A review on anisotropy analysis of spatial point patterns

T. Rajala[1], C. Redenbach[2], A. Särkkä[3], and M. Sormani[2,4]

[1] *Department of Statistical Science, University College London, Gower Street, London WC1E 6BT, UK*
[2] *Mathematics Department, University of Kaiserslautern, 67663 Kaiserslautern, Germany*
[3] *Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, 412 96 Gothenburg, Sweden*
[4] *Fraunhofer Institut für Techno- und Wirtschaftsmathematik, Fraunhofer-Platz 1, 67663 Kaiserslautern, Germany*

November 6, 2018

Abstract

A spatial point pattern is called anisotropic if its spatial structure depends on direction. Several methods for anisotropy analysis have been introduced in the literature. In this paper, we give an overview of nonparametric methods for anisotropy analysis of (stationary) point patterns in $\mathbb{R}^2$ and $\mathbb{R}^3$. We discuss methods based on nearest neighbour and second order summary statistics as well as spectral and wavelet analysis. All techniques are illustrated on both a clustered and a regular example. Finally, we discuss methods for testing for isotropy as well as for estimating preferred directions in a point pattern.

1 Introduction

In the early spatial point process literature, observed point patterns were typically small and no replicates were available. Hence, it was natural to assume that the patterns were realizations of stationary and isotropic point processes. In the more recent literature, large and complicated point pattern data with replicates are common and it is not as obvious that stationarity and isotropy hold. Therefore, the validity of these assumptions should be checked prior to the analysis. During the recent years, several authors have paid attention to non-stationarity and currently, it is straightforward to include non-stationarity in many point process models (Myllymäki and Penttinen, 2009; Rajala and Penttinen, 2014; Illian et al., 2012; Ang et al., 2012; Diggle et al., 2013; Baddeley et al., 2014). Isotropy of a point pattern, on the other hand, is often still assumed without further checking, and even though several tools have been suggested to detect anisotropy and test for it, they are not so widely used. To make such methods more easily accessible, we have, in this paper, collected methods that can be used to detect anisotropies, test for isotropy, and estimate preferred directions in point patterns. We restrict ourselves to unmarked point patterns and do not discuss orientation of marks.

Directional methods are especially useful for regular point patterns since it can be difficult to visually detect anisotropy in such patterns. One example is the amacrine cells data, see Figure 4 (right), which consist of locations of 'on' cells and 'off' cells. These data have been analyzed by several authors assuming stationarity and isotropy. However, it was recently detected by Wong and Chiu (2016) that both the marginal 'on' and 'off' patterns as well as the unlabeled pattern show some signs of anisotropy. Locations of air bubbles in polar ice are another example of regular anisotropic patterns (Redenbach et al., 2009; Rajala et al., 2016). Deep down in an ice sheet, the ice, and therefore the air bubble pattern, are deformed. By using directional analysis, we can learn more about the deformation and provide useful information to the glaciologists. Detecting anisotropies visually in the ice samples is especially hard since the air bubble patterns are not only regular but also in 3D. In clustered patterns, the shape and direction of clusters can reveal anisotropies but directional analysis is needed, for example, to estimate the preferred direction of the clusters. In a series of papers on spectral analysis (Renshaw and Ford, 1983; Ford and Renshaw, 1984; Renshaw and Ford, 1984), the authors emphasize that in ecological data, especially in growth processes, directional components are common and assuming isotropy is not acceptable.

Anisotropy can be caused by several mechanisms. In the existing literature, directional analysis has mainly focused on two types of anisotropy: geometric anisotropy, where anisotropy is caused by a linear transformation of a stationary and isotropic process, and increased intensity of points along directed lines. Geometric anisotropy has been considered both for clustered point patterns, such as the Welsh chapel data (Mugglestone and Renshaw, 1996b; Møller and Toftaker, 2014), and for regular point patterns, such as the earlier mentioned amacrine cells and air bubble data sets. The *Ambrosia dumosa* dataset is an example of an anisotropic point pattern with increased intensity along directed lines (Rosenberg, 2004). In this paper, we restrict ourselves to these two types.

We concentrate on reviewing methods that are generally applicable for any point pattern and do not require a specific model assumption. Therefore, we focus on methods that are based on spatial summary statistics, such as the nearest neighbor distance distribution function, Ripley's K function and the pair correlation function, spectral analysis, and wavelet analysis. Several tests for isotropy have been proposed based on these methods. Some of these tests are asymptotic, some based on Monte Carlo simulations and some rely on replicated data. Directional analysis based on summary statistics has been used both in 2D and 3D, whereas analysis based on spectral analysis and wavelets has so far been introduced only in 2D. Additionally, wavelet analysis discussed in the literature concentrates on situations where we have increased intensity along directed lines. We mention some models for anisotropic point patterns but will not discuss them thoroughly as they are typically tailor made for a specific situation and data, with particular models for the point location processes.

To illustrate how a typical result of each method looks like, we apply all presented methods to two simulated 2D data sets, a regular compressed point pattern and a clustered pattern with increased intensity along directed lines. Both point patterns are realizations of stationary point processes. The regular pattern is an example of geometric anisotropy and is comparable to the amacrine cells and the air bubble data. The clustered pattern has similar features as the *Ambrosia dumosa* data and the pyramidal cells data sets considered in Rafati et al. (2016). The Welsh chapels data set, an example of a geometrically anisotropic clustered pattern, is not covered by these two examples. Based on the analysis of the simulated example point patterns, we make some observations about how suitable the different methods seem to be in these particular cases. However, a much more thorough analysis of the methods would be needed to be able to give any general recommendations.

Introduction to our notation is given in Section 2. In Section 3, we describe the two mechanisms causing anisotropy included in this paper, geometric anisotropy and clustering along directed lines. Methods based on nearest neighbour and second order summary statistics are recalled in Section 4, spectral analysis in Section 5, and wavelets in Section 6. Finally, we give an overview of the tests for isotropy presented in the literature in Section 7 and conclude by discussing our findings and future work.

2 Definitions and notation

In this section, we give some basic definitions concerning point processes. Let $\mathbf{x} = \{x_1, ..., x_n\}$ be a point pattern observed in a window $W \subseteq \mathbb{R}^d$ with $n > 2$ points. We assume that $\mathbf{x}$ is obtained by intersecting a realization of a simple (no multiple points) point process X with W, and that W is bounded with volume $|W|$.

The point process X is stationary, if its distribution is invariant under shifts in $\mathbb{R}^d$. If its distribution is invariant under rotations around the origin, the process is isotropic. Unless stated otherwise, we will assume that the point process X is stationary. The task is then to detect and quantify anisotropies in the observed point pattern. Note that some of the methods described below can be generalized to second order intensity reweighted stationary point processes (Baddeley et al., 2000), see Section 4.3.4.

The point process X can be regarded as a random collection of points (as above) or as a counting measure on $\mathbb{R}^d$, also denoted by X. In the latter case, $X(B)$ is the random number of points of X in a Borel set B in $\mathbb{R}^d$. In particular, $X(W) = n$. The intensity function of X is defined as

$$\lambda(x) = \lim_{|dx| \to 0} \frac{\mathbb{E}[X(dx)]}{|dx|}.$$

In the stationary case, $\lambda(x) \equiv \lambda$, where $\lambda > 0$ is a constant. The Palm distribution of X will be denoted by P_x for any $x \in \mathbb{R}^d$. Heuristically, it can be interpreted as the conditional distribution of X given
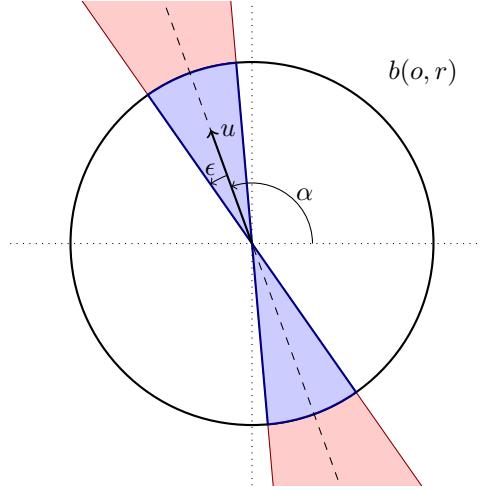
Figure 1: The sets $C(\alpha, \epsilon)$ (red and blue) and $S(\alpha, \epsilon, r)$ (blue) when $\alpha = \alpha(u) = 3\pi/5, \epsilon = \pi/12$, and $r = 1.5$.

$x \in X$. The corresponding expectation is denoted by $\mathbb{E}_x$. For stationary X, it suffices to consider only P_o and $\mathbb{E}_o$ where $o \in \mathbb{R}^d$ is the origin.

In the directional analysis, it is often convenient to use polar coordinates in 2D and spherical coordinates in 3D. The 2D Cartesian coordinates (x, y) can be represented in polar coordinates as

$$r = \sqrt{x^2 + y^2} \in [0, \infty) \quad \text{and} \quad \phi = \arctan(y/x) \in [0, 2\pi)$$

and the 3D coordinates (x, y, z) as

$$r = \sqrt{x^2 + y^2 + z^2} \in [0, \infty),$$
$$\phi = \arctan(y/x) \in [0, 2\pi), \quad \text{and}$$
$$\theta = \cos^{-1}\left(\frac{z}{\sqrt{x^2 + y^2 + z^2}}\right) = \cos^{-1}(z/r) \in [0, \pi].$$

We use the convention that the 2D angle goes anti-clockwise from the x-axis. Formally, we set $\arctan(y/x) = [(\mathrm{atan}(y/x) + 2\pi) \bmod 2\pi]$ where $\mathrm{atan}(y/x)$ is $\tan^{-1}(y/x)$ if $x > 0$, $\tan^{-1}(y/x) + \mathrm{sign}(y)\pi$ if $x < 0$, $\mathrm{sign}(y)\pi/2$ if $x = 0$, and left undefined if $x = y = 0$.

Let $b(x, r)$ denote a ball in $\mathbb{R}^d$ with center x and radius $r > 0$. For a unit vector $u \in \mathbb{R}^d$, i.e. an element of the unit sphere $S^{d-1} = \{x \in \mathbb{R}^d : ||x|| = 1\}$, the polar/spherical coordinates are $(r, \alpha) = (1, \alpha)$, where $\alpha = \alpha(u)$ gives the angle ϕ in 2D and angles ϕ and θ in 3D. The infinite double cone with a central axis spanned by u and with the opening half angle $\epsilon > 0$ is denoted by $C(u, \epsilon)$. Furthermore, we set $S(u, \epsilon, r) = C(u, \epsilon) \cap b(o, r)$. Equivalently, when considering angles we will write $C(\alpha, \epsilon)$ and $S(\alpha, \epsilon, r)$ with $\alpha = \alpha(u)$. An illustration is shown in Figure 1.

We denote the Minkowski sum of two sets A and B in $\mathbb{R}^d$ by

$$A \oplus B = \{a + b : a \in A, b \in B\}.$$

Then, $W_x = W \oplus \{x\}$ is the translation of the window W by a point $x \in \mathbb{R}^d$. The Minkowski difference of A and B is defined as

$$A \ominus B = \{x \in \mathbb{R}^d : B_x \subset A\}.$$

We will write $1(\mathcal{E})$ for the indicator function taking the value 1 if event $\mathcal{E}$ is true and 0 otherwise.

3 Anisotropy mechanisms

A point pattern can be anisotropic in different ways. Here, we recall two types of anisotropy, geometric anisotropy and anisotropy caused by oriented clusters, that are most common in the point process

literature.

3.1 Geometric anisotropy

Let X_0 be a stationary and isotropic point process. Define a point process X by the transformation $X = TX_0 = \{Tx : x \in X_0\}$, where $T : \mathbb{R}^d \to \mathbb{R}^d$ is an invertible linear mapping. The corresponding $d \times d$-matrix will also be denoted by T. Since X_0 is isotropic, we can decompose the mapping into two matrices, $T = RC$, where R is a rotation matrix and C is a diagonal scaling matrix that compresses and stretches the dimensions. Note that the definition excludes shear.

If C is not a multiple of the identity matrix, the resulting process X can be anisotropic. Furthermore, it has the following properties:

1. X is stationary.

2. The connection between the counting measures is $X(B) = X_0(T^{-1}B)$ for any Borel set $B \subset \mathbb{R}^d$.

3. The intensities of X and X_0 are related via $\lambda_X = \det(T^{-1})\lambda_{X_0}$.

4. If X_0 is a stationary Poisson process with intensity $\lambda > 0$, then X is a stationary Poisson process with intensity $\det(T^{-1})\lambda$. In particular, X is also isotropic.

In Møller and Toftaker (2014), the type of anisotropy defined above is called geometric anisotropy. The term is borrowed from geo-statistics: The distance $r = x - x'$ of two points is transformed into $r_T = (x - x')^T T^T T (x - x')$, so that, in analogy to the covariance function in geo-statistics, the second order properties of X depending on $r = x - x'$ can be expressed as the second order properties of X_0 depending on $r_{T^{-1}}$. An alternative name for the transformation is elliptical anisotropy: For any sphere $S = \{x : x^T x = m\}$ in $\mathbb{R}^d$ we have that $TS = \{x : x^T (TT^T)^{-1} x = m\}$ is an ellipsoid (ellipse in 2D). Note that Møller and Toftaker (2014) formulate the transformation in terms of $\Sigma = TT^T$.

