ORIGINAL ARTICLE

Multivariate geometric anisotropic Cox processes

James S. Martin[1] | **David J. Murrell[2]** | **Sofia C. Olhede[3]**

[1]Department of Mathematics, Imperial College London, London, UK

[2]Centre for Biodiversity and Environment Research, Department of Genetics, Evolution and Environment, University College London, London, UK

[3]Institute of Mathematics, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

Correspondence
James S. Martin, Department of Mathematics, Imperial College London, London, UK.
Email: james.martin@imperial.ac.uk

Abstract

This paper introduces a new modeling and inference framework for multivariate and anisotropic point processes. Building on recent innovations in multivariate spatial statistics, we propose a new family of multivariate anisotropic random fields, and from them a family of anisotropic point processes. We give conditions that make the proposed models valid. We also propose a Palm likelihood-based inference method for this type of point process, circumventing issues of likelihood tractability. Finally we illustrate the utility of the proposed modeling framework by analyzing spatial ecological observations of plants and trees in the Barro Colorado Island data.

KEYWORDS

forest ecology, intractable likelihood, multivariate point processes

1 | INTRODUCTION

In this paper, we introduce a new class of multivariate and heterogeneous point process models. In doing so, we address two fundamental problems in modern spatial statistics: we design valid and nontrivial models for multivariate point processes, an important outstanding problem; and we contribute tools for performing inference with respect to anisotropic point processes. We produce multivariate spatial models that can flexibly accommodate anisotropy in both the marginal and joint dependence structures.

We choose to build our models using a foundation of log-Gaussian Cox processes (Diggle & Milne, 1983; Møller et al., 1998). In this framework, the observed point pattern is modeled in terms

of a latent random intensity, generated by a random field. We build on progress made in the area of multivariate random fields, see for example, Gneiting et al. (2010), Apanasovich et al. (2012), and Genton and Kleiber (2015), by allowing for anisotropy in the second-order dependence structure of the underlying random field model.

Datasets that require anisotropic models have been common in the point process literature for many years, for example, chapels in Welsh valleys (Mugglestone & Renshaw, 1996; Rajala et al., 2016; Rajala, Redenbach, et al., 2018), the epicentral locations of earthquakes in California over a 20-year period (Veen & Schoenberg, 2006) and clustered locations of shrubs in dryland ecosystems (Haase, 2001). The Welsh chapels and Californian earthquakes both form elliptical clusters, whereas the dryland shrub data displays a directional preference in the interaction of points of different type: Haase (2001) found one species to grow more often to the east of a second species. Both phenomena could be modeled using an anisotropic model for the second-order dependence structure. In the point process literature, it is common to accommodate heterogeneities in the observed point pattern by using a spatially homogeneous random field model to specify an intensity process conditional upon some known covariates (see, e.g., Diggle et al., 2013; Waagepetersen, 2008; Waagepetersen & Guan, 2009). This approach is limited in its applicability, however, when faced with heterogeneous point pattern data with no given covariate measurements.

The need to develop anisotropic methodology is further highlighted by the real data depicted in Figure 1c,e. These plots show the estimated intensity of two species of tree within the Barro Colorado Island (BCI) tropical rainforest stand (Condit, 1998; Hubbell et al., 1999, 2010). The strongly anisotropic features of these species are clearly demonstrated in the elliptical regions of high estimated intensity. Inspired by this example, our chosen approach to summarizing anisotropy is based on the concept of geometric anisotropy from random field analysis (Goff & Jordan, 1988). Isotropic processes have spatial covariance functions with circular contours of equivariance, whereas those that drive geometric anisotropic processes have elliptical contours of equivariance. Geometric anisotropy in single-point processes was studied by Møller and Toftaker (2014), Rajala et al. (2016), Rajala, Murrell, and Olhede (2018), and Sormani et al. (2020). The aforementioned authors use integral transformations and the formalism of Fry points to characterise geometric anisotropy. An advantage of using geometric anisotropy to model anisotropy is that it can be used in conjunction with well-known covariance functions. There are few alternatives to geometric anisotropy: separable variation, which places a considerable emphasis on the axes of observation, and general anisotropy that is hard to both model, and to infer. We therefore restrict our attention to geometric anisotropy, and we briefly demonstrate the potential utility of geometric anisotropic point processes by plotting two simulated (independent) point patterns in Figure 1b; the simulated intensities for each of these point patterns are shown in Figure 1d,f, where we can clearly see elliptical regions of high intensity, partly replicating the behavior of the real data alongside.

We explore the restrictions that are naturally placed on all cross-pair correlation functions from the proposed form of geometric anisotropy, where we use recent results for isotropic multivariate random fields due to Apanasovich et al. (2012) and Gneiting et al. (2010) to understand the newly introduced anisotropic processes. By representing our multivariate process in both the spatial and spectral domains, we will also demonstrate that allowing for distinct geometric anisotropies in each marginal process may place further restrictions on valid forms of the cross-dependence structures, adding to existing isotropic understanding. This is an important result that yields unique insights into the possible variation of joint co-dependence in multivariate geometric anisotropic random fields, and by extension insights into the intrinsic properties of Cox processes.
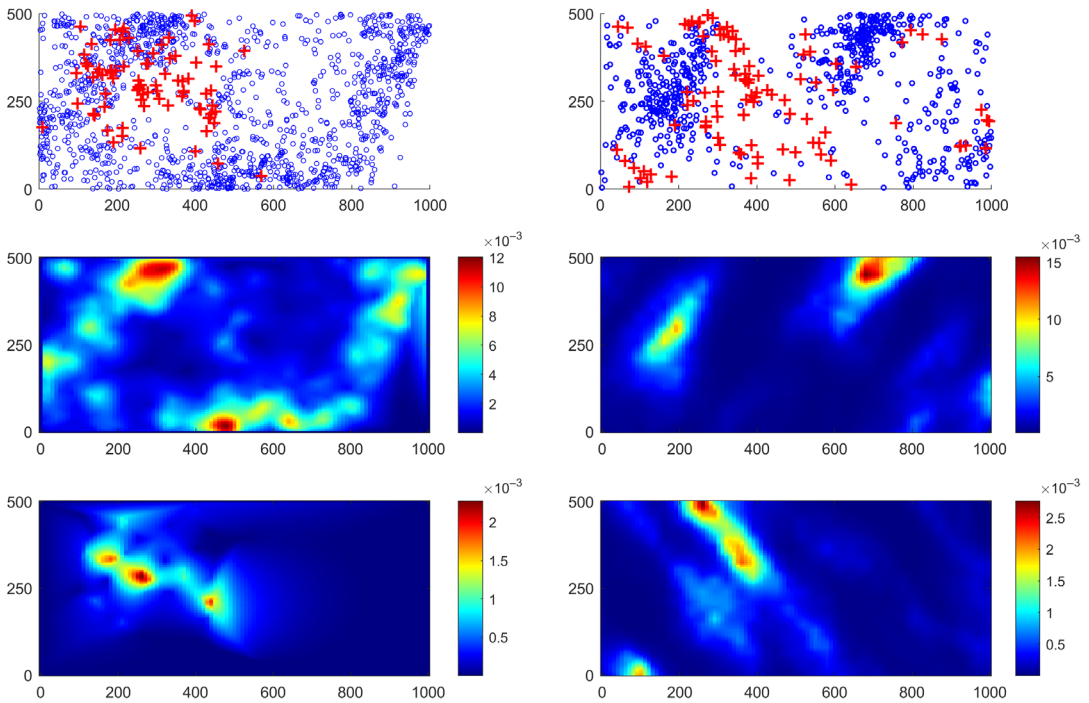
FIGURE 1 Point pattern data showing two species of tree from the Barro Colorado Island tropical rainforest (a; *Guatteria dumetorum*, blue circles; *Miconia hondurensis*, red crosses), along with their estimated intensity fields (c,e), and simulated point pattern data (b) from two independent univariate geometric anisotropic log-Gaussian Cox processes, with their corresponding simulated intensity fields (d,f)

Once we have understood the constraints on possible model forms, we turn to inference. Our problem is challenged by the presence of an intractable likelihood, an important topic of current interest (Bortot et al., 2009; Li & Fearnhead, 2018; Price et al., 2018; Sisson et al., 2007). We detail a two-stage estimation procedure in which we first estimate the anisotropy parameters, and then use these estimates to transform the data to be isotropic. The "isotropised" point pattern is then used to estimate the covariance parameters for the underlying random field model. For this second stage, Møller and Toftaker (2014) advocated the use of minimum contrast. We instead appeal to Cox's likelihood principle (Cox, 1958), and develop a maximum Palm likelihood-based approach to inference that builds on the work of Tanaka et al. (2008).

Traditional maximum likelihood estimates of the model parameters are not available, as the point process likelihood is intractable. Tanaka et al. (2008) showed that this intractability can be circumvented by considering the so-called Fry process (Fry, 1979). This is a secondary point pattern formed by the difference vectors of all point pairs in the original point pattern, and it can be treated as an inhomogeneous Poisson point process, with an associated (and tractable) likelihood. Tanaka et al. (2008) showed that the Fry process likelihood can be used to perform inference for univariate, isotropic point process models. In a novel extension of this work, we overcome the challenge of anisotropy and use the Fry process likelihood to perform inference for anisotropic, multivariate point processes.

Finally, we apply our newly developed methodology to real data from the BCI rainforest dataset, introduced above. Recent work by Waagepetersen et al. (2016) and Rajala, Murrell,

and Olhede (2018) has highlighted the importance of developing realistic multivariate point process models to aid the understanding of complex species interactions within this rainforest.

Thus, this paper gives a number of novel important insights for multivariate spatial processes. In particular, the paper describes the complex relationship between distinct geometric anisotropies in each univariate component, and the full multivariate model. Our understanding gives sufficient, but not necessary, conditions to yield valid multivariate random field models and, by extension, valid multivariate Cox processes. This addresses an important outstanding problem in spatial analysis. Additionally by extending intractable likelihood methods to this setting, we provide insights into the outstanding problem of inferring dependent point patterns, part of understanding spatial processes.

2 | NOTATION AND SETTING

2.1 | Log-Gaussian Cox processes

To establish notation, let $X = \{X_p; p = 1, \ldots, P\}$ denote a multivariate point process, where the index p is used to signal a univariate component of the multivariate process X. We consider each univariate process X_p to be a locally finite random subset of $\mathbb{R}^d$; that is, for the random variable $N_p(B) = |X_p \cap B|$ representing the number of points in X_p that lie in the region $B \subset \mathbb{R}^d$, $N_p(B) < \infty$ whenever B is bounded. We will be predominantly interested in planar point processes, where $d = 2$, though the work contained here can be extended to $d > 2$. For a bounded observation window $W \subset \mathbb{R}^2$, the observed point pattern corresponds to a realization x of $X \cap W$, where $x = \{x_{p,i} \in W; i = 1, \ldots, n_p, p = 1, \ldots, P\}$ and where $n_p \in \mathbb{N}$ is the realized value of $N_p(W)$. Henceforth, we will also use x_p to denote an arbitrary observed point pattern of type p.

We will primarily use the pair correlation function to describe the second-order properties of our multivariate point process X. For distinct locations $u, v \in \mathbb{R}^2$, the (cross-)pair correlation function $g_{pq}(u, v)$ of X_p with X_q is defined as $g_{pq}(u, v) = \rho_{pq}^{(2)}(u, v)/\rho_p(u)\rho_q(v)$ if $\rho_p(u)\rho_q(v) > 0$ and $g_{pq}(u, v) = 0$ otherwise, where $\rho_{pq}^{(2)}(u, v)$ is the second-order cross-intensity function between components p and q, and $\rho_p(u)$, $\rho_q(v)$ are the corresponding univariate (first order) intensity functions. For distinct locations $u, v \in \mathbb{R}^2$, we can interpret the pair correlation function as the probability of observing one point from process p in the infinitesimal region containing u, and one point from process q in the infinitesimal region containing v, divided by the corresponding probability under the assumption that X_p and X_q are independent Poisson processes.

Following Møller et al. (1998) we define X to be a multivariate log-Gaussian Cox process (LGCP), driven by a multivariate intensity *process* $\Lambda = \{\Lambda_p; p = 1, \ldots, P\}$; each subprocess X_p, conditional on $\Lambda_p(u)$ is a Poisson process with intensity at location $u \in \mathbb{R}^2$ given by a realization of the random variable

$$\Lambda_p(u) = \exp\{S_p(u)\}, \qquad u \in \mathbb{R}^2, \tag{1}$$

where $S(u) = \{S_p(u), p = 1, \ldots, P\}$ is a multivariate Gaussian random field (GRF). We will assume S_p, and therefore X_p, to be stationary for all $p = 1, \ldots, P$, with constant mean μ_p and variance σ_{pp}. The intensity process $\Lambda_p(u)$ will therefore also have a constant mean, which we denote $\lambda_p = \mathbb{E}\{\Lambda_p\} = \exp\{\mu_p + \sigma_{pp}/2\}$.

For a stationary multivariate point process, we can define the pair correlation function using a single argument, representing the vector distance between the two observation locations. We write $g_{pq}(h)$ for $h = v - u \in \mathbb{R}^2$, and for a stationary log-Gaussian Cox process we have that $g_{pq}(h) = \exp\{C_{pq}(h)\}$, where

$$C_{pq}(h) = cov\left\{S_p(v), S_q(u)\right\}, \qquad p, q = 1, \ldots, P, \quad u, v, h \in \mathbb{R}^2,$$

are the covariance functions for the multivariate GRF S. From this relationship, it is clear to see that $g_{pq}(h) = 1$ (no aggregation or segregation) is equivalent to $C_{pq}(h) = 0$ (no correlation), which indicates independence between processes p and q at any scale $h \in \mathbb{R}^2$. Thus, for a bivariate Poisson process $\{X_p, X_q\}$, that is, under an assumption of complete spatial randomness, we would expect $g_{pq}(h) = 1$, whereas significant departures from this would indicate either aggregation ($g_{pq}(h) > 1$) or segregation ($g_{pq}(h) < 1$) of points from processes p and q at separation $h \in \mathbb{R}^2$.

The second-order behavior of the multivariate GRF S can be equivalently described in both the spatial and spectral domains. We will make use of the (cross-)spectral density function $f_{pq}(\omega)$ of S_p with S_q, at frequency $\omega \in \mathbb{R}^2$, which forms a Fourier transform pair with the (cross-)covariance function $C_{pq}(h)$, $h \in \mathbb{R}^2$; see, for example, Sherman (2011). To describe between-component dependency we will also consider the complex coherence, $\gamma_{pq}(\omega)$: $\gamma_{pq}(\omega) = f_{pq}(\omega) / \left\{f_{pp}(\omega) f_{qq}(\omega)\right\}^{\frac{1}{2}}$. It is worth noting that the methodology presented in this paper does not require consideration of the point spectrum of Bartlett (1964), that is the spectral density of the point process X; all of our frequency-domain analysis concerns the underlying random field S.

We will also make use of the Palm intensity for the multivariate point process. For a given location $v \in \mathbb{R}^2$, the so-called reduced Palm distribution of a univariate point process X_p is the conditional distribution of the process $X_p \setminus \{v\}$, given the observation of a point at v. Without loss of generality, we can set v to be the spatial origin o, and heuristically define the corresponding *Palm intensity*, $\lambda_{o,p}$, as follows. For $u \in \mathbb{R}^2$ at a nonzero distance from the spatial origin o, the conditional probability of observing a point in the infinitesimal region containing u, given the observation of a point at the origin, is specified by $\lambda_{o,p}(u) du$, where du is the Lebesgue measure for the infinitesimal region containing u. A full, measure-theoretic treatment of Palm distributions for spatial point processes is given by Coeurjolly et al. (2017).

2.2 | Geometric anisotropic LGCPs

We construct a geometric anisotropic LGCP through the covariance structure of the underlying GRF. Given an isotropic covariance function $C_0(\|h\|)$, $h \in \mathbb{R}^2$, define a geometric anisotropic version as (Møller & Toftaker, 2014):

$$C(h) = C_0\left(\sqrt{h^T \Sigma^{-1} h}\right), \; h \in \mathbb{R}^2, \tag{2}$$

where $\Sigma = R_\theta \text{diag}(1, \zeta^2) R_\theta^T$, for $\theta \in [0, \pi)$ and $\zeta \in (0, 1]$, and where R_θ is the rotation matrix, with the restricted ranges in place to ensure identifiability. Geometric anisotropy in higher-dimensional Euclidean spaces can be defined through a similar construction, using a rotation matrix and diagonal scaling matrix of corresponding dimensions; see, for example, Rajala et al. (2016). For Σ defined as above, we also consider the following "square root" matrices: $\Sigma^{1/2} = \text{diag}(1, \zeta) R_\theta^T$, and $\Sigma^{-1/2} = \text{diag}(1, 1/\zeta) R_\theta^T$, such that $\Sigma = \left(\Sigma^{1/2}\right)^T \Sigma^{1/2}$, $\Sigma^{-1} = \left(\Sigma^{-1/2}\right)^T \Sigma^{-1/2}$ and $\left(\Sigma^{1/2}\right)^{-1} = \left(\Sigma^{-1/2}\right)^T$.

Under this parameterization, Σ is defined such that the ellipse $E = \{h \in \mathbb{R}^2 : h^T \Sigma^{-1} h = 1\}$ has a semi-major axis of unit length at angle θ, relative to the abscissa axis of the original coordinate system, and a semi-minor axis of length ζ at angle $\theta + \pi/2$. Accordingly, we can describe the covariance function defined in (2) as "elliptic," and the resulting LGCP will also display elliptic second-order behavior.

The corresponding pair correlation function and spectral density function will take the following forms:

$$g(h) = g_0 \left(\sqrt{h^T \Sigma^{-1} h} \right) = \exp \left\{ C_0 \left(\sqrt{h^T \Sigma^{-1} h} \right) \right\}$$
$$f(\omega) = |\Sigma|^{1/2} f_0 \left(\sqrt{\omega^T \Sigma \omega} \right),$$

for $h, \omega \in \mathbb{R}^2$, where $f_0(\|\omega\|)$ is the isotropic spectral density that forms a Fourier transform pair with $C_0(\|h\|)$, and $g_0(\|h\|)$ is the corresponding isotropic pair correlation function.

Our specification of geometric anisotropy differs slightly from that of Møller and Toftaker (2014), who include an additional scale parameter in their definition of the deformation matrix Σ; this is used to scale the axes in the resulting elliptical covariance structure. In practice, however, the majority of parametric covariance functions of interest incorporate a scale parameter that directly controls the correlation length, and so including a separate scale parameter in the definition of Σ creates nonidentifiability issues when performing parameter inference. We assume all scale information to be controlled by the parametric form of $C_0(\|h\|)$.

Since we are considering processes that display anisotropy, it will be useful for their analysis to be able to express their second-order properties in polar coordinates. We therefore define the anisotropic pair correlation function, replacing the vector $h \in \mathbb{R}^2$ with its length r and angle ϕ:

$$g^a(r, \phi) = g([r \cos \phi, r \sin \phi]) = g_0 \left(\frac{r}{\zeta} \sqrt{1 - (1 - \zeta^2) \cos^2(\phi - \theta)} \right). \tag{3}$$

3 | DEFINING THE MODEL

3.1 | Modeling multivariate geometric anisotropy

For a population of P LGCPs, we specify the multivariate dependence through the covariance structure of the P-variate GRF that drives the P conditionally independent intensity processes. We define the following family of geometric anisotropic auto- and cross-covariance functions:

$$C_{pq}(h) = C_{0,pq} \left(\sqrt{h^T \Sigma_{pq}^{-1} h} \right) = C_{0,pq} \left(\left\| \Sigma_{pq}^{-1/2} h \right\| \right), \quad p, q = 1, \ldots, P, h \in \mathbb{R}^2,$$

for some choice of isotropic covariance functions $\{C_{0,pq}; p, q = 1, \ldots, P\}$, and for a collection of deformation matrices $\{\Sigma_{pq}; p, q = 1, \ldots, P\}$, where Σ_{pq} is defined in terms of the parameter pair $(\theta_{pq}, \zeta_{pq})$ according to the specification of Σ that follows (2).

This framework will allow for the possibility of distinct geometric anisotropies in each of the marginal processes. This may be a valid modeling requirement; for ecological datasets for instance, there can be multiple anisotropic effects influencing the data. In this framework, we can reproduce bivariate point patterns in which each component displays elliptical clustering at

different orientations, or with differing degrees of ellipticity. Just as in the isotropic case, however, care must be taken in specifying the parameters for the cross-covariance functions C_{pq}, in order to ensure a valid multivariate model.

The matrix of covariance functions $(C_{pq}(h))_{p,q=1}^P$ must be nonnegative definite for all $h \in \mathbb{R}^2$. The equivalent requirement in the spectral domain is that the matrix of spectral densities $(f_{pq}(\omega))_{p,q=1}^P$ is nonnegative definite for all $\omega \in \mathbb{R}^2$. For the bivariate dependence structure between two processes X_p and X_q, this is equivalent to the magnitude squared coherence $|\gamma_{pq}(\omega)|^2$ being bounded above by 1; this corresponds to

$$0 \leq |f_{pq}(\omega)| \leq \left\{ f_{pp}(\omega) f_{qq}(\omega) \right\}^{\frac{1}{2}}, \qquad \omega \in \mathbb{R}^2. \tag{4}$$

This upper bound is displayed in Figure 2 for a bivariate process with distinct marginal geometric anisotropies. By considering the behavior of (4) over the two-dimensional Fourier domain, we can now make some general comments about the level of dependence between components in a bivariate geometric anisotropic LGCP. This discussion will be under the assumption that each autospectrum and cross-spectrum in the bivariate process is decreasing for increasing frequencies ω.

The inequality in (4) implies that, for any two processes, between-process dependence can only be nonnegligible at those frequencies that contribute significantly to the marginal dependence in *both* processes. For two processes with distinct marginal geometric anisotropies, this restriction impacts the high-frequency behavior of the bivariate process more than the low-frequency behaviour. Recall Figure 2: when constructing the upper bound for the cross-spectrum according to (4), the high-frequency contributions of each of the autospectra are killed by the negligible power at the same frequency in the other autospectrum; the contrasting behavior of the marginal processes at high frequencies kills any high-frequency dependence between the processes. As a result, for any two processes that display distinct anisotropic behavior, significant between-process dependence will be more evident at low frequencies, or large spatial scales.

