as

$$\hat{\tau}_p = \frac{1}{n(n-1)} \sum_{i \neq j} \operatorname{sgn}(C_{m+1}(y_i) - C_{m+1}(y_j)) \operatorname{sgn}(\tilde{s}(y_i) - \tilde{s}(y_j)).$$

We stress that independent sampling points need to be used in this case instead of simply using the observed points of $X \cap W$. In the latter case, preferential sampling issues could arise, resulting in biased estimates of the properties of the two random fields (Diggle et al., 2010; Dvořák et al., 2022a). Loosely speaking, if, for example, the sampling points $\{y_1, \ldots, y_n\}$ are more likely to be chosen in locations with high values of C_{m+1}, the sample mean and sample variance

of $C_{m+1}(y_1), \ldots, C_{m+1}(y_n)$ do not reflect well the true properties of C_{m+1}. This negatively affects all subsequent steps of the analysis.

We remark here that if no nuisance covariates are observed ($m = 0$), the correlation coefficient between the point process and the covariate of interest can be defined similarly as above, using the constant estimator of the intensity function given by the number of observed points per unit area. For details, see Section 3.2 of Dvořák and Mrkvička (2022b).

Covariate-weighted residual measure

While $\hat{\tau}_p$ is useful for quantifying the strength of dependence between X and the covariate of interest C_{m+1} after removing the influence of nuisance covariates $C_1, \ldots, C_m$, the random shift test using $\hat{\tau}_p$ as the test statistic turned out to have a rather low power in the simulation studies. The reason lies in the applied smoothing and the deliberate removal of the preferential sampling effects: the association between the points of X and the covariate C_{m+1} brings important information.

To overcome these issues, we define the following characteristic that we call the *covariate-weighted residual measure of* W:

$$\begin{aligned} CWR &= \int_W C_{m+1}(u)\tilde{\mathcal{R}}(\mathrm{d}u) \\ &= \sum_{x \in X \cap W} C_{m+1}(x) - \int_W C_{m+1}(u)\hat{\rho}(C_1(u), \ldots, C_m(u))\,\mathrm{d}u. \end{aligned} \tag{3.5}$$

This can be viewed as a generalization of the test statistic (3.1) from Section 3.3 of this thesis which also includes the sum of covariate values but does not take into account possible nuisance covariates. By sampling the values of C_{m+1} at the points of X we take advantage of any possible preferential sampling effects, and no smoothing is performed when computing the value of CWR, hence we avoid the problem of bandwidth selection. The expectation of CWR is close to 0 if the covariates $C_1, \ldots, C_m$ capture all variation in $\lambda(u)$, i.e. if $\hat{\rho}(C_1(u), \ldots, C_m(u))$ is close to $\lambda(u)$, and will differ from 0 otherwise. This enables testing the significance of C_{m+1} after removing the influence of $C_1, \ldots, C_m$.

Testing the covariate significance under the presence of nuisance covariates

Now we consider the null hypothesis that X and C_{m+1} are independent, conditionally on $C_1, \ldots, C_m$. We employ the random shift test described in Section 1.3.2 of this thesis, either with torus or variance correction. The test statistic can be $\hat{\tau}_p$ in which case the two spatial objects to be shifted against each other are the two random fields $\Phi = \tilde{s}$ and $\Psi = C_{m+1}$. Alternatively, we can use the covariate-weighted residual measure of W as a test statistic. In this case $\Phi = \tilde{\mathcal{R}}$ is a measure and $\Psi = C_{m+1}$ is a random field. If v_i is a shift vector, the shift of the random field Ψ should be interpreted in both cases as $(\Psi + v_i)(u) = \Psi(u - v_i)$.

The choice of the correction factors for the variance correction is discussed in the appendix of Dvořák and Mrkvička (2022b). The simulation experiments

reported there confirm that the given correction factors, in fact, correctly standardize the variance of the test statistics. Also, in a simplified setting, the variance of CWR can be determined, as given in the following theorem. The simplification lies in the assumption that the true intensity function $\lambda(u)$ is used to compute CWR and that the point process in question is Poisson.

Theorem 7. *Let X be a Poisson process on W with the intensity function $\lambda(u), u \in W$, and let $C(u), u \in W$, be a stationary random field with $\mathbb{E}C(u)^2 = K < \infty$. Denote by S the analogue of the covariate-weighted residual measure of W from* (3.5)*:*

$$S = \sum_{x \in X \cap W} C(x) - \int_W C(u)\lambda(u)\,\mathrm{d}u.$$

Then $\operatorname{var} S = K \int_W \lambda(u)\,\mathrm{d}u$.