Geometric anisotropy has been studied as a model of anisotropy for both clustered and regular point patterns. Cluster processes were considered in Møller and Toftaker (2014), Guan et al. (2006), and Wong and Chiu (2016). Møller and Toftaker (2014) consider log-Gaussian Cox processes and shot noise Cox processes, Guan et al. (2006) and Wong and Chiu (2016) use Poisson cluster processes with elliptic clusters generated by an anisotropic multivariate normal distribution, i.e. anisotropic Thomas processes. Note that, due to Property 4 above, linear transformations of Poisson cluster processes are Poisson cluster processes with transformed clusters. An example of a real data set with this structure is the Welsh chapel data discussed in Møller and Toftaker (2014) and Mugglestone and Renshaw (1996a). This data set as well as some model realisations are shown in Figure 2.
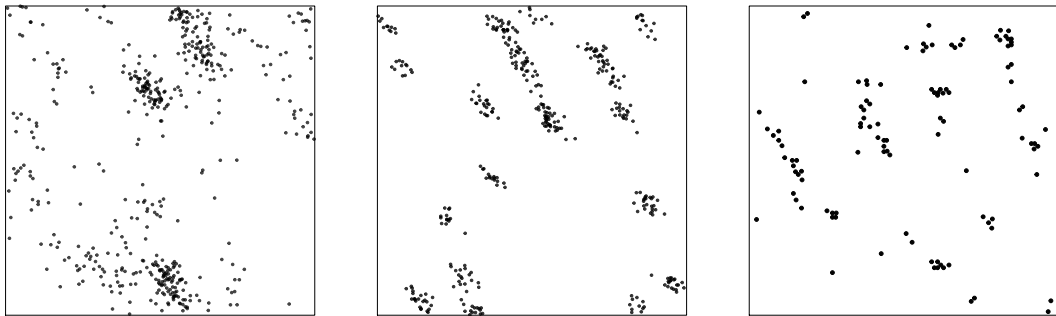

Figure 2: Linearly transformed realisations of a log-Gaussian Cox process (left) and a Thomas process (middle), and the Welsh chapels data (right).

Linear transformations of regular point patterns are studied in Redenbach et al. (2009), Rajala et al. (2016) and in Wong and Chiu (2016) (see Figure 3, left). In the simulation study in Wong and Chiu (2016), the regular case is represented by anisotropic Gibbs hard-core processes, in Rajala et al. (2016) by

transformed Strauss processes, and in Redenbach et al. (2009) by transformed Matern hard core processes. As an example of real data, Redenbach et al. (2009) and Rajala et al. (2016) study the locations of air bubbles in polar ice (see Figure 3, middle), while in Wong and Chiu (2016), amacrine cells in the retina of a rabbit are investigated (Diggle (1986), see Figure 3, right).
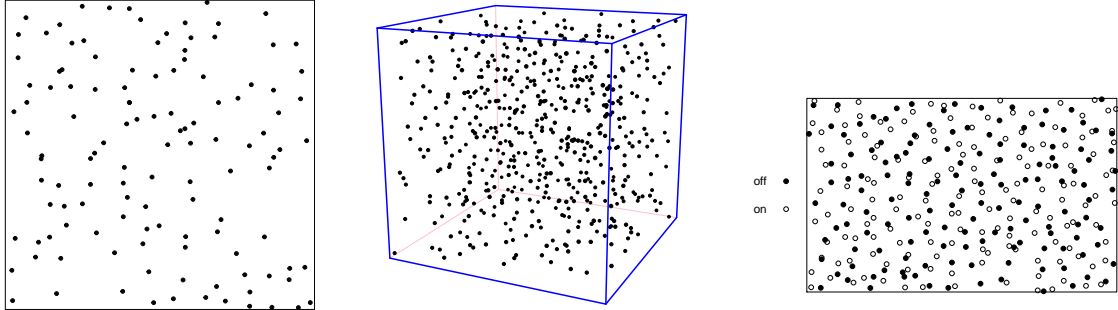

Figure 3: A realization of a compressed regular point process (left), locations of air bubbles in polar ice (middle), and locations of 152 amacrine cells labelled 'on' and 142 cells labelled 'off' (right).

3.2 Oriented clusters

Cluster processes with bounded anisotropic clusters were already considered in the last section. Further typical examples found in the anisotropy literature are Poisson processes with increased intensity along directed lines, see for instance Rosenberg (2004). These processes can be considered stationary if the distribution of line locations is stationary, e.g. given by a stationary Poisson line process (see Figure 4, left, middle). A real data set showing such structure is the *Ambrosia dumosa* data (Miriti et al., 1998), shown in Figure 4 (right). To model point patterns with clustering around oriented line segments, models introduced in Lawson et al. (2007) can be useful.

An example in 3D can be found in Rafati et al. (2016) where locations of pyramidal cells in the brain are investigated. The minicolumn hypothesis in neuroscience states that these cells are organized in parallel columns which results in an anisotropic arrangement. To mimic this structure, Møller et al. (2016) introduced a model called Poisson line cluster point process (PLCPP), where the points are clustered around the lines of a Poisson line process. If the directional distribution of this process differs from the uniform distribution on the sphere, the resulting point process is anisotropic. See Figure 5 for a sample of the minicolumn data and a realisation of a PLCPP model.
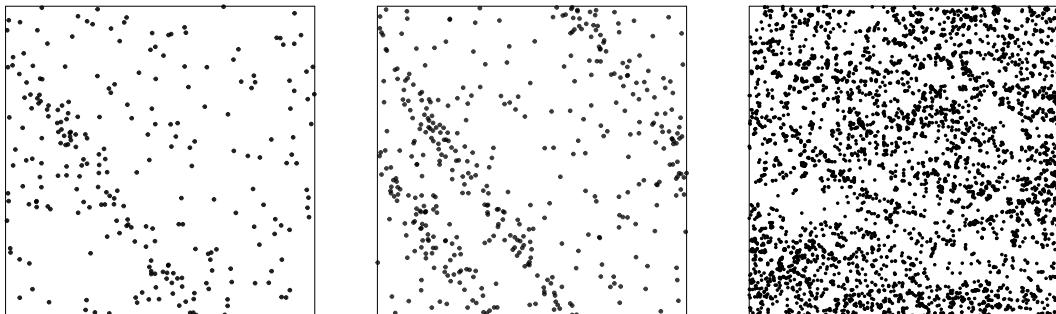

Figure 4: A superposition of a stationary Poisson process with a Poisson process whose intensity is concentrated around a line (left), and a version of this process with three lines which can be considered stationary (middle). Ambrosia dumosa data (right).
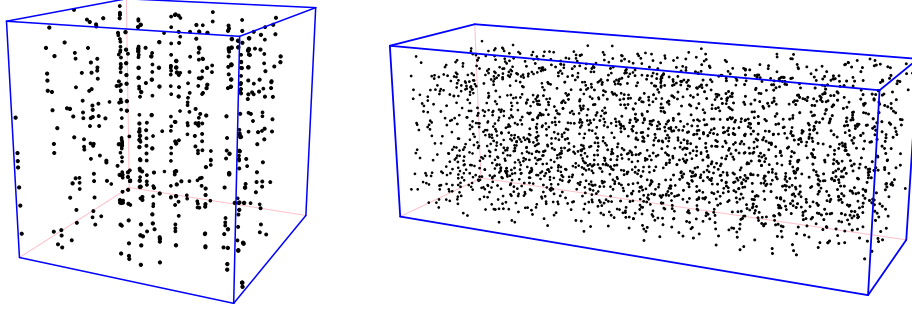
Figure 5: A realisation of a Poisson line cluster point process (left), and locations of nucleoli of pyramidal cells (right).

4 Analysis based on point process summary statistics

In this and the two following sections we will review non-parametric methods for anisotropy analysis. In order to understand how typical results of an analysis may look like, we apply the methods to the regular point pattern shown in Figure 3, left, as well as the clustered pattern shown in Figure 4, middle. The former is an example of geometric anisotropy and is obtained from an isotropic realization of a 2D Strauss process with range of interpoint interaction $R = 0.1$, strength of repulsive interaction $\gamma = 0.1$, and first order parameter $\beta = 100$. We then use the geometric anisotropy mechanism to compress in y-direction by a factor of 0.6 and stretch in x-direction by $1/0.6$, i.e. $C = diag(1/0.6, 0.6)$. Subsequently, the pattern is rotated clockwise $\pi/6$ radians. The analysis is based on a pattern observed inside the window $W = [-1, 1] \times [-1, 1]$. The clustered pattern is an example of clustering along directed lines and is obtained as a superposition of a stationary Poisson process with intensity $\lambda = 200$ with three parallel stripes formed by Poisson processes whose intensity functions are $\lambda_l(x) = 100\mathcal{N}(|x|_l; 0, 0.03^2), l = 1, 2, 3,$ where $\mathcal{N}$ stands for 1D Gaussian density function, and $|\cdot|_l$ is the distance from line $l = 1, 2, 3$. The lines, and therefore the stripes, form a fixed angle of $\pi/5$ w.r.t. the y-axis.

Classical summary statistics from spatial point process theory such as the nearest neighbour distance distribution function and Ripley's K-function were originally defined for isotropic point processes. Several directional versions of these summary statistics have been formulated and are suitable for detecting anisotropies. In a typical application, a summary statistic is estimated separately for different directions, and differences between the estimates for these directions indicate anisotropy of the point pattern. Below, we discuss such analyses based on the nearest neighbour distance distribution function, Ripley's K function and the pair correlation function.

4.1 Visualization of anisotropy: Fry plot

In clustered patterns, the shape and orientation of the clusters may reveal some anisotropies. To detect anisotropy in a regular pattern can be harder, see the point pattern in Figure 3, left. However, if we plot the pairwise difference vectors $x_i - x_j$ for all point pairs, as can be seen in Figure 6, left, we can see that the central area around the origin has fewer points, indicating less pairs at short distances than long distances, typical for a regular process. We can also see that the shape of the central area is elliptic, not circular. The pairwise difference vectors do not seem to have a rotationally invariant distribution, and therefore the pattern could be anisotropic. The clustered pattern shows a somewhat inverse structure: the central area is dense, typical for clustered processes (Figure 6, right). The elongated shapes taper off at the ends as we observe fewer and fewer long distance pairs due to a finite observation window.

Such plots, called *Fry plots* (Fry, 1979), are not only visually informative but can also be very helpful in the formal analysis as can be seen in Rajala et al. (2016), who fitted ellipsoids to the Fry plot to estimate the direction of the linear transformation in the case of geometric anisotropy (see Section 4.3.3). Nearest neighbor analysis can be based on a similar plot, where only the nearest neighbor vectors have

been plotted. The Fry plot was, in fact, an improved version of plots based on the nearest neighbor distance vectors which were first introduced by Ramsay (1967) who applied them to measure bulk strain in rocks with rigid objects. These plots are limited, however, to 2D only: In 3D it is hard to see any structure inside the point cloud.
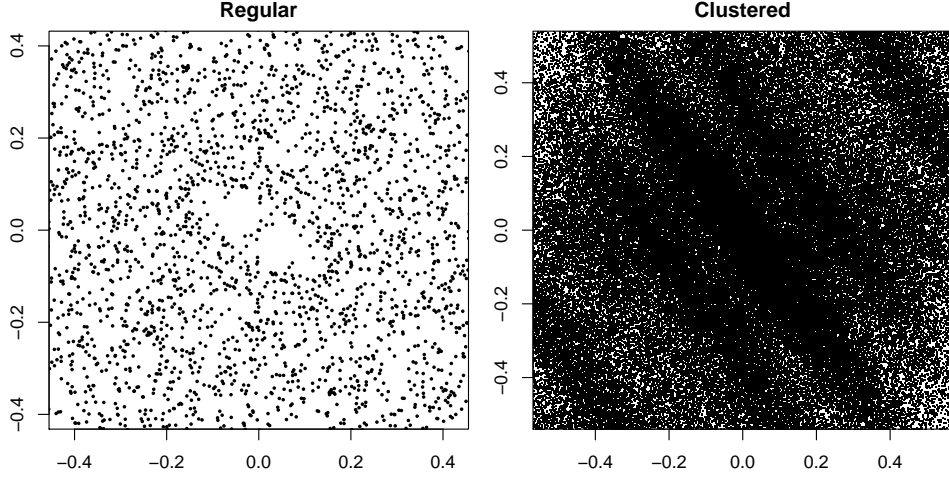

Figure 6: Fry points of the example patterns shown in Figures 3 and 4 : Regular (left) and clustered (right).

4.2 Nearest neighbour analysis

For $x \in \mathbf{x}$, let $\eta(x; \mathbf{x}) = \{y \in \mathbf{x} \setminus x : ||x - y|| \leq ||x - z|| \ \forall z \in \mathbf{x} \setminus x\}$ denote the nearest neighbour of x in the set $\mathbf{x}$. Let $d_i = d_i(\mathbf{x}) = ||x_i - \eta(x_i; \mathbf{x})||$ be the distance from $x_i \in \mathbf{x}$ to its nearest neighbour in $\mathbf{x}$, and let $\alpha_i = \alpha_i(\mathbf{x}) = \alpha(x_i - \eta(x_i; \mathbf{x}))$ be the angle(s) in the polar (spherical) representation of the vector from x_i to its nearest neighbour in $\mathbf{x}$. Recall that in 2D, we have $\alpha_i = \phi_i$ while in 3D, $\alpha_i = (\phi_i, \theta_i)$.

For an isotropic point process X, the nearest neighbour angles are uniformly distributed i.e. $\phi \sim U[0, 2\pi]$ in both 2D and 3D. In 3D, the additional angle θ is independent of ϕ with $\theta \sim \cos^{-1}(Z)$, $Z \sim U[-1, 1]$. The corresponding angle distribution functions are given by

$$F(\alpha) = F(\phi) = \phi/2\pi, \quad \phi \in [0, 2\pi] \tag{1}$$

and

$$F(\alpha) = F(\phi, \theta) = \phi(1 - \cos \theta)/4\pi, \quad (\phi, \theta) \in [0, 2\pi] \times [0, \pi]. \tag{2}$$

When testing isotropy via a uniformity test of the nearest neighbour angles (see Section 7), it should be noted that the nearest neighbour angles of an observed pattern are not independent. For example, when x_j is the nearest neighbour to x_i and vice versa, then $\alpha_i = \check{\alpha}_j$, where $\check{}$ denotes the antipodal direction of α.

Illian et al. (2008), Ch. 4.5.2. discuss the *nearest neighbour orientation density* $\varphi(\alpha)$, which is the density of the CDF $F(\alpha)$ in (1). Its edge corrected kernel estimator with some kernel k_h and some bandwidth h is given by

$$\hat{\varphi}(\alpha) = \sum_{i=1}^{n} \frac{1(d_i < e_i) k_h(\alpha - \alpha_i)}{|W \ominus b(0, d_i)|} / \hat{\lambda}_{nn}.$$

The edge correction reduces bias by considering only those points x_i whose distance e_i to the border of the observation window W is larger than d_i. This way, only points whose nearest neighbour is observed within W are included in the sampling. The intensity estimator $\hat{\lambda}_{nn}$ is adapted to this particular edge correction and is given by

$$\hat{\lambda}_{nn} = \sum_{i=1}^{n} \frac{1(d_i < e_i)}{|W \ominus b(o, d_i)|}.$$

7

Note that the original definition of $\varphi(\alpha)$ was for $\alpha = \phi \in [0, \pi]$ with antipodal flip when $\phi_i > \pi$. Furthermore, the kernel should be wrapped around the angle domain to avoid discontinuities.