Due to our modeling assumption of geometric anisotropy in the cross-dependence structure, the cross-spectrum will have elliptical contours of equal power density. Figure 2 illustrates that the elliptical geometries of the autospectra can dictate a nontrivial geometric structure for the upper bound of the cross spectrum. For any given pair of marginal spectra, and thus a given upper bound to the corresponding cross-spectrum, the ellipticity of the true cross-spectrum will impact its permissible coverage of the frequency space, as its elliptical structure must fit within
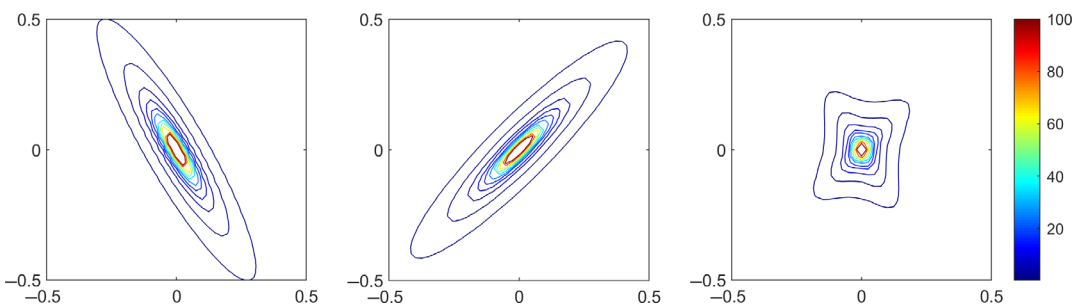

FIGURE 2 Geometric anisotropic autospectra (left and center) for a bivariate Gaussian random field, along with the upper bound on the corresponding cross-spectrum (right), as given in Equation (4)

the upper bound's nontrivial geometry. Indeed, as illustrated by Figure 2, in order for our elliptical cross-spectrum to extend further into the higher-frequency regions of the Fourier space, the ellipticity of the cross-spectrum must be more pronounced. Since the overall power in the cross-process dependence is obtained by integrating the cross-spectrum over the entire Fourier domain, this gives us a link between the power and the degree of anisotropy in the cross-process dependence.

3.2 | Multivariate Matérn correlation structure

The Matérn family of correlation functions (Guttorp & Gneiting, 2006; Stein, 1999) is a popular spatial modeling choice. For univariate random fields, one can use a single three-parameter covariance function to replicate dependence structures that act over any positive scale, whilst additionally controlling the smoothness of any realisations. The flexibility of this model has made it the preferred modeling choice for univariate processes in the spatial statistics literature. Gneiting et al. (2010), Apanasovich et al. (2012), and Kleiber and Nychka (2012) have recently extended this model to the multivariate case, proposing the use of a Matérn function to describe all auto- and cross-covariances for a multivariate isotropic stationary random field. We will incorporate elements from these approaches in our modeling framework, though we retain an assumption of stationarity unlike Kleiber and Nychka (2012).

We develop the stationary Matérn covariance structure for the multivariate GRF $S(u)$, ensuring that all auto- and cross-covariances are Matérn. The multivariate Matérn model was introduced by Gneiting et al. (2010) and extended by Apanasovich et al. (2012). The former established necessary and sufficient conditions for the validity of the bivariate model, and sufficient conditions for the validity of a restricted subclass of the multivariate ($P \geq 3$) model. We build on this, incorporating anisotropy when observing $P \geq 1$ processes. The results presented in this section are valid for all d-dimensional Matérn Gaussian processes, with $d \geq 2$, unless stated otherwise.

Following Gneiting et al. (2010), the isotropic multivariate Matérn covariance function is defined as:

$$C_{0,pq}(\|h\|; \alpha_{pq}, \nu_{pq}, \sigma_{pq}) = \frac{\sigma_{pq}}{2^{\nu_{pq}-1}\Gamma\left(\nu_{pq}\right)} \left(\frac{2\sqrt{\nu_{pq}}}{\alpha_{pq}}\|h\|\right)^{\nu_{pq}} \mathcal{K}_{\nu_{pq}}\left(\frac{2\sqrt{\nu_{pq}}}{\alpha_{pq}}\|h\|\right), \quad h \in \mathbb{R}^d, \quad (5)$$

where $\mathcal{K}_\nu(\cdot)$ is the modified Bessel function of the second kind (Abramowitz & Stegun, 1965, pp. 374–379). Here, $\sigma_{pq} \in \mathbb{R}$ ($\sigma_{pp} > 0$) is the zero-lag covariance between field components S_p and S_q, and $\alpha_{pq} > 0$ and $\nu_{pq} > 0$ are scale and smoothness parameters. The latter two parameters control the rate of decay of covariance between S_p and S_q with respect to distance. The form of the Matérn parameterization used in (5) differs from that used by Gneiting et al. (2010) and Apanasovich et al. (2012). We choose to use the Matérn parameterization introduced by Handcock and Wallis (1994), as it allows maximal separation of the roles of α_{pq} and ν_{pq} in determining the second-order behavior of $S(u)$ and, ultimately, the resulting point process X.

Having established the Matérn form of the auto- and cross-covariances for a multivariate isotropic GRF, we now generalize to allow for anisotropic multivariate covariance structures.

Recall from Section 3.1 that we obtain our geometric anisotropic (cross-)covariance function by applying the deformation matrix Σ_{pq}. We therefore define, for any $h \in \mathbb{R}^d$, $d \geq 2$,

$$C_{pq}(h; \alpha_{pq}, \nu_{pq}, \sigma_{pq}, \Sigma_{pq}) = \frac{\sigma_{pq}}{2^{\nu_{pq}-1}\Gamma\left(\nu_{pq}\right)} \left(\frac{2\sqrt{\nu_{pq}}}{\alpha_{pq}} \left\| \Sigma_{pq}^{-1/2} h \right\| \right)^{\nu_{pq}} \mathcal{K}_{\nu_{pq}} \left(\frac{2\sqrt{\nu_{pq}}}{\alpha_{pq}} \left\| \Sigma_{pq}^{-1/2} h \right\| \right). \quad (6)$$

Recall that we require the matrix $\left(C_{pq}(h; \alpha_{pq}, \nu_{pq}, \sigma_{pq}, \Sigma_{pq}) \right)_{p,q=1}^{P}$ to be nonnegative definite for all $h \in \mathbb{R}^d$, for validity of (6). Satisfaction of this requirement can be guaranteed by placing the following conditions on the cross-covariance parameters $\{\alpha_{pq}, \nu_{pq}, \sigma_{pq}, \theta_{pq}, \zeta_{pq}, p \neq q\}$, where we say that the real matrix $M \in \mathbb{R}^{P \times P}$ is *conditionally* nonnegative definite if and only if $y^T M y \geq 0$ for all $y \in \mathbb{R}^P$ such that $\sum_{p=1}^{P} y_p = 0$. Note that this is a weaker assumption than that of nonnegative definiteness, and it may be satisfied by a matrix containing only negative elements.

> **Condition 1.** There exists a nonnegative constant Δ_ν such that $\nu_{pq} - (\nu_{pp} + \nu_{qq})/2 = \Delta_\nu(1 - A_{\nu,pq})$, $p, q = 1,\ldots,P$, where A_ν is a valid $P \times P$ correlation matrix, with entries $0 \leq A_{\nu,pq} \leq 1$.

> **Condition 2.** The matrix with elements $-4\nu_{pq}/\alpha_{pq}^2$, $p, q = 1,\ldots,P$, is conditionally nonnegative definite.

> **Condition 3.** The matrix with elements
>
> $$\frac{|\Sigma_{pq}|^{1/2} \sigma_{pq} \Gamma(\nu_{pq} + d/2)}{\pi^{d/2} \Gamma\left(\frac{\nu_{pp}+\nu_{qq}}{2} + \frac{d}{2} \right) \Gamma(\nu_{pq})} \left(\frac{4\nu_{pq}}{\alpha_{pq}^2} \right)^{\Delta_\nu + \frac{\nu_{pp}+\nu_{qq}}{2}}, \qquad p, q = 1,\ldots,P,$$
>
> is nonnegative definite.

> **Condition 4.** The matrix with elements $-\|\Sigma_{pq}^{1/2}\omega\|^2$, $p, q = 1,\ldots,P$, is conditionally nonnegative definite for any $\omega \in \mathbb{R}^d$.

> **Proposition 1.** *For $p, q = 1,\ldots,P$, let $\alpha_{pq} > 0$, $\nu_{pq} > 0$, $\sigma_{pq} \in \mathbb{R}$, $\theta \in [0, \pi)$ and $\zeta \in (0, 1]$. Then the multivariate geometric anisotropic Matérn function (6) specifies a valid multivariate covariance model if Conditions 1–4 are met.*

The proof of Proposition 1 is given in the Appendix, and follows a similar argument to the proof of Theorem 1 of Apanasovich et al. (2012).

> *Remark* 1. If Condition 4 holds, then for $d = 2$, the $P \times P$ matrix with (p, q)-element $|\Sigma_{pq}|^{-1/2} = \zeta_{pq}^{-1}$, will be nonnegative definite; in particular, $\zeta_{pq}^2 \geq \zeta_{pp}\zeta_{qq}$, for all $p, q = 1,\ldots,P$.

Conditions 1–4 are similar in spirit to those placed by Apanasovich et al. (2012) on the Matérn parameters. In the simpler isotropic framework, the three conditions specified by Apanasovich et al. (2012) are sufficient to guarantee nonnegative definiteness of the resulting spectral density, and also to guarantee that all absolute zero-lag cross-correlations are bounded above by one. In the more general anisotropic setting, we (unsurprisingly) require a more extensive specification. Conditions 1–4, above, are sufficient to guarantee nonnegative definiteness of the geometric anisotropic spectral density, and are also sufficient for the absolute colocated cross-correlations to be bounded above by 1. These conditions constitute a set of implicit relationships that, between

them, specify a valid multivariate geometric anisotropic LGCP. We shall adopt a sequential approach to model construction, that will also reflect our model fitting procedures in Section 4.

Trivial rearrangement of Condition 1 yields an explicit expression for ν_{pq} in terms of the corresponding marginal values. In Remarks 2–5, below, we provide similar constructions for the cross-covariance parameters α_{pq}, σ_{pq}, θ_{pq}, and ζ_{pq}, such that Conditions 1–4 may be satisfied. The proofs for Remarks 2–4 are given in the Appendix.

Remark 2. Condition 2 is satisfied by the parameters $\{\nu_{pq}, \alpha_{pq}; p, q = 1, \ldots, P\}$ if

$$\frac{4\nu_{pq}}{\alpha_{pq}^2} = \frac{1}{2} \left(\frac{4\nu_{pp}}{\alpha_{pp}^2} + \frac{4\nu_{qq}}{\alpha_{qq}^2} \right) + \Delta_\alpha \left(1 - A_{\alpha,pq} \right),$$

for some constant $\Delta_\alpha \geq 0$ and for some $0 \leq A_{\alpha,pq} \leq 1$ that form a valid correlation matrix. This remark is also made by Apanasovich et al. (2012) in their chosen Matérn parameterization.

Remark 3. Condition 3 is satisfied by the parameters $\{\nu_{pq}, \alpha_{pq}, \zeta_{pq}, \sigma_{pq}; p, q = 1, \ldots, P\}$ if

$$\sigma_{pq} = \frac{\pi^{d/2} V_p V_q A_{\sigma,pq}}{\zeta_{pq}} \left(\frac{4\nu_{pq}}{\alpha_{pq}^2} \right)^{-\Delta_\nu - \frac{\nu_{pp} + \nu_{qq}}{2}} \frac{\Gamma \left(\frac{\nu_{pp} + \nu_{qq}}{2} + \frac{d}{2} \right) \Gamma(\nu_{pq})}{\Gamma \left(\nu_{pq} + \frac{d}{2} \right)} \qquad p, q = 1, \ldots, P,$$

for constants $V_p, V_q \geq 0$ and for some $A_{\sigma,pq} \in [-1, 1]$ that form a valid correlation matrix.

Remark 4. For $d = 2$, Condition 4 is satisfied by the deformation matrices $\{\Sigma_{pq}; p, q = 1, \ldots, P\}$ if their diagonal elements $\left[\Sigma_{pq} \right]_{ii}$, can be written as $\left[\Sigma_{pq} \right]_{ii} = \left[\Sigma_{pp} + \Sigma_{qq} \right]_{ii} / 2 + \Delta_\Sigma^{(i)} \left(1 - A_{\Sigma,pq}^{(i)} \right)$ for $i = 1, 2$.

Remark 5. For $d = 2$, and for small P, we can follow Apanasovich et al. (2012) and use equicorrelated matrices $A_\Sigma^{(i)}$, $i = 1, 2$, setting $A_{\Sigma,pq}^{(i)} = \rho_\Sigma^{(i)}$, $p \neq q$; in this scenario, for the sake of identifiability, we redefine $\Delta_\Sigma^{(i)} := \Delta_\Sigma^{(i)} (1 - \rho_\Sigma^{(i)})$, $i = 1, 2$.

Conditions 1–4, along with Remarks 2–4, indicate a sequential approach to specifying a valid multivariate geometric anisotropic Matérn covariance structure in two-dimensional Euclidean space. Condition 1 and Remarks 2–4 suggest that one must specify the parameters for the marginal covariance function before conditionally specifying the parameters for each cross-covariance function. These statements also indicate that, within each individual component of the joint model, that is, for fixed p, q, there is a particular order in which the five parameters $(\theta_{pq}, \zeta_{pq}, \alpha_{pq}, \nu_{pq}, \sigma_{pq})$ should be specified.

From Remark 3, we can see that, for each (p, q) pairing, the specification of the Matérn power parameter σ_{pq} is dependent upon the corresponding ratio of anisotropy ζ_{pq}, as well as the other Matérn parameters, α_{pq} and ν_{pq}. Remark 4 indicates that the anisotropy parameters $(\theta_{pq}, \zeta_{pq})$ should be jointly specified. In addition, Condition 1 and Remark 2 indicate that the smoothness parameter ν_{pq} should be specified before the scale parameter α_{pq}. We conclude that, for each marginal or bivariate component of the joint covariance model, the anisotropy parameters should be specified before the Matérn parameters, with the Matérn smoothness, scale and power

parameters being specified third, fourth and fifth, respectively. To be specific, we suggest the following sequential approach to specifying a multivariate geometric anisotropic Matérn covariance structure:

1. For each marginal process, we may specify the anisotropy parameters θ_{pp}, ζ_{pp} and the Matérn covariance parameters α_{pp}, ν_{pp}, and σ_{pp} independently of one another. For $p = 1, ..., P$, choose $\theta_{pp} \in [0, \pi), \zeta_{pp} \in (0, 1], \alpha_{pp}, \nu_{pp}, \sigma_{pp} > 0$.
2. For the joint dependence structure, first specify the anisotropy parameters $\theta_{pq} \in [0, \pi), \zeta_{pq} \in (0, 1]$ for $p, q = 1,...,P, p \neq q$; these should be jointly specified such that Remark 4 is satisfied.
3. Next, specify the smoothness parameter for the joint dependence structure, $\nu_{pq} > 0$. This should be specified conditional on ν_{pp} and ν_{qq}, such that Condition 1 is satisfied.
4. Given the chosen smoothness parameters $\nu_{pp}, \nu_{qq}, \nu_{pq}$, as well as the marginal scale parameters α_{pp}, α_{qq}, choose a scale parameter for the Matérn cross-covariance, α_{pq}, such that Remark 2 is satisfied.
5. Finally, we are free to specify the power parameter for the Matérn cross-covariance, σ_{pq}, conditional on the corresponding ratio of anisotropy ζ_{pq} and the Matérn parameters α_{pq} and ν_{pq}, such that Remark 3 is satisfied.

We conclude this section by considering the limitations placed on the zero-lag cross-correlation coefficients $\rho_{pq} := \sigma_{pq}/\sqrt{\sigma_{pp}\sigma_{qq}}$. By rearranging Condition 3, we can write:

$$\rho_{pq}^2 = \frac{\sigma_{pq}^2}{\sigma_{pp}\sigma_{qq}} \leq \prod_{i=1}^{4} \tau_{pq}^{(i)} \leq 1, \tag{7}$$

with

$$\tau_{pq}^{(1)} = \frac{\mathcal{B}^2\left(\nu_{pq}, \frac{d}{2}\right)}{\mathcal{B}^2\left(\frac{\nu_{pp}+\nu_{qq}}{2}, \frac{d}{2}\right)}, \quad \tau_{pq}^{(2)} = \left[\frac{\frac{4\nu_{pp}}{\alpha_{pp}^2}\frac{4\nu_{qq}}{\alpha_{qq}^2}}{\left(\frac{4\nu_{pq}}{\alpha_{pq}^2}\right)^2}\right]^{\Delta_\nu},$$

$$\tau_{pq}^{(3)} = \frac{\Gamma^2\left(\frac{\nu_{pp}+\nu_{qq}}{2}\right)\left(\frac{\alpha_{pq}^2}{4\nu_{pq}}\right)^{\nu_{pp}+\nu_{qq}}}{\Gamma(\nu_{pp})\left(\frac{\alpha_{pp}^2}{4\nu_{pp}}\right)^{\nu_{pp}}\Gamma(\nu_{qq})\left(\frac{\alpha_{qq}^2}{4\nu_{qq}}\right)^{\nu_{qq}}}, \quad \tau_{pq}^{(4)} = \frac{|\Sigma_{pp}|^{1/2}|\Sigma_{qq}|^{1/2}}{|\Sigma_{pq}|} = \frac{\zeta_{pp}\zeta_{qq}}{\zeta_{pq}^2},$$

where $\mathcal{B}(\cdot, \cdot)$ is the Beta function (Abramowitz & Stegun, 1965). The first inequality in (7) is directly implied by Condition 3. The second inequality in (7) can be shown componentwise: By Remark 1, Condition 4 ensures that $\tau_{pq}^{(4)} \leq 1$, and as noted by Apanasovich et al. (2012) in the isotropic framework, Conditions 1 and 2 are sufficient to guarantee that $\tau_{pq}^{(i)} \leq 1, i = 1, 2, 3$.

In the isotropic framework, $\tau_{pq}^{(4)} = 1$, and we are left with the limitations noted by Apanasovich et al. (2012): the zero-lag cross-correlation will be bounded above by 1 when the corresponding univariate isotropic processes share identical Matérn parameters. When the marginal parameter specifications differ, this upper bound will decrease as the smoothness and inverse correlation length of the cross-covariance structure depart from the arithmetic mean of the corresponding marginal quantities.

In our more general anisotropic framework, we can see from $\tau_{pq}^{(4)}$ in (7) that the upper bound on the colocated cross-correlations will also be affected by the relationship between the

cross-covariance ratio of anisotropy ζ_{pq} and the ratios of anisotropy in the corresponding marginal covariance structures. If we assume Condition 4 to hold, then by Remark 1, ζ_{pq} will be restricted to the closed interval $[\zeta_{pp}^{1/2}\zeta_{qq}^{1/2}, 1]$. If $\zeta_{pq} = \zeta_{pp}^{1/2}\zeta_{qq}^{1/2}$, then $\tau_{pq}^{(4)}$ will reduce to 1, and the upper bound of the colocated cross-correlation ρ_{pq} will behave as in the isotropic framework, that is, as described above. Increasing ζ_{pq} away from this geometric mean, however, will decrease $\tau_{pq}^{(4)}$, which will in turn shrink the upper bound on ρ_{pq}^2, given in (7). In other words, as the ellipticity of the cross-covariance function becomes less pronounced, the maximum possible degree of zero-lag correlation between the two components of the field will decrease. This formalizes the relationship between the power and the anisotropy of the cross-process dependence, discussed at the end of Section 3.1.

4 | FITTING THE MODEL

4.1 | Parameter estimation procedure

In order to fit our parametric model to an observed multivariate point pattern, we must estimate both the marginal and joint anisotropy parameters $\{\theta_{pq}, \zeta_{pq}; p, q = 1, \ldots, P\}$, as well as the parameters that specify the mean and Matérn covariance structure of the underlying GRF, $\{\mu_p, \alpha_{pq}, \nu_{pq}, \sigma_{pq}; p, q = 1, \ldots, P\}$. At a high level, we follow the approach of Møller and Toftaker (2014), who fit a univariate version of our model by first estimating the angle and ratio of anisotropy in the observed data, before using these estimates to back-transform the data into an isotropic framework. The resulting "isotropised" point pattern is then used to estimate the mean parameters and the Matérn parameters. Our approach to each component of this two-stage model fitting procedure will differ from the methods of Møller and Toftaker (2014). We use an approach to estimating anisotropy that is less sensitive to user-specified tuning parameters, which we adapt from the work of Rajala et al. (2016). We develop new methods from the work of Tanaka et al. (2008) to estimate the mean and Matérn parameters.

In developing new methodology, we are faced with the question of whether to put measures into place to guarantee that the fitted model satisfies Conditions 1–4, to ensure validity. This is the approach taken by Apanasovich et al. (2012) for fitting multivariate isotropic Matérn GRFs. Since Conditions 1–4 are sufficient, and not necessary conditions, the resulting restriction on the joint dependence structure could be overly conservative. Under the assumption that the smoothness is known, however, the power and scale parameters for a univariate Matérn covariance function cannot be consistently estimated under infill asymptotics (Zhang, 2004); consistency can only be achieved by increasing the observation window W. As noted by Apanasovich et al. (2012), constraining σ_{pq}^2 and α_{pq}^2 ($p \neq q$) conditional on their corresponding marginal values therefore provides no additional penalty in terms of estimator consistency when assuming a fixed observation window.

4.2 | Estimating the anisotropy parameters

We focus first on quantifying the anisotropy present in both the marginal and joint dependence structures in a multivariate point pattern. Møller and Toftaker (2014) estimate the angle of anisotropy in a univariate geometric anisotropic point pattern by finding the angle ϕ at which

the r-integrated difference between the anisotropic pair correlation function $g^a(r, \phi)$ and its phase-shifted self $g^a(r, \phi + \pi/2)$, is maximized. This is achieved by estimating $g^a(r, \phi)$ over a discrete lattice of polar coordinates (r, ϕ), and numerically approximating the required integral in r. Accuracy of the resulting estimator is therefore sensitive to the resolution of the polar lattice, as well as the choice of two bandwidth parameters used in estimating the anisotropic pair correlation function; for details of these bandwidth parameters, see Møller and Toftaker (2014). Finally, use of this estimation method is also dependent on the assumption that the isotropic pair correlation function is strictly decreasing. Whilst this assumption holds true for our assumed Matérn model, it can be violated by non-parametric estimators of the pair correlation function, when evaluated on real data. The approach we detail below is more widely applicable, as it does not depend on such an assumption, and it is also less sensitive to subjective choices of bandwidth parameters (Rajala et al., 2016).