The variance of S is proportional to $\int_W \lambda(u)\,\mathrm{d}u$ which is the expected number of points in W. In practical situations, this quantity is not known and can be estimated by the observed number of points $n(X \cap W)$. If the intensity function is bounded from above and from below by finite positive constants, $\int_W \lambda(u)\,\mathrm{d}u$ is of order $|W|$ for large observation windows.

In the simulation study in Dvořák and Mrkvička (2022b), the random shift tests match the nominal significance level correctly for all models. The tests based on CWR match it slightly more accurately than those based on $\hat{\tau}_p$. Both the torus correction and the variance correction perform well, with only a slight tendency toward liberality observed for the torus correction and the tests based on $\hat{\tau}_p$. Concerning the benchmark parametric tests considered in the simulation study, their performance illustrates that parametric tests are prone to perform poorly under model misspecification either in terms of the interaction structure or the intensity function. However, even when both of these model components are specified correctly, there is a risk of strong liberality of the parametric tests. From this point of view, the nonparametric tests are preferable, as they match the nominal significance level correctly for all models in this study.

Concerning the power of the tests, we make the following observations. The tests based on $\hat{\tau}_p$ have very low power due to the smoothing and removal of the preferential sampling effects. The tests based on CWR exhibit very high power comparable to parametric tests with the correct interaction model and the correct model for the intensity function. For some models, the tests based on CWR showed even higher power than the parametric tests. The torus correction and the variance correction perform nearly equivalently for tests based on CWR, while for tests based on $\hat{\tau}_p$ the torus correction shows slightly higher power, which can be explained by the small liberality of these tests.

3.7 Testing independence of animal trajectories

Another problem which can be solved using the random shift approach is testing of independence between a pair of trajectories, observed e.g. by GPS tracking of wild animals. This problem is studied in the master thesis Veselý (2021), which

analyzes the movement of wolves in the Voyageurs national park (USA). The data was kindly provided by the researches from the Voyageurs Wolf Project.[1]

Assume that the spatial position of an animal is recorded in the sequence of times $t_1, \ldots, t_n$ during the time interval $[0, T]$ in the spatial observation window W. Such a dataset can be represented as a temporal marked point pattern $\{(t_1, u_1), \ldots, (t_n, u_n)\}$ where $t_i \in [0, T]$ is the time of the ith observation and the mark $u_i \in W$ is the corresponding spatial location.

If a pair of (possibly interacting) trajectories is observed, we can obtain the Monte Carlo replications for the independence test by randomly shifting one of the trajectories along the time axis. In this setting, it is straightforward to apply the torus correction since the interval $[0, T]$ is a one-dimensional rectangle. If τ is the random variable with uniform distribution on the interval $[0, T]$, the shifted trajectory is obtained as $\{((t_1 + \tau) \bmod T, u_1), \ldots, ((t_n + \tau) \bmod T, u_n)\}$.

An appropriate test statistic must be chosen, depending on whether the observation times are equidistant and whether they are the same for the two trajectories. Several test statistics are proposed in Chapter 3 of Veselý (2021). The analysis of the wolf trajectories in Chapter 4 of Veselý (2021), where each tracked animal represents a different pack, identified no evidence of interactions between any pair of observed trajectories.

[1] `https://www.voyageurswolfproject.org`

4. Tests based on stochastic reconstruction

As discussed in Section 1.3.3 of this thesis, the stochastic reconstruction procedure can be used to obtain the replications required for performing a Monte Carlo test. While being computationally demanding if the desired number of replications N is large, the advantage of the procedure lies in its very broad applicability. However, caution is needed to make sure that the replications have the same distribution as the observed data (this is needed for achieving exchangeability) and that the replications fulfill the null hypothesis (this is needed for performing the test of the given hypothesis).

4.1 Tests for space-time point processes

In this section, we describe the possibility to use the stochastic reconstruction procedure for testing the first-order separability hypothesis for space-time point processes, as already discussed in Section 2.2 of this thesis in connection with the permutation tests. The stochastic reconstruction procedure was proposed in this context in Ghorbani et al. (2021), with a detailed illustration given in Section 7.2 of the paper. A brief summary is given below.