König and Schmidt (1992) define a summary for the directional distribution of the so-called s-nearest neighbour, i.e. the nearest neighbour outside some range s. For simplicity, we will restrict attention to the case $s = 0$. Let d_o and α_o denote the random distance and direction from the typical point o of X to its nearest neighbour, and consider the direction set $\mathcal{A} \subset S^{d-1}$. Then the *nearest neighbour directional distribution* is defined as the distribution on the unit sphere

$$D_r(\mathcal{A}) = P_o(\alpha_o \in \mathcal{A} | d_o < r), \quad \mathcal{A} \subset S^{d-1}. \tag{3}$$

Consequently, $D_r(\mathcal{A})$ can be interpreted as the probability that the nearest neighbour of the typical point of X is in $\mathcal{A}$ given that the distance to the nearest neighbour is at most r. A consistent and asymptotically unbiased estimator for the directional distribution is given by

$$\hat{D}_r(\mathcal{A}) = \frac{\sum_{i=1}^{n} 1(d_i < r)1(\alpha_i \in \mathcal{A})1(x_i \in W \ominus b(o, r))}{\sum_{i=1}^{n} 1(d_i < r)1(x_i \in W \ominus b(o, r))}. \tag{4}$$

Redenbach et al. (2009) consider a similar statistic where the roles of $\mathcal{A}$ and r are exchanged. The *global directional nearest neighbour distance distribution* function $G_{\mathrm{glob},\mathcal{A}}$ is defined as the distribution of d_o conditioned on $\alpha_o \in \mathcal{A}$,

$$G_{\mathrm{glob},\mathcal{A}}(r) = P_o(d_o < r | \alpha_o \in \mathcal{A}).$$

An estimator for this statistic is given by

$$\hat{G}_{\mathrm{glob},\mathcal{A}}(r) = \frac{\sum_{i=1}^{n} 1(d_i < r)1(\alpha_i \in \mathcal{A})1(x_i \in W \ominus b(o, d_i))}{\sum_{i=1}^{n} 1(\alpha_i \in \mathcal{A})1(x_i \in W \ominus b(0, d_i))}.$$

In Redenbach et al. (2009), the sets $\mathcal{A}$ are chosen such that the angle between elements of $\mathcal{A}$ and a prespecified direction α is less than a given ϵ. Hence, only points with the nearest neighbour in a double cone $C(\alpha, \epsilon)$ are considered in the estimation, which may drastically reduce the sample size. As an alternative statistic, Redenbach et al. (2009) also consider a local nearest neighbour distance distribution function $G_{\mathrm{loc},\alpha,\epsilon}$ which is the cumulative distribution function of the distance from a typical point to the nearest neighbour in the double cone $C(\alpha, \epsilon)$, or

$$G_{\mathrm{loc},\alpha,\epsilon}(r) = P_o\left[d_o\left(X \cap C(\alpha, \epsilon)\right) < r\right].$$

Writing $d_i^{\alpha,\epsilon} = d_i(\mathbf{x} \cap (x_i + C(\alpha, \epsilon)))$ for the $C(\alpha, \epsilon)$-nearest neighbour distance of x_i, Redenbach et al. (2009) propose the Hanisch type estimator

$$\hat{G}_{\mathrm{loc},\alpha,\epsilon}(r) = \frac{\hat{G}_{H,\mathrm{loc},\alpha,\epsilon}(r)}{\hat{G}_{H,\mathrm{loc},\alpha,\epsilon}(\infty)},$$

where

$$\hat{G}_{H,\mathrm{loc},\alpha,\epsilon}(r) = \sum_{i=1}^{n} \frac{1(d_i^{\alpha,\epsilon} < r)1(x_i \in W \ominus S(\alpha, \epsilon, d_i^{\alpha,\epsilon}))}{|W \ominus S(\alpha, \epsilon, d_i^{\alpha,\epsilon})|}.$$

This way, each point of the process contributes to the estimation of G_{loc} (unless disregarded by the edge correction).

Figure 7 depicts the four nearest neighbour summaries for the regular and clustered example patterns in Figure 3 and 4. The two leftmost plots in the top row show the orientation density. No guidelines were given in Illian et al. (2008) for the bandwidth, so the first estimate uses a bandwidth selected by eye and the second plot a bandwidth double that of the first estimate. The smoother curve for the regular pattern indicates two peaks around $\alpha = \pi/2 - \pi/6$ and $\alpha + \pi$, correctly identifying the rotation in the pattern. The two rightmost plots on the top row depict the directional distribution plots, which were computed with $\mathcal{A} = \mathcal{A}(a) = \{u \in S^1 : \alpha(u) \in [0, a]\}, a \in [0, 2\pi]$. No guidelines were given by König and Schmidt (1992) for choosing the ranges, so we chose $r = 0.1$ and $r = 0.3$. The former is the known interaction range for the regular process, and the latter shows the directional distribution for all points. The two
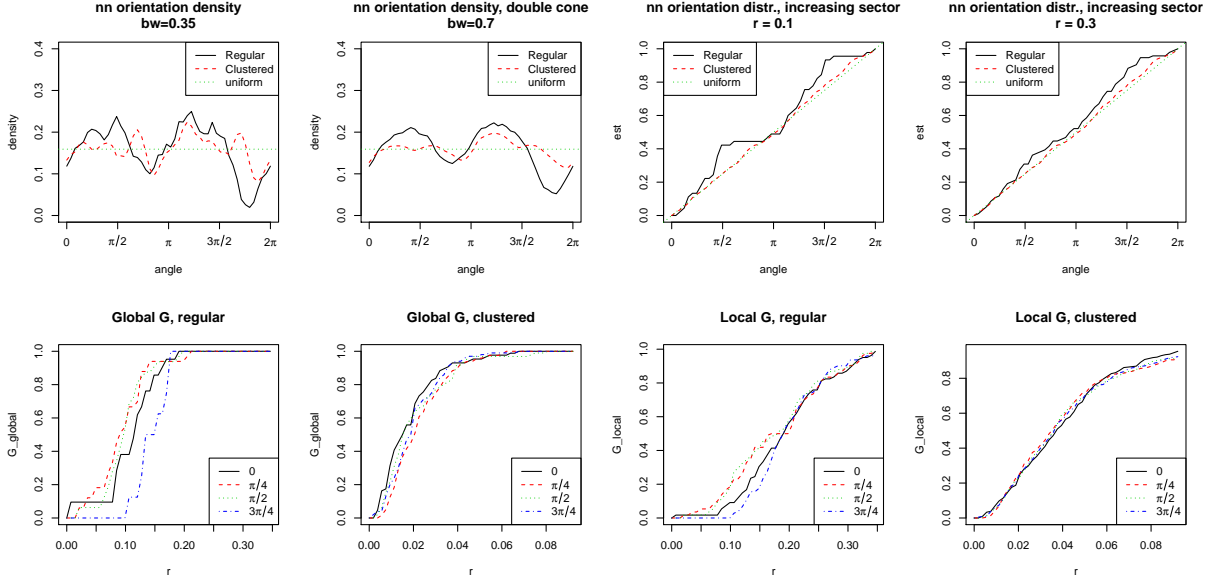
Figure 7: Directional nearest neighbour summaries for the two example patterns.

peaks for the regular pattern are visible at short range $r = 0.1$ but not at longer range $r = 0.3$, so the anisotropy information is present only in short nearest neighbour distances. The directed G-functions in directions $\alpha = 0, \pi/4, \pi/2$, and $3\pi/4$ are plotted in the bottom row. Redenbach et al. (2009) do not discuss the choice of the sector half-angle ϵ. To guarantee that there is no overlap between the double cones, $\pi/8$ is the largest half-angle that can be chosen. Here, we chose $\pi/4$ to double the sample sizes (we have only 122 points in the regular case, of which only ϵ/π on average contribute per direction). The G-function values in directions close to the correct rotation angle differ from those computed along the coordinate axes, indicating departure from anisotropy. All summaries detect poorly the anisotropy of the clustered pattern, as the pattern behaves like an isotropic Poisson process at the very small scales where the nearest neighbour information is concentrated.

4.3 Second order analysis: Ripley's K function and the pair-correlation function

Nearest-neighbour characteristics are known to be rather short-sighted as demonstrated by the examples discussed above. Therefore, Ripley's K function and the pair correlation function (pcf) may be better tools for directional analysis. There are several ways to derive a second order summary that takes direction into account. The first suggestion can be found in Ohser and Stoyan (1981). They define in 2D an angle dependent version of the K-function, such that the function is cumulative in both range and angle.

4.3.1 K-functions

Second order statistics are based on the second-order factorial moment measure

$$\alpha^{(2)}(A \times B) = \mathbb{E}\left(\sum_{x,y \in X}^{\neq} 1(x \in A)1(y \in B) \right)$$

for Borel sets A and B in $\mathbb{R}^d$, where the upper $\neq$ means that only pairs with $x \neq y$ are considered in the sum. We assume that the (second order) product density $\rho^{(2)}$ exists, i.e.

$$\alpha^{(2)}(A \times B) = \int_A \int_B \rho^{(2)}(x,y)dydx.$$

9

The value $\rho^{(2)}(x, y)dxdy$ is the probability that X has a point in each of the infinitesimally small discs with centers x and y and volumes dx and dy, respectively. For stationary X, we have that $\rho^{(2)}(x, y) = \rho^{(2)}(o, y - x) =: \rho^{(2)}(z)$ where $z = y - x$. Hence,

$$\alpha^{(2)}(A \times B) = \int_A \int_B \rho^{(2)}(o, y - x)dydx = \int_A \int_{B-x} \rho^{(2)}(z)dzdx.$$

Defining the reduced second-order moment measure $\mathcal{K}$ via

$$\lambda^2 \mathcal{K}(B) = \int_B \rho^{(2)}(z)dz,$$

we arrive at

$$\alpha^{(2)}(A \times B) = \lambda^2 \int_A \mathcal{K}(B - x)dx.$$

Alternatively, by the Campbell-Mecke formula (Illian et al., 2008, Eq.(4.1.8))

$$\alpha^{(2)}(A \times B) = \mathbb{E} \sum_{x \in X} 1(x \in A)X(B \setminus \{x\}) = \lambda \int_A \mathbb{E}_o[X((B - x) \setminus \{o\})]dx.$$

These two equations lead to

$$\lambda \mathcal{K}(B) = \mathbb{E}_o(X(B \setminus \{o\})),$$

and we can interpret $\lambda \mathcal{K}(B)$ as the expected number of (further) points in B conditioned on $o \in X$.

In the case of geometric anisotropy, i.e. when we consider $X = TX_0$, where T is an invertible linear mapping and X_0 a stationary and isotropic point process, the point process X has the following further property

5. The reduced second order moment measures are related by $\mathcal{K}_X(B) = \det(T)\mathcal{K}_{X_0}(T^{-1}B)$ for any Borel set $B \subset \mathbb{R}^d$.

An estimator for $\lambda^2\mathcal{K}(B)$ is given by

$$\widehat{\lambda^2\mathcal{K}}(B) = \sum_{x,y \in \mathbf{x}}^{\neq} \frac{1(y - x \in B)}{|W_x \cap W_y|}.$$

Unbiasedness of the estimator follows from the following generalization of the Campbell theorem

$$\mathbb{E} \sum_{x,y \in X}^{\neq} f(x, y) = \int \int f(x, y)\alpha^{(2)}(d(x, y)) = \lambda^2 \int \int f(x, x + h)dx\mathcal{K}(dh) \qquad (5)$$

for all nicely behaving f, see Illian et al. (2008, p. 228). The choice $K(r) = \mathcal{K}(b(o, r))$ yields Ripley's K-function. Note that $\rho^{(2)}$ is commutative, so all summaries based on the factorial moment measure are antipodally symmetric such that only the upper hemisphere of directions in S^{d-1} needs to be considered.

Anisotropic second order summaries can now be defined in terms of parametric, rotation variant test-sets $B = B(\psi)$, where ψ parameterizes the set in terms of direction and length-scale. An early 2D example of such a construction was given by Ohser and Stoyan (1981), who chose B to be the sector of radius r with angle γ from the positive x-axis and considered this as a function $K(r, \gamma)$. This has an interpretation as a cumulant in $\gamma \in [0, 2\pi]$. Stoyan (1991) then extended this idea to a more freely defined sector, where both bounding rays can have arbitrary angles γ and $\Gamma > \gamma$ w.r.t. the x-axis. Setting $\alpha = (\gamma + \Gamma)/2$ and $\epsilon = (\Gamma - \gamma)/2$ and using the sector $S(\alpha, \epsilon, r)$ as the test set, the 3D extension becomes obvious. Figure 8, left panel, illustrates the construction in 2D. The resulting version of the K-function is called the *conical K-function*, and was used by Redenbach et al. (2009) to assess anisotropy in 3D patterns by contrasting the z-direction to x- and y-directions.

An alternative choice is the *cylindrical K-function* (Møller et al., 2016). The test-set is the origin centred rectangle (2D) or cylinder (3D) $L(r, u, h_c)$ with major-axial direction unit vector $u \in S^{d-1}$, with height $2r$ and cross-section half-length $h_c, 0 < h_c < r$, i.e. $L(r, u, h_c) = \{x \in \mathbb{R}^d : d(x, l(u)) \leq$
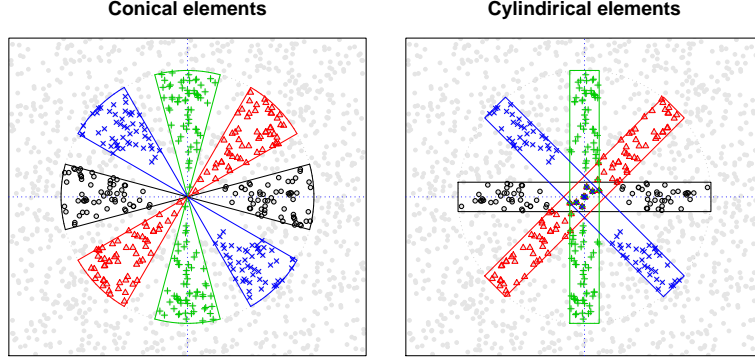
Figure 8: The anisotropic test sets for the conical (left) and cylindrical (right) K functions, overlaid on top of the Fry plot of the regular example pattern.

