Detecting anisotropy in spatial point patterns - a comparison of statistical indices

T. Rajala[a], C. Redenbach[b,], A. Särkkä[c], M. Sormani[b,d]

[a]*Natural Ressources Institute Finland, PL 2 00791 Helsinki, Finland*
[b]*Technische Universität Kaiserslautern, Mathematics Department, 67663 Kaiserslautern, Germany*
[c]*Chalmers University of Technology and University of Gothenburg, Department of Mathematical Sciences, 412 96 Gothenburg, Sweden*
[d]*Fraunhofer Institut für Techno- und Wirtschaftsmathematik, Fraunhofer-Platz 1, 67663 Kaiserslautern, Germany*

Abstract

Isotropy of a point process, defined as invariance of the distribution under rotation, is often assumed in spatial statistics. Formal tests for the hypothesis of isotropy can be created by comparing directional summary statistics in different directions. In this paper, the statistical powers of tests based on a variety of summary statistics and several choices of the test statistics are compared in a simulation study. Four models for anisotropic point processes are considered covering both regular and clustered cases. We discuss the robustness of the results to changes of the tuning parameters, and highlight the strengths and limitations of the methods.

Keywords: Fry plot, Monte-Carlo test, Power analysis, Second-order characteristics, Spatial summary statistics, Spectral analysis, Test statistics, Wavelets

1. Introduction

Non-parametric methods for the directional analysis of 2D and 3D stationary point processes were reviewed in Rajala et al. (2018). Common tools are based on nearest neighbour and second order summary statistics, spectral theory, and wavelets. In this follow-up paper, the performance of the methods in detecting anisotropy in regular and clustered point patterns is illustrated.

We present a thorough investigation of isotropy tests using test statistics based on the methods mentioned in Rajala et al. (2018). The main issue in practice is the choice of the particular null hypothesis of isotropy and how to determine, or numerically access, the distribution of the chosen test statistic under this null distribution. To this end, various techniques are used in the literature. Some tests are asymptotic, some are based on Monte Carlo simulations, and some rely on replicated data, see Rajala et al. (2018) and references therein. Here, we assume the ideal situation where the null model, consisting of an isotropic point process and an anisotropy mechanism, are known. In this case, we can determine any distributional property under the null hypothesis by using Monte Carlo simulations. This assumption may not be very realistic in practice, but it allows for a fair comparison of several approaches independently of additional error sources. We include tests based on nearest neighbour and second order summary statistics, spectral theory, and wavelets, as well as a Monte Carlo version of the test introduced in Wong & Chiu (2016). The often functional summary statistics are further combined with different test statistics, and special attention is paid to the choice of the required integration limits and other tuning parameters. Our focus will be on the sensitivity of the combinations of summary statistic and test statistic in terms of the power of the tests. For simplicity, we restrict our attention to 2D but we expect the results to transfer to 3D.

As outlined in Rajala et al. (2018), there are several mechanisms that can be used for introducing anisotropy into a point pattern. To explore the options we investigate four null models: the Strauss process, the Thomas process, points scattering around a single line (the so-called hump process), and the Poisson line cluster point process (Møller et al., 2016). The regular Strauss process and the Thomas cluster process are made anisotropic by using a linear transformation, resulting in the so called geometric

Email address: redenbach@mathematik.uni-kl.de (C. Redenbach)

anisotropy (Møller & Toftaker, 2014). The hump process, where points are concentrated around a horizontal line, is anisotropic by construction. Additionally, any single realisation is clearly not stationary. Nevertheless, it is included here since such processes often appear in the anisotropy literature, and some of the statistics are developed for such data. In the Poisson line cluster process, anisotropy is controlled by the direction distribution of the underlying random line process.

The paper is organised as follows. The directional summary statistics are briefly recalled in Section 2. Test statistics derived from them are introduced in Section 3. The set-up for the simulation study, including the different models, is given in Section 4, and the results are presented in Section 5. Finally, the findings are discussed in Section 6.

2. Directional methods

This section lists directional summary statistics that will be used in the comparison study in Section 5. For a more detailed description we refer to the review article Rajala et al. (2018). We restrict our attention to point patterns in $\mathbb{R}^2$.

We assume that the data is a set of distinct locations, i.e. a point pattern $\mathbf{x} = \{x_1, ..., x_n\}$, observed in a fixed window $W \subseteq \mathbb{R}^2$ with $n > 2$ points. Furthermore, we assume that $\mathbf{x}$ is obtained by intersecting a realisation of a simple (no multiple points) point process X with W, and that W is bounded with area $0 < |W| < \infty$.

The point process X is stationary, if its distribution is invariant under shifts in $\mathbb{R}^2$. If its distribution is invariant under rotations around the origin, the process is isotropic. With the exception of the hump process (see Section 4), we assume that the point process X is stationary with intensity $\lambda > 0$.

Let $b(x, r)$ denote a disc in $\mathbb{R}^2$ with center x and radius $r > 0$. A unit vector $u \in \mathbb{R}^2$ can be represented in polar coordinates via $u = (u_1, u_2)^T = (\cos(\alpha), \sin(\alpha))^T$ with $\alpha \in [0, 2\pi]$. The infinite double cone with a central axis spanned by $u = u(\alpha)$ and with the opening half angle $\epsilon > 0$ is denoted by $S(\alpha, \epsilon)$. Furthermore, we set $S(\alpha, \epsilon, r) = S(\alpha, \epsilon) \cap b(o, r)$, where o is the origin. Additionally, 2D cylinders (rectangles) are denoted by $C(\alpha, w, r)$, where α represents the main axis direction, w is the half-width and r is the length of the rectangle.

We will write $1_{\mathcal{E}}(x)$ for the indicator function taking the value 1 if $x \in \mathcal{E}$ and 0 otherwise.

2.1. Nearest neighbour methods

Let x_0 be the nearest neighbour of the typical point o of X which, by stationarity, can be assumed to be in the origin. Denote by $d_0 = ||x_0 - o||$ the distance from the typical point to its nearest neighbour x_0 and by θ the direction to the nearest neighbour given by the angle between the vector $x_0 - o$ and the positive x axis in anti-clockwise direction.

We assume that the distribution of θ has a density $p(\theta)$ called the *nearest neighbour orientation density*, which in 2D can be interpreted as a function on the interval $[0, 2\pi]$. For an isotropic point process X, the nearest neighbour directions follow a uniform distribution on the unit sphere, hence in 2D $p(\theta) = 1/2\pi$ for all θ. Note that the range can be reduced to $[0, \pi]$ if opposite directions are identified. For estimating p, a classical kernel density estimator with smoothing bandwidth $h > 0$ can be applied to the observed nearest neighbour angles.

The power of directional analysis based on nearest neighbours is improved if not only the direction but also the distance to the nearest neighbour is taken into account. To this end, a directional (local) version G_{loc} of the nearest neighbour distance distribution G can be defined as

$$G_{\text{loc},\alpha,\epsilon}(r) = P_o \left[d\left(o, X \cap S(\alpha, \epsilon)\right) < r \right],$$

where P_o refers to the Palm distribution of X, see Redenbach et al. (2009). That is, $G_{\text{loc},\alpha,\epsilon}$ is the cumulative distribution function of the distance from a typical point to the nearest neighbour in $\mathbf{x}$ restricted to the double cone $S(\alpha, \epsilon)$.

In addition, a global directional version of G has been introduced as the distribution of the distance from a typical point to its nearest neighbour in $\mathbf{x}$ conditionally on the event that this nearest neighbour is in the cone $S(\alpha, \epsilon)$, i.e.

$$G_{\text{glob},\alpha,\epsilon}(r) = P_o \left[d_0 < r | x_0 \in S(\alpha, \epsilon) \right],$$

Considering only points with nearest neighbour in the cone drastically reduces the sample size. Hence, this statistic may show an unstable behaviour when used for small point patterns (Redenbach et al., 2009).

2.2. Second order statistics

For Borel sets B, the reduced second-order moment measure $\mathcal{K}$ of the point process X is given by

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

where $\mathbb{E}_o$ refers to an expectation w.r.t. the Palm distribution P_o of X. The quantity $\lambda\mathcal{K}(B)$ can be interpreted as the expected number of further points within B conditioned on $o \in X$. Setting $K(r) = \mathcal{K}(b(o,r))$ yields Ripley's K-function. Directional K-functions can be defined by replacing $b(o,r)$ by a suitable scaling $B_{\alpha,r}$ of some directed set B_α with reference direction α. The choice $B_{\alpha,r} = S(\alpha,\epsilon,r)$ yields the sector or conical K-function $K_{\text{sect},\alpha,\epsilon}(r)$, see Redenbach et al. (2009). Setting $B_{\alpha,r} = C(\alpha,w,r)$ yields the cylindrical K-function $K_{\text{cyl},\alpha,w}(r)$ (Møller et al. (2016)). In practice, the half-width w of the cylinder can be fixed or can be chosen as a function of r, i.e. $w = w(r)$. For an illustration, see Figure A.11.

An edge corrected estimator of $K_{B,\alpha}(r)$ using translational edge correction weights can be defined by

$$\hat{K}_{B,\alpha}(r) := \frac{1}{\widehat{\lambda^2}} \sum_{x,y\in\mathbf{x}}^{\neq} \frac{1_{B_{\alpha,r}}(y-x)}{|W_y \cap W_x|}, \qquad \alpha \in [0,\pi], \tag{1}$$

where the upper $\neq$ means that only pairs with $x \neq y$ are included in the sum.

Choosing $B_{\alpha,r}$ to be an ellipse yields the elliptical K-function (Sormani et al., 2020). Here, we consider a smoothed version, where the indicator function in (1) is replaced by a Gaussian kernel with mean $\mu = 0$ and anisotropic covariance matrix $\Sigma_{\alpha,r}$ depending on a scaling parameter $r > 0$, to obtain

$$\hat{E}_{\text{gauss}}(r) = \frac{1}{\widehat{\lambda^2}} \sum_{x,y\in\mathbf{x}}^{\neq} \frac{1}{\sqrt{(2\pi)^2 \det(\Sigma_{\alpha,r})}} \frac{\exp\big(-\frac{1}{2}(x-y)^T \Sigma_{\alpha,r}^{-1}(x-y)\big)}{|W_y \cap W_x|}. \tag{2}$$

In our study, the covariance matrix is chosen as $\Sigma_{\alpha,r} = Q^T Q$ with $Q = R(\alpha)(r/2)\,\text{diag}(1,c)$, where $R(\alpha)$ is a rotation matrix, the standard deviation $r/2$ is set so that approximately 95% of density along the major axis is within distance r, and the parameter c controls the ellipticity of the kernel shape. An illustration of the resulting shapes can be found in Figure A.11.