We adopt and adapt the method introduced by Rajala et al. (2016) for estimating the angle of anisotropy. For the sake of generality, we describe the procedure for estimating θ_{pq}, $p \neq q$. We start by constructing the point pattern formed by the union of the difference vectors $\{x_{p,i} - x_{q,j}; i = 1, \ldots, n_p, j = 1, \ldots, n_q\}$ and their reflections $\{x_{q,i} - x_{p,j}; i = 1, \ldots, n_p, j = 1, \ldots, n_q\}$; this is the (bivariate) Fry process (Fry, 1979), and it will be rotationally symmetric of order 2, about the origin. The Fry process is useful here as its first-order properties will reflect the second-order properties of the original point pattern. We can therefore estimate any second-order anisotropy in the original bivariate point pattern by estimating the anisotropy in the intensity of the bivariate Fry process.

Dividing the polar plane into a selected number, n_F, of distinct sectors, and for $\ell \in L \subset \mathbb{N}$, we collect the ℓth nearest Fry point in each sector into a set, G_ℓ, of n_F points. Each G_ℓ sketches out a noisy contour around the origin, such that the intensity of the Fry process is reflected in the proximity of the G_ℓs to one another. For point patterns that display segregation, the anisotropy in the joint second-order dependence structure will be shared by the contours of the intensity field for the Fry process; for aggregated point patterns, the angle of anisotropy will be phase-shifted by $\pi/2$ in the Fry process. In order to quantify the anisotropy in the original point pattern then, we can treat the G_ℓs as sampled versions of the Fry intensity's contours, and assuming Gaussian measurement error we can infer the corresponding true contours using adjusted ordinary least squares (Kukush et al., 2004), and subsequently derive the angle of anisotropy in the original point pattern. In using *adjusted* least squares, we obtain unbiased estimates of the "true" elliptical contours. We also choose not to smooth the Fry contours before fitting these ellipses; this avoids a bias toward circularity of the fitted ellipse, though it has the effect of increasing the estimator variance. For full details of this approach to estimating the angle of anisotropy, see Rajala et al. (2016).

For each marginal process (Møller & Toftaker, 2014), we can transform the observed point pattern x_p and the corresponding observation window W by assuming fixed values for $\theta \in [0, \pi)$ and $\zeta \in (0, 1]$ and setting:

$$x_{p,\theta,\zeta} = \Sigma^{-1/2} x_p, \qquad W_{\theta,\zeta} = \Sigma^{-1/2} W.$$

If the chosen values of θ and ζ are equal to the values that describe the anisotropy of X_p, then the transformed point process $X_{p,\theta,\zeta}$ will be isotropic. Then the corresponding anisotropic pair correlation function $g^a_{pp,\theta,\zeta}(r, \phi)$ will be constant with respect to its second argument; this motivates the following approach to estimating the *marginal* anisotropy ratios ζ_{pp}.

Following Rajala et al. (2016), we define the following directional discrepancy statistic:

$$V_{pp,\theta}(\zeta) = \int_{b_1}^{b_2} \left[K_{pp,\theta,\zeta}^a(r,0) - K_{pp,\theta,\zeta}^a(r,\pi/2) \right] dr, \tag{8}$$

where

$$K_{pp,\theta,\zeta}^a(r,\phi) = \int_0^r g_{pp,\theta,\zeta}^a(s,\phi) ds, \tag{9}$$

is the sector-K-function, an anisotropic variant of Ripley's K-function, evaluated on the isotropized point pattern $x_{p,\theta,\zeta}$. To estimate the marginal ratio of anisotropy ζ_{pp}, we back-transform our observed point pattern using the estimated angle of anisotropy $\hat\theta_{pp}$ and a sequence of candidate ratios $\{\zeta_{pp,k} := k\zeta_{\max}/(1 + n_\zeta), k = 1, \ldots, n_\zeta\}$, for some user-defined upper bound $\zeta_{\max}$. We then choose $\hat\zeta_{pp} = \zeta_{pp,k} \in (0, \zeta_{\max})$ to be the candidate value that minimizes the estimate $\hat{V}_{pp,\hat\theta_{pp}}(\zeta_{pp,k})$. Note that, although we defined $\zeta_{pp} \in (0, 1]$ in Section 2.2, the sampling variance of the estimated sector-K-function can in some cases result in an estimate $\hat\zeta_{pp} > 1$.

Møller and Toftaker (2014) use a similar approach, in effect minimising the directional discrepancy statistic (8), but using the anisotropic pair correlation function in place of the sector-K-function. Indeed, any directional second-order statistic can be used in the integrand of (8). We choose to use $K_{pp,\theta,\zeta}^a$ for two reasons. Firstly, the analysis of Redenbach et al. (2009) suggests that the sector-K-function is better-suited to characterising anisotropy than nearest-neighbor statistics; the authors conclude that, for detecting anisotropy in point patterns, statistical tests based on the sector-K-function have greater power, in general, than those based on nearest-neighbor orientation statistics. Secondly, estimation of the sector-K-function requires the choice of only one tuning parameter, an angular bandwidth, whereas the use of the anisotropic pair correlation function would require the specification of both angular and radial bandwidths.

Our chosen approach to estimating ζ_{pp} can be extended to the multivariate scenario. We are interested in the geometric anisotropic cross-dependence exhibited by a given pair of Cox processes X_p and X_q. For each pair of processes, we once again define a discrete set of candidate multivariate anisotropy ratios $\{\zeta_{pq,k} \in (0, \zeta_{\max}), k = 1, \ldots, n_\zeta\}$, and we choose $\hat\zeta_{pq} = \zeta_{pq,k}$ for which the estimated value of $V_{pq,\hat\theta_{pq}}(\zeta_{pq,k})$ is minimized, where $V_{pq,\hat\theta_{pq}}(\zeta_{pq,k})$ is defined through transforming both x_p and x_q, along with their common observation window W.

The above approach to estimating the anisotropy parameters requires the selection of a number of control parameters. For estimating the angles of anisotropy, we must choose the number of sectors, n_F, into which we partition the Fry process, as well as the set L of contour levels, for determining the noisy contours of Fry points. As a rule of thumb, and for reasons outlined below, Rajala et al. (2016) suggest choosing $n_F \approx \lambda|W|/6$, where $\lambda|W|$ is the expected number of points in the original point process. We adopt this guideline for choosing n_F when estimating the anisotropy in the marginal processes. We derive a similar rule of thumb for n_F when estimating the between-process anisotropy, by following the same arguments as Rajala et al. (2016). For the bivariate Poisson process with intensity vector (λ_p, λ_q) in a circular spatial window W, the expected number of bivariate Fry points per sector is approximately $\lambda_p \lambda_q |W|^2/3n_F$. Each point in the bivariate process can be expected to contribute if there are at least $(\lambda_p + \lambda_q)|W|$ points per sector, and so we have a bivariate direction count rule of $n_F \approx \lambda_p \lambda_q |W|/3(\lambda_p + \lambda_q)$.

In specifying L, we choose to follow a data-driven approach, which is based on the work of Rajala et al. (2016), and which works well in practice. Suppose we have a set of Fry points, allocated to n_F sectors, and define $\ell_{\max}$ to be the smallest number of Fry points that lie in a single

sector. $\ell_{\max}$ is therefore the largest value of ℓ for which we are able to construct a complete "noisy contour" G_ℓ. As discussed by Rajala et al., the contours G_ℓ that result from clustered processes are expected to be roughly isotropic for small values of ℓ. We proceed by calculating, for a selection of evenly-spread values $\ell \in [1, \ell_{\max}]$, 95% Monte Carlo (MC) confidence intervals for the difference between the lengths of the semi-axes in the elliptical contour, of which G_ℓ is a noisy realization—see Rajala et al. (2016) for further details on these MC confidence intervals. For the smallest values of ℓ (i.e., for approximately isotropic contours), we expect the corresponding intervals to envelope zero; we therefore seek the first subsequent confidence interval that does not include zero, corresponding to $\ell = \ell_0$, say, and we set $L = \{\ell_0, \ldots, \ell_{\max}\}$.

For estimating the ratios of anisotropy, we must choose the number n_ζ of candidate ratios, as well as their upper bound $\zeta_{\max}$, and the limits of integration, b_1 and b_2 in (8); we use these values to calculate $\hat{V}_{pq,\hat{\theta}_{pq}}(\zeta_{pq,k})$ when estimating ζ_{pq}. Selection of both n_ζ and $\zeta_{\max}$ is straightforward: $\zeta_{\max}$ should be chosen such that $(0, \zeta_{\max})$ covers the majority of the sampling distribution of ζ_{pq}, and selection of n_ζ involves a trade-off between accuracy in the resulting estimates and computational expense of the estimation procedure. In Section 5, where we implement our model fitting procedure for both simulated data and tropical rainforest data, we use $\zeta_{\max} = 2$ and $n_\zeta = 199$ for estimating all marginal and joint ratios of anisotropy. Choice of the limits of integration, b_1 and b_2 in (8), is a more subjective task, and should be determined by the range of scales over which dependence (either within, or between processes) is sought to be characterized. In Section 5, we detail our choices of these limits of integration.

4.3 | Estimating the Matérn parameters

Once we have estimated our anisotropy parameters, we can isotropize the point pattern and its observation window. In order to ensure that the Matérn parameters satisfy Conditions 1–3, we define ν_{pq}, α_{pq}, and σ_{pq} according to the specifications in Condition 1, Remark 2, and Remark 3, respectively. Techniques for modeling the correlation matrices A_ν, A_α, and A_σ are discussed by Apanasovich and Genton (2010). When P is small, however, we can simplify our task by assuming the off-diagonal elements of $A_\alpha, A_\nu, A_\sigma$ to be constant (Apanasovich et al., 2012).

In order to estimate both the mean and Matérn parameters, we maximize the Palm log-likelihood, first proposed by Tanaka et al. (2008). We use the Palm likelihood of Dvořák and Prokešová (2012), where the inner region correction is proposed to deal with edge effects:

$$
\ell(\lambda_p, \alpha_{pp}, \nu_{pp}, \sigma_{pp}) \approx \sum_{\substack{x_{p,i} \in x_{p,\theta,\zeta} \cap W_{\theta,\zeta} \setminus R \\ x_{p,j} \in x_{p,\theta,\zeta} \cap W_{\theta,\zeta} \\ r_{ij} < R}}^{\neq} \log \left\{ \lambda_p g_{0,pp}(r_{ij}; \alpha_{pp}, \nu_{pp}, \sigma_{pp}) \right\}
$$
$$
= -\lambda_p |x_p \cap W \setminus R| K_p(R; \alpha_{pp}, \nu_{pp}, \sigma_{pp}),
$$
(10)

where $r_{ij} = \|x_{p,i} - x_{p,j}\|$, $K_p(r; \alpha_{pp}, \nu_{pp}, \sigma_{pp})$ is Ripley's univariate K-function, and $|x_{p,\theta,\zeta} \cap W \setminus R|$ denotes the number of points in the isotropized pattern $x_{p,\theta,\zeta}$ that lie further than a distance R from the boundary of $W_{\theta,\zeta}$. We approximate $K_p(r; \alpha_{pp}, \nu_{pp}, \sigma_{pp})$ by numerically integrating the corresponding isotropic pair correlation function $g_{0,pp}(r; \alpha_{pp}, \nu_{pp}, \sigma_{pp})$ over $r \in [0, R]$. R is a user-defined tuning parameter that can be objectively set based on the data; this is discussed further in Section 5. As is common in the point pattern literature, we use $\neq$ in the summation notation to indicate summation over pairs of distinct points.

The Palm log-likelihood (10) can be analytically maximized with respect to λ_p, yielding the maximum Palm likelihood estimate (MPLE) $\hat{\lambda}_p$, and we obtain MPLEs for the remaining marginal Matérn parameters by numerically maximizing $\ell(\hat{\lambda}_p, \alpha_{pp}, \nu_{pp}, \sigma_{pp})$. The MPLE for μ_p can be subsequently calculated as $\hat{\mu}_p = \log(\hat{\lambda}_p) - \hat{\sigma}_{pp}/2$. For univariate point processes, Prokešová and Vedel Jensen (2013) show that the MPLE estimators considered here are unbiased, strongly consistent and asymptotically normal. Their proof of strong consistency holds under the assumption that the point process is ergodic, that its log-Palm intensity is bounded and continuous with respect to both the spatial argument and any covariance parameters, and that the integrated Palm log-intensity (over $B(o, R)$) is also bounded and continous with respect to the spatial argument. The proof of asymptotic normality depends on the assumption that the point process is strongly mixing. As is also discussed by Prokešová and Vedel Jensen (2013), these conditions are all satisfied by the LGCP with Matérn covariance function.

We further develop the Palm log-likelihood approach, in order to estimate the parameters for the cross-covariance structure. First, we obtain the symmetric bivariate Fry process for components X_p and X_q, using the inner region correction to deal with edge effects. We then treat this Fry process as an inhomogeneous Poisson process, with intensity equal to a bivariate version of the Palm intensity (Daley & Vere-Jones, 2008; Prokešová & Vedel Jensen, 2013), which we define heuristically as follows: for u at distance $r > 0$ from the origin o, the occurrence rate of process q at u, assuming there to be a point of process p at the origin, is

$$\lambda_{o,pq}(u)du = \mathbb{P}\left(|X_q \cap du| = 1 \Big| |X_p \cap \{o\}| = 1\right),$$

where du is the Lebesgue measure for the infinitesimal region containing u.

Following this definition, we can relate the bivariate Palm intensity to the (isotropic) cross-pair correlation function for the original process: $\lambda_{o,pq}(u) = \lambda_q g_{0,pq}(\|u\|)$. This allows us to define the following bivariate extension to the Palm log-likelihood, which can be maximized to obtain estimates for $\alpha_{pq}, \nu_{pq}, \sigma_{pq}, p \neq q$:

$$
\ell(\lambda_p, \lambda_q, \alpha_{pq}, \nu_{pq}, \sigma_{pq}) \approx \sum_{\substack{x_{p,i} \in x_p \cap W_{\theta,\zeta} \\ x_{q,j} \in x_q \cap W_{\theta,\zeta} \\ r_{ij} < R}}^{\neq} \log\left\{(\lambda_p + \lambda_q)g_{0,pq}(r_{ij}; \alpha_{pq}, \nu_{pq}, \sigma_{pq})\right\}.
$$

$$
\qquad\qquad - \left(|x_q \cap W \setminus R|\lambda_p + |x_p \cap W \setminus R|\lambda_q\right)K_{pq}(R; \alpha_{pq}, \nu_{pq}, \sigma_{pq}). \tag{11}
$$

As usual $r_{ij} = \|x_{p,i} - x_{q,j}\|$, $K_{pq}(r; \alpha_{pq}, \nu_{pq}, \sigma_{pq})$ is Ripley's bivariate K-function, and $|x_p \cap W \setminus R|$ denotes the number of observed points in process p that lie further than a distance R from the boundary of the window R. By substituting our previous estimates of λ_p and λ_q into (11), we obtain an expression in terms of the Matérn cross-covariance parameters only. We numerically maximize this expression in $(\alpha_{pq}, \nu_{pq}, \sigma_{pq})$ over the constrained parameter space described by Condition 1, Remark 2 and Remark 3, and dependent on the corresponding estimated marginal Matérn parameters.

We would expect the maximum Palm likelihood estimators for the cross-covariance parameters to be both consistent and asymptotically normal, following similar arguments to those of Prokešová and Vedel Jensen (2013). As described in Section 4.1, the use of constrained optimisation should not affect the consistency of the cross-covariance parameter estimators, however, they may display some bias due to the truncation of their supports.

5 | IMPLEMENTATION

5.1 | Proof of concept simulations

We demonstrate the performance of the model fitting procedure described in Section 4, through simulation studies. Using the restrictions in Section 3, we define eight distinct bivariate geometric anisotropic LGCPs with valid Matérn covariance structures. In this section, we report on the simulation studies corresponding to the first four models; those corresponding to the second four models are detailed in the Appendix. The parameter values for each model are given in the corresponding results table. For all eight models, the parameters are chosen such that the expected log-intensity for each process component, $\log(\lambda_p) = 6.75$ ($p = 1, 2$), specifying point patterns with a similar intensity to the ecological data to be considered in Section 5.2. For each of the eight fully specified models, we simulate 500 distinct point patterns on the unit square, $W = [0, 1]^2$. One realisation from each of the first four models is shown in Figure 3.

For each model, and for $p, q = 1, 2$, we executed our parameter estimation procedure as described in Section 4.1, and a detailed algorithmic outline of this bivariate implementation is also provided in the Appendix. For both the marginal and cross-dependence relationships, we estimate θ_{pq} using Fry processes consisting of only those point pairs separated by $r \in (0, 0.25)$. Similarly, when estimating ζ_{pq}, we numerically approximate the integral $V_{pq,\hat{\theta}_{pq}}(\zeta)$ as defined in (8), using the limits of integration $b_1 = 0$, $b_2 = 0.25$. Approximation of $V_{pq,\hat{\theta}_{pq}}(\zeta)$ involves estimating the sector-K-function over a discrete, high-resolution set of distances r, using an angular bandwidth parameter which we choose to be $h_\phi = \pi/16$; details of the chosen sector-K-function estimator are given in the Appendix.

When estimating the Matérn parameters, despite using a favorable form of the Matérn parameterization as discussed in Section 3.2, there proved to be insufficient separation of the effects of ν_{pq} and α_{pq} in practice for both parameters to be allowed to vary freely during estimation. This lack of identifiability is often found in the Matérn parameterization (see, e.g., Handcock & Wallis, 1994). In order to avoid this issue, a common strategy (e.g., Diggle et al., 2013) is to restrict $\hat{\nu}_{pq}$ to a number of discrete candidate values. These values should represent sufficiently distinct levels of smoothness in the resulting random fields, so that they can be distinguished: we chose to seek $\hat{\nu}_{pq} \in \{0.05, 0.5, 5.0\}$, $p, q = 1, 2$. For the case $p \neq q$, this candidate vector was further restricted, to ensure that $\hat{\nu}_{12}$ satisfied Condition 1. The remaining Matérn parameters were allowed to vary on continuous bounded intervals: $\hat{\alpha}_{pq} \in (0, \alpha_{pq}^{UB})$ and $\hat{\sigma}_{pq} \in (0, \sigma_{pq}^{UB})$. In the marginal cases, $\alpha_{pp}^{UB} = 10$ and $\sigma_{pp}^{UB} = 50$, $p = 1, 2$, were chosen such that these constituted generous intervals around the corresponding true values. For estimating the cross-covariance parameters, α_{12}^{UB} and σ_{12}^{UB} were chosen to ensure compliance with Conditions 2 and 3.

We used an interior-point algorithm (Byrd et al., 1999) to carry out constrained maximization of the Palm-log likelihood with respect to $(\alpha_{pq}, \sigma_{pq})$, for each candidate value of ν_{pq}. We initialised the parameters using a computationally inexpensive version of the widely used minimum contrast method, minimizing the difference between the estimated (isotropic) pair correlation function and its closed-form expression across a coarse grid of parameter pairs $(\alpha_{pq}, \sigma_{pq})$.

As discussed by Dvořák and Prokešová (2012), the accuracy of the MPLE estimates are sensitive to the choice of tuning parameter R. We therefore repeat our parameter estimation exercise for multiple values of R. We followed the guidance of Prokešová and Vedel Jensen (2013) in considering a suitable model-based value of R, in addition to considering $R \in \{0.1, 0.2, 0.3\}$. Prokešová and Vedel Jensen (2013) suggest setting R equal to the estimated "range of interaction" in the data. This is a well-defined concept for data from cluster models and Gibbs models, but is less
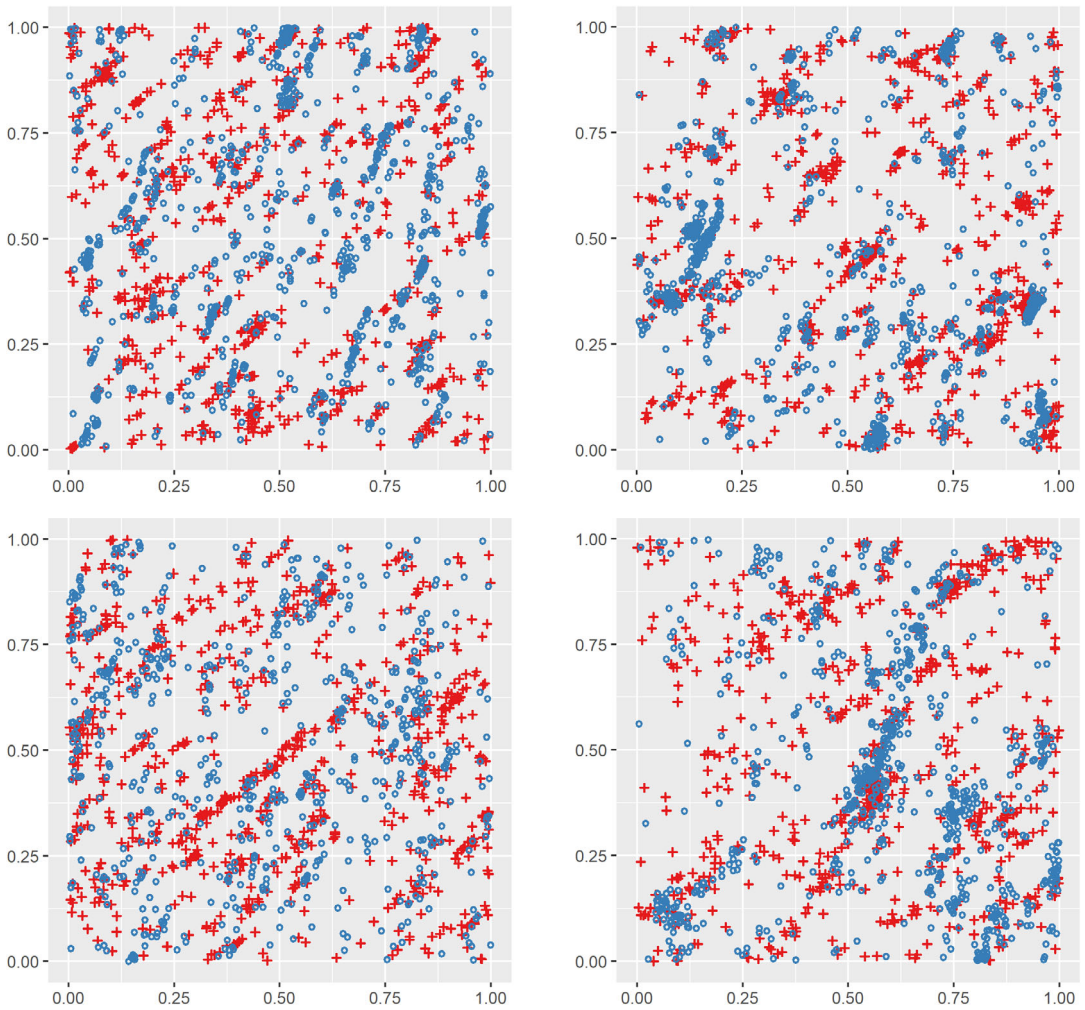
FIGURE 3 A two-type point pattern, simulated from each of Models 1–4 (left to right, top to bottom), which are specified in Tables 1 and 2. In each realisation, points of type 1 are shown as red crosses and points of type 2 are shown as blue circles.

well-defined for Cox processes, which do not *directly* model the pairwise interaction between points. We therefore choose to set R_{model} approximately equal to the "practical range" of dependence, which is defined in the geostatistics literature to be the distance at which the spatial auto- or cross-correlation decays to 0.05. We allow for distinct scales in the multivariate dependence structure by calculating R_{model} separately for each marginal and joint component of the covariance model.