$h_c, d(x, l(u)^\perp) \leq r\}$, where $l(u)$ is the line spanned by u, $l(u)^\perp$ is the orthogonal hyperplane, and $d(x, l(u))$ and $d(x, l(u)^\perp)$ denote the distances from a point x to the line and the hyperplane, respectively. Figure 8, right panel, illustrates this construction. Two notable differences to the conical case are that the local area of the test set is not increasing in r, and that the test-sets in different directions always overlap at short ranges.

Illian et al. (2008) Ch. 4.5.3. define an orientation distribution in 2D using another version of the K function. Let $T(\alpha, r)$ be the r-sector formed by the positive x-axis and angle α, and let $T(\alpha, r_1, r_2) = T(\alpha, r_2) \setminus T(\alpha, r_1)$ for $r_1 < r_2$. Then we set $K(r_1, r_2, \alpha) = \mathcal{K}(T(\alpha, r_1, r_2))$. For fixed ranges r_1 and r_2,

$$F_K(\alpha) = \frac{K(r_1, r_2, \alpha)}{K(r_1, r_2, \pi)}$$

is an orientation distribution function with a density $f_{r_1, r_2}(\alpha)$, which we call *second order orientation density*. Its estimator is, up to the constant $K(r_1, r_2, \pi)$,

$$\hat{f}_{r_1, r_2}(\alpha) \propto \sum_{x, y \in \mathbf{x}}^{\neq} \frac{1(r_1 < ||x - y|| < r_2) k_{h_\alpha}(\alpha(x, y) - \alpha)}{|W_x \cap W_y|}$$

with some kernel k_{h_α} and bandwidth h_α. For isotropic processes $f_{r_1, r_2} \equiv 1/\pi$. The definition can clearly be generalized to 3D. However, its application is complicated since in this case a function of two angles has to be considered.

Figure 9 depicts the second order summaries for the example patterns in directions $\alpha = 0, \pi/4, \pi/2, 3\pi/4$. For the conical K, no guidelines for selecting the sector half-angle are available, so we chose the maximal angle with no overlaps, $\pi/8$. Guidelines for the cylinder cross-section half-lengths under specific volume and shape constraints are discussed by Safavimanesh and Redenbach (2016), who derive equations connecting the cylinder height and half-length to a corresponding conical shape. For simplicity, we fixed $h_c = 0.03$ by eye (as shown in Figure 8). The conical and cylindrical K-functions show decreased amount of pairwise directions near the true stretching direction and increased amount near the direction of compression of the regular pattern when compared to alternative directions. In the clustered pattern, there are most pairs in the direction of the clusters. For the orientation density, in Illian et al. (2008) the range interval $[r_1, r_2]$ was "found by experimentation". For the example, we selected two intervals corresponding to "short" and "long" ranges. The level of smoothing was not discussed, so by some trial and error we decided on $h_\alpha = 3/\sqrt{\lambda}$ for the Epanechnikov kernel, leading to stable looking estimates. The orientation density captures the anisotropy of the regular process well when using a short range interval $[r_1, r_2]$ but is practically uniform for long ranges, and vice versa for the clustered pattern.

An alternative visualisation of any density function over angles in 2D is given by the rose-of-directions (Mecke and Stoyan, 1980). It depicts the density over a rotation, emphasizing departures from a circle. Figure 10 illustrates the rose for the short range 2nd order orientation densities of our examples.
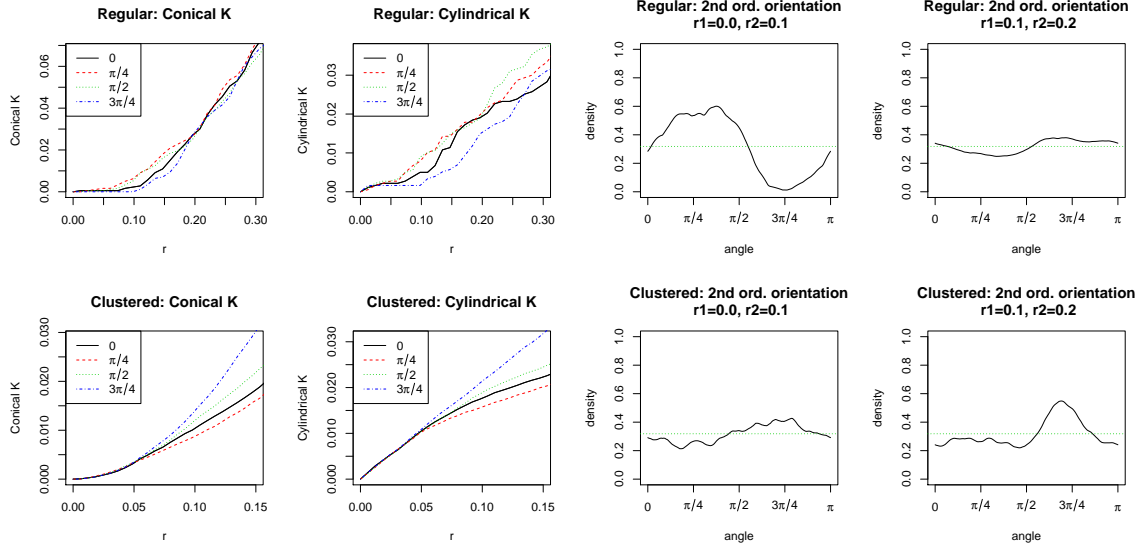
Figure 9: The conical K with $\epsilon = \pi/8$, cylindrical K with $h_c = 0.03$, and orientation densities with $(r_1 = 0, r_2 = 0.1)$ and $(r_1 = 0.1, r_2 = 0.2)$ and bandwidth $h_\alpha = 3/\sqrt{\lambda}$. Top row: Regular pattern. Bottom row: Clustered pattern.
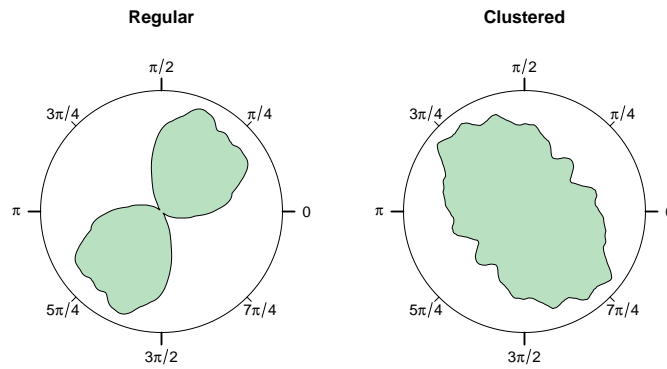

Figure 10: The rose-of-directions plots of the 2nd order orientation densities in Figure 9 with $r_1 = 0$ and $r_2 = 0.1$.

All summary statistics introduced in this section will most clearly reveal the anisotropy if r is chosen according to relevant scales in the point pattern. For instance, the hardcore or cluster radius R of geometric anisotropic point processes may strongly vary in different directions. Hence, obvious differences in the directional K-function can be expected for $r \approx R$.

4.3.2   Pair correlation functions (pcf)

The cumulative nature of the $\mathcal{K}$-measures can sometimes obscure fine details, and their derivatives offer a more detailed option. We can rewrite the definition of the $\mathcal{K}$-measure as

$$\mathcal{K}(B) = \lambda^{-2} \int_B \rho^{(2)}(z)dz = \int_B g(z)dz,$$

where the integrand

$$g(z) = \frac{\rho^{(2)}(z)}{\lambda^2}$$

is called the pair correlation function (pcf). The pcf is more practical than the product density as it is independent of intensity and the Poisson process has $g \equiv 1$. Note that it is, however, not a correlation in the conventional statistical sense as it takes values on $[0, \infty)$. In the case of geometric anisotropy, the process X has the following further properties

6. The second order product densities (if they exist) fulfill $\rho_X^{(2)}(z) = \det(T)^{-2}\rho_{X_0}^{(2)}(T^{-1}z)$ for all $z \in \mathbb{R}^d$.

7. The pair correlation functions (if they exist) are related via $g_X(z) = g_{X_0}(T^{-1}z)$ for all $z \in \mathbb{R}^d$.

Given a parametric set $B = B(\psi)$ for the $\mathcal{K}$-measure, we then define the anisotropic pair correlation function as the Radon-Nikodym derivative

$$\mathcal{K}(B(\psi)) = \int_{B(\psi)} g(y)dy.$$

For a set B, where the integral has a simple geometrical decomposition, the corresponding derivatives have simple interpretations as anisotropic pcf's. For example, the conical set gives

$$\mathcal{K}(S(u, \epsilon, r)) = \int_0^r \int_{v \in S^{d-1} : cos^{-1}(v \cdot u) < \epsilon} g(v, t)dvdt$$

with $g(v, t) = g(vt)$ for unit vectors $v \in S^{d-1}$. The change to polar coordinates in 2D leads to the simple form

$$\mathcal{K}(S(u, \epsilon, r)) = \mathcal{K}(S(\alpha(u), \epsilon, r)) = 2 \int_0^r \int_{\alpha(u)-\epsilon}^{\alpha+\epsilon} g(a, t)t \, dadt,$$

where $g(a, t) = g(u, t)$ for direction u with $\alpha(u) = a$. The expression is multiplied by 2 as $S(u, \epsilon, r)$ is defined as a double cone. The cylindrical case is more complicated.

A simplification of the three parameter case is to assume that $g(v, t)$ is constant over the arc of the r-sector or the rotation cap of the r-cone. Then for a fixed ϵ

$$\mathcal{K}(S(u, \epsilon, r)) = \int_0^r g_u(t)dt$$

with

$$\frac{g_u(t)}{v(\epsilon, r)} = g(ut),$$

where the length or area $v(\epsilon, r)$ of the cap is $2r\epsilon$ when $d = 2$ and $2\pi r^2(1 - cos(\epsilon))$ when $d = 3$. We call the function $g_u(t)$ the *conical pcf* (Stoyan, 1991). Note that assuming that the pcf is isotropic is the same as setting $\epsilon = \pi$.

For the cylinder element $L(r, u, h_c)$ we can also simplify the corresponding pcf by assuming that for each r the g is approximately constant over the cylinder cross section, and then, by holding the direction u and the half-width of the cylinder h fixed, define the *cylindrical pcf* as a function of range for which

$$\mathcal{K}(L(r, u, h_c)) = \int_0^r g_u^c(t) dt$$

so that

$$\frac{g_u^c(t)}{b_{d-1} h_c^{d-1}} = g(ut) \quad \text{for } ut \in \{x : ||x - (x \cdot u)u|| < h_c\},$$

where b_d is the volume of d-dimensional unit ball.

The estimator for the anisotropic pcf in 2D as given by Stoyan (1991) and also replicated by Møller and Toftaker (2014) is

$$\hat{g}(a, r) = \frac{1}{4r\hat{\lambda}^2} \sum_{x,y \in \mathbf{x}}^{\neq} \frac{w_h(y - x, (r, a)) + w_h(y - x, (r, \check{a}))}{|W_x \cap W_y|}$$

with intensity estimator $\hat{\lambda}$, kernel function

$$w_h(y - x, (r, a)) = k_{h_r}(||y - x|| - r) k_{h_a}(\alpha(y - x) - a),$$

smoothing parameters $h = (h_r, h_a)$, and the antipodal direction $\check{a}$ of a. For 3D with spherical coordinates (r, ϕ, θ), the analogous estimator is

$$\hat{g}(\phi, \theta, r) = \frac{1}{4r^2 \sin(\theta)\hat{\lambda}^2} \sum_{x,y \in \mathbf{x}}^{\neq} \frac{w_h(y - x, (r, \phi, \theta)) + w_h(y - x, (r, \check{\phi}, \check{\theta}))}{|W_x \cap W_y|}$$

with an additional kernel element for the second angle. Note that the estimator behaves poorly for θ close to 0 or π, which corresponds to the unit vectors $(0,0,\pm1)$.

An alternative estimator for the anisotropic pcf is given by

$$\hat{g}(u, r) = \frac{1}{\hat{\lambda}^2} \sum_{x,y \in \mathbf{x}}^{\neq} \frac{k_{h_d}(x - y - ru)}{|W_x \cap W_y|}$$

with some $\mathbb{R}^d$ kernel k_{h_d} (Guan, 2007). This estimator does not suffer from the small denominator problem, but it introduces bias near the origin as the $\mathbb{R}^d$ smoothing kernels for different directions might overlap at small r.