A directional version of the pair correlation function can be derived from the sector K-function via

$$K_{\text{sect},\alpha,\epsilon}(r) = \int_0^r g_{\text{sect},\alpha,\epsilon}(t)dt$$

and can be estimated via

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

with intensity estimator $\hat{\lambda}$ and where $\check{\alpha}$ denotes the antipodal direction of α. The kernel w_h with bandwidths $h = (\sigma,\epsilon)$ is chosen as

$$w_h(v,(r,\alpha)) = \frac{3}{4\sigma^2}(1 - (||v|| - r)^2/\sigma^2) \cdot \frac{1}{2\epsilon} 1_{[0,\epsilon)}(|\alpha(v) - \alpha|).$$

2.3. Spectral methods

In spectral analysis, the *spectrum* $\mathcal{F}(\omega)$ of a stationary point process X is defined as the Fourier transform of the complete covariance density function $\kappa(z) = \lambda\delta(z) + \rho^{(2)}(z) - \lambda^2$, where δ is the Dirac delta function and $\rho^{(2)}$ denotes the second-order product density or second-order intensity function (assuming that it exists). That is,

$$\mathcal{F}(\omega) := \int_{\mathbb{R}^2} \kappa(z) e^{-i\omega^T z} dz. \tag{3}$$

The estimator of the spectrum $\widehat{\mathcal{F}}(\omega)$ is called periodogram. It is based on the discrete Fourier transform of the pattern $\mathbf{x}$ (for details see Rajala et al. (2018, 2022); Bartlett (1964)).

Assume that the observation window is a square with edge length L. We define a 2D grid of frequencies $\omega \in 2\pi\{-15L/2, -14L/2, ..., 15L/2\}^2$. Let u_θ be the unit vector in direction θ and denote the angle between ω and u_θ by angle(w, u_θ). A marginal summary in direction only, the so called Θ *spectrum*, can be derived from the periodogram by

$$\widehat{\mathcal{F}}_\Theta(\theta) = \frac{\sum_\omega 1_{[0,h)}(\text{angle}(w, u_\theta))\widehat{\mathcal{F}}(\omega)}{\sum_\omega 1_{[0,h)}(\text{angle}(w, u_\theta))}, \quad \theta \in [0, 180^\circ).$$

The angle calculation should adjust for antipodal symmetry. In the examples, we used a smoothing bandwidth of $h = 7.5^\circ$ and $\theta = 0^\circ, 3.6^\circ, ..., 176.4^\circ$.

The Θ spectrum summarises average periodogram values for ordinates with similar values of θ and should be constant under isotropy.

2.4. Wavelet methods

An idea closely related to spectral analysis is to use 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 & Nicolis, 2015; D'Ercole & Mateu, 2013a,b, 2014).

We first describe the approach of Rosenberg (2004). Given a realisation $\mathbf{x} = \{x_1, \ldots, x_n\}$ 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 $\theta_i = \frac{\pi i}{180}$. 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 $\eta(x, \theta_i)$, $i = 0, \ldots, 179$. 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

$$W(x, \theta, b_k) = \frac{1}{b_k} \sum_{i=0}^{179} \eta(x, \theta_i)\psi\left(\frac{\theta_i - \theta}{b_k}\right). \tag{4}$$

We will follow the choice of Rosenberg (2004) who uses the French Top Hat wavelet for ψ. When applying Equation (4) one should treat η as a periodic function in θ_i to avoid edge effects in the angle domain. The overall variance of the wavelet transform for given focal point x and direction θ is defined as

$$P(x, \theta) = \frac{1}{m} \sum_{k=1}^{m} W^2(x, \theta, b_k), \tag{5}$$

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

D´Ercole, Mateu, and Nicolis suggest to apply a directional continuous wavelet transform (CWT) to the intensity function restricted to W obtaining the coefficients

$$V(a, b, \theta) = a^{-1} \int_W \overline{\psi_{a,b,\theta}(x)}\lambda(x)dx, \quad x \in \mathbb{R}^2. \tag{6}$$

Here, $\psi_{a,b,\theta}(x) := \psi(a^{-1}R_{-\theta}(x - b))$, where ψ is the mother wavelet, $a > 0$ is the scaling parameter, $b \in \mathbb{R}^2$ the translation parameter, and $R_{-\theta}$ represents a clockwise rotation by an angle θ. The mother wavelet is chosen as the Morlet wavelet

$$\psi(x) = \sqrt{D/\pi} \exp\left(ik_0^T x - 0.5\ x^T Q x\right)$$

with $D = 0.1, k_0 = (0, 5.5)^T$ and $Q = diag(D^2, 1)$.

An estimator for $V(a, b, \theta)$ is given by

$$\hat{V}(a, b, \theta) = a^{-1} \sum_{x \in \mathbf{x}} \overline{\psi_{a,b,\theta}(x)}, \tag{7}$$

which can be shown to be unbiased by using the Campbell theorem, see D'Ercole & Mateu (2013a,b, 2014).

The square modulus of $\hat{V}(a,b,\theta)$ is integrated over b, obtaining the *scale-angle energy density* (D'Ercole & Mateu, 2013a; Mateu & Nicolis, 2015)

$$\nu(a,\theta) = \int |\hat{V}(a,b,\theta)|^2 db. \tag{8}$$

Additionally, we consider an *angle energy density* where (8) is averaged over scale, that is

$$\bar{\nu}(\theta) = \int\int |\hat{V}(a,b,\theta)|^2 db\, da. \tag{9}$$

In practice, both integrals are approximated by summing over an equidistant grid of values for a and b.

3. Anisotropy tests

The directional summary statistics defined in the previous section can be transformed into a test statistic for an anisotropy test in several ways. In the following, let S_{α_1} and S_{α_2} denote the same directional summary statistic S for two selected angles (directions) α_1 and α_2. The test is based on the idea that the contrast $v(r) = S_{\alpha_1}(r) - S_{\alpha_2}(r)$ should be small for any r under the null hypothesis of isotropy.

For a given point pattern $\mathbf{x}$, we estimate S_{α_1} and S_{α_2} on an equidistant grid of distance values $r_0 = 0, \ldots, r_m = r_{\max}$. Thus, we obtain estimates $\hat{S}_{\alpha_1}(r_i)$ and $\hat{S}_{\alpha_2}(r_i)$ which are used to form a contrast vector

$$\mathbf{v} = (v_0, \ldots, v_m) = \left(\hat{S}_{\alpha_1}(r_0) - \hat{S}_{\alpha_2}(r_0), \ldots, \hat{S}_{\alpha_1}(r_m) - \hat{S}_{\alpha_2}(r_m)\right).$$

Additionally, we consider the mean contrast under H_0

$$\mathbf{m} = (\mathbb{E}(S_{\alpha_1}(r_0) - S_{\alpha_2}(r_0)), \ldots, \mathbb{E}(S_{\alpha_1}(r_m) - S_{\alpha_2}(r_m)))$$

which can either be known or estimated from Monte Carlo simulations of the null model. We will do the latter as it allows to take possible estimation bias for the different directions into account. Additionally, we consider the empirical covariance matrix

$$\hat{C} = (\widehat{\mathrm{cov}}(v_i, v_j))_{i,j}.$$

We use the following test statistics, where $\hat{\mathbf{m}}$ is a Monte Carlo estimate of $\mathbf{m}$:

1. a χ^2-style statistic $T_{\chi^2} = (\mathbf{v} - \hat{\mathbf{m}})^T \hat{C}^{-1} (\mathbf{v} - \hat{\mathbf{m}})$,
2. a mean squared deviation statistic $T_{ms} = ||\mathbf{v} - \hat{\mathbf{m}}||^2 = (\mathbf{v} - \hat{\mathbf{m}})^T (\mathbf{v} - \hat{\mathbf{m}})$ as a discrete approximation of the integral statistic
$$T_{L^2} = \int_0^{r_{\max}} (v(r) - \hat{m}(r))^2 dr,$$
3. a range-wise standardised version of T_{ms} defined as
$$T_{ms_{st}} = \sum_{i=0}^{m} \frac{(v_i - \hat{m}_i)^2}{\hat{C}_{ii}} = (\mathbf{v} - \hat{\mathbf{m}})^T \hat{C}_{diag}^{-1} (\mathbf{v} - \hat{\mathbf{m}}),$$
where $\hat{C}_{diag}$ is $\hat{C}$ with off-diagonals set to 0, and
4. the maximum absolute deviation $T_{mad} = \max_i |v_i - \hat{m}_i|$. Note that for summary statistics that are distribution functions this equals the Kolmogorov-Smirnov test statistic.

Some summary statistics take an angle θ rather than a range r as argument, namely, the nearest neighbour orientation density, the spectral and the wavelet statistics. In these cases, we simply treat θ like r and use the summary itself in place of the contrast in $\mathbf{v}$ above.

Additionally, we apply the test by Wong & Chiu (2016) which is based on the ratio

$$F_{r,\theta}(\epsilon) = K_{\text{sect},\theta+\frac{\epsilon}{2},\frac{\epsilon}{2}}(r)/K_{\text{sect},\frac{\pi}{2},\frac{\pi}{2}}(r),$$

so we consider sectors of radius r centred at the origin enclosed between the lines making angles θ and $\theta + \epsilon$ w.r.t. the x-axis. Under the null hypothesis of isotropy, $F_{r,\theta}(\epsilon)$ is the distribution function of the uniform distribution on $[0, \pi]$. The test is based on the statistic

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

for fixed r.

The main challenge in practice is to determine the distribution of the test statistic under the null hypothesis. Here, we assume an idealised setting where both the null model X_0 and the optimal choice of the angles α_1 and α_2 are known. In this case, a Monte Carlo test using simulations of the null model can be applied.

In Sormani et al. (2020), an isotropy test based on a projection of the Fry points on the unit sphere is introduced. This test, however, is only valid for regular point patterns. Hence, it is not included in our study.

4. Simulation study

4.1. Models

We consider four point process models covering regular processes and clustered processes with different cluster shapes and arrangements. All models are simulated with a fixed intensity $\lambda \equiv 0.005$ in a square window $W = [-L/2, L/2]$. We chose L to be either $100, 141.42$ or 244.95, so the expected number of points was $n = 50, 100$ or 300, respectively. Additionally, we introduce the following set of parameters

- a range parameter $R > 0$,
- a parameter $\gamma \in [0, 1]$ controlling the strength of structural features, and
- the degree of anisotropy $a \in [0, 1]$.

The explicit choice of parameters for the different models will be detailed below. We simulated 500 model data patterns for each parameter combination (L, R, γ, a). For each data pattern we simulated further 199 isotropic realisations with the same (L, R, γ) to compute the empirical null distributions of summaries and test statistics.

4.1.1. Strauss process

To represent regular point processes we consider a geometric anisotropic version of the Strauss process. The range parameter R is the interaction range of the process. We use the values $R = 4$ and $R = 8$. The parameter γ is the strength of repulsive interaction which is set to $\gamma = 0, 0.2$, and 0.4, where 0 is the strongest and 0.4 the weakest, as 1 would equal Poisson process. Anisotropy is induced by compressing the point pattern in y-direction by a factor of $a = 0.7$ and stretching in x-direction by $1/a$. The case $a = 1$ yields the null model of isotropy. Variation of all parameters yields 36 combinations. Some examples are shown in Figure 1.