In a small proportion of runs, the MPLE procedure returned seemingly degenerate estimates of either $\hat{\alpha}_{pp}$ or $\hat{\sigma}_{pp}$, $p = 1, 2$, with one or the other being returned equal to their upper bound. In this scenario, the number of points contributing to the Palm log-likelihood (10) is reduced, leading to a loss of accuracy in the MPLE procedure. We therefore counter this phenomenon by decreasing R when necessary. When the initial attempt returns estimates of any of the marginal Matérn scale or power parameters greater than 95% of their corresponding upper bound, we

iteratively repeat the MPLE procedure, reducing R by 0.01 each time, until all marginal scale and power estimates are below this 95% threshold. We found this iterative procedure to be an adequate, if somewhat ad-hoc remedy to the problem. After applying our iterative solution, for each of the four models considered, fewer than one in fifty MC runs returned any marginal Matérn scale or power estimates greater than 50% of their corresponding upper bound.

In Tables 1 and 2, we summarize the MC sampling distributions of the parameters in Models 1–4. For the smoothness parameters, we report the modal estimate from our MC simulations, as we consider only three potential values for these parameters. For the estimated scales of anisotropy, we provide the median of the MC samples, since their sampling distributions display evidence of skewness, along with the root-mean-square error (RMSE). For the estimated angles of anisotropy, as well as the Matérn scale and power parameters, we provide the MC sample mean and the corresponding RMSE.

For all models, the estimated anisotropy parameters are in reasonable agreement with their corresponding true values. The accuracy of these estimates can be improved by reducing the range of distances r, over which we seek to characterize anisotropy; their broad accuracy, however, suggests that our bivariate generalization of Rajala et al.'s method of estimating anisotropy is competitive.

Broadly speaking, we found the accuracy of the Matérn parameter estimates to improve as R decreases, and this was particularly evident when the true Matérn scale parameter was larger. Using a fixed value of $R \in \{0.1, 0.2, 0.3\}$ for estimating both the marginal and joint Matérn parameters was not found to offer a consistent improvement in estimator accuracy. In the remaining analysis, we therefore focus on the simulations that used $R = R_{\text{model}}$. The sampling distributions of the parameter estimates for Model 1, using $R = R_{\text{model}}$, are depicted in Figure 4, and the corresponding figures for Models 2–4 are provided in the Appendix.

For these four models, the scale parameters, α_{pq}, have been satisfactorily estimated, and we have also recovered the correct values of the smoothness parameters ν_{pq}. For all four models, the power parameter estimates σ_{pq}, $p, q = 1, 2$ show reasonable accuracy, though there is often

T A B L E 1 Monte Carlo (MC) estimates and root-mean-square errors (RMSEs) for the anisotropy parameters in four distinct models (Models 1–4).

	θ_{11}	θ_{22}	θ_{12}	ζ_{11}	ζ_{22}	ζ_{12}
Model 1	36°	72°	54°	0.20	0.20	0.35
MC estimate	40.02°	74.99°	63.10°	0.25	0.24	0.40
RMSE	21.86°	18.85°	35.01°	0.35	0.34	0.52
Model 2	36°	72°	54°	0.40	0.40	0.60
MC estimate	43.17°	75.56°	74.83°	0.41	0.41	0.59
RMSE	29.74°	21.52°	48.84°	0.33	0.31	0.45
Model 3	36°	72°	54°	0.20	0.20	0.35
MC estimate	37.15°	71.61°	57.55°	0.24	0.22	0.37
RMSE	11.97°	8.85°	24.74°	0.19	0.10	0.32
Model 4	36°	72°	54°	0.40	0.40	0.60
MC estimate	38.93°	72.73°	62.19°	0.40	0.40	0.53
RMSE	19.41°	12.22°	35.33°	0.21	0.14	0.27

TABLE 2 Monte Carlo (MC) estimates and root-mean-square errors (RMSEs) for the mean and Matérn parameters in four distinct models (Models 1–4).

	R	α_{11}	α_{22}	α_{12}	ν_{11}	ν_{22}	ν_{12}	σ_{11}	σ_{22}	σ_{12}	μ_1	μ_2
Model 1		0.045	0.065	0.050	0.5	0.5	0.5	4.00	4.50	1.97	4.75	4.50
MC estimate	$R = R_{\text{model}}$	0.040	0.103	0.050	0.5	0.5	0.5	5.26	5.52	1.43	2.78	2.50
	$R = 0.1$	0.043	0.068	0.046	0.5	0.5	0.5	4.56	4.35	1.29	3.13	2.95
	$R = 0.2$	0.052	0.103	0.052	0.5	0.5	0.5	6.23	5.54	1.55	2.30	2.47
	$R = 0.3$	0.083	0.065	0.054	0.5	0.5	0.5	7.37	6.21	1.69	1.67	2.10
RMSE	$R = R_{\text{model}}$	0.020	0.572	0.021	NaN	NaN	NaN	4.15	4.29	3.54	2.74	2.92
	$R = 0.1$	0.032	0.083	0.024	NaN	NaN	NaN	2.05	1.66	1.16	2.35	3.79
	$R = 0.2$	0.239	0.525	0.026	NaN	NaN	NaN	4.91	4.66	4.10	3.30	3.01
	$R = 0.3$	0.612	0.100	0.032	NaN	NaN	NaN	6.63	5.08	4.02	4.08	3.38
Model 2		0.045	0.065	0.050	0.5	0.5	0.5	4.00	4.50	2.30	4.75	4.50
MC estimate	$R = R_{\text{model}}$	0.068	0.082	0.056	0.5	0.5	0.5	4.86	5.20	1.29	3.41	2.98
	$R = 0.1$	0.051	0.080	0.051	0.5	0.5	0.5	4.21	4.22	1.40	3.74	3.38
	$R = 0.2$	0.066	0.097	0.057	0.5	0.5	0.5	5.89	5.22	1.52	2.88	2.93
	$R = 0.3$	0.086	0.074	0.057	0.5	0.5	0.5	6.27	5.73	1.57	2.70	2.72
RMSE	$R = R_{\text{model}}$	0.410	0.127	0.028	NaN	NaN	NaN	2.90	2.48	1.56	2.40	2.92
	$R = 0.1$	0.071	0.111	0.029	NaN	NaN	NaN	2.25	2.07	1.27	1.99	3.85
	$R = 0.2$	0.345	0.354	0.037	NaN	NaN	NaN	4.86	3.31	2.92	2.92	3.22
	$R = 0.3$	0.572	0.133	0.036	NaN	NaN	NaN	5.98	4.23	3.72	3.25	3.27
Model 3		0.090	0.120	0.100	0.5	0.5	0.5	2.00	2.25	0.98	5.75	5.62
MC estimate	$R = R_{\text{model}}$	0.129	0.177	0.107	0.5	0.5	0.5	3.01	2.87	0.59	3.64	3.64
	$R = 0.1$	0.113	0.221	0.084	0.5	0.5	0.5	2.35	2.72	0.62	3.87	3.33
	$R = 0.2$	0.100	0.181	0.103	0.5	0.5	0.5	2.69	2.62	0.68	3.85	3.73
	$R = 0.3$	0.156	0.163	0.105	0.5	0.5	5.0	3.00	2.77	0.68	3.68	3.70
RMSE	$R = R_{\text{model}}$	0.465	0.697	0.049	NaN	NaN	NaN	2.32	2.22	0.80	2.95	2.43
	$R = 0.1$	0.272	0.719	0.045	NaN	NaN	NaN	1.50	1.08	0.72	3.80	5.18
	$R = 0.2$	0.146	0.588	0.048	NaN	NaN	NaN	2.29	1.62	0.94	2.45	2.66
	$R = 0.3$	0.699	0.561	0.054	NaN	NaN	NaN	2.68	2.62	1.89	2.47	2.42
Model 4		0.090	0.120	0.100	0.5	0.5	0.5	2.00	2.25	1.15	5.75	5.62
MC estimate	$R = R_{\text{model}}$	0.143	0.261	0.121	0.5	0.5	5.0	2.74	2.64	0.68	4.34	4.12
	$R = 0.1$	0.115	0.194	0.087	0.5	0.5	0.5	2.08	2.47	0.68	4.68	4.18
	$R = 0.2$	0.122	0.164	0.111	0.5	0.5	0.5	2.62	2.55	0.70	4.38	4.23
	$R = 0.3$	0.159	0.201	0.115	0.5	0.5	5.0	2.85	2.63	0.66	4.28	4.22
RMSE	$R = R_{\text{model}}$	0.464	0.928	0.146	NaN	NaN	NaN	1.93	1.94	1.40	1.90	3.08
	$R = 0.1$	0.341	0.600	0.067	NaN	NaN	NaN	1.32	2.99	0.81	2.13	3.93
	$R = 0.2$	0.201	0.322	0.069	NaN	NaN	NaN	3.43	1.31	1.06	2.54	2.69
	$R = 0.3$	0.596	0.620	0.081	NaN	NaN	NaN	2.01	1.85	1.18	1.96	2.14
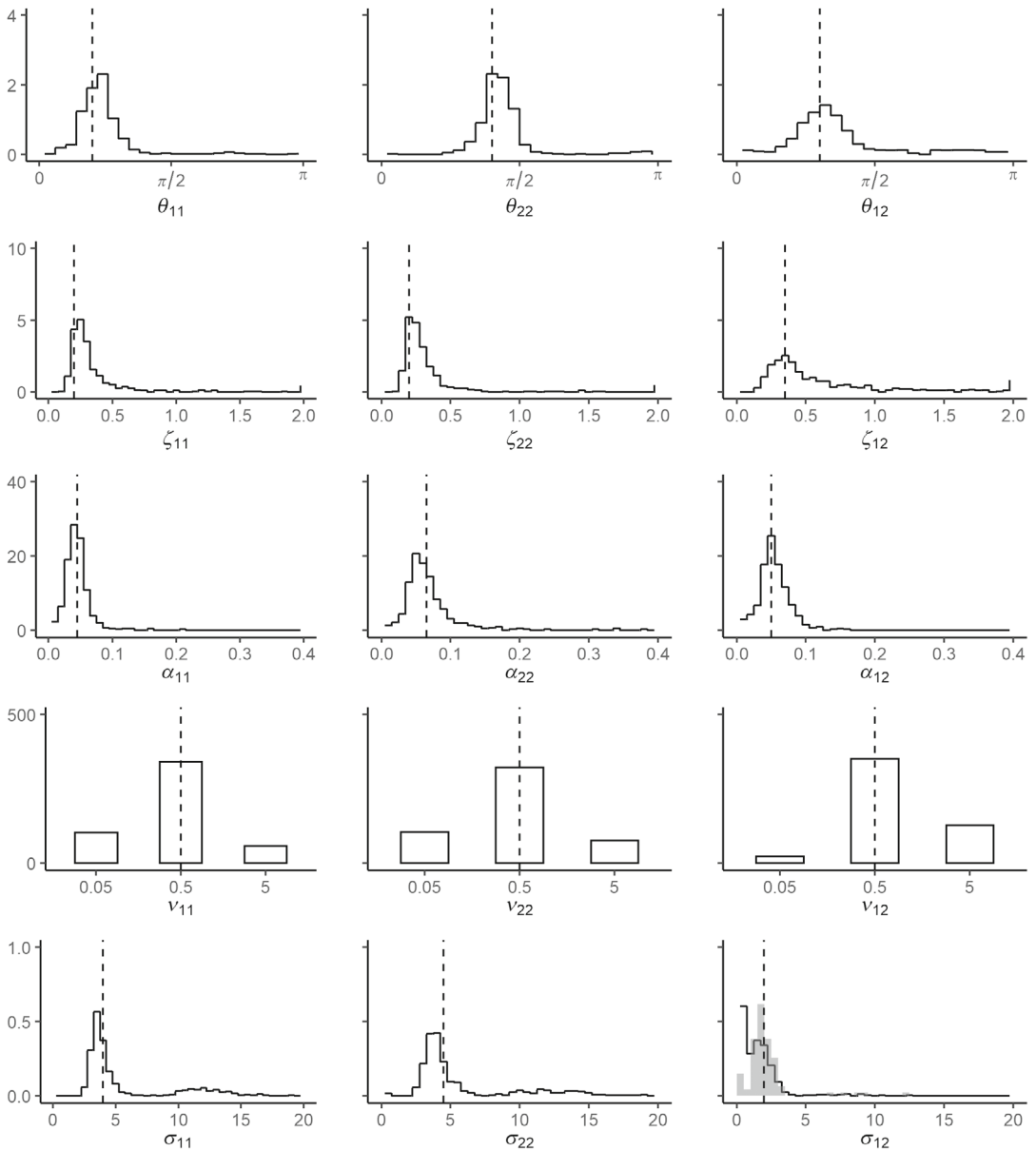
FIGURE 4 Histograms of the parameter distributions for the synthetic bivariate geometric anisotropic log-Gaussian Cox process with Matérn covariance structure specified by Model 1. The parameter values used to generate each dataset are marked by vertical dashed lines. The grey histogram in the final panel shows the empirical distribution for $\hat{\sigma}_{12}$, restricted to simulations where $\hat{\sigma}_{12}$ was not equal to the upper bound dictated by $\hat{\sigma}_{11}$ and $\hat{\sigma}_{22}$.

underestimation of the joint dependence power parameter. In the final panel of Figures 4 and A2–A8, we have overlaid the empirical parameter distribution for $\hat{\sigma}_{12}$, restricted to those MC simulations where $\hat{\sigma}_{12}$ was not equal to the upper bound dictated by $\hat{\sigma}_{11}$ and $\hat{\sigma}_{22}$. This suggests that our use of constrained optimization limits the accuracy of the estimated power parameter; this is the cost of ensuring that each fitted parameter vector specifies a valid multivariate dependence structure.

We also corroborate the observation of Dvořák and Prokešová (2012): that the MPLE estimates of μ_1, μ_2 are particularly poor, and that ultimately, a more reliable estimator for μ_p can be obtained by using the classical intensity estimator, $\hat{\lambda}_p = n_p/|W|$, and combining this with $\hat{\sigma}_{pp}$. When performing inference with respect to the Matérn parameters, however, we jointly maximize the (bivariate) Palm likelihood with respect to both the mean and covariance parameters, so substitution of the classical intensity estimator is only performed after this optimization exercise.

We make two final notes regarding the accuracy of the Palm likelihood estimation procedure. When the value of R is smaller than the true scale parameter α_{pq}, then estimation of $\hat{\alpha}_{pq}$ was found to be poor. This is to be expected, as R is equal to the maximum distance between point pairs that contribute to the Palm likelihood. Finally, and despite the measures taken to ameliorate the lack of identifiability in our Matérn parameterization, our procedure was observed to misspecify some of the models in a relatively small number of iterations. For Model 6, we found that the iterations for which σ_{11} was overestimated, corresponded to an underestimation of the corresponding smoothness parameter, ν_{11}; this can be seen in the bottom two panels of the left-hand column in Figure A6. Similarly, in Model 5, we verified that the overestimation of ν_{11} corresponded to iterations where σ_{11} was underestimated (see Figure A5). Overall, however, these results indicate reasonable success for our model fitting procedure, and motivates its usage in practice.

5.2 | Application to ecological data

We fit our multivariate Matérn geometric anisotropic LGCP to a bivariate point pattern from a 50ha plot in the BCI forest stand in Panama. We study two tree species, *Cecropia obtusifolia* and *Spondias radlkoferi*. To ease comparison with the studies in the previous section, we rescale the coordinates to the half-unit window $[0, 1] \times [0, 0.5]$; this rescaled bivariate point pattern is displayed in Figure 5. *C. obtusifolia* and *S. radlkoferi* were chosen as a preliminary study of the data revealed empirical evidence of between-process anisotropy at a range of 50 m. This is demonstrated in Figure 6, which we describe below. The anistropic interaction between these two species is likely explained by the shared preference for moist soils, and in the BCI site these can be found on the steepest slopes that run in the middle (near to a swamp) and towards the eastern edge of the 50 ha plot (Kupers et al., 2019).

As above, we estimate θ_{pq} using Fry processes consisting of only those point pairs separated by $r \in (0, 0.25)$, and we estimate ζ_{pq} using $b_1 = 0$ and $b_2 = 0.25$ as the limits of integration in $\hat{V}_{pq,\hat{\theta}_{pq}}(\zeta)$. In Section 5.1, we used the known specifications of the Matérn covariance model to determine a suitable *model*-based choice of R. Here, we replicate this approach by using the data to estimate a suitable *data*-based choice of R. We do so by consulting the estimated marginal and cross-pair correlation functions for the isotropized data, once the marginal and between-process anisotropy parameters have been estimated, setting R equal to the smallest distance at which the corresponding estimated pair correlation function is less than 1.05. This resulted in the use
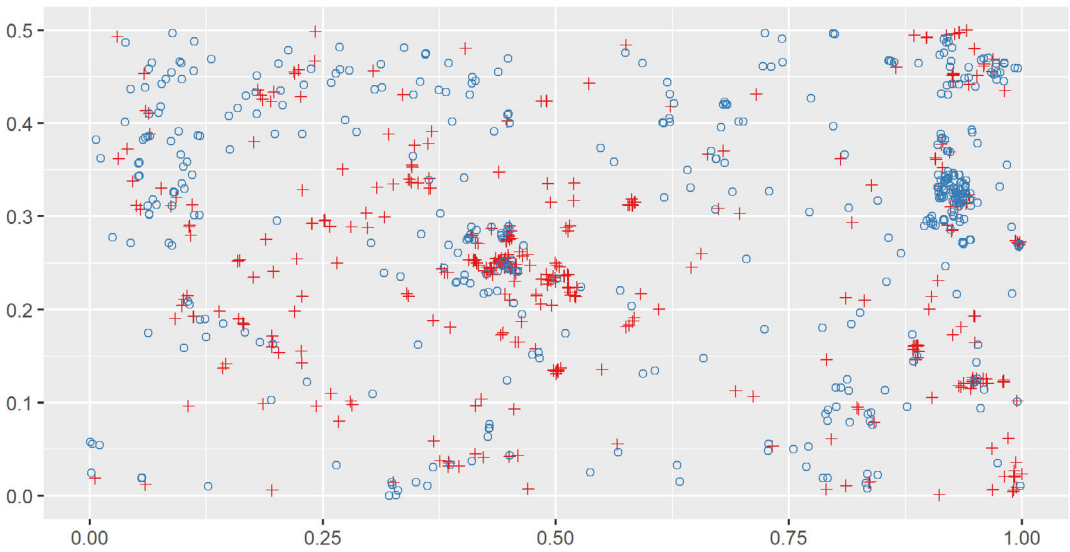
FIGURE 5 Rescaled point pattern data from the 50 ha tropical rainforest census plot on Barro Colorado Island. Two species are shown: *Cecropia obtusifolia* (red crosses) and *Spondias radlkoferi* (blue circles).
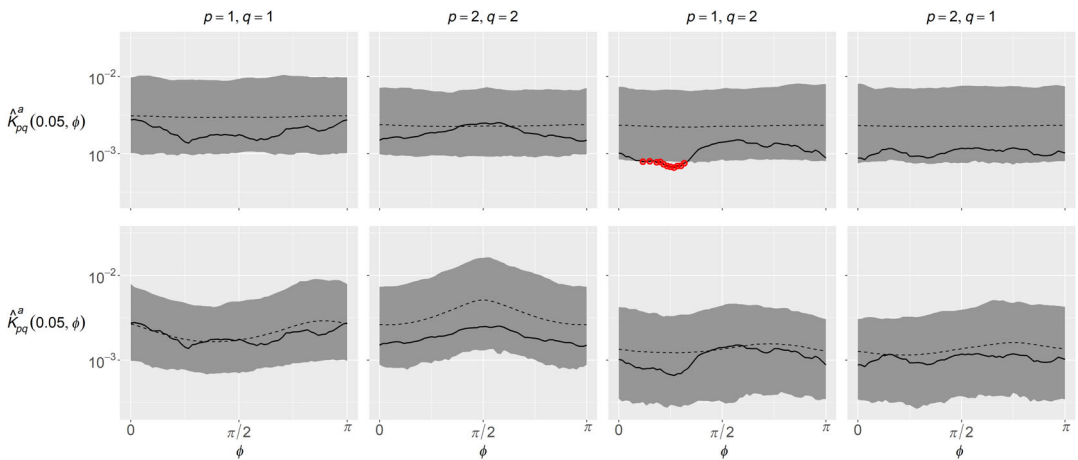

FIGURE 6 Estimates of the sector-K-function, at a fixed range of 50 m, for the observed bivariate Barro Colorado Island point pattern (black line), along with the corresponding 90% directional-quantile maximum absolute difference envelopes obtained from a fitted multivariate geometric anisotropic log-Gaussian Cox process (LGCP; bottom row) and from a fitted multivariate isotropic LGCP (top row). Departure of the data from the adopted model is highlighted with red circles

of $R = 0.113$ for estimating the marginal Matérn parameters for *C. obtusifolia*, $R = 0.214$ for the corresponding marginal parameters for *S. radlkoferi*, and $R = 0.092$ for their interspecific dependence. We also implement the model fitting procedure using fixed $R \in \{0.1, 0.15, 0.2\}$. When reporting our fitted models, we use the classical intensity estimate $\hat{\lambda}_p = n_p/|W|$, $p = 1, 2$, in place of the corresponding Palm likelihood estimate, following the discussion in the previous section.

For the proof-of-concept studies in Section 5.1, we were able to avoid constraining the anisotropy parameters during the estimation procedure, and we picked values satisfying the validity conditions of Section 3.2. When fitting the model to observed data, however, we have no assurance that the true values satisfy these validity conditions. Instead of introducing any new constraints on the anisotropy parameters here, we acknowledge this uncertainty by checking each fitted model against Conditions 1–4; all of the fitted models we present here were found to satisfy the required validity conditions. In Section 5.1, we also found that estimating the Matérn parameters via constrained optimisation can result in underestimation of the overall power in the between-process covariance. This occurs when the estimated value of σ_{12} is equal to the upper bound specified by the marginal dependence structures. By calculating this upper bound explicitly, and comparing with $\hat{\sigma}_{12}$, we can therefore ascertain whether each fitted model accurately represents the between-species dependence structure; this is important, as it describes the interspecific interaction between individual trees in our dataset.