The user chooses a set of summary characteristics that should be preserved by the reconstruction procedure. For testing the first-order separability hypothesis, the paper suggests using the square root of a nonparametric estimator of the inhomogeneous space-time K-function (Gabriel and Diggle, 2009; Dvořák and Prokešová, 2016), together with the separable estimator of the intensity function $\hat{\rho}_{\text{sep}}$ from (2.2). Using the separable estimator secures the separable structure of the intensity function of the output patterns, meaning the replications are indeed generated under the null hypothesis.

Based on the experience from (Koňasová and Dvořák, 2021a), we also suggest using as further summaries the values $\hat{D}_k(r,t)$ giving the fraction of observed events that have at least k neighbours within distance r (in the spatial domain) and within lag t (in the temporal domain). These are considered only to be empirical characteristics describing the inter-event distances rather than being estimators of some theoretical quantities. However, they are closely related to the raw estimates of the kth nearest neighbour distribution functions in a stationary space-time point process.

The energy functional, quantifying the dissimilarity between the input pattern ζ (observed pattern) and another pattern ξ, is then constructed:

$$\begin{aligned} E(\zeta,\xi) = {} & w_K \int_0^{T_K} \int_0^{R_K} \left[\sqrt{\widehat{K}(\zeta;r,t)} - \sqrt{\widehat{K}(\xi;r,t)} \right]^2 \mathrm{d}r\,\mathrm{d}t \\ & + \sum_{k=1}^{k_{max}} w_{D_k} \int_0^{T_D} \int_0^{R_D} \left[\widehat{D}_k(\zeta;r,t) - D_k(\xi;r,t) \right]^2 \mathrm{d}r\,\mathrm{d}t \\ & + w_\Delta \sum_{i=1}^{I} a \left[\hat{\rho}_{\text{sep}}(\zeta;v_i,s_i) - \hat{\rho}_{\text{sep}}(\xi;v_i,s_i) \right]^2, \end{aligned}$$

where w_K, w_{D_k}, w_Δ are the weights determining the relative importance of the individual terms, $\{(v_1, s_1), \ldots, (v_I, s_I)\}$ are the center locations of the cells of a regular grid covering $W \times T$, a is the volume of the grid cell and T_K, R_K, T_D, R_D and k_{max} are user-selected tuning constants.

The stochastic reconstruction procedure starts with a binomial pattern ξ_0 generated as a collection of n independent points (the same as the number of observed points in the input pattern ζ) following a probability density function proportional to $\hat{\rho}_{\text{sep}}(\zeta; u, t)$. Then the iteration steps are repeated in which a new pattern ξ^{new} is proposed by randomly deleting one point from the current pattern, say ξ_m, and generating a new point in $W \times T$ with density again proportional to $\hat{\rho}_{\text{sep}}(\zeta; u, t)$. The proposal is accepted if $E(\zeta, \xi^{new}) \leq E(\zeta, \xi_m)$, otherwise it is rejected. The algorithm stops when a user specified stopping rule is met, e.g. after performing a maximum allowed number of iterations or after rejecting a certain amount of proposals in a row. By minimization of the energy functional, the output pattern ξ^{out} is forced to have approximately the same interaction structure as the input pattern ζ (as described by the K- and D_k-functions) while having a separable first-order structure (as described by $\hat{\rho}_{\text{sep}}$).

After a large number of independent output patterns is generated, these can be used to perform a Monte Carlo test of the separability hypothesis. The outputs can be considered to be independent replicates of the data obtained under the null hypothesis. The performance of the Monte Carlo test of course relies on the interaction structure of the observed data being correctly captured by the reconstruction procedure.

Using the stochastic reconstruction procedure requires some tuning of the parameters. It is also strongly suggested to verify on simulated data that the outputs of the reconstruction algorithm have the same properties as the simulations from the correct model. This can be done following the suggestions in Koňasová and Dvořák (2021a).

Bibliography

Adler, R. J. and Taylor, J. E. (2007) *Random Fields and Geometry*. Springer Monographs in Mathematics. Springer, New York, 1 edn.

Anselin, K. (1995) Local indicators of spatial association-LISA. *Geographical Analysis*, **27**, 93–115.

Baddeley, A., Chang, Y.-M., Song, Y. and Turner, R. (2012) Nonparametric estimation of the dependence of a spatial point process on spatial covariates. *Statistics and Its Interface*, **5**, 221–236.

Baddeley, A. and Gill, R. D. (1997) Kaplan-Meier estimators of distance distributions for spatial point processes. *Annals of Statistics*, **25**, 263–292.