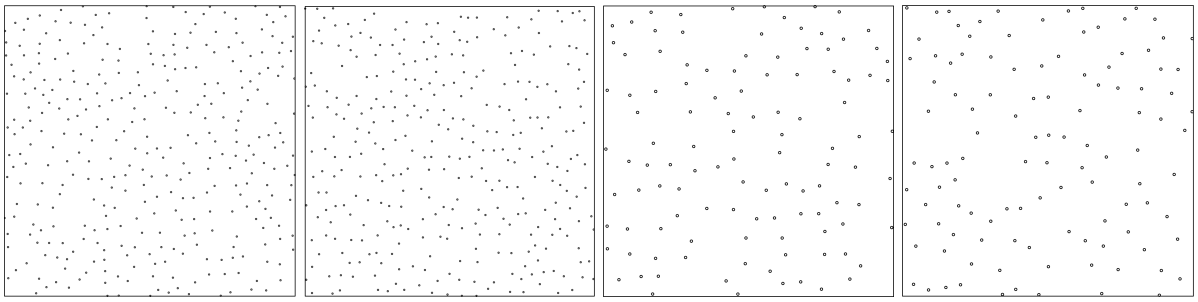

Figure 1: Realisations of the Strauss process with $\gamma = 0$, $R = 8$ and: $n = 300$, $a = 1$; $n = 300$, $a = 0.7$; $n = 100$, $a = 1$; $n = 100$, $a = 0.7$.

4.1.2. Thomas process

A Thomas process subject to geometric anisotropy is used as an example of a cluster process with spherical clusters under isotropy. The Thomas process is a Cox process which initially samples a homogeneous Poisson process of parent points. The observed part of the process is formed by the offsprings of these points which are simulated by random displacements of the parents that are drawn from an isotropic 2D Gaussian distribution with mean 0 and covariance matrix $\sigma^2 I$. In our parametrisation framework, we set $R = \sigma$ and γ is chosen such that $n\gamma$ is the number of clusters, bounded below by 5. The special case $\gamma = 0$ refers to the most clumped cases while $\gamma = 1$ yields a Poisson process. We use the same values as for the Strauss process, namely $R = 4$ and 8, and $\gamma = 0, 0.2$, and 0.4. Geometric anisotropy is modelled as described for the Strauss process with parameters $a = 0.7$ and $a = 1$ with the exception that compression is now in the x-direction. This yields 36 parameter combinations. Some examples are shown in Figure 2.

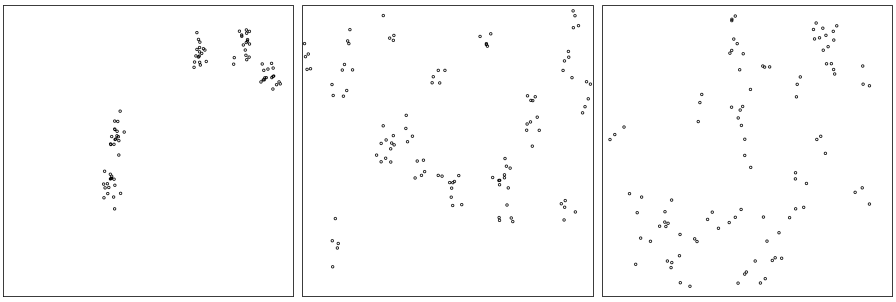

Figure 2: Realisations of the Thomas process, $n = 100$, $R = 8$ and $a = 0.7$, and from left to right: $\gamma = 0$; $\gamma = 0.2$; $\gamma = 0.4$.

4.1.3. Hump process

The name hump process refers to points accumulating along a horizontal line in the center of the window. Points are displaced in vertical direction with a distance drawn from a Gaussian distribution with mean 0 and variance σ^2. This horizontal cluster is superposed with background noise from a homogeneous Poisson process. We use $R = \sigma$ as a range parameter. The parameter γ is the fraction of points in the background process. That is $\gamma = 0$ has no background noise, while $\gamma = 1$ yields a homogeneous Poisson process and, hence, the null model. We consider $R = 2$ and 6, and $\gamma = 0.5$ and 1. In total, this yields 12 parameter combinations for the hump process. Examples are shown in Figure 3.

Obviously, this model is not stationary. Nevertheless, estimates of all summary statistics introduced in Section 2 can be computed.

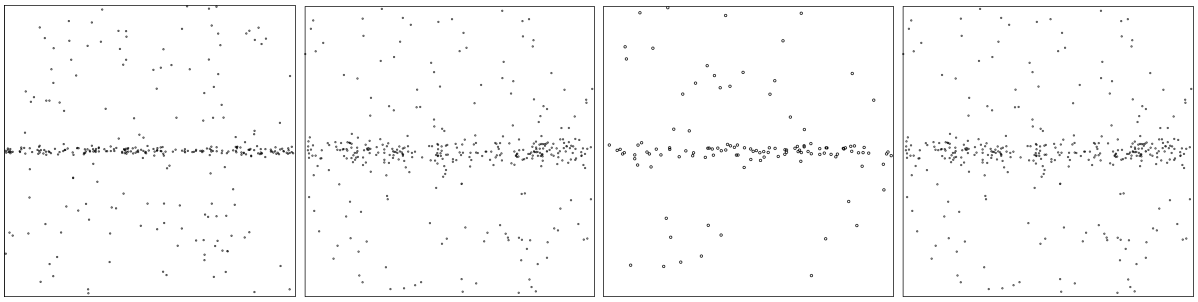

Figure 3: Realisations of the hump process for $\gamma = 0.5$ and from left to right: $n = 300$, $R = 2$; $n = 300$, $R = 6$; $n = 100$, $R = 2$; $n = 100$, $R = 6$.

4.1.4. Poisson line cluster point process

As the fourth model, we use a Poisson line cluster point process (Møller et al., 2016). As in the hump process, points are randomly displaced from lines in space. Again, shifts are drawn from a Gaussian distribution with mean 0 and variance σ^2. However, here, there are several lines drawn from a stationary Poisson line process. The number of lines follows a Poisson distribution with mean $\gamma = n/15$. Patterns are restricted to contain at least five lines. We set $R = \sigma$ which is chosen as either 0 (no displacement) or 1. Anisotropy is introduced when choosing a non-uniform distribution of the line direction. We model line directions by a von Mises distribution with preferred direction along the x-axis and concentration parameter $\kappa = 5 \cdot (1 - \exp(1 - \frac{1}{a}))$. The choice $a = 1$ yields a uniform distribution of line directions, hence an isotropic pattern. Additionally, we consider anisotropic patterns with $a = 0.6$ which yields $\kappa = 0.49$. Here, we have 12 parameter combinations. Some realisations are shown in Figure 4.

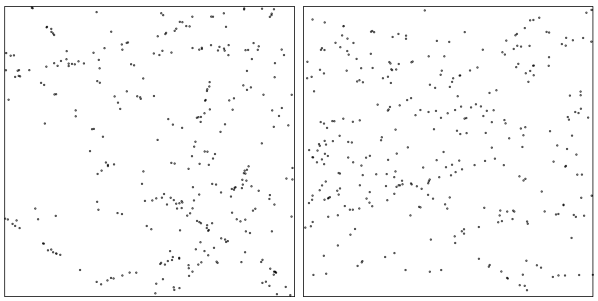

Figure 4: Realisations of the Poisson line cluster process for $n = 300$ with $R = 1$ and $a = 1$(left); $a = 0.6$ (right).

4.2. Summaries and choice of parameters

When formulating an isotropy test, both the summary statistic S and the test statistic T can be varied. An overview of the summaries and their parameters used in our study are given in Table 1.

For second order summary statistics, the grid of evaluation points for r is chosen from 0 to 20 in steps of 0.5, to ensure that estimates are free of edge effects. For computing the test statistics, we consider $25, 50, 75$, and 100% of the interval $[0, 20]$, such that $r_{\max} \in \{5, 10, 15, 20\}$. For nearest neighbour summaries with a range argument, the upper limit was set to 10, and only the full length interval was tested. The contrasted directions are chosen as $\alpha_1 = 0$ (x-axis) and $\alpha_2 = \frac{\pi}{2}$ (y-axis). For the cylindrical K-function K_{cyl}, two scaling regimes of the cylinder are considered. In the first case, the half-width w of the cylinder is fixed. In the second one, we fix the aspect ratio such that the area of the cylinder (rectangle) equals that of the double sector used in K_{sect}. An illustration can be found in Figure A.11.

To explore the effect of the two bandwidths σ and ϵ in g_{sect}, we will either fix σ and vary ϵ or the other way round. The two cases are denoted by $g^{\epsilon}_{\text{sect}}$ and g^{σ}_{sect}, respectively.

5. Results

Below, we summarise the results for the four models, Strauss, Thomas, Hump, and linear Cox. Under isotropy, all the tests have the nominal significance level of 5%. We discuss only a selection of parameter combinations per model and provide comprehensive results as supplementary material.

Summary/statistic	Main parameters	Tuning parameters
Nearest neighbour orientation density p	θ	$h = 0.35 \cdot c, \quad c \in \{0.8, 1, 1.2\}$.
G_{loc}	r, α	$\epsilon = \frac{\pi}{4}$
G_{glob}	r, α	$\epsilon = \frac{\pi}{4}$
K_{sect}	r, α	$\epsilon = c \cdot \frac{\pi}{4}, \quad c \in \{0.5, 1.0, 1.5\}$
K_{cyl} (fixed width)	r, α	$w = 1.5, 3, 4.5$
K_{cyl} (fixed aspect ratio)	r, α	$w(r) = c \cdot r\pi/8 \quad c \in \{0.5, 1, 1.5\}$
$\hat{E}_{\text{gauss}}$	r, α	$c = 0.25, 0.5, 0.75$
$g^{\epsilon}_{\text{sect}}$	r, α	$\epsilon = c \cdot \frac{\pi}{4}, \quad c \in \{0.5, 1.0, 1.5\}$; $\sigma = 0.15\lambda^{-1/2}$
g^{σ}_{sect}	r, α	$\epsilon = \frac{\pi}{4}$; $\sigma = 0.15\lambda^{-1/2} \cdot c, \quad c \in \{1.5, 3\}$
Direction spectrum $\mathcal{F}_{\Theta}$	θ	$h = 7.5^{\circ}$
Rosenberg $\bar{P}$	θ	$b_k = \frac{k\pi}{180}$, $k = 1 \ldots, 45$
Continuous wavelet ν, $\overline{\nu}$	θ	$a \in \{0, 10, ..., 100\}$, $b = \{0, 10, ..., 100\} \times \{0, 10, ..., 100\}$
Wong-Chiu T_r	$r \in \{5, 10, 15\}$	-

Table 1: Summary statistics, their main parameters of interest, and their tuning parameters.