To begin with, we applied our model fitting procedure as in Section 5.1, with one change: we choose not to implement the ad-hoc sequential reduction of R, as a guard against degeneracy in the marginal scale and power parameters. Since this approach involves adjusting the estimation algorithm, based on the values of the observed data, this is likely to introduce a bias into the resulting estimates. In Section 5.1, this is implemented for only a small proportion of runs and so the effect of this bias is not evident in the results. The magnitude of this bias remains unquantified, however, and so it would be inappropriate to use this approach when performing inference based on a single realization.

The first four rows of Table 3 give the resulting parameter estimates, with each row specifying a multivariate geometric anisotropic Matérn model. When using $R = 0.15$ and $R = R_{data}$, we see that the method returned values of $\hat{\alpha}_{11}$ equal to $\alpha_{11}^{UB} = 10$, suggesting degeneracy in these fitted models. Furthermore, we see that for $R = 0.1$ and $R = 0.2$, the estimated value of $\hat{\sigma}_{12}$ is equal to the upper bound specified by the marginal parameter estimates, suggesting that the optimization has been over-constrained here. In all four cases, we conclude that the resulting fitted model does not accurately represent the interspecific interaction in our dataset. Motivated by the observation that distinct values of the marginal smoothness parameters lead to prohibitively small values of σ_{12}^{UB}, we next proceeded to fix the smoothness parameters, $\nu_{11} = \nu_{22} = \nu_{12} = 0.5$, resulting in a geometric anisotropic bivariate exponential covariance model. Crucially, all of the discussion from Sections 3 and 4 is valid for fixed values of ν_{pq}, $p, q = 1, 2$. The resulting parameter estimates for this model are given in rows 4–6 of Table 3. In order to demonstrate the utility of our multivariate anisotropic framework, we also fit an isotropic version of the multivariate exponential LGCP to the same data, for the purpose of comparison. In practice, we achieve this by fixing $\zeta_{pq} = 1$ and $\theta_{pq} = 0$ for $p, q = 1, 2$, and implementing the MPLE portion of the model fitting procedure as described above, using $R = 0.1, 0.15, 0.2$ and $R = R_{data}$; under an assumption of isotropy, the appropriate values of R_{data} were found to be $R = 0.138$ for *C. obtusifolia*, $R = 0.085$ for *S. radlkoferi*, and $R = 0.07$ for their interspecific dependence. The resulting four sets of estimated scale and power parameters for this model are given in the bottom rows of Table 3. As is shown in this table, the interspecific interaction is well-represented in only one of the anisotropic models, and in two of the isotropic exponential models. In order to compare the fit of two models obtained with similar estimation methods, we now restrict our attention to the two models that used constant R across the marginal and bivariate Palm likelihoods.

In order to assess the ability of our fitted models to replicate the bivariate anisotropy present in the data, we use global envelope tests (GETs; Myllymäki et al., 2017), in which we compare

estimates of the multivariate sector K-function for the observed data with those for M bivariate point patterns, each independently simulated from the fitted model. We estimate the multivariate sector-K-function $K_{pq}^a(r, \phi)$ at $r = 0.05$, $\phi = k\pi/60$, $k = 0, \ldots, 60$, for $p, q \in \{1, 2\}$. We choose to use the sector-K-function as our test statistic, as it has been shown to result in tests with higher statistical power than those based on comparable directional test statistics, when testing for anisotropy in point patterns (Redenbach et al., 2009).

The envelopes provided by the GETs describe a proper statistical test: if the observed test statistic lies outside the simulated envelope *at any instance*, then the null hypothesis that the observed data belong to the fitted model may be rejected. We construct the envelopes by using the scaled directional quantile maximum absolute difference (MAD) to define the critical bounds; full details of this approach are provided by Myllymäki et al. (2017). We configure our tests such that they have a global type I error probability of 0.1, using $M = 499$.

The estimated sector-K-functions for the observed BCI data are presented in Figure 6, along with their corresponding directional quantile MAD envelopes. This figure demonstrates that the two chosen species in the BCI forest stand exhibit evidence of anisotropic interspecific interaction at a range of 50 m: for several values of $\phi \in [7\pi/60, 19\pi/60]$, the estimated sector-$K$-function $\hat{K}_{12}^a(0.05, \phi)$ lies outside of the envelope generated by a multivariate isotropic LGCP. The p-value for the global envelope test that was carried out for this statistic was 0.058, indicating departure from the isotropic model when using a global type I error probability of 0.1. From the bottom row of panels, we can see that our multivariate geometric anisotropic LGCP can comfortably replicate this observed heterogeneity. Finally, in Figure 7 we provide

TABLE 3 Parameter estimates for three bivariate log-Gaussian Cox processes (LGCPs), fitted to the tropical rainforest data described in the text.

$\hat{\theta}_{11}$	$\hat{\theta}_{22}$	$\hat{\theta}_{12}$		$\hat{\xi}_{11}$	$\hat{\xi}_{22}$	$\hat{\xi}_{12}$	$\hat{\lambda}_1$	$\hat{\lambda}_2$
158.90°	87.64°	127.39°		0.51	0.39	0.53	824	878
Covariance model	R	$\hat{\alpha}_{11}$	$\hat{\alpha}_{22}$	$\hat{\alpha}_{12}$	$\hat{v}_{11}$	$\hat{v}_{22}$	$\hat{v}_{12}$	$\hat{\sigma}_{11}$	$\hat{\sigma}_{22}$	$\hat{\sigma}_{12}$
---	---	---	---	---	---	---	---	---	---	---
Anisotropic Matérn	0.1	0.03	0.71	0.15	0.05	0.5	0.5	11.54	13.65	1.16*
Anisotropic Matérn	0.15	10.00*	0.08	0.11	0.05	5.0	5.0	19.05	3.09	1.01e-07*
Anisotropic Matérn	0.2	0.07	0.10	0.14	0.05	5.0	5.0	12.50	2.00	0.03*
Anisotropic Matérn	R_{data}	10.00*	0.08	0.12	0.05	0.05	0.05	17.95	9.16	8.80*
Anisotropic exponential	0.1	0.03	0.71	0.04	—	—	—	3.09	13.65	1.61*
Anisotropic exponential	**0.15**	**0.10**	**0.18**	**0.12**	—	—	—	**3.47**	**5.22**	**2.45**
Anisotropic exponential	0.2	0.07	0.17	0.08	—	—	—	3.30	3.03	1.96*
Anisotropic exponential	R_{data}	0.05	0.08	0.05	—	—	—	3.03	2.41	1.89*
Isotropic exponential	0.1	0.03	0.66	0.04	—	—	—	3.32	7.42	1.45*
Isotropic exponential	0.15	0.08	0.13	0.07	—	—	—	3.34	2.79	2.14*
Isotropic exponential	**0.2**	**0.12**	**0.09**	**0.10**	—	—	—	**4.28**	**3.40**	**2.62**
Isotropic exponential	R_{data}	0.06	0.23	0.07	—	—	—	3.24	5.01	2.34

Notes: The anisotropy parameters in the top table apply to both anisotropic models described in the bottom table. Those values of $\hat{\alpha}_{11}$ and $\hat{\sigma}_{12}$ marked with an asterisk (*) are equal to the corresponding upper bounds, α_{11}^{UB} and σ_{12}^{UB}, respectively. The two models that we choose to assess using global envelope tests, are highlighted in bold.
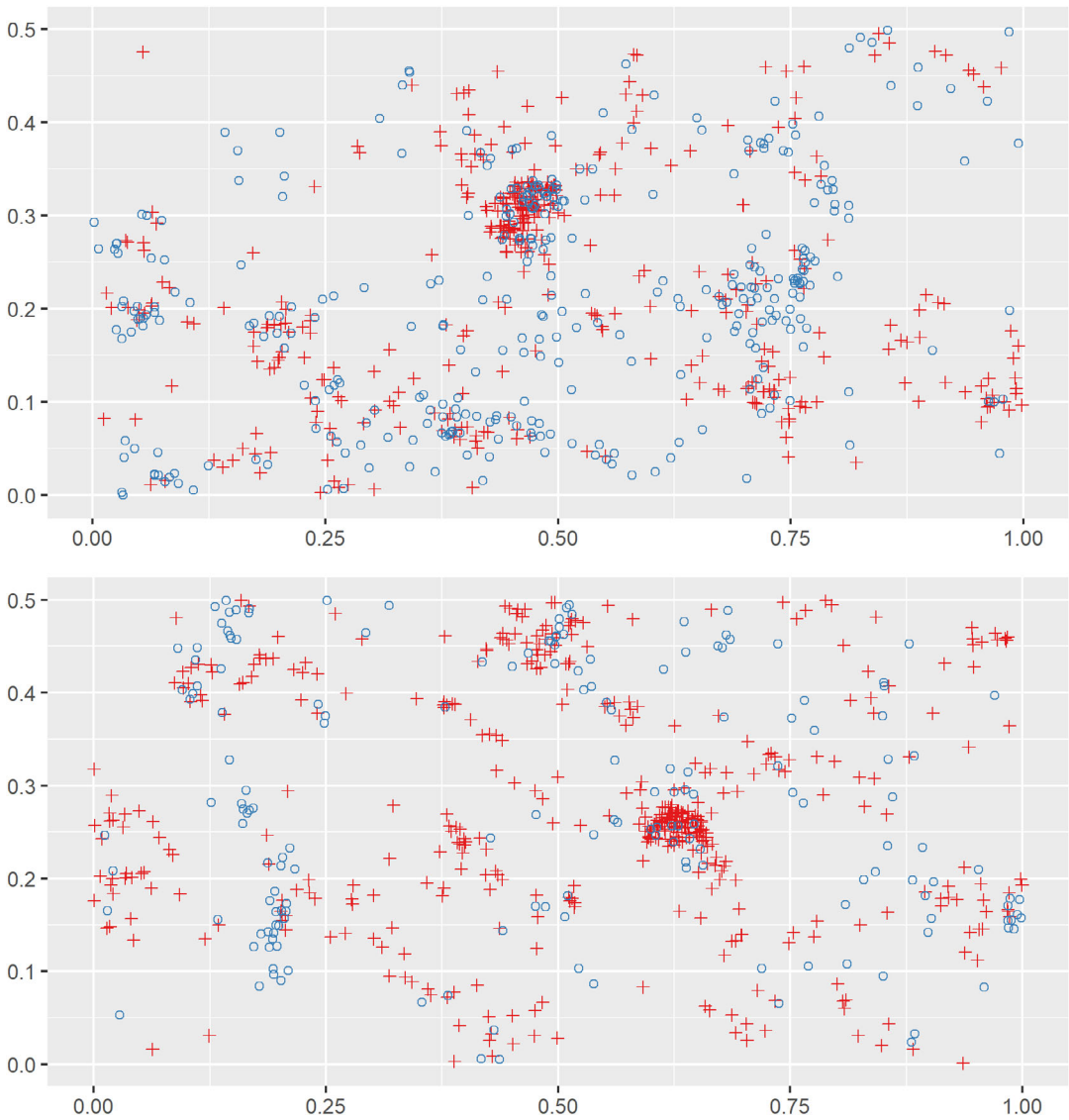
FIGURE 7   Single realizations from each of the fitted isotropic (top) and anisotropic (bottom) log-Gaussian Cox processes (LGCPs), for which the global envelope tests are implemented. The marginal processes represented here correspond to the two BCI species we consider: *Cecropia obtusifolia* (red crosses) and *Spondias radlkoferi* (blue circles).

one realization from each of the fitted isotropic and anisotropic models for which we carry out the GETs.

6 | DISCUSSION

We have shown that by incorporating geometric anisotropy into the between-process dependence, as well as the marginal dependence, we can construct a log-Gaussian Cox process (LGCP) that

more accurately replicates any rotationally heterogeneous interaction between points in a multivariate point pattern. We have focused on a covariate-free approach, motivated in part by the desire to allow the description of anisotropic between-process dependence in data for which there are no explanatory spatial variables. Nevertheless, the models presented are flexible enough to use (potentially incomplete) covariate information where it is available. Indeed, an interesting first extension of this work would be to incorporate covariates into the first-order description of the GRF underlying our LGCPs. For instance, the expected value of the GRF could be specified through a linear regression model, and inference with respect to the regression parameters may be achievable through the use of estimating functions (e.g. Waagepetersen, 2008; Waagepetersen & Guan, 2009). Such an approach would allow the user to exploit any knowledge of spatial covariates while being confident that any residual heterogeneity in the data would be accounted for by the increased flexibility of the multivariate geometric anisotropic second-order dependence structure.

ACKNOWLEDGMENTS

ORCID

James S. Martin https://orcid.org/0000-0002-2555-2026

REFERENCES

Abramowitz, M., & Stegun, I. (1965). *Handbook of mathematical functions with fomulas, graphs and mathematical tables*. Dover Publications Inc.

Apanasovich, T. V., & Genton, M. G. (2010). Cross-covariance functions for multivariate random fields based on latent dimensions. *Biometrika*, *97*, 15–30.

Apanasovich, T. V., Genton, M. G., & Sun, Y. (2012). A valid Matérn class of cross-covariance functions for multivariate random fields with any number of components. *Journal of the American Statistical Association*, *107*, 180–193.

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

Binder, B. J., & Simpson, M. J. (2015). Spectral analysis of pair-correlation bandwidth: Application to cell biology images. *Royal Society Open Science*, *2*, 140494.

Bortot, P., Coles, S. G., & Sisson, S. A. (2009). Inference for stereological extremes. *Journal of the American Statistical Association*, *102*, 84–92 http://www.jstor.org/stable/27639822

Byrd, R. H., Hribar, M. E., & Nocedal, J. (1999). An interior point algorithm for large-scale nonlinear programming. *SIAM Journal on Optimization*, *9*, 877–900.

Coeurjolly, J.-F., Møller, J., & Waagepetersen, R. (2017). A tutorial on palm distributions for spatial point processes. *International Statistical Review*, *85*, 404–420.

Condit, R. (1998). *Tropical forest census plots: Methods and results from Barro Colorado Island, Panama and a comparison with other plots*. Springer Science & Business Media.

Cox, D. R. (1958). Some problems connected with statistical–inference. *Annals of Mathematical Statistics*, *29*, 357–372.

Cramér, H. (1945). *Mathematical methods of statistics*. Almqvist & Wiksells.

Daley, D. J., & Vere-Jones, D. (2008). *An introduction to the theory of point processes General theory and structure* (Vol. *II*). Springer.

Diggle, P. J., & Milne, R. K. (1983). Bivariate Cox processes: Some models for bivariate spatial point patterns. *Journal of the Royal Statistical Society: Series B (Methodological)*, *45*, 11–21.

Diggle, P. J., Moraga, P., Rowlingson, B., & Taylor, B. M. (2013). Spatial and spatio-temporal log-Gaussian Cox processes: Extending the geostatistical paradigm. *Statistical Science*, *28*, 542–563.

Dvořák, J., & Prokešová, M. (2012). Moment estimation methods for stationary spatial Cox processes - A comparison. *Kybernetika*, *48*, 1007–1026.

Fry, N. (1979). Random point distributions and strain measurements in rocks. *Tectonophysics*, *60*, 89–105.

Genton, M. G., & Kleiber, W. (2015). Cross-covariance functions for multivariate Geostatistics. *Statistical Science*, *30*, 147–163.

Gneiting, T., Kleiber, W., & Schlather, M. (2010). Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association*, *105*, 1167–1177.

Goff, J. A., & Jordan, T. H. (1988). Stochastic modeling of seafloor morphology: Inversion of sea beam data for second-order statistics. *Journal of Geophysical Research*, *93*, 13589–13608.

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

Guttorp, P., & Gneiting, T. (2006). Studies in the history of probability and statistics XLIX: On the Matérn correlation family. *Biometrika*, *93*, 989–995.

Haase, P. (2001). Can isotropy vs. anisotropy in the spatial association of plant species reveal physical vs. biotic facilitation? *Journal of Vegetation Science*, *12*, 127–136.

Handcock, M. S., & Wallis, J. R. (1994). An approach to statistical spatial-temporal Modeling of meteorological fields. *Journal of the American Statistical Association*, *89*, 368–378.

Hubbell, S. P., Condit, R., & Foster, R. (2010). *Barro Colorado forest census plot data*. http://ctfs.si.edu/webatlas/datasets/bci

Hubbell, S. P., Foster, R. B., O'Brien, S. T., Harms, K., Condit, R., Wechsler, B., Wright, S. J., & De Lao, S. L. (1999). Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. *Science*, *283*, 554–557.

Illian, J., Penttinen, A., Stoyan, H., & Stoyan, D. (2008). *Statistical analysis and modelling of spatial point patterns* (Vol. *70*). John Wiley & Sons.

Jalilian, A., & Waagepetersen, R. (2018). Fast bandwidth selection for estimation of the pair correlation function. *Journal of Statistical Computation and Simulation*, *88*, 2001–2011.

Kleiber, W., & Nychka, D. (2012). Nonstationary modeling for multivariate spatial processes. *Journal of Multivariate Analysis*, *112*, 76–91.

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

Kupers, S. J., Wirth, C., Engelbrecht, B. M. J., & Rüger, N. (2019). Dry season soil water potential maps of a 50 hectare tropical forest plot on Barro Colorado Island, Panama. *Scientific Data*, *6*, 63.

Li, W., & Fearnhead, P. (2018). On the asymptotic efficiency of approximate Bayesian computation estimators. *Biometrika*, *105*, 285–299.

Møller, J., Syversveen, A. R., & Waagepetersen, R. P. (1998). Log Gaussian Cox processes. *Scandinavian Journal of Statistics*, *25*, 451–482.

Møller, J., & Toftaker, H. (2014). Geometric anisotropic spatial point pattern analysis and Cox processes. *Scandinavian Journal of Statistics*, *41*, 414–435.

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

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

Price, L. F., Drovandi, C. C., Lee, A., & Nott, D. J. (2018). Bayesian synthetic likelihood. *Journal of Computational and Graphical Statistics*, *27*, 1–11.

Prokešová, M., & Vedel Jensen, E. (2013). Asymptotic palm likelihood theory for stationary point processes. *Annals of the Institute of Statistical Mathematics*, *65*, 387–412.

Rajala, T. A., Murrell, D. J., & Olhede, S. C. (2018). Detecting multivariate interactions in spatial point patterns with gibbs models and variable selection. *Journal of the Royal Statistical Society Series C (Applied Statistics)*, 1237-1273.

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

Rajala, T. A., Särkkä, A., Redenbach, C., & Sormani, M. (2016). Estimating geometric anisotropy in spatial point patterns. *Spatial Statistics*, *15*, 100–114.

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

Sherman, M. (2011). *Spatial statistics and spatio-temporal data: Covariance functions and directional properties*. John Wiley & Sons.

Sisson, S. A., Fan, Y., & Tanaka, M. M. (2007). Sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences*, *104*, 1760–1765.

Sormani, M., Redenbach, C., Särkkä, A., & Rajala, T. (2020). Second order analysis of geometric anisotropic point processes revisited. *Spatial Statistics*, *38*, 100456.

Stein, M. L. (1999). *Interpolation of spatial data: Some theory for kriging*. Springer.

Tanaka, U., Ogata, Y., & Stoyan, D. (2008). Parameter estimation and model selection for Neyman-Scott point processes. *Biometrical Journal*, *50*, 43–57.

Veen, A., & Schoenberg, F. P. (2006). *Assessing spatial point process models using weighted K-functions: Analysis of California earthquakes*. In A. Baddeley, P. Gregori, J. Mateu, R. Stoica, & D. Stoyan (Eds.), *Case studies in spatial point process modeling* (pp. 293–306). Springer.

Waagepetersen, R. (2008). Estimating functions for inhomogeneous spatial point processes with incomplete covariate data. *Biometrika*, *95*, 351–363.

Waagepetersen, R., & Guan, Y. (2009). Two-step estimation for inhomogeneous spatial point processes. *Journal of the Royal Statistical Society: Series B (Methodological)*, *71*, 685–702.

Waagepetersen, R., Guan, Y., Jalilian, A., & Mateu, J. (2016). Analysis of multispecies point patterns by using multivariate log-Gaussian Cox processes. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *65*, 77–96.

Zhang, H. (2004). Inconsistent estimation and asymptotically equal interpolations in model-based Geostatistics. *Journal of the American Statistical Association*, *99*, 250–261.

APPENDIX

A.1 Proof of Proposition 1

In Proposition 1, we state that Conditions 1–4 are sufficient for the geometric anisotropic Matérn function in (6) to specify a valid multivariate covariance model, and we sketch the proof here. This proof is similar to that of Theorem 1 of Apanasovich et al. (2012), with additional consideration required to account for geometric anisotropy. As such, our proof depends on the following lemmas, due to Apanasovich et al. (2012), proofs for which can be found in that paper.

Lemma 1. *(Apanasovich et al., 2012). Let $0 < b_p < \infty$, $p = 1,\ldots,P$, $\delta \geq 0$, and $B_{pq} > 0$, $p, q = 1,\ldots,P$, be such that the matrix $(-B_{pq})_{p,q=1}^{P}$ is conditionally nonnegative definite. Then the $P \times P$ matrix with entries*

$$\frac{\Gamma(b_p + b_q + \delta)}{B_{pq}^{b_p+b_q+\delta}} \qquad p, q = 1,\ldots P,$$

is nonnegative definite.