Baddeley, A., Møller, J. and Waagepetersen, R. (2000) Non- and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, **54**, 329–350.

Baddeley, A., Turner, R., Møller, J. and Hazelton, M. (2005) Residual analysis for spatial point processes (with discussion). *Journal of the Royal Statistical Society (Series B)*, **67**, 617–666.

Cressie, N. and Collins, L. B. (2001a) Analysis of spatial point patterns using bundles of product density LISA functions. *Journal of Agricultural, Biological, and Environmental Statistics*, **6**, 118–135.

— (2001b) Patterns in spatial point locations: local indicators of spatial association in a minefield with clutter. *Naval Research Logistics*, **48**, 333–347.

Cressie, N. A. C. (1993) *Statistics for Spatial Data*. Wiley Series in Probability and Mathematical Statistics. Wiley, New York, revised edn.

Daley, D. J. and Vere-Jones, D. (2008) *An Introduction to the Theory of Point Processes. Volume II: General Theory and Structure*. Springer, New York, 2nd edn.

Davison, A. C. and Hinkley, D. V. (1997) *Bootstrap Methods and their Application*. Cambridge University Press, Cambridge.

Diggle, P. J., Menezes, R. and Su, T.-l. (2010) Geostatistical inference under preferential sampling. *Journal of the Royal Statistical Society (Series C)*, **59**, 191–232.

Dvořák, J. and Mrkvička, T. (2022a) Graphical tests of independence for general distributions. *Computational Statistics*, **37**, 671–699.

— (2022b) Nonparametric testing of the covariate significance for spatial point patterns under the presence of nuisance covariates. *Preprint*. URL: `https://arxiv.org/abs/2210.05424`.

Dvořák, J., Mrkvička, T., Mateu, J. and González, J. A. (2022a) Nonparametric testing of the dependence structure among points–marks–covariates in spatial point patterns. *International Statistical Review*, **90**, 592–621.

— (2022b) Nonparametric tests of dependence between a spatial point process and a covariate. In *METMA X, Proceedings of the 10th International Workshop on Spatio-Temporal Modelling* (eds. C. Comas and J. Mateu), 55–59. Universitat de Lleida, Lleida.

Dvořák, J. and Prokešová, M. (2016) Parameter estimation for inhomogeneous space-time shot-noise Cox point processes. *Scandinavian Journal of Statistics*, **43**, 939–961.

Fortin, M.-J. and Payette, S. (2002) How to test the significance of the relation between spatially autocorrelated data at the landscape scale: A case study using fire and forest maps. *Ecoscience*, **9**, 213–218.

Gabriel, E. and Diggle, P. J. (2009) Second-order analysis of inhomogeneous spatio-temporal point process data. *Statistica Neerlandica*, **63**, 43–51.

Ghorbani, M., Vafaei, N., Dvořák, J. and Myllymäki, M. (2021) Testing the first-order separability hypothesis for spatio-temporal point patterns. *Computational Statistics and Data Analysis*, **161**, 107245.

González, J. A., Dvořák, J., Mrkvička, T. and Mateu, J. (2022) Testing independence between functional marks and covariates in spatial point processes. In preparation.

Illian, J., Penttinen, A., Stoyan, H. and Stoyan, D. (2008) *Statistical Analysis and Modelling of Spatial Point Patterns*. Wiley, Chichester.

Koňasová, K. and Dvořák, J. (2021a) Stochastic reconstruction for inhomogeneous point patterns. *Methodology and Computing in Applied Probability*, **23**, 527–547.

— (2021b) Techniques from functional data analysis adaptable for spatial point patterns. In *Proceedings of the 22nd European Young Statisticians Meeting* (eds. A. Makridis, F. Milienos, P. Papastamoulis, C. Parpoula and A. Rakitzis), 56–60. Panteion University of Social and Political Sciences, Athens.

Lotwick, H. W. and Silverman, B. W. (1982) Methods for analysing spatial processes of several types of points. *Journal of the Royal Statistical Society (Series B)*, **44**, 406–413.

Marozzi, M. (2016) Multivariate tests based on interpoint distances with application to magnetic resonance imaging. *Statistical Methods in Medical Research*, **25**, 2593–2610.

Møller, J. and Waagepetersen, R. P. (2004) *Statistical Inference and Simulation for Spatial Point Processes*. Chapman & Hall/CRC, Boca Raton.