5.1. Strauss process

Figure 5 shows the powers of the anisotropy tests for the different parameter sets of the Strauss process. The results are shown only for the best choices of the integration limit and the tuning parameters. The larger the sample size n, the stronger the interaction (smaller γ), and the larger the interaction range R, the easier it is to detect anisotropy. Over all, methods based on second order summary statistics (K functions, pair correlation functions g and $\hat{E}_{\text{gauss}}$) show the best performance. The Wong and Chiu test whose test statistic T_r is also based on second order information mostly yields comparable results. In the class of second order methods, the cylindrical K-function with fixed aspect ratio tends to give the highest powers. The local version G_{loc} is the best performing summary statistic among the nearest neighbour methods. The spectral approach only works well for the larger interaction range and the highest sample size considered. The use of the wavelet methods is not recommended.

When comparing the different test statistics, the standardised mean square statistic $T_{ms_{st}}$ has the tendency to produce the best results while the mean absolute deviation statistic T_{mad} behaves worst. In particular, for $\hat{E}_{\text{gauss}}$ the choice of the test statistic is crucial.

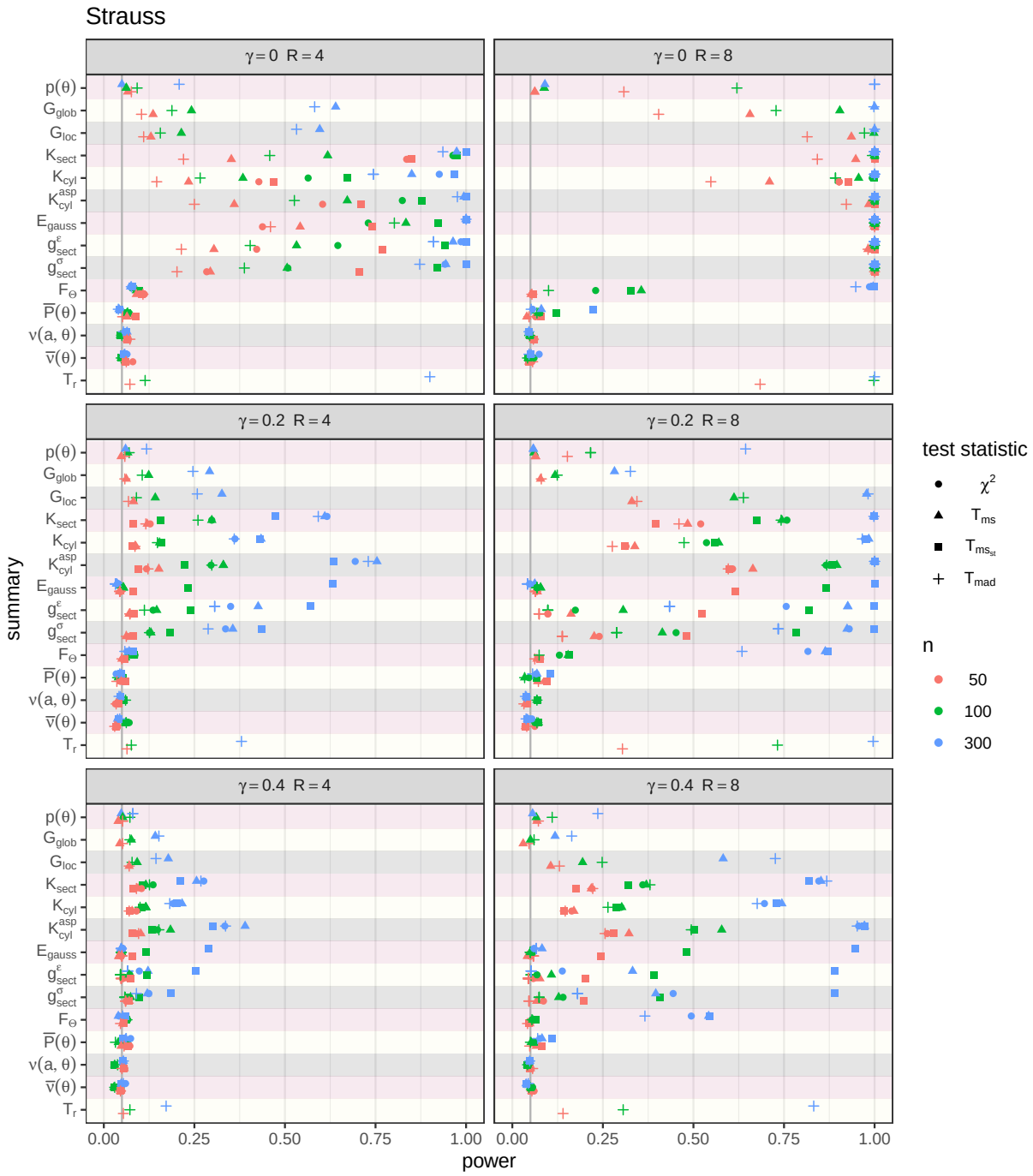

Figure 5: Powers of the tests for the Strauss process. The colour of the marker indicates the sample size and the shape the test statistic T. The results are shown only for the best choices of the integration limit and the tuning parameters.

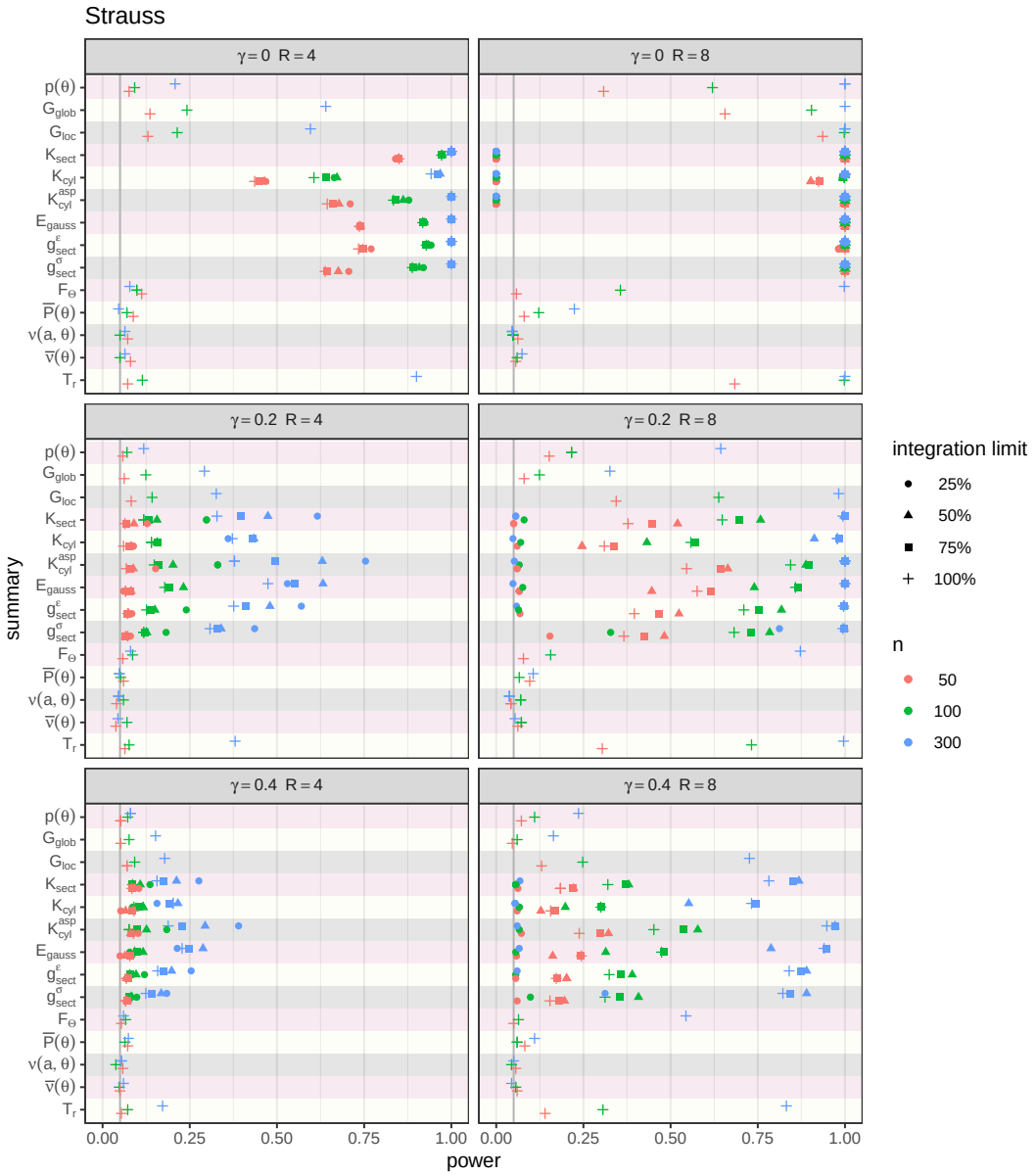

Figure 6: Powers of the tests for the Strauss process. The colour of the markers indicates the sample size and the shape the integration limit. The results are shown only for the best choices of the test statistic and the tuning parameters.

Additionally, the power of the test turned out to be sensitive to the choice of the upper limit $r_{\max}$ of the grid of evaluation points, see Figure 6. In general, $r_{\max}$ should be slightly larger than the interaction range R for all the second order methods. Choosing $r_{\max}$ too large reduces the power of the test. However, the best choice may depend on the other parameters and the sample size.

The pcf and the sector K function are not so sensitive to the choice of the opening angle but the best choice may depend on the parameters of the process, see Figures A.12 and A.13. Opening angles close to $\frac{\pi}{4}$ seem always to give good results. The aspect ratio does not play such a big role for the cylinder K function, although the larger aspect ratios give slightly better results when the interaction range is small and interaction weak. For the elliptical K function, the optimal aspect ratio depends on the other parameters but in general, the larger aspect ratios give always rather good results.

5.2. Thomas process

The powers of the tests for the Thomas process are shown in Figure 7 for the best choice of integration interval and tuning parameters. The general trend is that the power of the test increases with increasing

sample size, decreasing number of clusters (controlled by γ), and decreasing cluster radius R which yields more concentrated clusters.