Lemma 2. *(Apanasovich et al., 2012). Let $\delta \geq 0$ and B_{pq}, $p, q = 1,\ldots,P$ be as in Lemma 1. Then the $P \times P$ matrix with (p,q)th entry $(B_{pq} + \delta)^{-r}$ is nonnegative definite, for any $0 < r < \infty$.*

Proof of Proposition 1. We operate in the spectral domain: by Cramér's generalization of Bochner's Theorem (Cramér, 1945), the covariance matrix $\left(C_{pq}(h)\right)_{p,q=1}^{P}$ is nonnegative definite if and only if the corresponding matrix of spectral densities $\left(f_{pq}(\omega)\right)_{p,q=1}^{P}$ is also nonnegative definite. We therefore consider the form of the multivariate spectral density function, corresponding to (6):

$$f_{pq}(\omega) = |\Sigma_{pq}|^{1/2} f_{0,pq}\left(\Sigma_{pq}^{1/2}\omega\right)$$

$$= \frac{|\Sigma_{pq}|^{1/2}\sigma_{pq}\Gamma(\nu_{pq} + d/2)}{\pi^{d/2}\Gamma(\nu_{pq})}\left(\frac{4\nu_{pq}}{\alpha_{pq}^2}\right)^{\nu_{pq}}\left(\frac{4\nu_{pq}}{\alpha_{pq}^2} + \|\Sigma_{pq}^{1/2}\omega\|^2\right)^{-\nu_{pq}-d/2},$$

where each anisotropic deformation matrix Σ_{pq} is defined in terms of θ_{pq} and ζ_{pq}. We can decompose this spectrum as follows, in the process defining four terms numbered I to IV:

$$f_{pq}(\omega) = \overbrace{\frac{\Gamma\left(\frac{\nu_{pp}+\nu_{qq}}{2} + \frac{d}{2}\right)}{\left(\frac{4\nu_{pq}}{\alpha_{pq}^2} + \|\Sigma_{pq}^{1/2}\omega\|^2\right)^{\frac{\nu_{pp}+\nu_{qq}}{2}+\frac{d}{2}}}}^{\text{Term I}} \times \overbrace{\left(\frac{\frac{4\nu_{pq}}{\alpha_{pq}^2}}{\frac{4\nu_{pq}}{\alpha_{pq}^2} + \|\Sigma_{pq}^{1/2}\omega\|^2}\right)^{-\Delta_\nu A_{\nu,pq}}}^{\text{Term II}}$$

$$\times \underbrace{\frac{1}{\left(\frac{4\nu_{pq}}{\alpha_{pq}^2} + \|\Sigma_{pq}^{1/2}\omega\|^2\right)^{\Delta_\nu}}}_{\text{Term III}} \times \underbrace{\frac{|\Sigma_{pq}|^{1/2}\sigma_{pq}\Gamma(\nu_{pq} + d/2)}{\pi^{d/2}\Gamma\left(\frac{\nu_{pp}+\nu_{qq}}{2} + \frac{d}{2}\right)\Gamma(\nu_{pq})}\left(\frac{4\nu_{pq}}{\alpha_{pq}^2}\right)^{\Delta_\nu + \frac{\nu_{pp}+\nu_{qq}}{2}}}_{\text{Term IV}}, \quad \text{(A1)}$$

where $A_{\nu,pq} = 1 - \left\{\nu_{pq} - \left(\nu_{pp} + \nu_{qq}\right)/2\right\}/\Delta_\nu$ is the (p,q)-element of a valid nonnegative correlation matrix; nonnegative definiteness of the spectral matrix follows from nonnegative definiteness of the matrices formed from these constituent terms.

As noted by Apanasovich et al. (2012), a Hermitian matrix $A \in \mathbb{C}^{P \times P}$ is conditionally nonnegative definite if and only if the matrix $\left(\exp\left\{rC_{pq}\right\}\right)_{p,q=1}^{P}$ is nonnegative definite for all $r \geq 0$. Therefore, using the Schur product theorem,

we can combine the matrices in Conditions 2 and 4, to state that the matrix $\left(\exp\left\{ -r\left(\frac{4v_{pq}}{\alpha_{pq}^2} + \|\Sigma_{pq}^{1/2}\omega\|^2 \right) \right\} \right)_{p,q=1}^{P}$ is nonnegative definite for all $r \geq 0$, and thus the matrix $\left(-\frac{4v_{pq}}{\alpha_{pq}^2} - \|\Sigma_{pq}^{1/2}\omega\|^2 \right)_{p,q=1}^{P}$ is conditionally nonnegative definite. This allows us to use Lemma 1, with $b_p = v_{pp}/2$, $b_q = v_{qq}/2$ and $\delta = d/2$ to guarantee nonnegative definiteness of the matrix with elements given by the first term in (A1).

Conditions 1,2, and 4 are sufficient to guarantee nonnegative definiteness of the matrix with elements given by the second term of (A1). To see this, we first rewrite the second term in (A1) as

$$
\left(\frac{\frac{4v_{pq}}{\alpha_{pq}^2}}{\frac{4v_{pq}}{\alpha_{pq}^2} + \|\Sigma_{pq}^{1/2}\omega\|^2} \right)^{-\Delta_v A_{v,pq}} = \exp\left\{ \Delta_v A_{v,pq}\left[-\log\left(1 - \frac{\|\Sigma_{pq}^{1/2}\omega\|^2}{\frac{4v_{pq}}{\alpha_{pq}^2} + \|\Sigma_{pq}^{1/2}\omega\|^2} \right) \right] \right\}
$$

$$
= \prod_{k=1}^{\infty} \exp\left\{ \frac{\Delta_v A_{v,pq}}{k}\left[\frac{\|\Sigma_{pq}^{1/2}\omega\|^2}{\frac{4v_{pq}}{\alpha_{pq}^2} + \|\Sigma_{pq}^{1/2}\omega\|^2} \right]^{k} \right\},
$$

where we note that the infinite expansion of the logarithm is valid when

$$
\frac{\|\Sigma_{pq}^{1/2}\omega\|^2}{\frac{4v_{pq}}{\alpha_{pq}^2} + \|\Sigma_{pq}^{1/2}\omega\|^2} < 1,
$$

and this is satisfied at all times, since $4v_{pq}/\alpha_{pq}^2 > 0$.

Now, consider the matrices B and C with elements $B_{pq} > 0$, $C_{pq} > 0$, $p, q = 1, \ldots, P$ and suppose that both $-B$ and $-C$ are conditionally nonnegative definite. By applying Lemma 2 (with $\delta = 0$, $r = 1$), we have that the matrix with elements $1/B_{pq}$ is nonnegative definite, and therefore by the Schur product theorem, we have that the matrix with elements $-C_{pq}/B_{pq}$ is conditionally nonnegative definite. Now, using the matrices in Conditions 2 and 4 in place of the matrices $-C$ and $-B$, respectively, we can state that the matrix with elements

$$
-\frac{4v_{pq}/\alpha_{pq}^2}{\|\Sigma_{pq}^{1/2}\omega\|^2},
$$

is conditionally nonnegative definite for all $\omega \in \mathbb{R}^d \setminus \{0\}$, and we note that for $\omega = 0$, the second term collapses to 1. By applying Lemma 2 once more (this time with $\delta = 1$), we therefore have that the matrix with elements

$$
\left(\frac{4v_{pq}/\alpha_{pq}^2}{\|\Sigma_{pq}^{1/2}\omega\|^2} + 1 \right)^{-r},
$$

is nonnegative definite for all $r > 0$. It is now clear that, since A_ν is nonnegative definite and $\Delta_\nu \geq 0$ (both by Condition 1), each exponential argument within the product above specifies a nonnegative definite matrix. Repeated further use of the Schur product theorem therefore allows us to conclude that the matrix with elements given by the second term in (A1) is indeed nonnegative definite.

Recall from our examination of the first term that the matrix with (p, q)-element $-\frac{4\nu_{pq}}{\alpha_{pq}^2} - \|\Sigma^{1/2}\omega\|^2$ is conditionally nonnegative definite. For $\Delta_\nu > 0$, we can therefore use Lemma 2 to guarantee the nonnegative definiteness of the third term, and we note that for $\Delta_\nu = 0$, the third term trivially collapses to 1. Finally, Condition 3 states the nonnegative definiteness of the matrix with entries specified by the fourth term of (A1), and so we may conclude the stated result. ∎

A.2 Proofs of Remarks 2–4

In Remarks 2–4, we provide definitions of the correlation length, smoothness parameter and spatial deformation matrix for the geometric anisotropic Matérn cross-covariance function $C_{pq}\left(h \,|\, \alpha_{pq}, \nu_{pq}, \sigma_{pq}, \Sigma_{pq}\right)$, in terms of the corresponding marginal quantities. In this subsection, we prove that these definitions satisfy Conditions 2–4, respectively. Recall that a matrix $A \in \mathbb{C}^{P \times P}$ is conditionally nonnegative definite if, for all $x \in \mathbb{C}^P$ such that $\sum_{p=1}^P x_p = 0$, $\sum_{p,q=1}^P x_p A_{pq} x_q^* \geq 0$, where x_p^* is the complex conjugate of x_p.

Proof of Remark 2. This proof is given in the appendix of Apanasovich et al. (2012) for a different Matérn parameterization; we translate it to the current Matérn parameterization here. Suppose that

$$\frac{4\nu_{pq}}{\alpha_{pq}^2} = \frac{1}{2}\left(\frac{4\nu_{pp}}{\alpha_{pp}^2} + \frac{4\nu_{qq}}{\alpha_{qq}^2}\right) + \Delta_\alpha\left(1 - A_{\alpha,pq}\right), \quad p, q = 1,\ldots,P, \tag{A2}$$

with $\Delta_\alpha \geq 0$ and $0 \leq A_{\alpha,pq} \leq 1$ that form a valid correlation matrix. Consider $x \in \mathbb{C}^P$ such that $\sum_{p=1}^P x_p = 0$. Using (A2),

$$\begin{aligned}
\sum_{p,q} x_p \frac{4\nu_{pq}}{\alpha_{pq}^2} x_q^* &= \frac{1}{2}\left\{\left(\sum_p x_p \frac{4\nu_{pp}}{\alpha_{pp}^2}\right)\left(\sum_q x_q^*\right) + \left(\sum_p x_p\right)\left(\sum_q \frac{4\nu_{qq}}{\alpha_{qq}^2} x_q^*\right)\right\} \\
&\quad + \Delta_\alpha \sum_p x_p \sum_q x_q^* - \Delta_\alpha \sum_{pq} x_p A_{\alpha,pq} x_q^* \\
&= -\Delta_\alpha \sum_{pq} x_p A_{\alpha,pq} x_q^* \leq 0, \quad \text{as } A_{\alpha,pq} \text{ is nonnegative definite.}
\end{aligned}$$

Hence, the matrix with (p, q)-element $-4\nu_{pq}/\alpha_{pq}^2$ is conditionally nonnegative definite. ∎

Proof of Remark 3. Through straightforward manipulation of the expression in Remark 3, we see that the matrix in Condition 3 is equal to the matrix with (p, q)-element given by $V_p V_q A_{\sigma,pq}$, where V_p, V_q, and A_σ are defined in Remark 3. Since A_σ is a (nonnegative definite) correlation matrix, and since $V_p, V_q \geq 0$, this is also nonnegative definite. ∎

Proof of Remark 4. We also give motivation for the chosen construction of Σ_{pq}. We wish to have Σ_{pq} such that the $P \times P$ matrix with (p, q)-element $-\omega^T \Sigma_{pq} \omega$ is conditionally nonnegative definite. Now, for a $P \times P$ matrix with (p, q)-element $-C_{pq}$ to be nonnegative definite, a necessary condition is for $C_{pq} \geq \left(C_{pp} + C_{qq}\right)/2$, for $p, q = 1,\ldots, P$. It therefore follows that for Condition 4 to hold, we need

$$\omega^T \Sigma_{pq} \omega \geq \frac{1}{2} \left(\omega^T \Sigma_{pp} \omega + \omega^T \Sigma_{qq} \omega\right), \qquad \forall \omega \in \mathbb{R}^2.$$

Since this must hold for all $\omega \in \mathbb{R}^2$, we can consider the particular case for $\left\{\omega \in \mathbb{R}^2 : \omega_2 = 0\right\}$, from which we can deduce $\left[\Sigma_{pq}\right]_{11} \geq \left(\left[\Sigma_{pp}\right]_{11} + \left[\Sigma_{qq}\right]_{11}\right)/2$. Similarly, we can deduce $\left[\Sigma_{pq}\right]_{22} \geq \left(\left[\Sigma_{pp}\right]_{22} + \left[\Sigma_{qq}\right]_{22}\right)/2$. This motivates the construction of the diagonal elements of Σ_{pq} in Remark 4:

$$\left[\Sigma_{pq}\right]_{ii} = \frac{1}{2} \left[\Sigma_{pp} + \Sigma_{qq}\right]_{ii} + \Delta_\Sigma^{(i)} \left(1 - A_{\Sigma, pq}^{(i)}\right), \qquad i = 1, 2,$$

where each $A_\Sigma^{(i)}$ is a $P \times P$ correlation matrix and each $\Delta_\Sigma^{(i)}$ is a nonnegative constant. Now, consider $x \in \mathbb{C}^P$ such that $\sum_{p=1}^P x_p = 0$. We wish to show that

$$\sum_{p,q} x_p \left(\omega^T \Sigma_{pq} \omega\right) x_q^* = \omega^T \left(\sum_{p,q} x_p \Sigma_{pq} x_q^*\right) \omega \leq 0 \qquad \forall \omega \in \mathbb{R}^2.$$

By expanding the above quadratic in ω, and then substituting our chosen construction for the diagonal elements, we can simplify to obtain

$$\omega^T \left(\sum_{p,q} x_p \Sigma_{pq} x_q^*\right) \omega = -\Delta_\Sigma^{(1)} B_\Sigma^{(1)} \omega_1^2 - \Delta_\Sigma^{(2)} B_\Sigma^{(2)} \omega_2^2 + 2\omega_1 \omega_2 \left(\sum_{p,q} x_p \left[\Sigma_{pq}\right]_{12} x_q^*\right),$$

where, for $i = 1, 2$, $B_\Sigma^{(i)} = \left(\sum_{pq} x_p A_{\Sigma, pq}^{(i)} x_q^*\right)$ is nonnegative, as $A_\Sigma^{(i)}$ is a correlation matrix. In order for this quadratic term to maintain the same sign for all $\omega \in \mathbb{R}^2$, we must be able to factorise it further, that is, we must be able to write

$$\omega^T \left(\sum_{p,q} x_p \Sigma_{pq} x_q^*\right) \omega = k_1 (\omega_1 \pm k_2 \omega_2)^2.$$

for some $k_1, k_2 \in \mathbb{R}$. By expanding and equating terms, it is straightforward to show that this form can be obtained: we can write

$$\omega^T \left(\sum_{p,q} x_p \Sigma_{pq} x_q^*\right) \omega = -\Delta_\Sigma^{(1)} B_\Sigma^{(1)} \left(\omega_1 \pm \omega_2 \sqrt{\frac{\Delta_\Sigma^{(2)} B_\Sigma^{(2)}}{\Delta_\Sigma^{(1)} B_\Sigma^{(1)}}}\right)^2,$$

if and only if the off-diagonal elements of Σ_{pq} satisfy the relationship

$$\left(\sum_{p,q=1}^P x_p \left[\Sigma_{pq}\right]_{12} x_q^*\right)^2 = \Delta_\Sigma^{(1)} B_\Sigma^{(1)} \Delta_\Sigma^{(2)} B_\Sigma^{(2)}. \tag{A3}$$

Note that this specifies a relationship between the diagonal and off-diagonal elements of the set of matrices $\{\Sigma_{pq}, p, q = 1, \dots, P\}$, which must be satisfied in order for the $P \times P$ matrix with (p, q)-element $-\omega^T \Sigma_{pq} \omega$ to be conditionally nonnegative definite.

Note that, since each Σ_{pq} is a deformation matrix Σ, its diagonal and off-diagonal elements must be consistent with the same choice of $(\theta_{pq}, \zeta_{pq})$. This places a fundamental restriction on the form of each Σ_{pq}. We can circumvent this apparent incompatibility of restrictions on the set of deformation matrices $\{\Sigma_{pq}, p, q = 1, \dots, P\}$ by writing the off-diagonal elements in the form

$$\left[\Sigma_{pq}\right]_{12} = \left[\Sigma_{pq}\right]_{21} = b_p + c_q + A^{(3)}_{\Sigma,pq}, \tag{A4}$$

where $b, c \in \mathbb{R}^P$ are constant P-length vectors, $\Delta^{(3)}_\Sigma$ is a nonnegative constant, and $A^{(3)}_\Sigma \in \mathbb{R}^{P \times P}$ is a $P \times P$ real matrix that satisfies

$$\left(\sum_{p,q=1}^P x_p A^{(3)}_{\Sigma,pq} x^*_q\right)^2 = \Delta^{(1)}_\Sigma B^{(1)}_\Sigma \Delta^{(2)}_\Sigma B^{(2)}_\Sigma.$$

By specifying the off-diagonal elements of Σ_{pq} in this way, we have that our conditional nonnegative definiteness restriction (A3) reduces to a restriction on $A^{(3)}_\Sigma$, which is unaffected by the need for Σ_{pq} to maintain the form of a valid deformation matrix, specified by Σ; since there are no further restrictions on the form of $A^{(3)}_\Sigma$, such a matrix will certainly exist.

Therefore, if the diagonal elements of the deformation matrix Σ_{pq} are specified as in Remark 4, the resulting off-diagonal elements will always satisfy a valid decomposition (A4), guaranteeing satisfaction of the relationship (A3). This allows us to conclude that, if the diagonal elements of the deformation matrix Σ_{pq} are specified as in Remark 4, the $P \times P$ matrix with (p, q)-element $-\omega^T \Sigma_{pq} \omega$ will be conditionally nonnegative definite. ∎

A.3 An estimation algorithm

In Section 4 we detail a two-stage procedure for estimating the anisotropy parameters and Matérn parameters in a multivariate geometric anisotropic LGCP. In Algorithm 1, we provide a detailed outline of this model fitting procedure, as applied to a bivariate geometric anisotropic LGCP. This description includes details of the bounds that are required in order to guarantee model validity when estimating the cross-dependence parameters, as well as a list of the tuning parameters that should be established prior to execution.

A.4 Estimators of second-order summary statistics

We present details of two estimators of second-order summary statistics that are used in our parameter estimation procedure. The first estimator we consider is for the isotropic cross-pair correlation function $g_{0,pq}(r)$, used in initializing the Matérn power and scale parameters:

$$\hat{g}_{0,pq}(r) = \sum_{\substack{x_p \in X_p \cap W \\ x_q \in X_q \cap W}}^{\neq} \frac{\kappa_{h_r}(\|x_p - x_q\| - r)}{2\pi r \hat{\lambda}_p \hat{\lambda}_q |W \cap W_{x_p - x_q}|}, \tag{A5}$$

Algorithm 1. An end-to-end procedure for estimating the parameters in a bivariate geometric anisotropic LGCP

Tuning Parameters: A maximum distance over which to characterise dependence, $r_{\max}$; an edge-correction parameter R; a set of contour levels L; the number of candidate anisotropy ratios n_ζ; and their maximum value $\zeta_{\max}$

Data: A bivariate point pattern (x_1, x_2), and an observation window W.

Output: Estimates of the following anisotropy and Matérn parameters, which describe the covariance structure of the multivariate GRF underlying the LGCP. For the pth marginal process ($p = 1, 2$),

- θ_{pp}, ζ_{pp} describe the angle and ratio of anisotropy, respectively;
- α_{pp}, ν_{pp} are the Matérn scale and smoothness parameters, resp., which jointly control the range of autocorrelation;
- σ_{pp} is the Matérn power parameter, that is, the variance of the pth marginal process; and
- μ_{pp} is the first moment of the corresponding marginal GRF.

In addition, we obtain the corresponding cross-covariance parameters:

- θ_{12}, ζ_{12} describe the angle and ratio of anisotropy, resp., for the cross-covariance;
- α_{12}, ν_{12} are the Matérn scale and smoothness parameters, resp., which jointly control the cross-correlation range;
- σ_{12} is the Matérn power parameter, that is, the colocated cross-covariance in the bivariate process.

1 **for** $p \in \{1, 2\}$ **do**
2   **for** $q \in \{p, 2\}$ **do**
3   Use the two-type point pattern $\{x_p, x_q\}$ to construct the symmetric Fry pattern
        $\{x_{p,i} - x_{q,j} ; \ i = 1, \ldots, n_p, j = 1, \ldots, n_q\} \cup \{x_{q,i} - x_{p,j} ; \ i = 1, \ldots, n_p, j = 1, \ldots, n_q\}$;
4   Dividing the polar plane into n_F sectors, use this Fry pattern to construct the ℓ-contours
        $\{G_\ell ; \ \ell = 1, \ldots, \ell_{\max}\}$, each of which contains the ℓth-nearest point to the origin, from each polar sector;
5   Using adjusted ordinary least squares, estimate the ellipse A_ℓ that fits each of the ℓ-contours G_ℓ;
6   Establish a subset of contour levels $L \subseteq \{1, \ldots, \ell_{\max}\}$ such that the fitted ellipses $\{\hat{A}_\ell, \ell \in L\}$ can be trusted to accurately describe the anisotropy in the original point pattern; see Rajala et al. (2018b) for an approach based on constructing MC confidence intervals.
7   Sample 5000 points from the surfaces of the *fitted* ellipses $\{\hat{A}_\ell, \ell \in L\}$; fit an ellipse to these points, and use the semi-axes of this final fitted ellipse to calculate $\hat{\theta}_{pq}$.
8   **for** $k \in \{1, \ldots, n_\zeta\}$ **do**
9   Using $\hat{\theta}_{pq}$ and the candidate value $\zeta_{k,pq} := k\zeta_{\max}/(1 + \zeta_{\max})$, isotropize x_p, x_q and the observation window W.
10   Using the isotropized point pattern(s) and observation window, estimate the directional discrepancy statistic $V_{pq, \hat{\theta}_{pq}}(\zeta_{k,pq})$, given in (8), using the tuning parameters b_1, b_2.
11   Set $\hat{\zeta}_{pq} = \underset{k \in 1, \ldots, n_\zeta}{\arg\min} \ \hat{V}_{pq, \hat{\theta}_{pq}}(\zeta_{k,pq})$

12 **for** $p \in \{1, 2\}$ **do**
13   Using $\hat{\theta}_{pp}$ and $\hat{\zeta}_{pp}$, isotropize x_p and the observation window W.
14   Estimate λ_{pp} as a function of $x_p, \alpha_{pp}, \nu_{pp}, \sigma_{pp}$ (and the tuning parameter R), by analytically maximising the Palm-log-likelihood in (10).
15   Substitute $\hat{\lambda}_{pp}(x_p, \alpha_{pp}, \nu_{pp}, \sigma_{pp})$ into the Palm-log-likelihood in (10), and numerically maximise the resulting expression to estimate α_{pp}, ν_{pp} and σ_{pp}. To avoid identifiability issues, we advise restricting $\hat{\nu}_{pp} \in \mathcal{N} := \{0.05, 0.5, 5.0\}$. α_{pp} and σ_{pq} can be optimized over bounded intervals: $\hat{\alpha}_{pp} \in (0, \alpha_{pp}^{UB})$, $\hat{\sigma}_{pq} \in (0, \sigma_{pp}^{UB})$. The numerical optimisation procedure can be initialized using moment-based method, for example, minimum contrast.
16   Evaluate $\hat{\lambda}_{pp}(x_p, \hat{\alpha}_{pp}, \hat{\nu}_{pp}, \hat{\sigma}_{pp})$, and hence estimate $\hat{\mu}_{pp} = \log(\hat{\lambda}_{pp}) - \hat{\sigma}_{pp}/2$.

