Modeling and simulating spatio-temporal, multivariate and nonstationary Gaussian Random Fields: a Gaussian mixtures perspective

Denis Allard, Lionel Benoit, Said Obakrim

Modeling and simulating spatio-temporal, multivariate and nonstationary Gaussian Random Fields: a Gaussian mixtures perspective

Denis Allard[1], Lionel Benoit[1], Said Obakrim[2]
[1]Biostatistics and Spatial Processes (BioSP), INRAE, 84914 Avignon CEDEX 9, France
[2]Institute of Earth Surface Dynamics, University of Lausanne, Switzerland

April 15, 2025

Abstract

Gaussian Random Fields (GRFs) play a critical role in modeling and simulating environmental and climate-driven processes. The simulation of GRFs enables the representation of the variability of the process under study through the generation of multiple equally plausible realizations.

Gaussian vectors corresponding to a sample of moderate size of a GRF can easily be simulated using the Cholesky decomposition of the associated covariance matrix, but this approach is limited to vectors of moderate size. To overcome this limitation, an interesting alternative is to rely on spectral methods that are based on the decomposition of the target GRF into spectral waves. This approach has been recently extended in various directions in order to make it more versatile, including in spatial, multivariate and spatio-temporal settings. To further increase the versatility of spectral simulation methods, we propose to revisit them adopting a Gaussian mixture perspective.

This paper has a threefold aim: firstly, it offers a literature review leading to a unified view of the spectral simulation approach; secondly, it provides a revisited perspective based on Gaussian mixtures; thirdly, it leverages the Gaussian mixture perspective to propose extensions covering new classes of covariance functions for nonstationary (univariate or multivariate) spatio-temporal GRFs, as well as simulation algorithms for those that are currently missing in the framework of spectral simulation. All simulation methods are translated into pseudo-code algorithms, and an illustration is provided for a bivariate nonstationary spatio-temporal example.

Keywords: Matérn covariance; Cauchy covariance; Gneiting class of covariance; spectral methods; stochastic simulation; importance sampling

1 Introduction

Random Fields (RFs) are mathematical objects that represent random functions defined over multidimensional spaces, see for example Yaglom (1987), Adler (2010) or Adler and Taylor (2009). In this paper, we focus on RFs defined on subspaces of $\mathbb{R}^d$, $d \geq 1$, such as spatial domains (two-dimensional) and space-time domains (three-dimensional), commonly referred to as spatial random fields and spatio-temporal random fields, respectively. RFs can be either univariate, modeling a single variable, or multivariate, modeling multiple dependent variables. They play a critical role in modeling and simulating environmental and climate-driven processes by capturing their spatio-temporal and inter-variable dependence structures. Consequently, they have become indispensable tools in diverse applications such as remote sensing (Stein et al., 1999; Zakeri and Mariethoz, 2021), in geosciences (Christakos, 2013; Chilès and Delfiner, 2012; Linde et al., 2015), or in climate studies for the quantification of climate uncertainties and local weather variability (Peleg et al., 2017; Benoit et al., 2018; Obakrim et al., 2025).

A key research topic in spatial statistics is the modeling of the dependence structure of random fields within their domain of definition. This involves making assumptions about properties such as stationarity, isotropy, and symmetry (Li et al., 2007; Bagchi and Dette, 2020). For Gaussian Random Fields (GRFs), which have gained particular interest due to their characterization by their first two moments, the dependence structure is encapsulated in covariance functions. These covariance functions can vary based on the domain (spatial or spatio-temporal), dimensionality (univariate or multivariate), and specific assumptions about the dependence (e.g., separable or nonseparable, stationary or nonstationary). Numerous parametric covariance functions have been proposed in the literature, see for example Genton and Kleiber (2015) for a review on multivariate covariances as well as Chen et al. (2021) and Porcu et al. (2021) for recent reviews on spatio-temporal covariances. However, as the complexity of the domain (e.g., space-time), dimensionality (e.g., multivariate), and dependence assumptions (e.g., nonstationarity or nonseparability) increase, and even more importantly when they interact, the range of available models becomes more limited. Current research in spatial statistics aims to develop valid covariance functions, estimate their parameters, and simulate random fields that conform to specific covariance functions. In this work, we will focus on the first and third aspects because, as will be demonstrated below, the constructive definition entailed by the simulation algorithms that we develop in this paper leads to the formulation of valid covariance models, even in the complex setting of spatio-temporal multivariate and nonstationary GRFs.