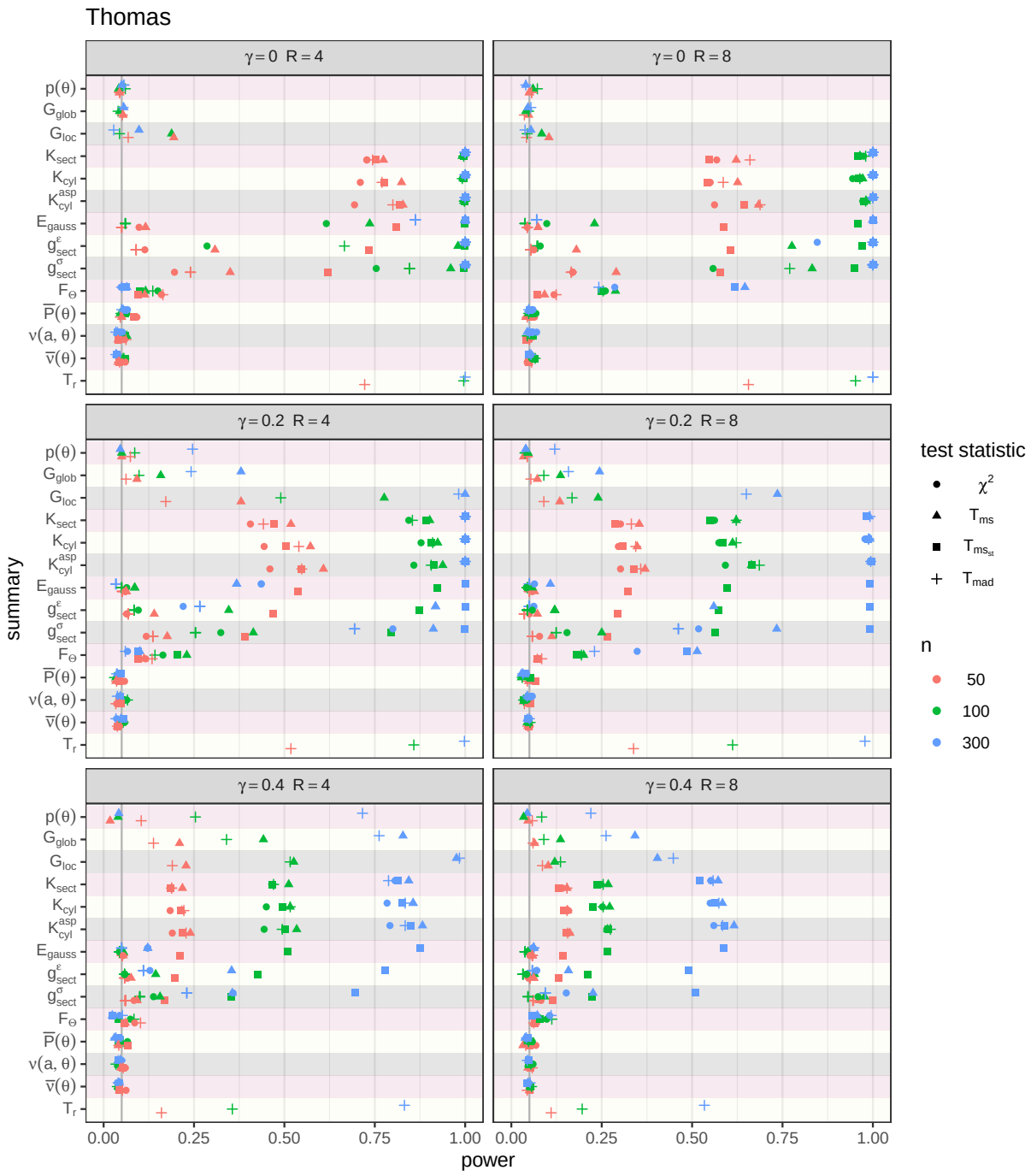

Figure 7: Powers of the tests for the Thomas process. The colour of the marker indicates the sample size and the shape the test statistic T. The results are shown only for the best choices of the integration limit and the tuning parameters.

In general, the second order methods (including Wong-Chiu) show the best results. Among those, using any of the K functions as summary statistics should be preferred to the pair correlation function g or $\hat{E}_{\text{gauss}}$. Nearest neighbour methods do not work very well in general. An exception is the case of small range clusters with only few points where G_{loc} shows results comparable to or even better than the K functions. The wavelet and spectral methods do not seem to be appropriate for the Thomas process.

The tests based on K functions are rather robust to the choice of the test statistic T. Typically, T_{ms} yields the highest powers. For g and $\hat{E}_{\text{gauss}}$ the differences between the statistics are much more pronounced, and $T_{ms_{st}}$ usually gives the best result. The test statistic T_{ms} is the best when the local nearest neighbour method is used.

The integration upper limit should be slightly larger than the range but not too large. The same is true for the range parameter r in the Wong-Chiu test. The other tuning parameters of the summary statistics in general do not have much effect on the results, see Figures A.14 and A.15. The summaries $\hat{E}_{\text{gauss}}$ and the cylindrical K function with fixed aspect ratio are exceptions.

5.3. Hump process

Powers for the anisotropy tests for the hump process are shown in Figure 8. The smaller the range (standard deviation of the point displacements from the line) and the larger the sample size, the easier it is to detect anisotropy. The fraction of background noise points γ was fixed in our study. It seems reasonable that larger noise levels should decrease the powers.

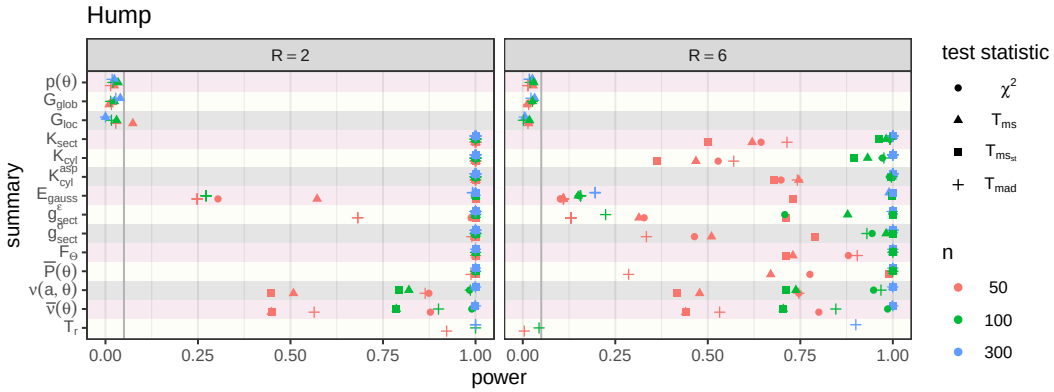

Figure 8: Powers of the tests for the hump process. The colour of the marker indicates the sample size and the shape the test statistic T. The results are shown only for the best choices of the integration limit and the tuning parameters.

All summary statistics, except the nearest neighbour statistics, work very well for the hump process, especially when the range is small. If the sample size is small and range large, the Rosenberg test is the best test, followed by the spectral test. The Wong-Chiu test is not recommended for sample sizes as low as 100 when the range is large.

Any test statistic can be chosen when the sample size is reasonably large and range small, independently of the summary statistic. For smaller sample sizes, there are significant differences between the powers for the different test statistics T when using the same summary statistic S. There is no uniformly optimal choice of test statistic across the different summaries.

A large integration limit, clearly larger than the range, is recommended. The same holds for the parameter r in the Wong-Chiu test. The remaining tuning parameters do not have much effect on the results, again with the exception of $\hat{E}_{\text{gauss}}$ and the cylindrical K function with fixed aspect ratio, see Figure A.16.

5.4. Poisson line cluster point process

The results obtained for the Poisson line cluster point process are shown in Figure 9. The process has two parameters, range and degree of anisotropy. For smaller range points lie closer to the lines while for smaller anisotropy parameter the lines are more parallel. It is more difficult to detect anisotropy when the sample size decreases and when the range parameter increases, as can be seen in Figure 9.

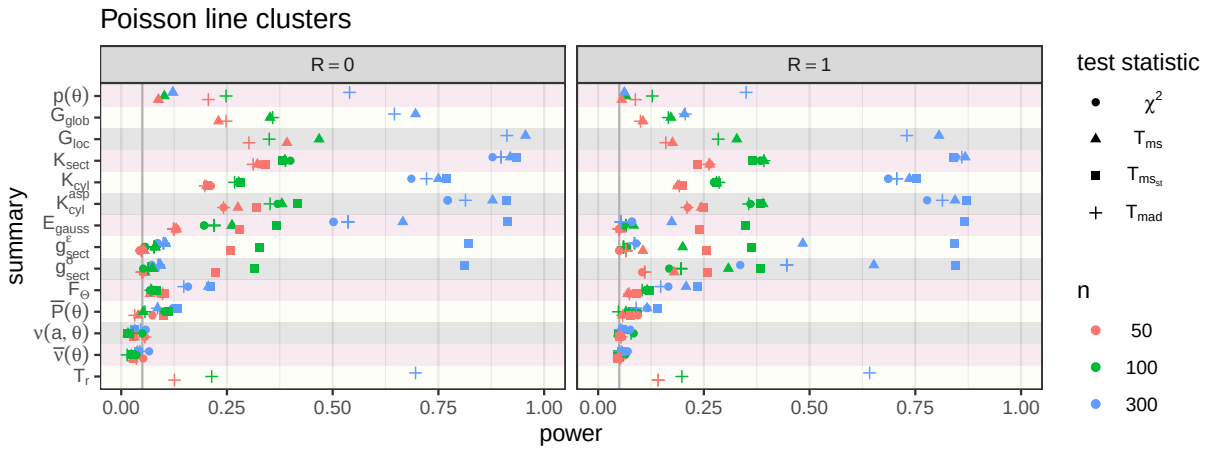

Figure 9: Powers of the tests for the Poisson line cluster process. The colour of the marker indicates the sample size and the shape the test statistic T. The results are shown only for the best choices of the integration limit and the tuning parameters.

All pcf and K function based methods work well for large sample sizes. For smaller sample sizes, the K function based methods work slightly better than the pcf method. However, if anisotropy is weak, parameter 0.6, the methods do not work so well except for the largest sample size. If the sample size is large enough, also the local G function works well. Spectral and wavelet methods are not recommended.

The test statistic $T_{ms_{st}}$ works well with all the pcf and K function based methods. However, the cylinder and sector K functions work well even with the other test statistics when the sample size is large. The local G function gives best results combined with T_{ms} and the nearest neighbour orientation density combined with T_{mad}. E_{gauss} shows high differences between the powers obtained with different test statistics.

A large integration limit (or r in the Wong-Chiu test) is always recommended, see Figure A.17. In most cases, the medium choice of the tuning variant gave the best results.

6. Discussion

Previous work on anisotropy analysis of point processes has mainly concentrated on detecting anisotropy and finding preferred directions in regular and clustered point patterns, see Rajala et al. (2018). Here, our emphasis was on isotropy tests and we performed a thorough comparison of tests based on nearest neighbour or second order summary statistics, spectral theory or wavelets combined with different test statistics. In addition, we investigated how the parameters chosen by the user affect the performance of the tests.