The conical pcf with central half-angle ϵ and direction u can be estimated by

$$\hat{g}_u(r) = \frac{1}{v(\epsilon, r)\hat{\lambda}^2} \sum_{x,y \in \mathbf{x}}^{\neq} \frac{1(\alpha(y - x, u) < \epsilon) k_{h_r}(||y - x|| - r)}{|W_x \cap W_y|}$$

with $\alpha(v, u) = acos(\frac{v \cdot u}{||v|| ||u||})$. Note that the conical estimator is the same as the anisotropic estimator with a box kernel for the angles. Similarly, the cylindrical pcf with a cylinder half-height r, half-diameter h_c and direction u can be estimated by

$$\hat{g}_u^c(r) = \frac{1}{b_{d-1} h_c^{d-1} \hat{\lambda}^2} \sum_{x,y \in \mathbf{x}}^{\neq} \frac{1(||(y - x) - [(y - x) \cdot u]u|| < h_c) k_{h_r}(|(y - x) \cdot u| - r)}{|W_x \cap W_y|}.$$

These estimators have generally two tuning parameters. Several methods for choosing the range smoothing bandwidth have been proposed for isotropic pcfs, with Stoyan's rule of thumb $h_r \approx 0.15/\sqrt{\lambda}$ (Illian et al., 2008, Sec.4.3.3) still being the most used one. Guan et al. (2006) suggest a block-sampling based optimisation approach for the alternative anisotropic pcf estimator's only bandwidth. Further guidelines for anisotropic smoothing have not been presented.
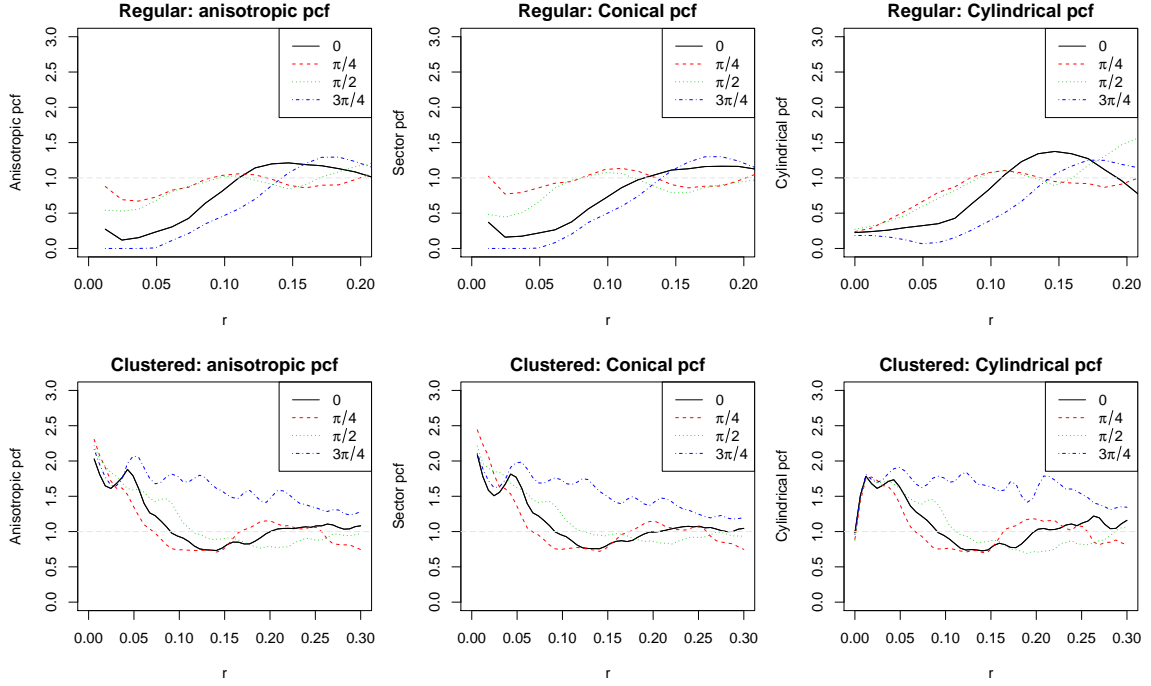
Figure 11: The anisotropic, conical and cylindrical pcf estimates for the example patterns, with $h_\alpha = \epsilon = \pi/8$, cylindrical width $h_c = 0.03$, range bandwidth $h_r = 0.3/\sqrt{\lambda}$. Top row: Regular pattern. Bottom row: Clustered pattern.

Figure 11 depicts the anisotropic, conical and cylindrical pcf estimates for the two example patterns. The sector half-angle for the conical pcf and the angular smoothing bandwidth for the anisotropic pcf were set to $\pi/8$. The cylinder half-width was set to $h = 0.03$, the same as for the cylindrical K-function in the previous section. For the range bandwidth we increased Stoyan's rule of thumb to $0.3/\sqrt{\lambda}$ to account for reduced sample sizes due to directional sub-sampling. The anisotropic and conical pcfs are very similar, depicting differences between directions for the regular pattern at short ranges and for the clustered pattern at longer ranges. The cylindrical sets overlap at short ranges so that the estimates in different directions are similar, but the longer range differences in the clustered pattern are captured well.

4.3.3 Estimation of preferred directions: Fitting ellipsoids to Fry points

To detect geometric anisotropy, Rajala et al. (2016) suggested to fit ellipsoids (ellipses in 2D) to the contours of directed cumulants of the pairwise difference vectors, i.e. unscaled sector-K functions. The fitted ellipsoid not being a ball is an indication of anisotropy and the favored direction can be estimated based on the orientation of the ellipsoid. They define a set of so-called pseudo Fry points $G_l := \{g_u = r_l(u)u : u \in U\}$, where $r_l(u)$ is the distance to the lth nearest Fry point in the sector $S(u, \epsilon)$ and U is the set of directions of interest. For a fixed set $L \subset \mathbb{N}$ of contour levels l, one has then a collection of pseudo-Fry point sets $\{G_l : l \in L\}$, from which the contour ellipsoids $\{E_l : l \in L\}$ can be estimated. More specifically, for each $l \in L$, the observed points $g_u \in G_l$ are assumed to follow a measurement error model

$$g_u = e_u + \varepsilon_u, \qquad u \in U,$$

where $e_u \in E_l$ are the true contour points, fulfilling the origin centered quadratic equation $e_u^T A_l e_u = m_l$ with a semi-definite matrix A_l and scale parameter m_l, and ε_u's are independent measurement errors with Gaussian distribution $\varepsilon_u \sim N(0, \sigma^2 I)$ with I being the identity matrix. It is enough to consider origin centered ellipsoids since the underlying process is stationary. The scale parameters can be fixed to $m_l = 1$. The model can be fitted using penalized and adjusted ordinary least squares to obtain $\hat{A}_l$

(Kukush et al., 2004). A rotation matrix estimate can then be derived using an eigenvalue decomposition of $\hat{A}_l$. To produce an estimate of the average rotation, Rajala et al. (2016) sample a set of contour ellipsoids with noise and fit another ellipsoid to the superimposed samples.

The method is sensitive to the choice of the contour levels, the number of directions and the sector angles. The set of useful contours depends on whether the process is regular or clustered, and on the intensity. In addition, if too many narrow sectors are used, they will not capture enough Fry points and the ellipses represent poorly the contours. Too few, wide sectors, on the other hand, bias towards circular contours. Further technical issues involving border correction and the double-effect of each Fry point pair are omitted here.

Figure 12 illustrates the rotation estimation for our two example patterns. For the regular pattern the low-level contours, $l < 10$, are most informative as the relevant shape is close to the origin. The first few contours have high variance due to the "noise" pairs at short distances, typical to a Strauss process. For the clustered process the low-level contours $l < 10$ are not informative due to the local independence in the process, but at larger contours, $l > 100$, the deformation of the cluster shape is captured. Notably, the clustering induces ellipses that have their main-axis perpendicular to the direction of interest.
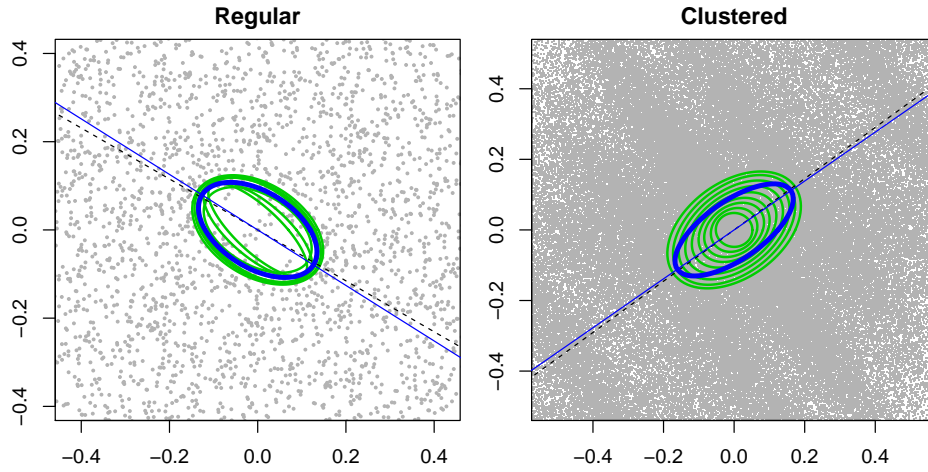

Figure 12: Rotation estimation by Fry point ellipsoid fitting for the regular (left) and clustered (right) examples. Figures superimpose over the Fry points the estimated contour ellipses (green), an average ellipse (blue) and its main-axis direction (blue line), and the target direction (dashed line). The target direction for the clustered process is rotated 90 degrees clockwise.

4.3.4 Inhomogeneous patterns

If we cannot assume that the intensity function $\lambda(x)$ is constant, the pair correlation needs to be defined as

$$g(x, y) = \frac{\rho^{(2)}(x, y)}{\lambda(x)\lambda(y)}.$$

A relaxation of the notion of stationarity is obtained by only assuming that the second-order properties of the point process are translation invatiant, i.e. $g(x, y) = g(x - y)$ does not depend on exactly where x and y are. Then, the summaries can be estimated using the second-order intensity reweighted stationary (SOIRS) process methodology as described by Baddeley et al. (2000) and van Lieshout (2011). Given X is a SOIRS process, the $\mathcal{K}$-measure is defined as

$$\lambda^2 \mathcal{K}_{soirs}(B) := \frac{1}{|W|} \mathbb{E} \left(\sum_{x \in \mathbf{x}, y \in X}^{\neq} \frac{1_B(y - x)}{\lambda(x)\lambda(y)} \right).$$

16

The corresponding pair correlation function is again defined as the Radon-Nikodym derivative

$$\mathcal{K}_{soirs}(B) = \int_B g_{soirs}(u)du,$$

and the SOIRS extension of the classical K-function would be obtained by setting $B = b(o,r)$. Using e.g. a cone or a cylinder, we arrive at the corresponding anisotropic versions of the second order summaries. The estimators are modified by including the varying intensities in the double sums, for example the SOIRS conical pcf estimator becomes

$$\hat{g}_{u,soirs}(r) = \frac{1}{v(\epsilon,r)} \sum_{x,y \in \mathbf{x}}^{\neq} \frac{1(\alpha(y-x,u) < \epsilon)k_h(||y-x||-r)}{|W_x \cap W_y|\hat{\lambda}(x)\hat{\lambda}(y)},$$

see also Häbel et al. (2017).

5 Spectral analysis

Fourier transformations and spectral analysis techniques can be used to determine the presence of periodic structures in a spatial point pattern. This yields information on both characteristic scales and directions. The theory for spectral analysis of point patterns was first introduced by Bartlett (1964). However, at that time the applicability of the method suffered from a lack of efficient tools for computation and visualisation of sample spectra, e.g., Bartlett simply tabulates estimated periodogram values. Consequently, 'virtually no development of spectral methods' was achieved until Mugglestone and Renshaw (1996a) published their 'Practical guide to the spectral analysis of spatial point processes'. They argue that the advent of powerful computers and graphics packages has overcome computational limitations such that the method deserves being taken into consideration again. It should be noted that notation and normalization of the Fourier transform in Bartlett (1964) and Mugglestone and Renshaw (1996a) differ. The following presentation is closer to Mugglestone and Renshaw (1996a) than to Bartlett (1964).

We consider a point process X with intensity function λ and second-order product density $\rho^{(2)}$. The *covariance density function* is defined as

$$\gamma(x,y) = \rho^{(2)}(x,y) - \lambda(x)\lambda(y), \quad x,y \in \mathbb{R}^d, x \neq y.$$

Because X is assumed to be simple with no duplicate points the function above is not well defined at $x = y$, and is set to 0. The *complete covariance density function* is defined as

$$\kappa(x,y) = \lambda(x)\delta(y-x) + \gamma(x,y),$$

where $\delta(\cdot)$ denotes the d-dimensional Dirac delta function. If the process is stationary, the complete covariance density function reduces to

$$\kappa(z) = \lambda\delta(z) + \gamma(z),$$

where $z = y - x$.

The (Bartlett) *spectral density function* is defined as the Fourier transform of the complete covariance density function, namely

$$\mathcal{F}(\eta,\omega) = \mathcal{F}[\kappa](\eta,\omega) = \int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \kappa(x,y)e^{-i(\eta^T x + \omega^T y)}dxdy, \quad \eta,\omega \in \mathbb{R}^d,$$

where T denotes the transpose and $i = \sqrt{-1}$. The values of $\mathcal{F}$ are called the *spectrum* of X, and the arguments η,ω are called *frequencies*. In the stationary case, $\mathcal{F}(\eta,\omega)$ should depend on x and y only via $z = y - x$. Hence, the equation can be simplified to

$$\mathcal{F}(\omega) := \mathcal{F}(-\omega,\omega) = \int_{\mathbb{R}^d} \kappa(z)e^{-i\omega^T z}dz = \lambda + \int_{\mathbb{R}^d} \gamma(z)e^{-i\omega^T z}dz. \tag{6}$$

Estimation of $\mathcal{F}$ has only been discussed in 2D and when the window W is rectangular with side lengths l_1 and l_2. In general dimension d, let W be rectangular with side lengths $l_j > 0, j = 1, ..., d$. The spectrum of the pattern can be estimated by a periodogram (Bartlett, 1964) which is based on the discrete Fourier transform of the pattern $\mathbf{x}$

$$DFT[\mathbf{x}](\omega) = |W|^{-\frac{1}{2}} \sum_{x \in \mathbf{x}} e^{-i\omega^T x} = A(\omega) + iB(\omega).$$

The *periodogram estimator* of the spectrum, and hence the spectral density function, is given by

$$\widehat{\mathcal{F}}(\omega) = DFT[\mathbf{x}](\omega)\overline{DFT[\mathbf{x}]}(\omega) = A(\omega)^2 + B(\omega)^2.$$

This estimator has a bias term which for a rectangular window can be written as

$$\lambda^2 |W|^{-1} \prod_{j=1}^{d} \frac{\sin^2(\frac{1}{2}l_j\omega_j)}{(\frac{1}{2}l_j\omega_j)^2}, \tag{7}$$

where $\omega = (\omega_1, ..., \omega_d)$. Estimating the spectrum is therefore recommended on frequencies that are integer multiples of $2\pi/l_j$ so that the sine term and hence the bias term become 0. Additional bias comes from taking only a finite integral in (6) (see Section 4 in Bartlett (1964)). The periodogram is asymptotically unbiased for $\omega \neq 0$ when $|W| \to \infty$ (see also Diggle et al., 1987). To avoid the δ-function spike at $\omega = 0$, Diggle et al. (1987) suggest truncation for small ω. They also report that more sophisticated methods for extrapolating the periodogram to $\omega = 0$ were tried without success.

Bartlett (1964) as well as Mugglestone and Renshaw (1996a) suggest standardizing the coordinates by replacing each point $x = (x_1, ..., x_d)$ by $x' = (nx_1/l_1, ..., nx_d/l_d)$ which is claimed to reduce bias for $\omega \approx 0$. However, the benefits of the division by l_j are unclear: for non-cubical windows, it produces an artificial geometric anisotropy which will change the spectra as both the intensity and the product density are affected (see the discussion above as well as Mugglestone (1990)). Hence, this step should be used cautiously.