Møller, J., Christensen, H. S., Cuevas-Pacheco, F. and Christoffersen, A. D. (2019) Structured space-sphere point processes and K-functions. *Methodology and Computing in Applied Probability*, **23**, 569–591.

Møller, J. and Ghorbani, M. (2012) Aspects of second-order analysis of structured inhomogeneous spatio-temporal point processes. *Statistica Neerlandica*, **66**, 472–491.

Mrkvička, T., Myllymäki, M., Jílek, M. and Hahn, U. (2020) A one-way ANOVA test for functional data with graphical interpretation. *Kybernetika*, **56**, 432–458.

Mrkvička, T., Dvořák, J., González, J. A. and Mateu, J. (2021) Revisiting the random shift approach for testing in spatial statistics. *Spatial Statistics*, **42**, 100430.

Myllymäki, M., Mrkvička, T., Grabarnik, P., Seijo, H. and Hahn, U. (2017) Global envelope tests for spatial processes. *Journal of the Royal Statistical Society (Series B)*, **79**, 381–404.

Myllymäki, M. and Mrkvička, T. (2019) GET: Global envelopes in R. *Preprint*. URL: `https://arxiv.org/abs/1911.06583`.

Narisetty, N. N. and Nair, V. J. (2016) Extremal depth for functional data and applications. *Journal of the American Statistical Association*, **111**, 1705–1714.

Oden, N. L. (1991) Allocation of effort in Monte Carlo simulation for power of permutation tests. *Journal of the American Statistical Association*, **86**, 1074–1076.

Pawlasová, K. and Dvořák, J. (2022) Supervised nonparametric classification in the context of replicated point patterns. *Image analysis and Stereology*, **41**, 57–74.

Prokešová, M. and Dvořák, J. (2014) Statistics for inhomogeneous space-time shot-noise Cox processes. *Methodology and Computing in Applied Probability*, **16**, 433–449.

Rajala, T., Olhede, S. C. and Murrell, D. J. (2019) When do we have the power to detect biological interactions in spatial point patterns? *Journal of Ecology*, **107**, 711–721.

Stuart, A. and Ord, K. (1994) *Kendall's Advanced Theory of Statistics*. Arnold, London.

Tscheschel, A. and Stoyan, D. (2006) Statistical reconstruction of random point patterns. *Computational Statistics & Data Analysis*, **51**, 859–871.

Upton, G. J. G. and Fingleton, B. (1985) *Spatial Data Analysis by Examples, Vol. I. Point Pattern and Quantitative Data*. Wiley, New York.

van der Vaart, A. (1998) *Asymptotic Statistics*. Cambridge series on statistical and probabilistic mathematics. Cambridge University Press, Cambridge.

Veselý, M. (2021) *Nonparametric tests of independence between animal movement trajectories*. Master thesis, Faculty of Mathematics and Physics, Charles University, Prague, Czech Republic. URL: `http://hdl.handle.net/20.500.11956/152530`.

Wiegand, T., He, F. and Hubbell, S. (2013) A systematic comparison of summary characteristics for quantifying point patterns in ecology. *Ecography*, **36**, 92–103.

Wong, K. Y. and Chiu, S. N. (2016) Isotropy test for spatial point processes using stochastic reconstruction. *Spatial Statistics*, **15**, 56–69.

Appendix: Attached papers

[1] T. Mrkvička, J. Dvořák, J.A. González, J. Mateu (2021): Revisiting the random shift approach for testing in spatial statistics. *Spatial Statistics* 42, 100430. doi: 10.1016/j.spasta.2020.100430

[2] M. Ghorbani, N. Vafaei, J. Dvořák, M. Myllymäki (2021): Testing the first-order separability hypothesis for spatio-temporal point patterns. *Computational Statistics and Data Analysis* 161, 107245. doi: 10.1016/j.csda.2021.107245

[3] J. Dvořák, T. Mrkvička (2022): Graphical tests of independence for general distributions. *Computational Statistics* 37, 671–699. doi: 10.1007/s00180-021-01134-y

[4] J. Dvořák, T. Mrkvička, J. Mateu, J.A. González (2022): Nonparametric testing of the dependence structure among points-marks-covariates in spatial point patterns. *International Statistical Reviews* 90, 592–621. doi: 10.1111/insr.12503

[5] J. Dvořák, T. Mrkvička (2022): Nonparametric testing of the covariate significance for spatial point patterns under the presence of nuisance covariates. Preprint. doi: 10.48550/arXiv.2210.05424