This resulted in a large number of combinations of different summary statistics, test statistics, and tuning parameters. Figure 10 illustrates the best case performance of each summary over the simulation trials. In general, most of the combinations work well if the number of points is large and the pattern strongly anisotropic and clearly regular or clustered. More specifically, the pair-correlation and K function based tests work well for all point process models, especially when combined with the standardised mean squared deviation statistic $T_{ms_{st}}$. In the Strauss and Thomas processes, the optimal integration limit should always be chosen only slightly larger than the interaction/cluster range. Too large values tend to decrease the power. In contrast, large integration ranges are to be preferred in the models with linear clustering of points. The remaining tuning parameters, such as the cylinder width or the opening angle of the sector, do not have much effect in most cases. The Wong-Chiu test generally behaves similarly to the other second order methods combined with T_{mad}. An exception is the hump process with large range.

The local nearest neighbour distribution function G_{loc} works well for the Strauss process if the sample size is large, interaction strong and interaction range large, and for the linear Poisson line cluster process but not for the two other processes. Wavelet based tests perform well only for the hump process. Spectral methods work also well for the hump process and even for the Strauss process if the sample size is large, interaction strong and interaction range large.

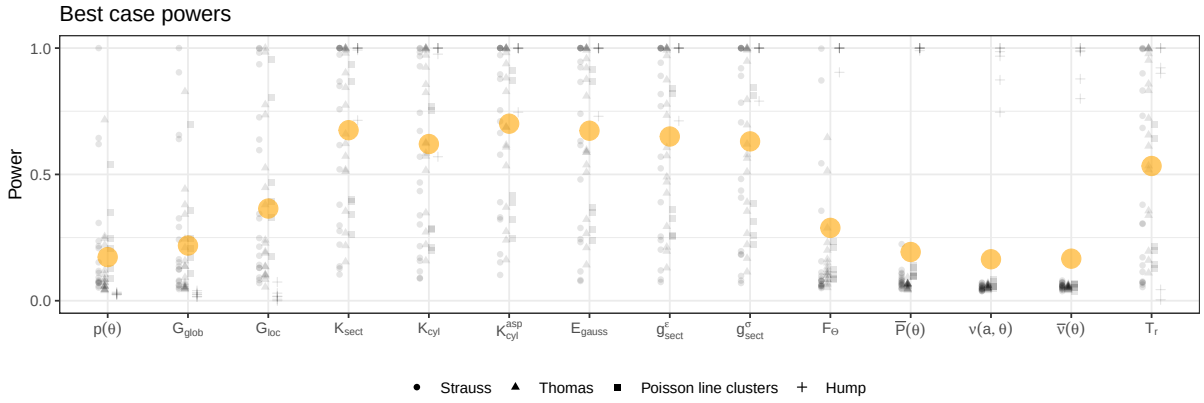

Figure 10: Overall summary of the simulation trials. For each anisotropy summary, the gray points, with shape by model, give the best power over tuning parameters, integration limits, and tests. The large dot is the mean over all powers.

In our study, we have assumed an idealised setting where both the directions of anisotropy and the null model are known and which, therefore, gives the maximum powers. In practice, directions of anisotropy can be estimated, e.g., by fitting ellipsoids to the Fry plots Rajala et al. (2016) or by the projection method introduced in Sormani et al. (2020).

7. Acknowledgement

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) in the framework of the priority programme "Antarctic Research with comparative investigations in Arctic ice areas" by grant RE 3002/3-1 and by grant VR 2018-03986 from the Swedish Research Council.

References

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

D'Ercole, R., & Mateu, J. (2013a). A continuous wavelet-based approach to detect anisotropic properties in spatial point processes. *International Journal of Wavelets, Multiresolution and Information Processing*, *11*, 1350017–1–1350017–29. arXiv:http://www.worldscientific.com/doi/pdf/10.1142/S0219691313500173.

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

D'Ercole, R., & Mateu, J. (2014). A wavelet-based approach to quantify the anisotropy degree of spatial random point configurations. *International Journal of Wavelets, Multiresolution and Information Processing*, *12*, 1450017--1--1450017--22. arXiv:http://www.worldscientific.com/doi/pdf/10.1142/S0219691314500374.

Mateu, J., & Nicolis, O. (2015). Multiresolution analysis of linearly oriented spatial point patterns. *Journal of Statistical Computation and Simulation*, *85*, 621--637. arXiv:http://dx.doi.org/10.1080/00949655.2013.838565.

Mateu, J., Nicolis, O., & D'Ercole, R. (2010). Testing for anisotropy in spatial point processes. In G.-M. et al. (Ed.), *Proceedings of the Fifth International Workshop on Spatio-Temporal Modelling*. Unidixital.

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

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

Rajala, T., Redenbach, C., Särkkä, A., & Sormani, M. (2018). A review on anisotropy analysis of spatial point patterns. *Spatial Statistics*, *28*, 141--168.

Rajala, T., Särkkä, A., Redenbach, C., & Sormani, M. (2016). Estimating geometric anisotropy in spatial point patterns. *Spatial Statistics*, *15*, 139--155.

Rajala, T. A., Olhede, S. C., & Murrell, D. J. (2022). What is the Fourier transform of a spatial point process? ArXiv Preprint, arXiv:2009.01474.

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

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

Sormani, M., Redenbach, C., Särkkä, A., & Rajala, T. (2020). Second order analysis of geometric anisotropic point processes revisited. *Spatial Statistics*, *38*, 100456. URL: http://www.sciencedirect.com/science/article/pii/S2211675320300506. doi:https://doi.org/10.1016/j.spasta.2020.100456.

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

Appendix A. Additional figures

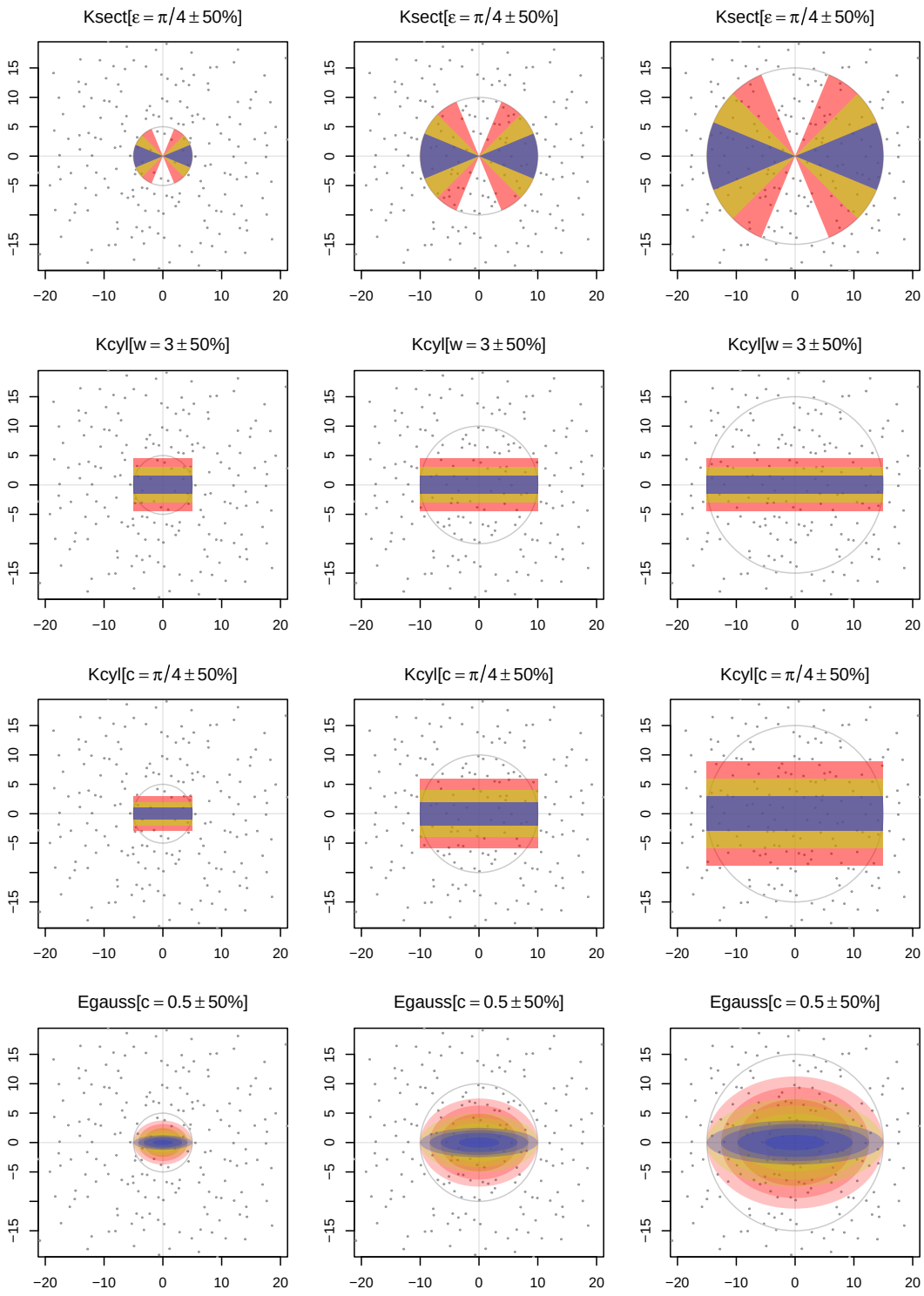

Figure A.11: Illustration of the directed templates used in K_{sect}, K_{cyl} (fixed width and fixed aspect ratio) and $\hat{E}_{\text{gauss}}$. Columns show different the cases $r = 5, 10$, and 15. The colours refer to the three choices of opening angle, cylinder width, and kernel ellipticity, respectively, see Table 1.