The estimates $\widehat{\mathcal{F}}(\omega)$ are often displayed as a pixel image using the grid of integer factors of the bias-canceling frequencies (Figure 13, left column). The range of suitable integers depends on the number of points n. As n increases, we expect to obtain more reliable information on the interactions over smaller distances and therefore, higher frequencies. According to Renshaw and Ford (1984) and Mugglestone (1990), in 2D a reasonable range of frequencies is covered by integer factors $\{0, 1, ..., 16\} \times \{-16, -15, ..., 15\}$, with only the positive values needed for the first (or second) argument as the periodogram estimator is symmetric in the origin. The sizes of the point patterns studied in these papers were roughly between 50 and 300 points.

Furthermore, to obtain a consistent estimator of the spectral density the periodogram needs to be smoothed (Bartlett, 1964). Concerning time series data, Chatfield (1989) discusses the advantages and disadvantages of two types of smoothing techniques, namely smoothing the periodogram using the fast Fourier transform or smoothing the covariance function using lag windows and calculating the periodogram based on the smoothed covariance function. He concludes that the smoothed periodogram has better theoretical properties than the one based on the smoothed covariance function. Kanaan (2000) gives a detailed description of an approach where the periodogram is smoothed by a weighted moving average technique. The same technique is used by Mugglestone and Renshaw (1996b) with the modification that the moving average procedure is repeated several times resulting in a smoother surface, similar to what would be obtained by using a Gaussian kernel with bandwidth equal to the number of repetitions of the moving average.

Assuming the biases are not detrimental, the 2D polar coordinate representation of $\widehat{\mathcal{F}}(\omega)$ can be useful in anisotropy analysis. It gives two one-dimensional summaries of the periodogram, called the *R spectrum*, $\widehat{\mathcal{F}}_R(r)$, and the Θ *spectrum*, $\widehat{\mathcal{F}}_\Theta(\theta)$. The R spectrum summarises average periodogram values for ordinates with similar values of r and is used to investigate scales of the pattern under the assumption of isotropy. The Θ spectrum summarises average periodogram values for ordinates with similar values of θ and is used to investigate directional features. When evaluating the periodogram on a grid $\{2\pi(p_1/l_1, p_2/l_2)\}$ where p_1 and p_2 are integers, we consider $r = \sqrt{p_1^2 + p_2^2}$ and $\theta = \arctan(p_2/p_1)$ and define the R spectrum as

$$\widehat{\mathcal{F}}_R(r) = \frac{1}{n_r} \sum_{r-1 < r' \leq r} \sum_{\theta} \hat{s}(\omega_{r'}, \omega_\theta), \quad r = 1, 2, \ldots$$

and the Θ spectrum as

$$\widehat{\mathcal{F}}_\Theta(\theta) = \frac{1}{n_\theta} \sum_r \sum_{\theta-5°<\theta'\leq\theta+5°} \hat{s}(\omega_r, \omega_{\theta'}), \quad \theta = 0°, 10°, ..., 170°.$$

Here, $\hat{s}(\omega_r, \omega_\theta)$ is $\widehat{\mathcal{F}}(\omega)$ in polar coordinates, and n_r and n_θ are the (chosen) numbers of the periodogram ordinates. Note that $\widehat{\mathcal{F}}(0)$ is not included in the averaging since its asymptotic distribution differs from that of the rest of the periodogram ordinates (Mugglestone and Renshaw, 1996a).

With the number of points going to infinity, estimates of the spectrum as well as the R and Θ spectrum are asymptotically distributed as χ^2 random variables. For $\omega \neq 0$, we have:

$$\frac{2\widehat{\mathcal{F}}(\omega)}{\mathcal{F}(\omega)} \sim \chi_2^2,$$

and periodogram ordinates for different ω are asymptotically independent. Additivity of independent χ^2 random variables implies

$$\frac{1}{n_r} \sum_{r'} \sum_\theta \frac{\hat{s}(\omega_{r'}, \omega_\theta)}{s(\omega_{r'}, \omega_\theta)} \sim \frac{1}{2n_r} \chi_{2n_r}^2$$

and

$$\frac{1}{n_\theta} \sum_r \sum_{\theta'} \frac{\hat{s}(\omega_r, \omega_{\theta'})}{s(\omega_r, \omega_{\theta'})} \sim \frac{1}{2n_\theta} \chi_{2n_\theta}^2.$$

Under complete spatial randomness (CSR) the periodogram is constant, $\mathcal{F}(\omega) = \lambda$ for all ω, which implies

$$\widehat{\mathcal{F}}_R(r)/\lambda \sim \frac{1}{2n_r} \chi_{2n_r}^2$$

and

$$\widehat{\mathcal{F}}_\Theta(\theta)/\lambda \sim \frac{1}{2n_\theta} \chi_{2n_\theta}^2.$$

In practice, λ is replaced by $\hat{\lambda} = n/|W|$ which does not change the asymptotic distribution due to consistency of $\hat{\lambda}$ and Slutsky's Lemma. Sometimes it is suggested to generally standardize the periodogram ordinates by dividing by the intensity such that the theoretical value under CSR becomes 1. Based on these asymptotics, preferred directions (compared to CSR) in the spectrum can be detected by comparing the scaled estimated Θ spectrum with appropriate quantiles of the $\chi_{2n_\theta}^2$ distribution.

Figure 13 shows the results of a spectral analysis of our two sample patterns. The raw periodograms and their values after smoothing by an isotropic Gaussian filter are shown together with the resulting R and Θ spectra. For the linear pattern the standard deviation of the kernel was chosen as $\sigma = 1$, so smoothing was done more cautiously than for the regular pattern where we used $\sigma = 2$. For the compressed regular pattern, the void ellipse is clearly visible in the periodogram. As compression of a point pattern decreases the period in this direction, the orientation of the ellipse is rotated by 90° compared to the Fry plot. The compression direction at $\pi/2 - \pi/6$ is found as the minimum of the Θ spectrum while the stretch direction is indicated by the maximum at about $\pi - \pi/6$. For the clustered pattern, the periodogram shows a bright line perpendicular to the main direction of the clusters. The Θ spectrum contains a clear peak at the corresponding angle $\pi/5$. The direction of the lines is obtained by adding $\pi/2$ to this angle.

6   Wavelet analysis

An approach closely related to spectral analysis is obtained by using a wavelet transform instead of a Fourier transform. Wavelet analysis has been used to detect directional properties in 2D spatial point processes first in Rosenberg (2004) and later by D´Ercole, Mateu, and Nicolis in a series of papers (Mateu et al., 2010; Mateu and Nicolis, 2015; D'Ercole and Mateu, 2013a,b, 2014). All these papers consider anisotropy caused by an increased intensity along directed lines in square windows, the type of anisotropy observed in the Ambrosia dumosa data set (Figure 4).
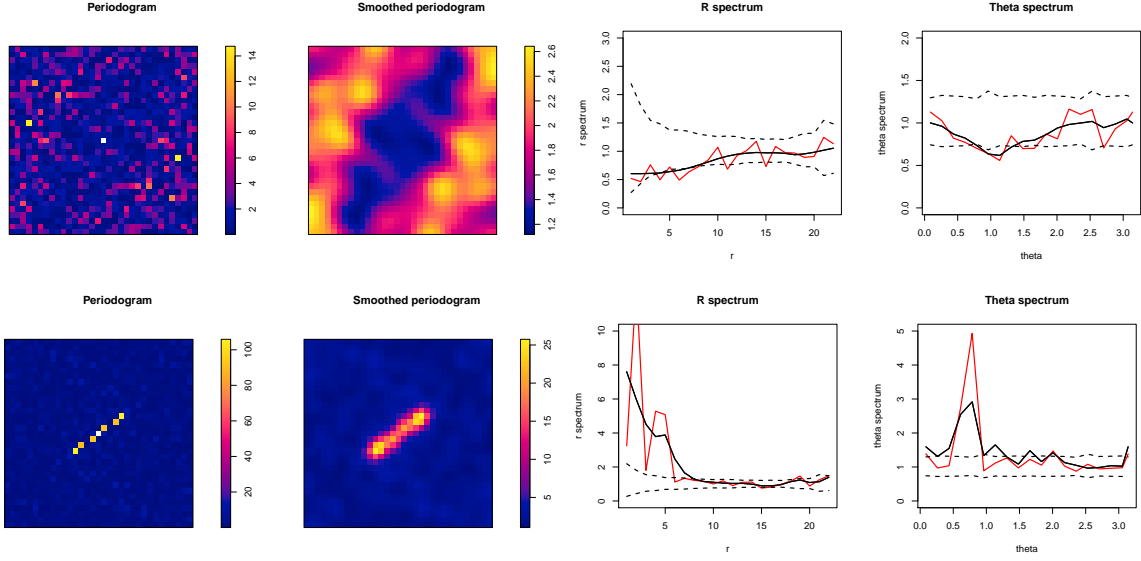
Figure 13: Spectral summaries for the regular (top row) and clustered (bottom row) example patterns. From left to right: Bartlett's periodogram estimator for the spectrum; Smoothed periodogram; The R-spectrum with 95% confidence intervals (dashed) for the raw (red) and the smoothed (black) periodogram; The θ-spectrum with 95% confidence intervals (dashed) for the raw (red) and the smoothed (black) periodogram.

We first describe the approach of Rosenberg (2004). Given a realization $\mathbf{x} = \{x_1, ..., x_n\}$ of a spatial point process X, Rosenberg (2004) starts by selecting one of the points of $\mathbf{x}$ as "specific focal point". The space around the focal point is divided into 360 angular sectors of width $1°$ and directions $\theta_i = i°$. Then, the number of further points of $\mathbf{x}$ in each sector is counted and counts from opposite sectors are combined.

Finally, the count is divided by the area of the intersection of the sector and the observation window yielding the point intensities $\eta(x, \theta_i)$, $\theta_i = 0° \ldots 179°$. The 1D discrete wavelet transform of $\{\eta(x, \theta_i)\}$ in direction θ, with some scale parameters $b_k, k = 1 \ldots m$, and a 1D wavelet function ψ is defined as

$$W(x, \theta, b_k) = \frac{1}{b_k} \sum_{i=1}^{180} \eta(x, \theta_i) \psi \left(\frac{\theta_i - \theta}{b_k} \right). \tag{8}$$

In Rosenberg (2004), ψ is chosen as the French Top Hat wavelet. According to the paper, the choice of ψ should only weakly affect the results of the analysis. When applying Equation (8) one should treat η as a periodic function in θ to avoid edge effects in the angle domain.

The overall variance of the wavelet transform for a given focal point x and direction θ is defined as

$$P(x, \theta) = \frac{1}{m} \sum_{k=1}^{m} W^2(x, \theta, b_k), \tag{9}$$

which are then averaged over the data points to a directional summary $\bar{P}(\theta)$. To avoid edge effects, points near the border are excluded from the average.

The plot of $\bar{P}(\theta)$ is useful when detecting preferred directions. For example, in the case of directed lines, the peaks of $\bar{P}(\theta)$ correspond to the directions of the lines. The wavelet analysis as described above is implemented in the software package PASSaGE (Rosenberg and Anderson, 2011). Figure 14 shows the average overall variance of Equation (9) for our two sample patterns with $b_k = 1° \ldots 45°$ as used in Rosenberg (2004). In the clustered case, the plot on the right shows a clear peak in the direction of the clusters around $\frac{\pi}{2} + \frac{\pi}{5}$, while in the regular case (plot on the left), no clear direction is detectable.

In generalization of Equation (8), the 2D directional continuous wavelet transform (CWT) of a square
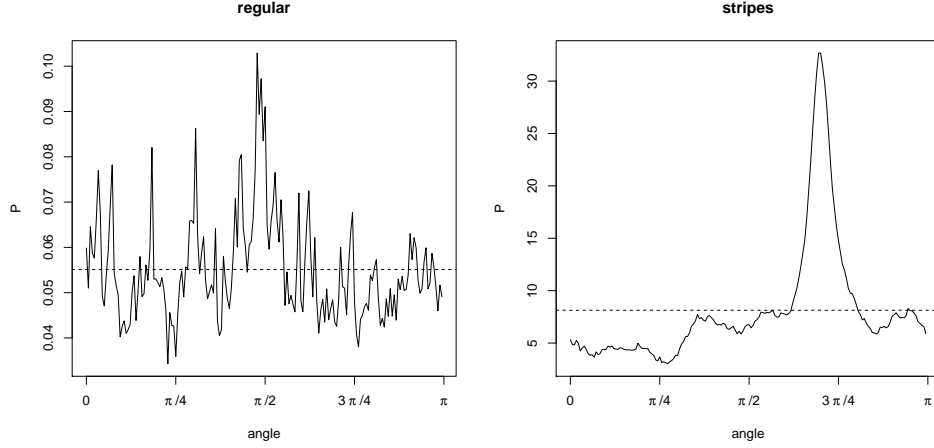
Figure 14: The overall variance of Equation (9) for the regular (left) and the clustered (right) example patterns.

integrable function $f \in L^2(\mathbb{R}^2)$ (with respect to the Lebesgue measure) is defined as

$$F(a, b, \theta) = a^{-1} \int_{\mathbb{R}^2} \overline{\psi(a^{-1}R_{-\theta}(x-b))} f(x) dx =: \int_{\mathbb{R}^2} \overline{\psi_{a,b,\theta}(x)} f(x) dx, \quad x \in \mathbb{R}^2, \tag{10}$$

where $\psi \in L^2(\mathbb{R}^2)$ is the mother wavelet of the transform, $a > 0$ the scaling parameter, $b \in \mathbb{R}^2$ the translation parameter, and $R_{-\theta}$ represents a clockwise rotation by an angle $-\theta$. The wavelet transform is called directional since the mother wavelet ψ is not only translated and rescaled, but also rotated.