17 **for** $p = 1, q = 2$ **do**
18   Use Condition 1 to calculate a lower bound for $\hat{\nu}_{pq}: \hat{\nu}_{pq} \geq \nu_{pq}^{LB} = (\hat{\nu}_{pp} + \hat{\nu}_{qq})/2$.
19   **for** *each candidate value* $\nu_{pq}^* \in \mathcal{N}$, *if* $\nu_{pq}^* \geq \nu_{pq}^{LB}$ **do**
20   Use Remark 2 to establish an upper bound for α_{pq} in terms of $\nu_{pq}^*, \hat{\alpha}_{pp}, \hat{\alpha}_{qq}, \hat{\nu}_{pp}$ and $\hat{\nu}_{qq}$; denote this α_{pq}^{UB}.
21   Writing $\Delta_\nu := \nu_{pq}^* - \nu_{pq}^{LB}$, use Remark 3 to evaluate $\hat{V}_p$ in terms of $\nu_{pq}^*, \hat{\sigma}_{pp}, \hat{\zeta}_{pp}, \hat{\alpha}_{pp}$; similarly, evaluate $\hat{V}_q$.
22   Use Remark 3 once more to establish an upper bound for σ_{pq} in terms of $\nu_{pq}^*, \hat{V}_p, \hat{V}_q, \hat{\zeta}_{pq}$ and $\hat{\alpha}_{pq}$.
23   By substituting $\hat{\lambda}_{pp}, \hat{\lambda}_{qq}$, and the candidate ν_{pq}^*, numerically maximise the bivariate Palm-log-likelihood in (11) over the constrained space $(0, \alpha_{pq}^{UB}) \times (0, \sigma_{pq}^{UB})$.
24   Set $\hat{\nu}_{pq}, \hat{\alpha}_{pq}, \hat{\sigma}_{pq}$ equal to the values that maximise the bivariate Palm-log-likelihood.

where κ_{h_r} is a radial kernel function with bandwidth h_r, $\hat{\lambda}_p$ is an estimator for the constant expected intensity of component X_p, λ_p, and $|W \cap W_u|$ is an edge correction factor, defined as the area of overlap between the observation window W and its translation by $u \in \mathbb{R}^2$; without such a correction, due to the finite observation region, the estimator would underestimate the number of point pairs that lie within distance r of each other. In (A5), and in the remainder of the paper, the notation $\Sigma^{\neq}$ indicates summation over all point pairs formed of distinct points; for bivariate definitions such as (A5), this is clearly only relevant for the case where $p = q$. For component p of our multivariate LGCP, we choose to estimate the expected intensity parameter $\hat{\lambda}_p$ using the classical global intensity estimator, $\hat{\lambda}_p = n_p/|W|$. The choice of kernel function κ_{h_r} is discussed by Illian et al. (2008) and common choices include the Epanechnikov kernel and the box kernel; we make use of the latter as it can be shown to minimize the variance of (A5):

$$\kappa_{h_r}(s) = \begin{cases} 1/2h_r & -h_r \leq s \leq h_r \\ 0 & \text{otherwise.} \end{cases}$$

In practice, we are faced with the choice of a suitable value for the bandwidth h_r, in order to evaluate (A5). Briefly, the bandwidth parameter controls the degree of smoothing in the corresponding kernel function; increasing the bandwidth will therefore reduce its variability, though at the cost of inducing a bias in the resulting estimator. Bandwidth selection for such estimators is an important topic, and is covered in depth by Illian et al. (2008). Indeed, methods for automatically selecting the bandwidth parameter are an area of active research, with recent contributions including the work of Guan (2007), Binder and Simpson (2015), and Jalilian and Waagepetersen (2018). In this work, we have chosen not to optimize the value of h_r, using a reasonably small value of $h_r = 0.01$. We have made this choice for the sake of simplicity; the resulting estimate of the isotropic pair correlation function is used only in initialising the search for the Matérn power and scale parameters, and we have found this choice of bandwidth to work sufficiently well in practice.

The second estimator that we detail here corresponds to the anisotropic sector-K-function $K_{pq}^a(r, \phi)$:

$$\hat{K}_{pq}^a(r, \phi) = \hat{K}_{pq}^a(r, \phi + \pi) = \sum_{\substack{x_p \in X_p \cap W \\ x_q \in X_q \cap W}}^{\neq} \frac{H(x_p - x_q, (r, \phi))}{\hat{\lambda}_p \hat{\lambda}_q |W \cap W_{x_p - x_q}|}, \tag{A6}$$

where

$$H(x_1 - x_2, (r, \phi)) = \mathbb{I}(\|x_1 - x_2\| \leq r)\kappa_{h_\phi}(\psi(x_1, x_2) - \phi),$$

with $\mathbb{I}(\cdot)$ the indicator function, κ_{h_ϕ} an angular kernel function with bandwidth h_ϕ, and $\psi(x_1, x_2)$ the angle between the directed line from x_1 to x_2 and the abscissa-axis. In our implementation, we will use a box kernel for κ_{h_ϕ}, defined analogously to the radial kernel function κ_{h_r} above. As discussed in Section 5 of the manuscript, we choose to use a angular bandwidth of $h_\phi = \pi/16$; we found that this choice of bandwidth worked well in practice, balancing the aforementioned tradeoff between estimator bias and variance.

A.5 Additional simulation studies

In Section 5 of the article, we provide proof-of-concept results for our model fitting procedure. There, we have given numerical summaries of the estimated parameter distributions

for four distinct model specifications, along with an illustration, in Figure 4, corresponding to one of these models. Here, we provide details of simulation studies corresponding to four further model specifications. We provide numerical summaries of the estimated parameter distributions for Models 5–8, in Tables A1 and A2, and one realization from each of these four further models is shown in Figure A1. Finally, we provide illustration of the estimated parameter distributions for each of Models 2–8. Figures A2–A8 correspond to Models 2–8, respectively.

The first marginal process for each of Models 5 and 7 corresponds to the two distinct univariate geometric anisotropic LGCPs considered by Møller and Toftaker (2014); this is with the exception of the Matérn scale parameter α_{11} in Model 7, which had to be decreased for computational reasons. For each of these models, the second marginal process was obtained by retaining the anisotropy and smoothness of the first marginal process, and significantly reducing (resp. increasing) the Matérn scale (resp. power) parameter. This has the effect of replicating bivariate point patterns in which one point type has relatively large, disperse clusters, and the second has small, concentrated clusters. The specification of the cross-covariance parameters was arrived at through the sequential procedure detailed in Section 3.2.

Models 5 and 7 demonstrate consistent anisotropic behavior across both marginals, while differing in their Matérn specification. Models 6 and 8 were constructed to demonstrate the opposite scenario. The Matérn parameters for both marginal processes in Model 6 (resp. Model 8) are the same as those used for the second marginal process in Model 5 (resp. Model 7), displaying relatively small, concentrated clusters. For both Model 6 and Model 8, the anisotropy specifications are such that one pattern contains strongly anisotropic features in the direction parallel to the y-axis, and the other contains weakly anisotropic features at an angle of $\pi/10$ from the x-axis. It is perhaps less realistic to consider two marginal processes with such contrasting anisotropies, however these examples serve to illustrate the flexibility of the modeling framework.

T A B L E   A1   Monte Carlo estimates and root-mean-square errors for the anisotropy parameters in four distinct models (Models 5-8).

	θ_{11}	θ_{22}	θ_{12}	ζ_{11}	ζ_{22}	ζ_{12}
Model 5	45°	45°	45°	0.20	0.20	0.35
MC estimate	45.19°	48.86°	69.18°	0.22	0.32	0.41
RMSE	8.21°	23.99°	49.44°	0.10	0.42	0.59
Model 6	18°	90°	45°	0.60	0.20	0.70
MC estimate	53.69°	89.83°	66.66°	0.59	0.30	0.60
RMSE	65.35°	22.64°	49.04°	0.32	0.49	0.37
Model 7	45°	45°	45°	0.60	0.60	0.60
MC estimate	50.41°	51.92°	83.33°	0.54	0.57	0.34
RMSE	30.46°	27.35°	62.06°	0.17	0.19	0.55
Model 8	18°	90°	45°	0.60	0.20	0.70
MC estimate	43.44°	90.25°	67.70°	0.56	0.24	0.61
RMSE	57.37°	17.57°	50.70°	0.20	0.31	0.24

TABLE A2 Monte Carlo estimates and root-mean-square errors for the mean and Matérn parameters in four distinct models (Models 5-8).