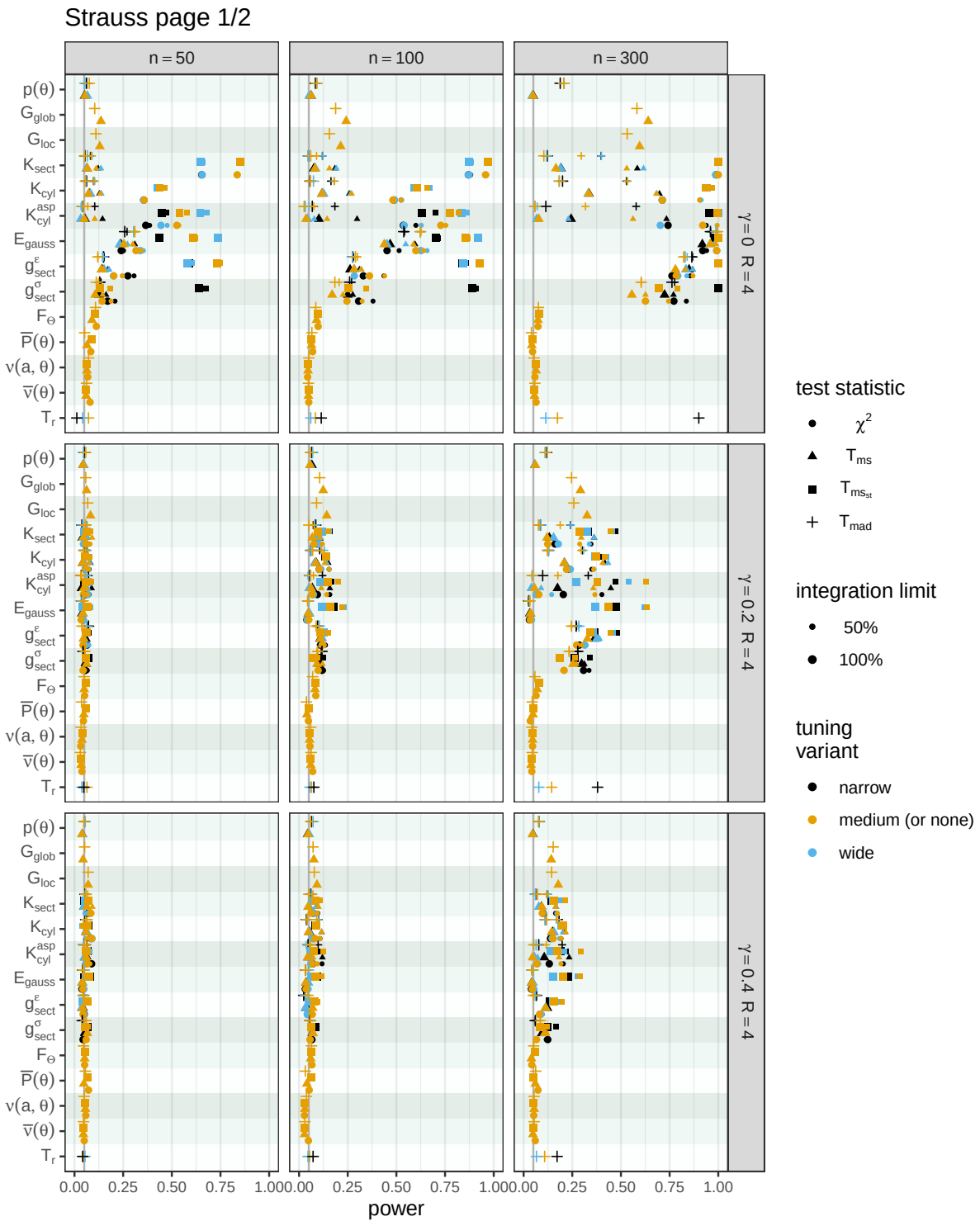

Figure A.12: Statistical powers in Strauss model trials. Included are all tuning and test statistic combinations, at two integration limit levels of 50% and 100%. Continues in Figure A.13.

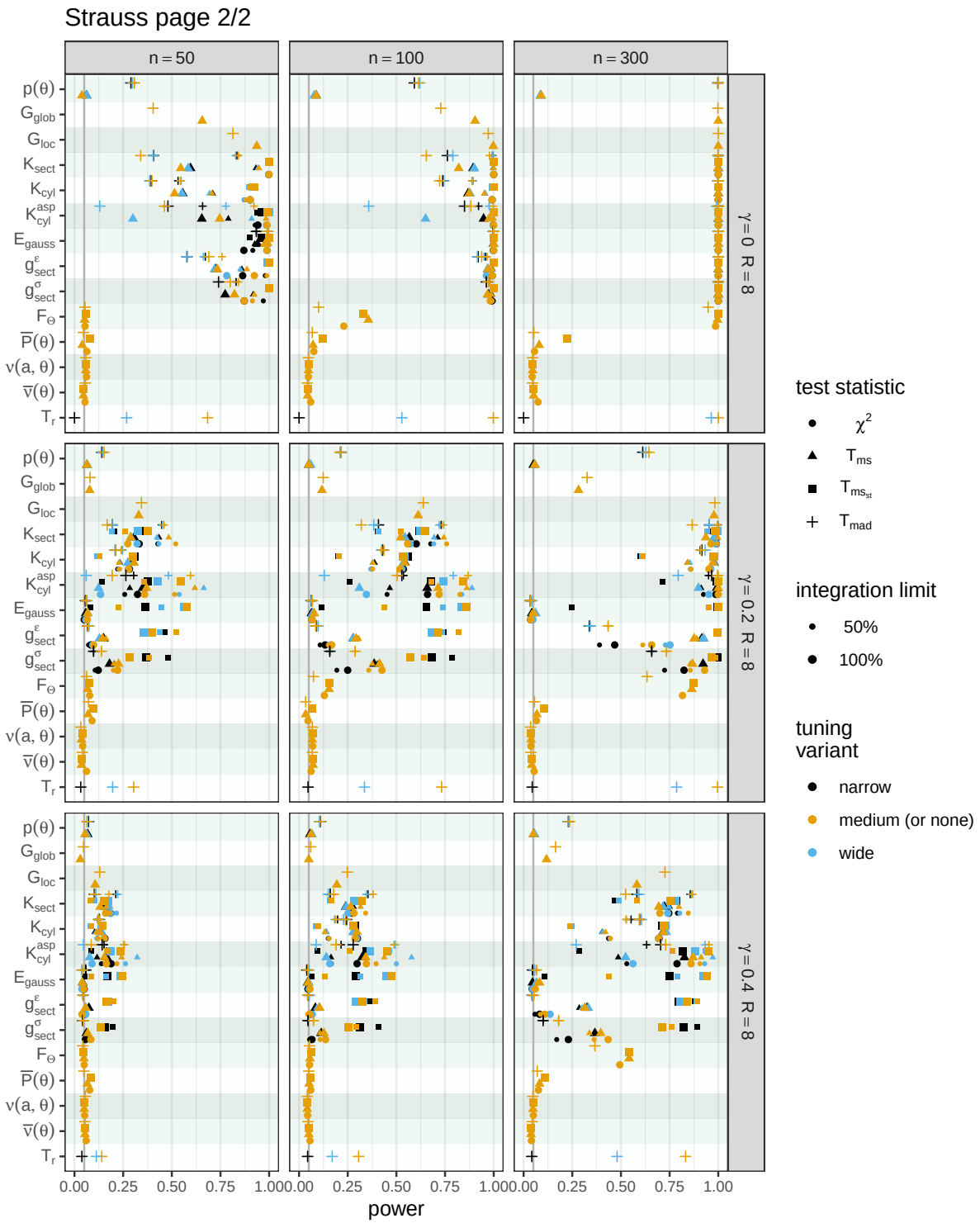

Figure A.13: Statistical powers in Strauss model trials. First part in Figure A.12.

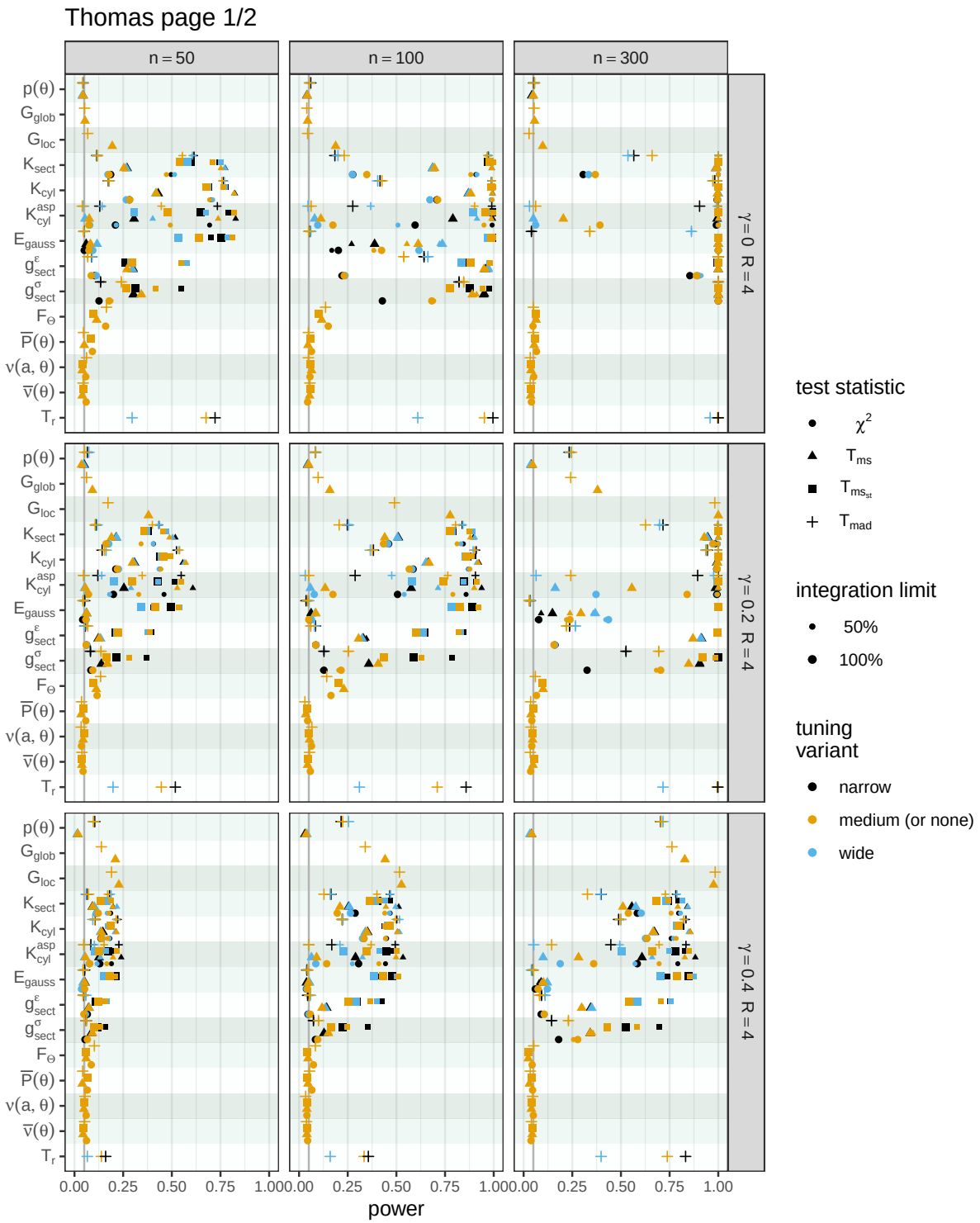

Figure A.14: Statistical powers in Thomas model trials. Continues in Figure A.15.