Let now X be a (possibly non-stationary) point process with intensity function λ observed in a finite window W. D´Ercole, Mateu, and Nicolis suggest to apply the directional CWT of equation (10) to the intensity function restricted to W obtaining the coefficients

$$S(a, b, \theta) = a^{-1} \int_W \overline{\psi_{a,b,\theta}(x)} \lambda(x) dx, \quad x \in \mathbb{R}^2, \tag{11}$$

which can be estimated by estimating λ and by discretizing the integral (Mateu et al., 2010; Mateu and Nicolis, 2015). An alternative estimator used in D'Ercole and Mateu (2013a,b, 2014) is given by

$$\hat{S}(a, b, \theta) = a^{-1} \sum_{x \in X \cap W} \overline{\psi_{a,b,\theta}(x)}, \tag{12}$$

which can be shown to be unbiased by using the Campbell theorem.

If X is a stationary point process (as considered throughout this paper), the intensity function λ is constant. Hence, the transform (10) will not contain information on possible anisotropies. However, one can consider the squared modulus of $\hat{S}(a, b, \theta)$ given in Equation (12) (see D'Ercole and Mateu (2013a,b, 2014)). In fact, by applying the Campbell formula and its generalisation given in Equation (5), it is possible to prove that the expectation of $|\hat{S}(a, b, \theta)|^2$ depends both on the first and second order properties of X. As discussed in Section 4.3, the latter contain directional information also in the stationary case.

Using a similar idea as in Equation (9), the square modulus of $\hat{S}(a, b, \theta)$ can be integrated over b, obtaining the *scale-angle energy density* (D'Ercole and Mateu, 2013b; Mateu and Nicolis, 2015)

$$\nu(a, \theta) = \int |\hat{S}(a, b, \theta)|^2 db. \tag{13}$$

It can be useful in detecting anisotropies since, under isotropy, the energy should be equally distributed in different directions. In D'Ercole and Mateu (2013a,b) other types of energy densities are described.

However, in the stationary case, the scale-angle energy density is the most suitable choice since it exploits the translational invariance of the distribution of X.

In Equation (11) the mother wavelet ψ has to be chosen. Clearly, to be useful in directional analysis, ψ has to be anisotropic. D´Ercole, Nicolis, and Mateu suggest using the Morlet (or Gabor) mother wavelet, for which the normalized version is given by

$$\psi(x) = \sqrt{D}(\sqrt{\pi})^{-1} \exp\left(ik_0^T x\right) \exp\left(-\frac{1}{2}x^T A^T A x\right),$$

where k_0 with $||k_0|| \geq 5.5$ is the wave vector, $A = \text{diag}(D, 1)$ denotes a diagonal matrix, and D is the anisotropy ratio. Note that D'Ercole and Mateu (2013a,b, 2014) use a slightly different version of the wavelet which is adjusted to integrate to zero.

Figure 15 is based on the approach presented by D'Ercole, Mateu and Nicolis and shows the estimated scale-angle energy density. For the analysis, we used the Morlet mother wavelet with $D = 0.1$ and $k_0 = (0, 5.5)$, as suggested in Mateu and Nicolis (2015). The scale parameter a was chosen between 0 and the dimension of the window edge length. The energy map for the clustered pattern (right) clearly shows the direction of the clusters around $\pi/2 + \pi/5$ as a bright spot at small scale. However, the energy map for the regular pattern (left) does not give a clear indication of the anisotropy. A different choice of the mother wavelet or adjusting the parameters D and k_0 might improve the results.
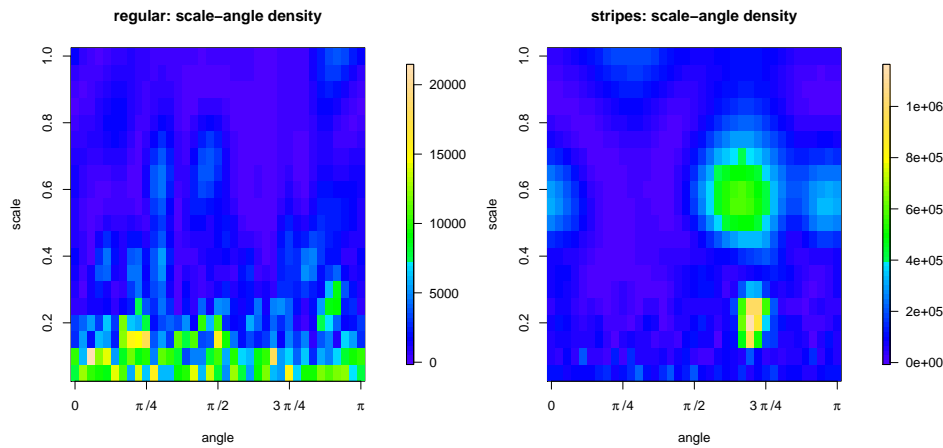

Figure 15: Heat-map of the angle-scale energy density map of Equation (13) for the regular (left) and clustered (right) example patterns.

7 Testing isotropy

In many practical situations, testing the null hypothesis of isotropy versus anisotropy is of interest. However, the variety of anisotropy mechanisms listed in Section 3 shows that there might be no generally best test for this purpose. Most tests for point processes found in the literature consider testing the null hypothesis of complete spatial randomness. Typically, only this case allows for a derivation of theoretical values or distributions of test statistics. As a stationary Poisson process is automatically isotropic, anisotropy in the pattern may result in a rejection of the CSR hypothesis. However, also other deviations from CSR such as regularity or clustering will result in a rejection. Hence, isotropy tests which do not require CSR are of general interest. For many of the statistics considered above it will be impossible to determine the asymptotic distribution of the statistic under a general isotropy hypothesis. In these cases, bootstrapping or the use of replicated data are suitable alternatives to determine critical values of the test statistics.

7.1 Nearest neighbour angle uniformity

A very simple anisotropy test can be based on the observation that under isotropy directions to nearest neighbours should be uniformly distributed on the unit sphere. Hence, a test for this hypothesis could be applied to the nearest neighbour angles. However, as mentioned earlier, the nearest neighbour angles in an observed pattern are not independent. This has to be taken into account when choosing the test. König and Schmidt (1992) state that the directional distribution (3) (or its estimator (4)) can be compared to the uniform distribution on the sphere, large deviations indicating anisotropy of the underlying point pattern, but they do not provide any formal test.

As indicated in Section 4.2, nearest neighbour statistics tend to be short sighted. Here, even the distance information is discarded. Hence, tests based on nearest neighbour directions can be expected to have a very small power in general (see also Redenbach et al., 2009).

7.2 Tests based on summary statistics

A more promising approach for testing isotropy is to compare directional summary statistics for a set of directions. Such tests are introduced in the papers Guan et al. (2006), Wong and Chiu (2016), Redenbach et al. (2009), and Rajala et al. (2016). All approaches are based on similar ideas, but a lot of details are solved in different ways. One of the main questions is how to get hold of the distribution of the test statistics under the null hypothesis. The paper Guan et al. (2006) presents a nonparametric approach based on the asymptotic joint normality of the sample second-order intensity function (i.e. unnormalised pcf). From that, a test statistic with asymptotic χ^2 distribution is constructed.

Let $\hat{G} = (\hat{\rho}^{(2)}(z_1), \ldots, \hat{\rho}^{(2)}(z_k))^T$ denote a vector of estimates of the second order intensity function $\rho^{(2)}$ at lags $z_1, \ldots, z_k \in \mathbb{R}^d$ with $||z_i|| = ||z_j||$ for all $1 \leq i, j \leq k$ but representing different directions. As estimator for $\rho^{(2)}$, they use $\hat{\rho}^{(2)}(z_i) = \hat{\lambda}^2 \hat{g}(z_i)$, where $\hat{g}$ is a kernel estimator of the pcf as described in Section 4.3.2.

Under the null hypothesis of isotropy, $E[\hat{\rho}^{(2)}(z_i)] = E[\hat{\rho}^{(2)}(z_j)]$ for any pair z_i, z_j of lags. The estimated second-order intensities for the different lags are compared via a set of contrasts formed by a full row rank matrix A. Then, $AE[\hat{G}] = 0$ under H_0. Guan et al. (2006) consider the test statistic

$$TS = |W|h^2(A\hat{G})^T(A\hat{\Sigma}A^T)^{-1}(A\hat{G}),$$

where $\hat{\Sigma}$ is an estimate of the asymptotic variance Σ of

$$|W|^{1/2}h\left(\hat{\rho}^{(2)}(z_1) - E[\hat{\rho}^{(2)}(z_1)], \ldots, \hat{\rho}^{(2)}(z_k) - E[\hat{\rho}^{(2)}(z_k)]\right)^T,$$

which can be obtained using a subsampling approach. Under H_0 and additional suitable assumptions, the statistic TS is asymptotically χ_r^2 distributed for increasing windows, where r is the rank of A.

The method requires user input on the lags to compare, the bandwidth for $\hat{g}$, and the subblock size for estimating Σ. For practical applications, Guan et al. (2006) give the following recommendations: The subblock size is chosen such that each block contains approximately $cn^{1/4}$ points where n denotes the total sample size. They found $c \approx 0.8$ to be a suitable choice. The simulation study presented in Guan et al. (2006) indicates that the choice of the lags and the bandwidth affects the test size only slightly but has a greater influence on the power. A recommendation by Guan et al. (2006) in the clustered case is to use $||z||$ which is between $1/3$ and $1/2$ of the dependence range (4 times the cluster standard deviation) of X, and a recommendation by Wong and Chiu (2016) is to use $||z||$ which is between $1/2$ and $2/3$ of the dependence range (2 or 3 times the hard-core distance in the regular case) while the directions of the lags should be chosen evenly spaced on $[0, \pi]$. The bandwidth h should be chosen such that at least 200 pairs of points are considered in the estimation and disjoint sets of pairs are used to estimate $\rho^{(2)}$ for different directions. They also introduce a data driven method for selecting the bandwidth h. However, they admit that this method is computationally intensive and requires a large sample size.

The test presented in Wong and Chiu (2016) is based on the ratio $F_{r,\psi}(\theta) = K_r(\theta, \psi)/K_r(\pi, 0)$, where $K_r(\theta, \psi)$ is the reduced second moment measure of a sector of radius r centred at the origin enclosed between the lines making angles ψ and $\psi + \theta$ w.r.t. the x-axis. For $\psi = 0$, $\theta = \alpha$, $r_1 = 0$, and $r_2 = r$ this corresponds to the second order orientation distribution $F_K(\alpha)$ discussed in Section 4.2. Under the null

hypothesis of isotropy, $F_{r,\psi}(\theta)$ is the distribution function of the uniform distribution on $[0, \pi]$. The test is based on the Kolmogorow-Smirnow statistic

$$d_{r,\psi} = \sup_{\theta \in [0,\pi)} \left| \frac{\theta}{\pi} - \hat{F}_{r,\psi}(\theta) \right|,$$

whose maximum w.r.t. ψ is considered, i.e.

$$T_r = \sup_{\psi \in [0,\pi)} d_{r,\psi}.$$

As a suitable choice of r, Wong and Chiu (2016) recommend a half to two thirds of the dependence range of the point process, see the comment above. In contrast to the approach by Guan et al. (2006), the asymptotic distribution for the test statistic T_r under the null hypothesis of isotropy is not known. Hence, a critical value of the test statistic is determined by using the reconstruction algorithm of Tscheschel and Stoyan (2006) to simulate isotropic versions of the data such that some predefined isotropic summary statistics are close to the ones of the real data. In a simulation study, the approach is applied to simulated realisations of regular and clustered anisotropic point patterns and compared to the asymptotic approach by Guan et al. (2006). An advantage of the bootstrap test is that there is no need to choose a bandwidth. Additionally, according to the authors, the power of the bootstrap test is more robust to the choice of the user-specified parameters than the power of the test by Guan et al. (2006).

While the two tests discussed so far have been introduced for the 2D case, testing anisotropy in 3D point patterns has been considered in Redenbach et al. (2009). The test statistics used there are based on the conical K-function and the directional nearest neighbour distance distribution (see Section 4). Again, the test is nonparametric in the sense that no parametric model is assumed for the data. To derive critical values for the test, replicated data are used. The particular type of anisotropy investigated in the paper is geometric anisotropy generated by a transformation of an isotropic pattern with a matrix $T = \mathrm{diag}(1/\sqrt{c}, 1/\sqrt{c}, c)$. Hence, the difference of the summary statistics for the x- and y-direction can be used to describe the behaviour under the null hypothesis. Test statistics used are

$$T_{xy,i} = \int_{r_1}^{r_2} |\hat{S}_{x,i}(r) - \hat{S}_{y,i}(r)| dr, i = 1, \ldots, n$$

and

$$T_{z,i} = \min \left(\int_{r_1}^{r_2} |\hat{S}_{x,i}(r) - \hat{S}_{z,i}(r)| dr, \int_{r_1}^{r_2} |\hat{S}_{y,i}(r) - \hat{S}_{z,i}(r)| dr \right), i = 1, \ldots, n.$$

Here, for instance, $\hat{S}_x$ is an estimate of a directional summary statistic pointing to the x-axis. When using these test statistics there is no need for choosing a bandwidth. Instead, the opening angle of the cone or the cross-section half-length of the cylinder needs to be chosen. Furthermore, the method is sensitive to the choice of the integration limits which is equivalent to the choice of the lags in the two previous methods. For hardcore processes, Redenbach et al. (2009) recommend choosing the right limit r_2 slightly larger than the hardcore distance. They perform a simulation study comparing the powers of tests based on the conical K-function as well as the local and global directional nearest neighbour distance distributions. When using the optimal integration range, the tests based on the conical K-function had the highest powers. In contrast, the test based on the local nearest neighbour distance distributions was more robust to the choice of the integration range than the test based on the K-function. The test based on the global nearest neighbour distance distributions behaved poorly.

Ellipsoids fitted to the Fry plot can also be used to construct tests against isotropy as discussed in Rajala et al. (2016). The least squares approach provides asymptotically normal estimators of the coefficients in the quadratic model space, by which one can simulate and compute Monte Carlo confidence intervals for chosen functions of the parameters. Especially, the equality of the semi-axes of the ellipsoid can be assessed. The contrast of interest in 2D is "$a_1 - a_2 = 0$", where a_i is the ith semi-axis length. In 3D, one can check either the contrasts "$a_1 - \frac{1}{2}(a_2 + a_3) = 0$", "$a_2 - \frac{1}{2}(a_1 + a_3) = 0$" or "$a_3 - \frac{1}{2}(a_1 + a_2) = 0$".