|	R	α_{11}	α_{22}	α_{12}	ν_{11}	ν_{22}	ν_{12}	σ_{11}	σ_{22}	σ_{12}	μ_1	μ_2
Model 5		0.200	0.050	0.065	0.5	0.5	0.5	1.99	2.62	0.80	5.76	5.44
MC estimate	$R = R_{\text{model}}$	0.552	0.064	0.094	5.0	0.5	5.0	2.24	3.50	0.35	3.72	3.90
	R=0.1	0.294	0.049	0.067	0.5	0.5	5.0	2.34	3.55	0.34	3.44	3.86
	R=0.2	0.314	0.060	0.072	0.5	0.5	0.5	2.32	4.53	0.40	3.50	3.42
	R=0.3	0.309	0.080	0.080	0.5	0.0	5.0	2.20	5.17	0.35	3.74	3.07
RMSE	$R = R_{\text{model}}$	1.620	0.411	0.268	NaN	NaN	NaN	2.22	2.93	1.67	3.12	2.13
	R=0.1	0.766	0.102	0.057	NaN	NaN	NaN	2.76	2.74	0.68	6.26	2.28
	R=0.2	0.773	0.423	0.050	NaN	NaN	NaN	1.29	4.30	1.02	4.75	2.63
	R=0.3	0.835	0.550	0.056	NaN	NaN	NaN	1.53	5.24	1.40	3.94	3.18
Model 6		0.050	0.050	0.050	0.5	0.5	0.5	2.62	2.62	1.25	5.44	5.44
MC estimate	$R = R_{\text{model}}$	0.076	0.037	0.042	0.5	0.5	0.5	3.05	3.13	0.85	4.61	4.15
	R=0.1	0.076	0.037	0.041	0.5	0.5	0.5	2.99	3.11	0.82	4.65	4.17
	R=0.2	0.054	0.092	0.046	0.5	0.5	0.5	3.96	4.81	1.25	4.19	3.28
	R=0.3	0.115	0.091	0.049	0.5	0.5	0.5	4.19	5.55	1.31	4.06	2.81
RMSE	$R = R_{\text{model}}$	0.408	0.021	0.021	NaN	NaN	NaN	2.62	2.39	1.04	2.33	1.66
	R=0.1	0.408	0.021	0.020	NaN	NaN	NaN	2.60	2.27	1.02	2.31	1.64
	R=0.2	0.052	0.704	0.024	NaN	NaN	NaN	3.15	5.24	3.09	2.01	2.77
	R=0.3	0.720	0.622	0.027	NaN	NaN	NaN	4.49	6.17	3.21	2.24	3.36
Model 7		0.250	0.100	0.130	5.0	5.0	5.0	1.06	2.00	0.15	6.22	5.75
MC estimate	$R = R_{\text{model}}$	1.252	0.140	0.132	5.0	5.0	5.0	1.75	1.93	0.06	3.92	4.81
	R=0.1	1.695	0.151	0.090	5.0	5.0	5.0	2.74	2.35	0.04	2.83	4.22
	R=0.2	0.732	0.144	0.122	5.0	5.0	5.0	2.08	1.92	0.07	3.60	4.80
	R=0.3	0.847	0.174	0.120	5.0	5.0	5.0	1.37	1.85	0.06	4.63	4.86
RMSE	$R = R_{\text{model}}$	2.628	0.189	0.240	NaN	NaN	NaN	1.76	0.97	0.31	5.88	1.95
	R=0.1	3.072	0.238	0.111	NaN	NaN	NaN	2.19	3.50	0.16	12.94	4.60
	R=0.2	1.704	0.240	0.164	NaN	NaN	NaN	1.40	0.81	0.15	8.50	2.08
	R=0.3	1.969	0.699	0.135	NaN	NaN	NaN	2.10	1.70	0.24	4.57	2.72
Model 8		0.100	0.100	0.100	5.0	5.0	5.0	2.00	2.00	0.97	5.75	5.75
MC estimate	$R = R_{\text{model}}$	0.165	0.116	0.105	5.0	5.0	5.0	2.02	2.08	0.58	4.66	4.10
	R=0.1	0.177	0.167	0.108	5.0	5.0	5.0	2.34	2.45	0.60	4.23	3.47
	R=0.2	0.161	0.117	0.107	5.0	5.0	5.0	2.02	2.06	0.57	4.64	4.09
	R=0.3	0.233	0.103	0.109	5.0	5.0	5.0	2.00	1.94	0.47	4.66	4.26
RMSE	$R = R_{\text{model}}$	0.427	0.140	0.038	NaN	NaN	NaN	0.58	1.35	0.69	3.71	2.70
	R=0.1	0.414	0.449	0.063	NaN	NaN	NaN	1.66	0.93	0.90	4.95	4.52
	R=0.2	0.442	0.145	0.042	NaN	NaN	NaN	0.63	1.18	0.70	4.10	2.84
	R=0.3	0.847	0.054	0.061	NaN	NaN	NaN	0.99	0.55	0.74	3.47	1.65
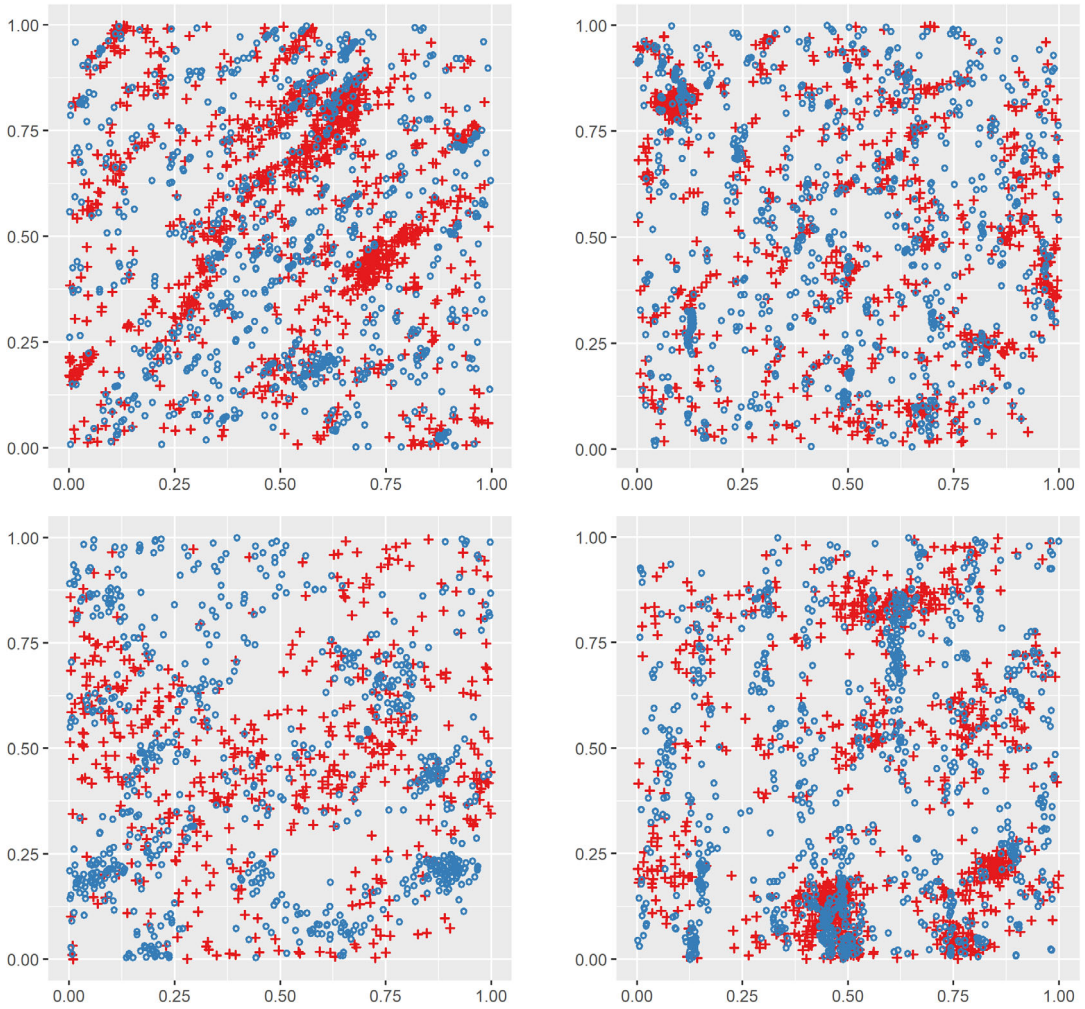
F I G U R E A1   A two-type point pattern, simulated from each of Models 5–8 (left to right, top to bottom), which are specified in Tables A1 and A2. In each realization, points of type 1 are shown as red crosses and points of type 2 are shown as blue circles.
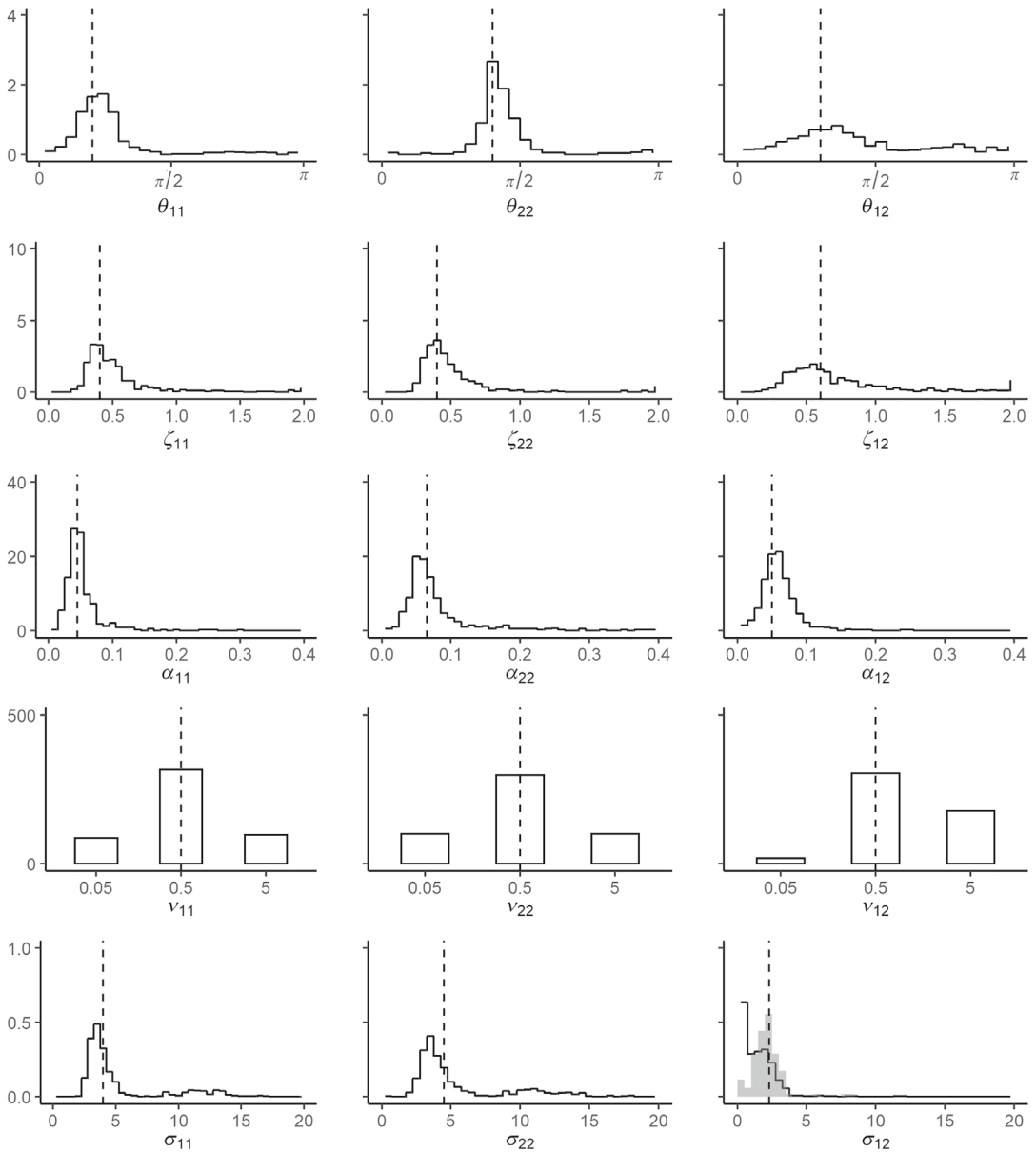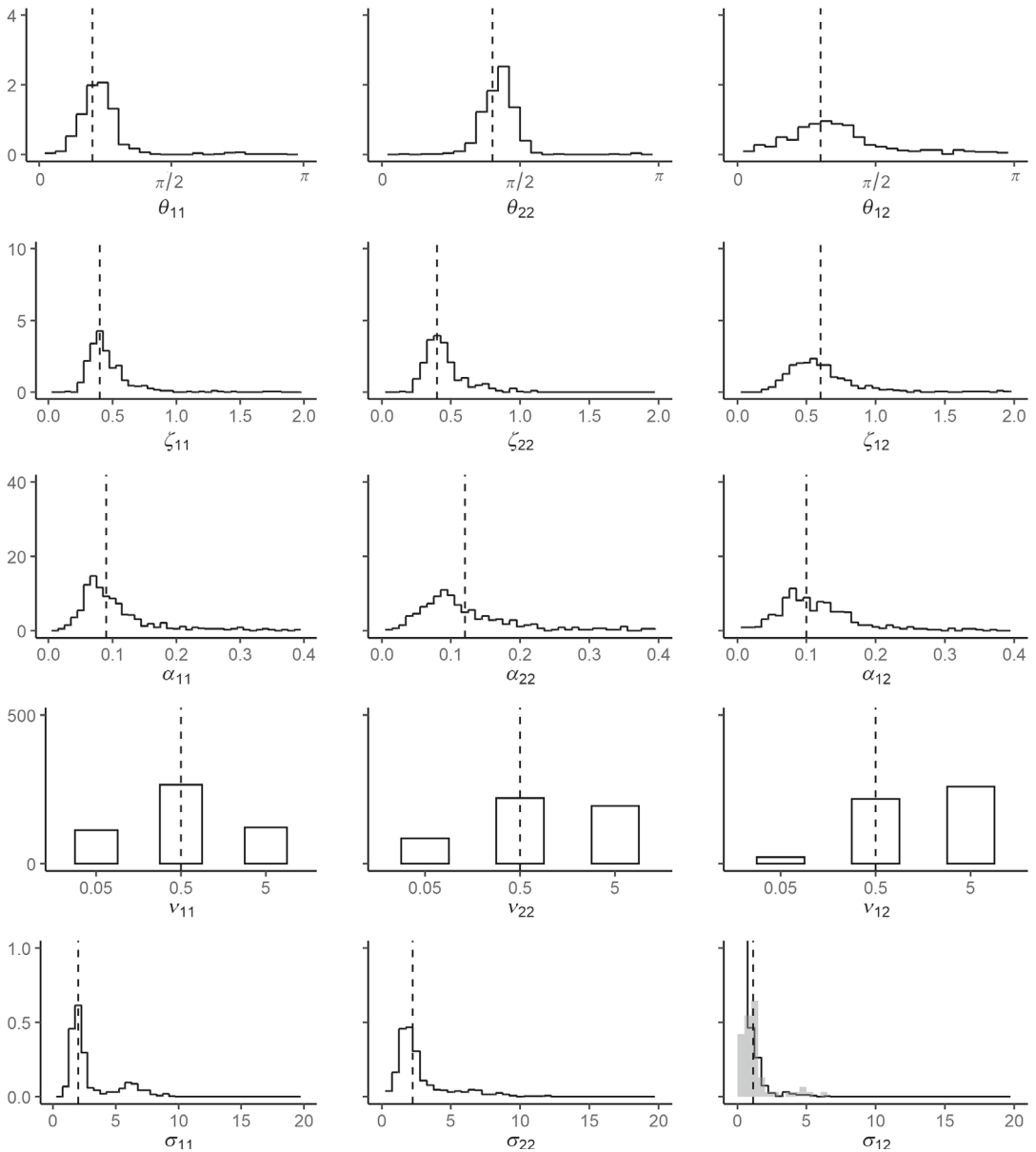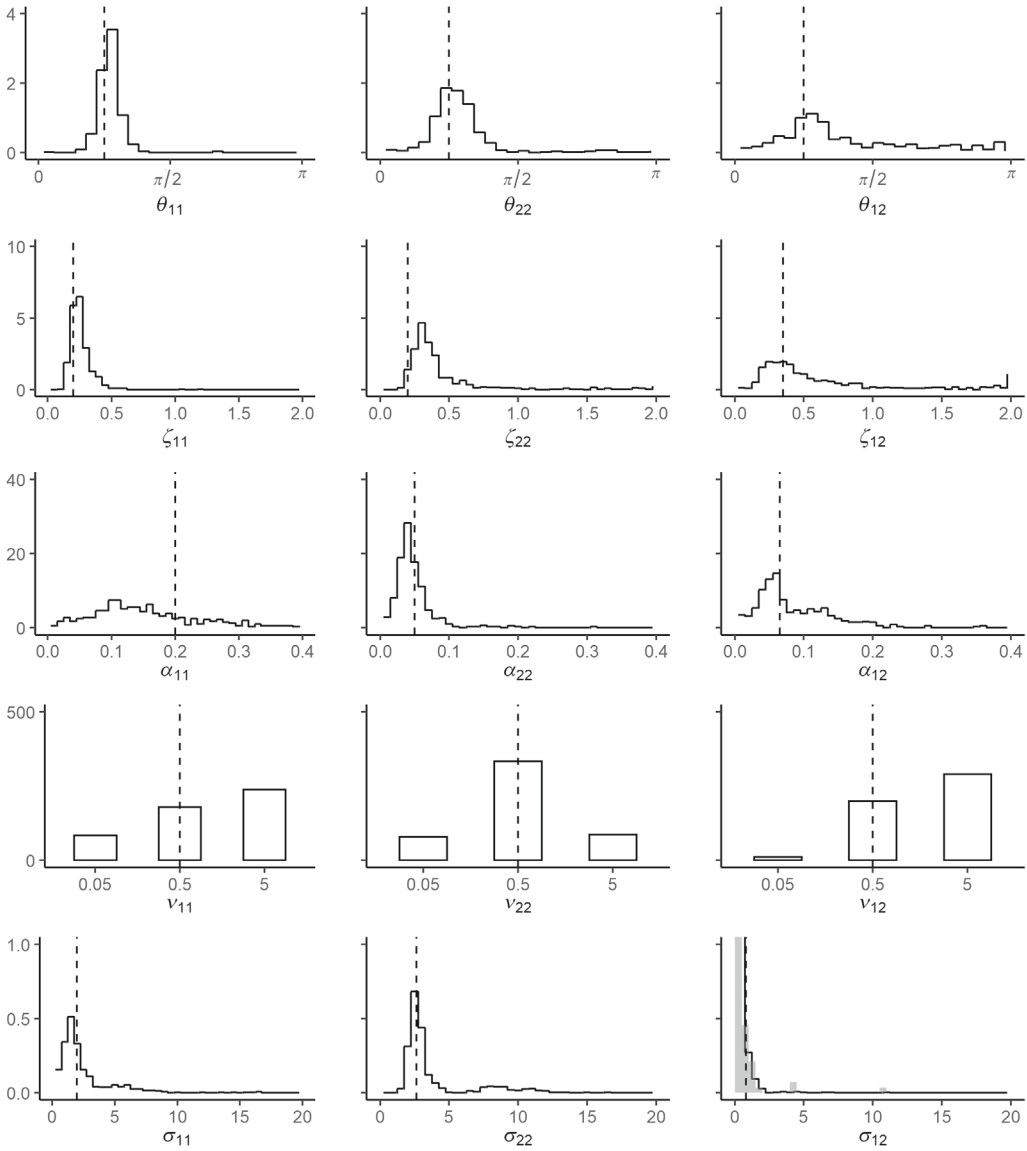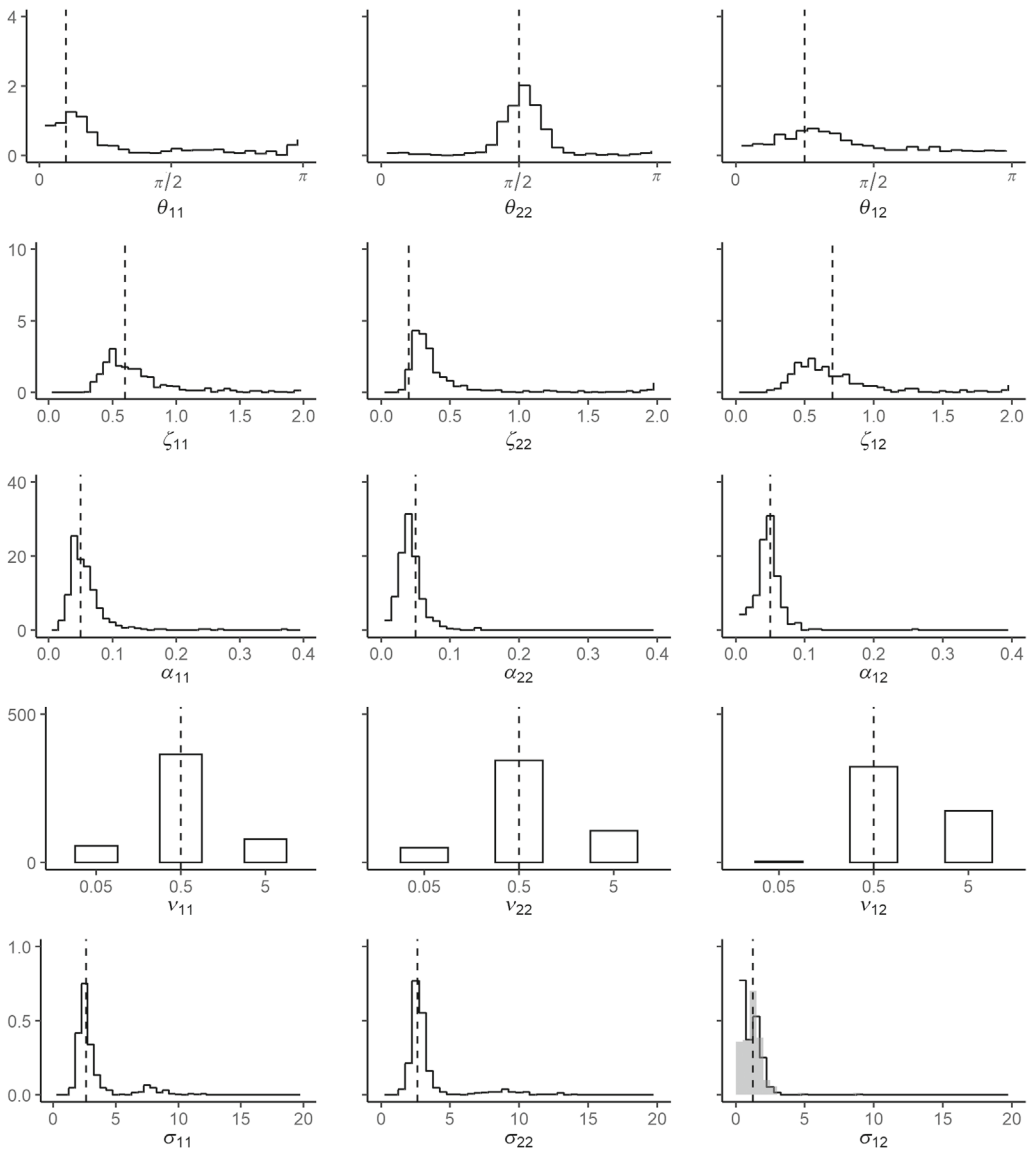
FIGURE A2 Histograms of the parameter distributions for the synthetic bivariate geometric anisotropic log-Gaussian Cox process with Matérn covariance structure specified by Model 2. The parameter values used to generate each dataset are marked by vertical dashed lines. The grey histogram in the final panel shows the empirical distribution for $\hat{\sigma}_{12}$, restricted to simulations where $\hat{\sigma}_{12}$ was not equal to the upper bound dictated by $\hat{\sigma}_{11}$ and $\hat{\sigma}_{22}$.
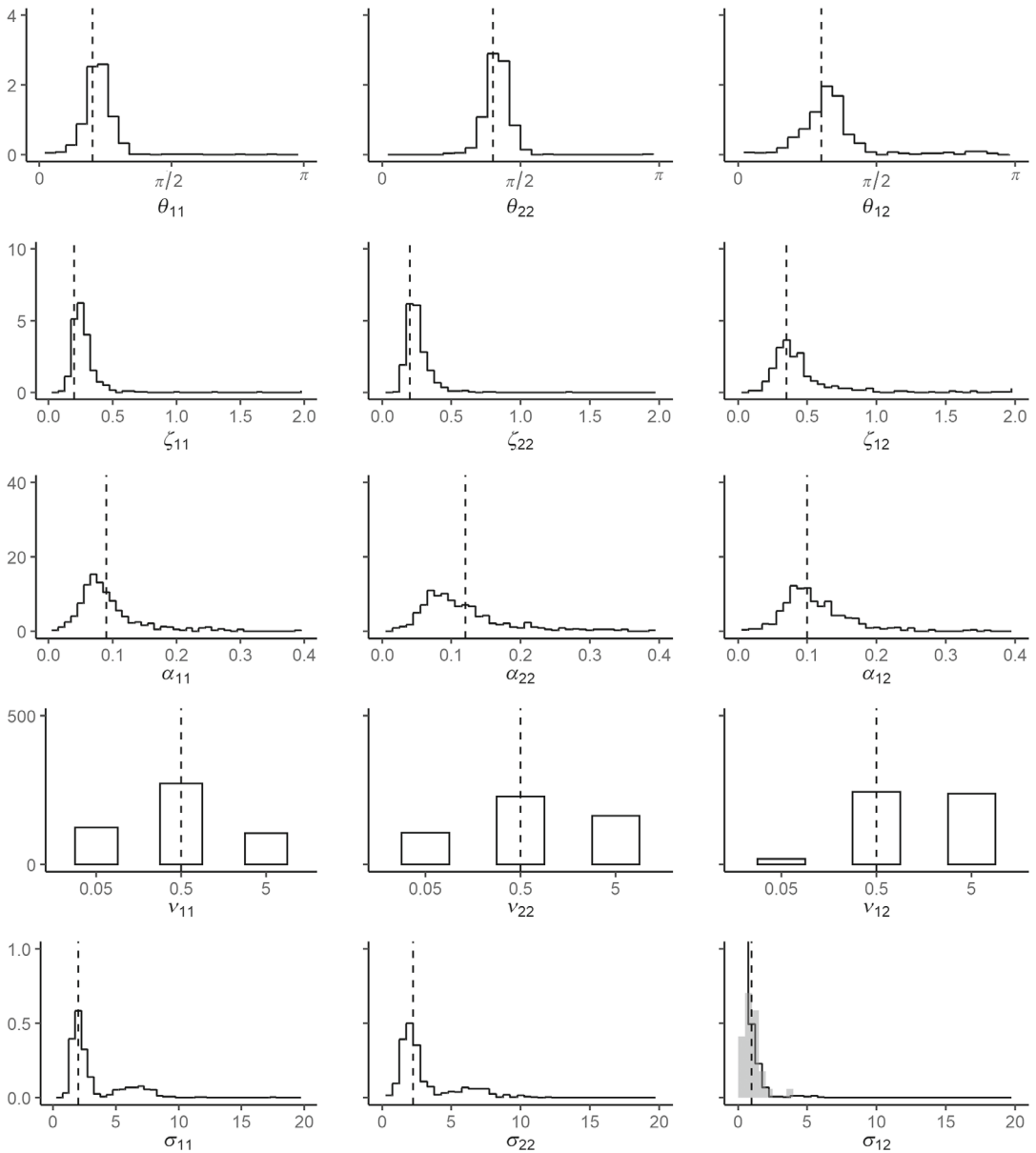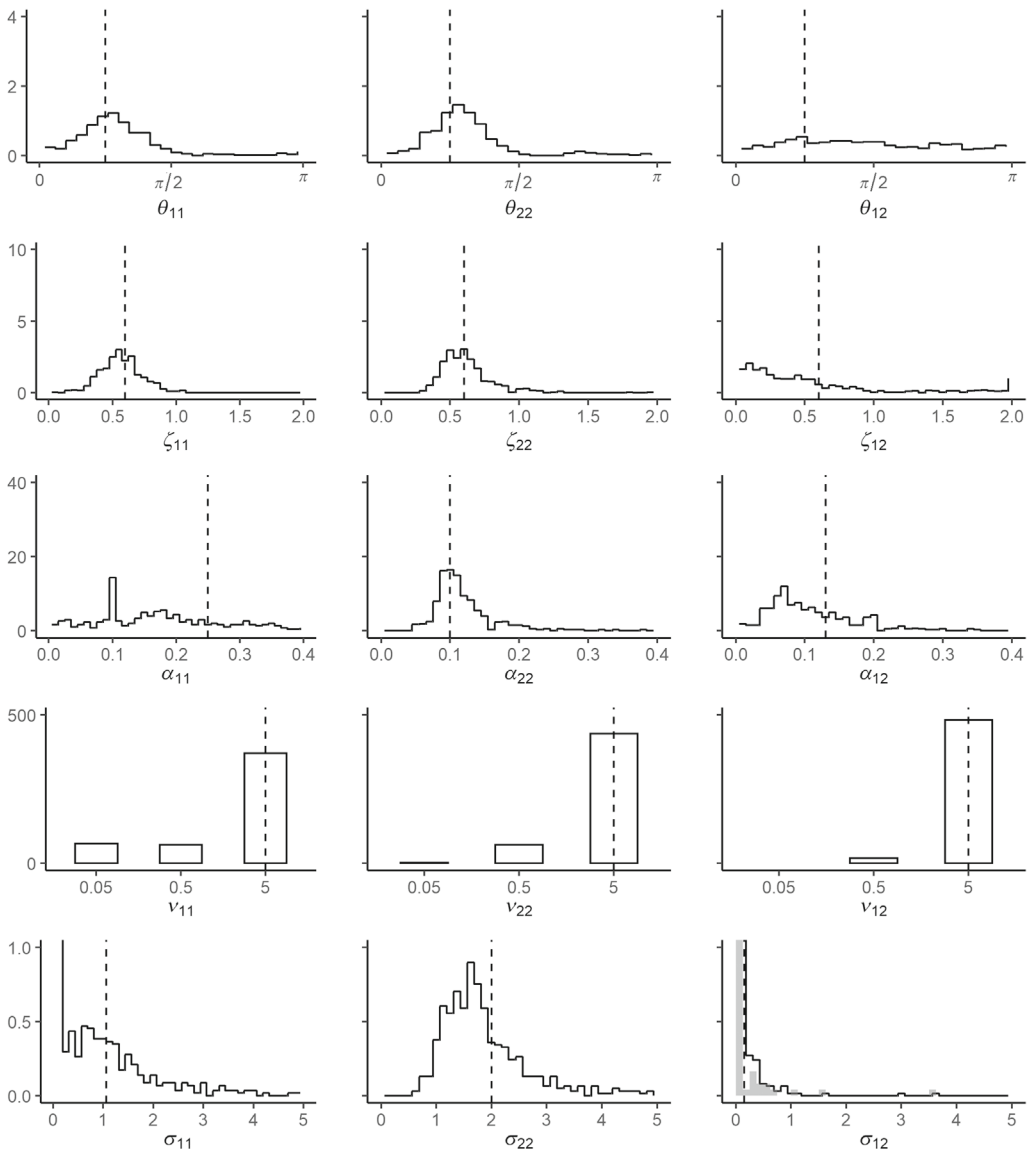
FIGURE A3   Histograms of the parameter distributions for the synthetic bivariate geometric anisotropic log-Gaussian Cox process with Matérn covariance structure specified by Model 3. The parameter values used to generate each dataset are marked by vertical dashed lines. The grey histogram in the final panel shows the empirical distribution for $\hat{\sigma}_{12}$, restricted to simulations where $\hat{\sigma}_{12}$ was not equal to the upper bound dictated by $\hat{\sigma}_{11}$ and $\hat{\sigma}_{22}$.

FIGURE A4 Histograms of the parameter distributions for the synthetic bivariate geometric anisotropic log-Gaussian Cox process with Matérn covariance structure specified by Model 4. The parameter values used to generate each dataset are marked by vertical dashed lines. The grey histogram in the final panel shows the empirical distribution for $\hat{\sigma}_{12}$, restricted to simulations where $\hat{\sigma}_{12}$ was not equal to the upper bound dictated by $\hat{\sigma}_{11}$ and $\hat{\sigma}_{22}$.
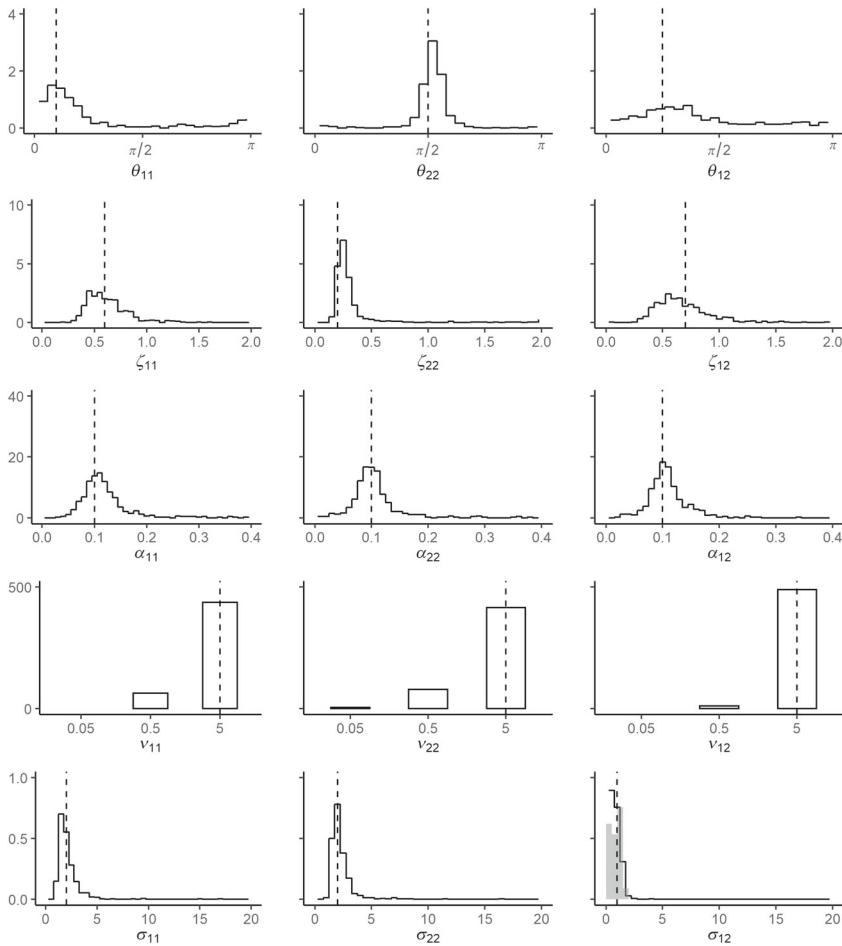
FIGURE A5 Histograms of the parameter distributions for the synthetic bivariate geometric anisotropic log-Gaussian Cox process with Matérn covariance structure specified by Model 5. The parameter values used to generate each dataset are marked by vertical dashed lines. The grey histogram in the final panel shows the empirical distribution for $\hat{\sigma}_{12}$, restricted to simulations where $\hat{\sigma}_{12}$ was not equal to the upper bound dictated by $\hat{\sigma}_{11}$ and $\hat{\sigma}_{22}$.

FIGURE A6 Histograms of the parameter distributions for the synthetic bivariate geometric anisotropic log-Gaussian Cox process with Matérn covariance structure specified by Model 6. The parameter values used to generate each dataset are marked by vertical dashed lines. The grey histogram in the final panel shows the empirical distribution for $\hat{\sigma}_{12}$, restricted to simulations where $\hat{\sigma}_{12}$ was not equal to the upper bound dictated by $\hat{\sigma}_{11}$ and $\hat{\sigma}_{22}$.

F I G U R E A7 Histograms of the parameter distributions for the synthetic bivariate geometric anisotropic log-Gaussian Cox process with Matérn covariance structure specified by Model 7. The parameter values used to generate each dataset are marked by vertical dashed lines. The grey histogram in the final panel shows the empirical distribution for $\hat{\sigma}_{12}$, restricted to simulations where $\hat{\sigma}_{12}$ was not equal to the upper bound dictated by $\hat{\sigma}_{11}$ and $\hat{\sigma}_{22}$.

FIGURE A8 Histograms of the parameter distributions for the synthetic bivariate geometric anisotropic log-Gaussian Cox process with Matérn covariance structure specified by Model 8. The parameter values used to generate each dataset are marked by vertical dashed lines. The grey histogram in the final panel shows the empirical distribution for $\hat{\sigma}_{12}$, restricted to simulations where $\hat{\sigma}_{12}$ was not equal to the upper bound dictated by $\hat{\sigma}_{11}$ and $\hat{\sigma}_{22}$.