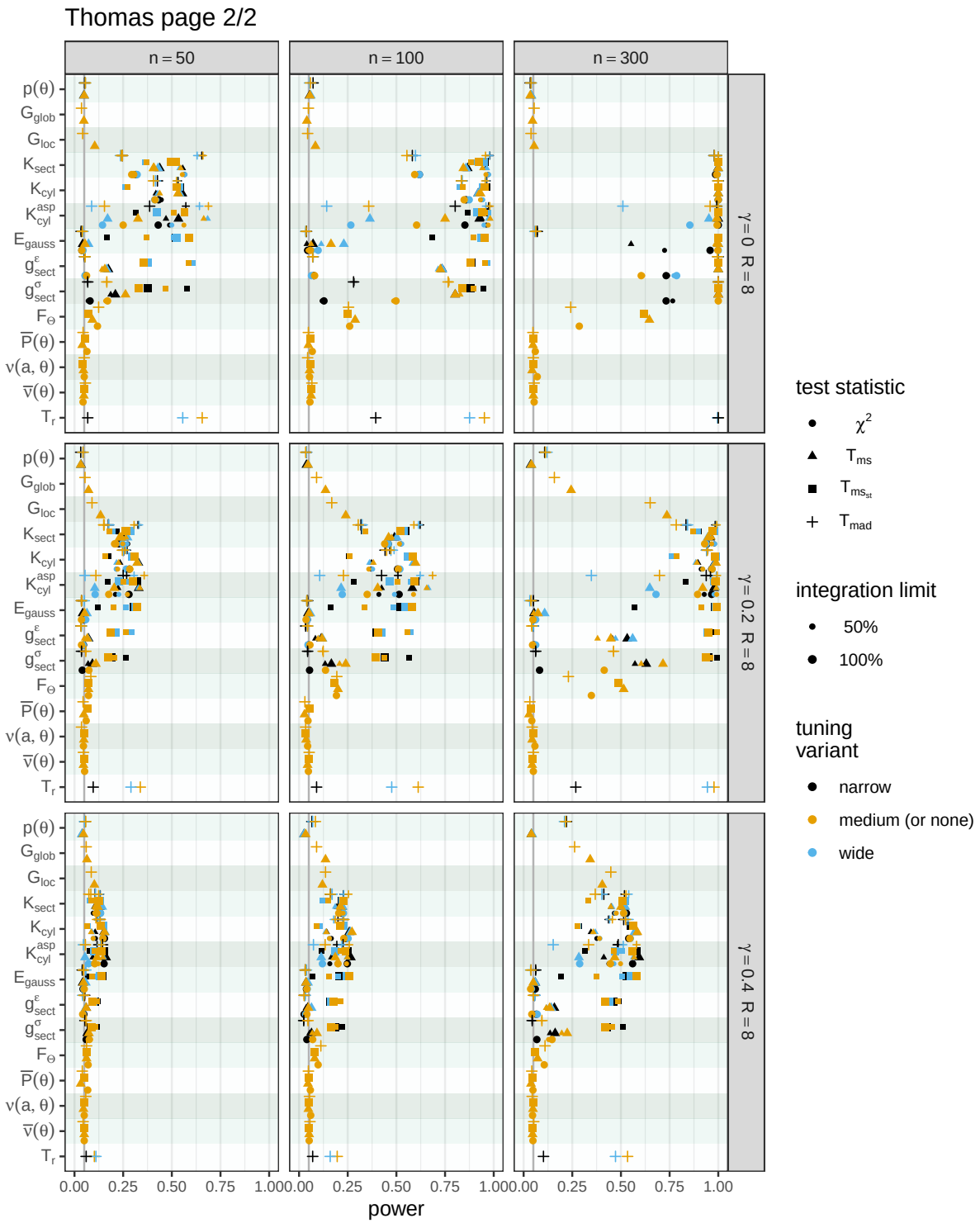

Figure A.15: Statistical powers in Thomas model trials. First part in Figure A.14.

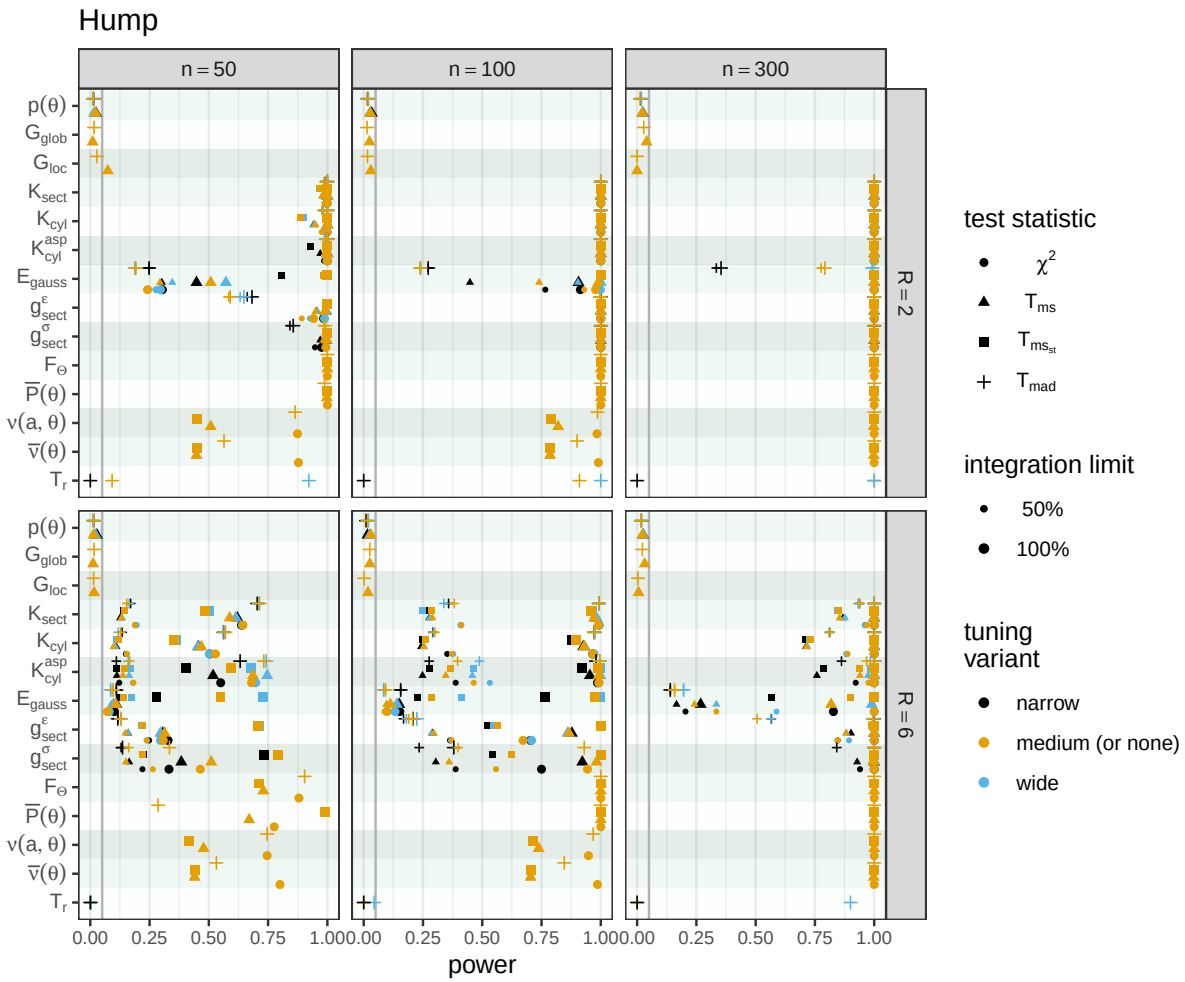

Figure A.16: Statistical powers in Hump model trials.

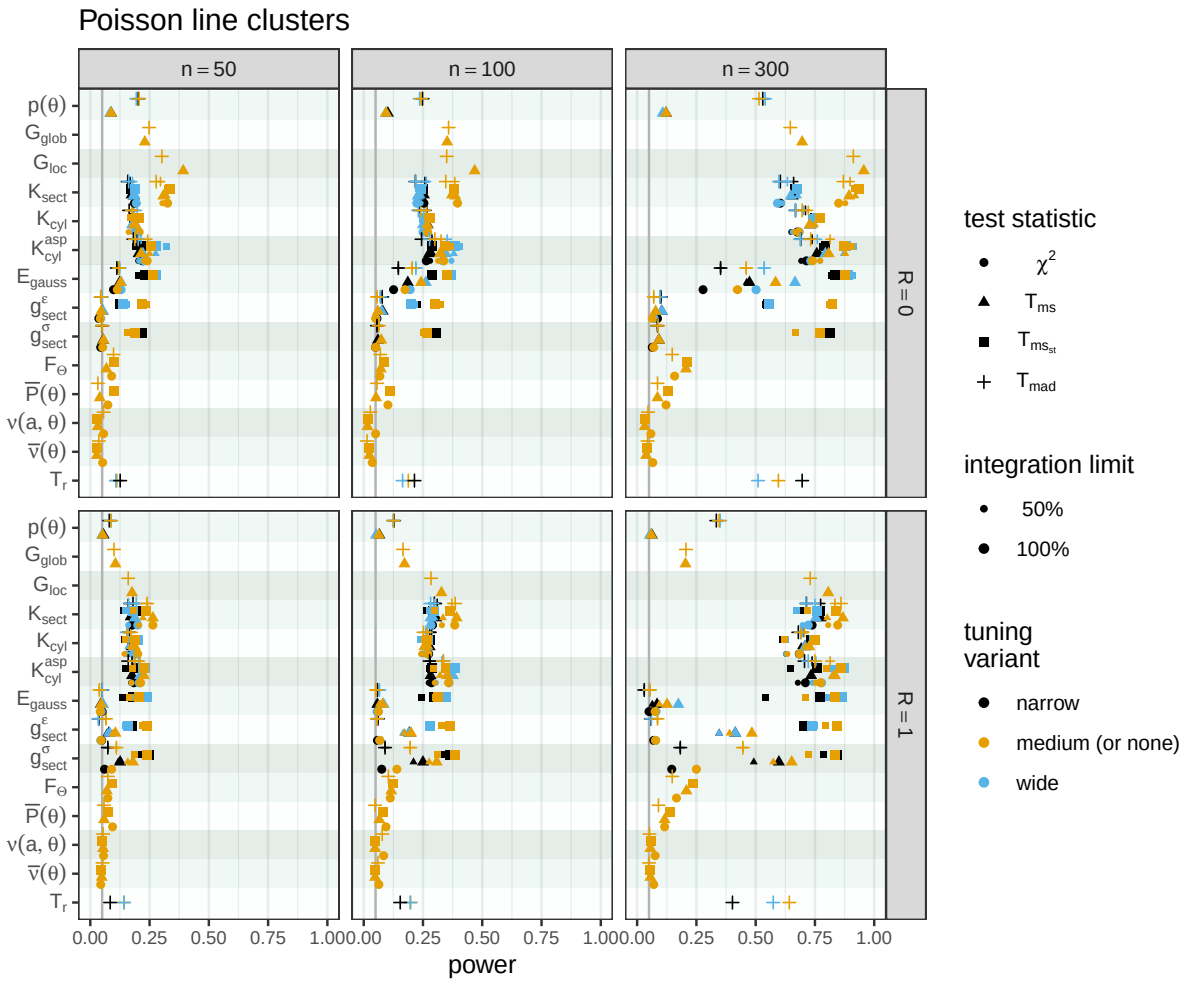

Figure A.17: Statistical powers in Poisson line clusters model trials.