As reported in Rajala et al. (2016), in the case of a regular process the contour ellipses E_l are converging towards a circle/sphere with increasing l and therefore, E_l overestimates the "roundness" of T as l increases. The test is expected to be conservative, i.e. the test based on E_l's with large l is not expected to be very powerful. In the clustered case, the ellipsoids tend to be circular for small l.

Therefore, the test should be based on a reasonably small l in the regular case and reasonably large l in the clustered case. How large l should be depends on the process and its intensity. The recommendation in Rajala et al. (2016) is to choose three or four different values for l and to check the confidence intervals for these values. Note that the results need to be adjusted due to the multiple testing.

Performance of the isotropy test was investigated in a simulation study both in 2D and in 3D. First, realizations of a stationary and isotropic Strauss process were simulated and then, the realizations were compressed to obtain anisotropic patterns. Several regularity parameters and compression strengths were used. As expected, the power of the test increased with regularity and compression both in 2D and in 3D. The test has a very small power when the compression factor is larger than 0.9, i.e. the pattern is very close to being isotropic.

7.3 Spectral tests

Mugglestone and Renshaw (2001) propose five tests against CSR based on the periodogram. None of them is, however, based on the Θ spectrum that could show deviations from isotropy. According to Mugglestone and Renshaw (1996a), for general testing of isotropy without assuming CSR, one could scale the Θ spectrum by the average periodogram value rather than λ to investigate whether spectral power is distributed evenly across the frequency angles. Details are not given.

7.4 Tests based on wavelet analysis

An anisotropy test based on wavelet transforms is proposed in Mateu and Nicolis (2015). They estimate the wavelet coefficients $S(a, b, \theta)$ of Equation (10) (Section 6) for a given selection of directions $\theta_i \in (0, \pi]$, $i = 1, \ldots, m$, scales a_j, $j = 1, \ldots, L$, and locations b_k, $k = 1, \ldots, N$, in the domain of λ. They then compute a discretized version of the scale-angle energy density $\nu(a, \theta)$ in Equation (13), summing up the modules of the wavelet coefficients over the positions b_k. Finally, they consider the following sums:

$$T(\theta_i) = \frac{1}{L} \sum_{j=1}^{L} \log(\nu(\theta_i, a_j)), \quad i = 1, \ldots, m.$$

Under isotropy, this test statistic should not depend on the direction. However, the exact distribution of $T(\theta_i)$ is unknown. In practice, Mateu and Nicolis (2015) approximate the distribution by the empirical distribution of the $T(\theta_i)$ for an isotropic parametric model fitted to the data. For the directions, they choose $\theta_i = i°$, $i = 1, \ldots, 180$. The grids for a_j and b_k are not explicitly specified. In a simulation study, to test the null hypothesis of isotropy they introduce separate tests for each direction. Anisotropy tests based on energy densities different from the scale-angle energy density are described in D'Ercole and Mateu (2014).

8   Discussion

In this paper, we have given an overview of methods for directional analysis of unmarked stationary point processes covering methods based on Fry points and summary statistics, spectral analysis, and wavelet analysis. We have illustrated how the methods can be used to detect isotropy, to estimate favorable directions in the data and to test for isotropy. The methods based on the nearest neighbor and second-order summary statistics are described in 2D and 3D and the methods based on wavelet analysis in 2D. Spectral analysis, except the polar coordinate representation, is presented in general dimension even though in the literature the presentation has been limited to 2D. We have considered two simple simulated examples, a compressed and rotated realization of a Strauss process to represent geometric anisotropy, and a pattern with clustering along parallel stripes to represent 1st order anisotropy, to illustrate how the results of the analyses based on the different methods would typically look like. Based on the plots in Sections 4, 5, and 6, we can make some observations concerning the suitability of the methods in these two archetypal cases. However, a much more thorough study than the one presented here would be needed to draw any general conclusions.

Our review does not cover all aspects of anisotropy that have been discussed in the literature. First, we have concentrated on point patterns that are realizations of stationary point processes. In the literature,

directional analysis based on wavelets has been focused on situations where anisotropy can be detected from the intensity. Typically, the patterns that have been analyzed have a higher intensity of points along a line or along two perpendicular lines, and stationarity has not been assumed. Second, we have not given a thorough overview of anisotropic stationary point process models even though some of them are mentioned in the paper. Finally, there is some literature on anisotropy in marked point processes, namely orientation of marks and anisotropic distribution of points (see e.g. Penttinen and Stoyan 1989; Stoyan 1991; Stoyan and Benes 1991) which is not covered here.

We would like to finish by pointing out some directions for future research. Extension of the methods to higher dimensions, for example finding a polar coordinate representation in spectral analysis and a proper treatment of rotations in the wavelet approach, would be desirable. More in-depth analysis of the statistics of the summary functions would be beneficial as directional sub-sampling effectively reduces sample sizes, making user's input on estimation parameters more influential. Related to this, testing for isotropy could be made more objective as the current tests depend heavily on the user providing good input parameters, for example the bandwidth and lag vectors in the asymptotic test by Guan et al. (2006), distance r in the test based on stochastic reconstruction by Wong and Chiu (2016), opening angle and integration limits in the test based on replicates by Redenbach et al. (2009), and opening angle and contour level in the ellipsoid based test by Rajala et al. (2016). The relevant mathematics might turn out to be overtly challenging, in which case guidelines could be improved by diverse simulation experiments. The current tests based on spectral and wavelet analyses require some specified isotropic null model, which can be difficult to specify especially in 3D. Therefore, general tests that do not require any pre-specified model for the location process would be welcome. An interesting question related to the modelling of anisotropic point patterns is whether the linear transformation is the only useful, generally applicable anisotropy mechanism if the location process and anisotropy are considered separable.

9   Acknowledgement

References

Q. W. Ang, A. Baddeley, and G. Nair. Geometrically corrected second order analysis of events on a linear network, with applications to ecology and criminology. *Scandinavian Journal of Statistics*, 39 (4):591–617, 2012.

A. Baddeley, J. Møller, and R. Waagepetersen. Non-and semi-parametric estimation of interaction in inhomogeneous point patterns. *Statistica Neerlandica*, 54(3):329–350, 2000.

A. Baddeley, J-F. Coeurjolly, E. Rubak, and R. Waagepetersen. Logistic regression for spatial gibbs point processes. *Biometrika*, 101(2):377, 2014.

M.S. Bartlett. The spectral analysis of two-dimensional point processes. *Biometrika*, 51(3):299–311, 1964.

C. Chatfield. *The Analysis of Time Series: An Introduction*. Chapman and Hall, 4th edition, 1989.

R. D'Ercole and J. Mateu. On wavelet-based energy densities for spatial point processes. *Stochastic Environmental Research and Risk Assessment*, 27(6):1507–1523, 2013a.

R. D'Ercole and J. Mateu. A continuous wavelet-based approach to detect anisotropic properties in spatial point processes. *International Journal of Wavelets, Multiresolution and Information Processing*, 11(2): 1350017–1–1350017–29, 2013b.

R. D'Ercole and J. Mateu. A wavelet-based approach to quantify the anisotropy degree of spatial random point configurations. *International Journal of Wavelets, Multiresolution and Information Processing*, 12(06):1450017–1–1450017–22, 2014.

P. Diggle. Displaced amacrine cells in the retina of a rabbit: analysis of a bivariate spatial point pattern. *Journal of Neuroscience Methods*, 18:115–125, 1986.

P. Diggle, D. Gates, and A. Stibbard. A nonparametric estimator for pairwise-interaction point processes. *Biometrika*, 74(4):763–770, 1987.

P. Diggle, P. Moraga, B. Rowlingson, and B. M. Taylor. Spatial and Spatio-Temporal Log-Gaussian Cox Processes: Extending the Geostatistical Paradigm. *Statistical Science*, 28(4):542–563, 11 2013.

E.D. Ford and E. Renshaw. The interpretation of process from pattern using two-dimensional spectral analysis: modelling single species patterns in vegetation. *Vegetatio*, 56(2):113–123, 1984.

N. Fry. Random point distributions and strain measurement in rocks. *Tectonophysics*, 60:89–105, 1979.

Y. Guan. A composite likelihood cross-validation approach in selecting bandwidth for the estimation of the pair correlation function. *Scandinavian Journal of Statistics*, 34(2):336–346, 2007.

Y. Guan, M. Sherman, and J. Calvin. Assessing isotropy for spatial point processes. *Biometrics*, 62(1): 119–25, mar 2006.

H. Häbel, T. Rajala, M. Marucci, C. Boissier, K. Schladitz, C. Redenbach, and A. Särkkä. A three-dimensional anisotropic point process characterization for pharmaceutical coatings. *Spatial Statistics*, 2017. (in press).

J. Illian, A. Penttinen, H. Stoyan, and D. Stoyan. *Statistical analysis and modelling of spatial point patterns*. John Wiley & Sons Ltd., 2008.

J. B. Illian, S. H. Sorbye, and H. Rue. A toolbox for fitting complex spatial point process models using integrated nested Laplace approximation (INLA). *Annals of Applied Statistics*, 6(4):1499–1530, 2012.

M. N. Kanaan. *Cross-Spectral Analysis For Spatial Point-Lattice Processes*. PhD thesis, The Open University, 2000.

D. König and V. Schmidt. Directional distributions for multi-dimensional random point processes. *Communications in Statistics. Stochastic Models*, 8(4):617–636, 1992.

A. Kukush, I. Markovsky, and S. Van Huffel. Consistent estimation in an implicit quadratic measurement error model. *Computational Statistics and Data Analysis*, 47(1):123–147, 2004.

A.B. Lawson, S. Simeon, M. Kulldorff, A. Biggeri, and C. Magnani. Line and point cluster models for spatial health data. *Computational Statistics and Data Analysis*, 51:6027–6043, 2007.

J. Mateu and O. Nicolis. Multiresolution analysis of linearly oriented spatial point patterns. *Journal of Statistical Computation and Simulation*, 85(3):621–637, 2015.

J. Mateu, O. Nicolis, and R. D'Ercole. Testing for anisotropy in spatial point processes. In González-Manteiga et al., editor, *Proceedings of the Fifth International Workshop on Spatio-Temporal Modelling*. Unidixital, 2010. ISBN 978-84-934272-8-3.

J. Mecke and D. Stoyan. Formulas for stationary planar fibre processes I - General theory. *Series Statistics*, 11:267–279, 1980.

M. N. Miriti, H. F. Howe, and S. J. Wright. Spatial patterns of mortality in a colorado desert plant community. *Plant Ecology*, 136(1):41–51, 1998.

J. Møller and H. Toftaker. Geometric Anisotropic Spatial Point Pattern Analysis and Cox Processes. *Scandinavian Journal of Statistics*, 41(2):414–435, 2014.

J. Møller, F. Safavimanesh, and J. G. Rasmussen. The cylindrical K-function and Poisson line cluster point processes. *Biometrika*, 103(4):937–954, 2016.

M.A. Mugglestone. *Spectral analysis of spatial point processes*. PhD thesis, University of Edinburgh, 1990.

M.A. Mugglestone and E. Renshaw. A practical guide to the spectral analysis of spatial point processes. *Computational Statistics & Data Analysis*, 21(1):43–65, 1996a.

M.A. Mugglestone and E. Renshaw. The explanatory analysis of bivariate spatial point patterns using cross-spectra. *Environmetrics*, 7(4):361–377, 1996b.

M.A. Mugglestone and E. Renshaw. Spectral tests of randomness for spatial point patterns. *Environmental and Ecological Statistics*, 8(3):237–251, 2001.

M. Myllymäki and A. Penttinen. Conditionally heteroscedastic intensity-dependent marking of log Gaussian Cox processes. *Statistica Neerlandica*, 63(4):450–473, 2009.

J. Ohser and D. Stoyan. On the second-order and orientation analysis of planar stationary point processes. *Biometrical Journal*, 23(6):523–533, 1981.

A. Penttinen and D. Stoyan. Statistical Analysis for a Class of Line Segment Processes. *Scandinavian Journal of Statistics*, 16:153–168, 1989.

A.H. Rafati, F. Safavimanesh, K.-A. Dorph-Petersen, J.G. Rasmussen, J. Møller, and J.R. Nyengaard. Detection and spatial characterization of minicolumnarity in the human cerebral cortex. *Journal of Microscopy*, 261(1):115–126, 2016.

T. Rajala and A. Penttinen. Bayesian analysis of a Gibbs hard-core point pattern model with varying repulsion range. *Computational Statistics & Data Analysis*, 71:530 – 541, 2014.

T. Rajala, A. Särkkä, C. Redenbach, and M. Sormani. Estimating geometric anisotropy in spatial point patterns. *Spatial Statistics*, 15:139–155, 2016.

J.G. Ramsay. *Folding and Fracturing of Rocks*. McGraw Hill, San Francisco, 1967.

C. Redenbach, A. Särkkä, J. Freitag, and K. Schladitz. Anisotropy analysis of pressed point processes. *Advances in Statistical Analysis*, 93(3):237–261, 2009.

E. Renshaw and E.D. Ford. The interpretation of process from pattern using two-dimensional spectral analysis: Methods and problems of interpretation. *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 32(1):51–63, 1983.

E. Renshaw and E.D. Ford. The description of spatial pattern using two-dimensional spectral analysis. *Vegetatio*, 56(2):75–85, 1984.

M.S. Rosenberg. Wavelet analysis for detecting anisotropy in point patterns. *Journal of Vegetation Science*, 15(2):277–284, 2004.

M.S. Rosenberg and C.D. Anderson. PASSaGE: Pattern Analysis, Spatial Statistics and Geographic Exegesis. Version 2. *Methods in Ecology and Evolution*, 2(3):229–232, 2011.

F. Safavimanesh and C. Redenbach. A comparison of functional summary statistics to detect anisotropy of three-dimensional point patterns. *https://arxiv.org/abs/1604.04211*, 2016.

D. Stoyan. Describing the anisotropy of marked planar point process. *Statistics*, 22(3):449–462, 1991.

D. Stoyan and V. Benes. Anisotropy analysis for particle systems. *Journal of Microscopy*, 164:159–168, 1991.

A. Tscheschel and D. Stoyan. Statistical reconstruction of random point patterns. *Computational Statistics and Data Analysis*, 51(2):859–871, 2006.

M. N M van Lieshout. A J-function for inhomogeneous point processes. *Statistica Neerlandica*, 65(2): 183–201, 2011.

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