The simulation of GRFs enables the representation of the variability of the process under study through the generation of multiple equally plausible realizations. Gaussian vectors corresponding to a sample of moderate size, say n, of a GRF can easily be simulated using the

Cholesky decomposition of the associated covariance matrix, which is an exact method. But since it has a computational cost to the order of $O(n^3)$, this approach is limited to vectors of moderate size (say, smaller than 10^4 elements). To overcome this limitation, several approaches such as the sparse approximation (Jurek and Katzfuss, 2022) or the low rank approximation (Romary, 2013) of the covariance matrix have been proposed to boost GRF simulation based on the Cholesky decomposition method. Another way to reduce the computational burden consists in using block circulant matrices and the efficiency of the Fast Fourier Transform to replace the Cholesky decomposition of the covariance matrix when performing simulations (Wood and Chan, 1994; Chan and Wood, 1999). This approach requires working on regular grids, which is a strong constraint when space-time simulations are intended. Indeed, conducting simulations over extensive regions and prolonged timeframes results in extremely large space-time grids, which often hinder or even prevent the simulation due to memory issues. In addition, the simulation of GRF using block circulant matrices is limited to stationary covariance functions, which is often an overly simplistic assumption when large domains are considered. More recently, the Stochastic Partial Differential Equation (SPDE) approach has been proposed to simulate a GRF at a discrete set of locations, which under certain conditions can be represented by a Gaussian Markov Random Field (GMRF). If one can find a SPDE that links the target GRF to a GMRF representation, then the covariance matrix of the GRF can be substituted by the sparse precision matrix of the GMRF, which enables the use of computationally efficient numerical methods for simulation (Lindgren et al., 2011). This approach has the advantage to permit the efficient simulation of spatial or space-time (Clarotto et al., 2024), univariate or multivariate (Bolin and Wallin, 2020), as well as stationary or nonstationary GRFs (Pereira et al., 2022), but it is currently restricted to Matérn covariance functions and some of their space-time extensions (Lindgren et al., 2022).

An interesting alternative to the aforementioned simulation approaches is to rely on spectral methods that are based on the decomposition of the target GRF into spectral waves. This approach was pioneered about 50 years ago independently by Shinozuka (1971) and Matheron (1973) and more recently referred to as *Random Fourier Features* in the Machine Learning literature (Rahimi and Recht, 2007). It was initially designed for univariate, spatial and stationary random fields. Briefly sketched (more details are given in Section 2.5), the simulation algorithm associated with spectral methods is the following: (i) draw a sample of size L of independent random frequencies according to an appropriate probability density function (pdf); (ii) project the cosine waves originating from the above random frequencies (associated with random phases) onto the locations where the simulated values are needed; (iii) add and properly re-scale the cosine waves contributions. Thanks to the central limit theorem, the marginals are

approximately Gaussian. In addition, the covariance function of the resulting random field is the Fourier transform of the pdf used to draw the random frequencies, which is equal to the target covariance function.

A compelling feature of this approach is that the stochastic part is run very quickly, and can be saved as vectors of random frequencies and phases of limited size (say, a few thousands scalars). The realization (i.e., one actual simulation of the random field) can subsequently be computed in a separate step, using massive parallel computing, only at those locations where it is needed in order to improve computational efficiency. Until recently, simulation algorithms based on this approach were easily accessible in the `R` package `Randomfields` (Schlather et al., 2015), which is unfortunately no longer available.

More recently, the spectral approach has been extended in various directions in order to make it more versatile. In an univariate and stationary context, Emery et al. (2016) proposed an importance sampling approach allowing to simulate from spatial Gaussian Random Fields (GRFs) characterized by virtually any covariance function and variograms (Arroyo and Emery, 2017). Then, Emery and Arroyo (2018) made multivariate and nonstationary extensions of the spectral simulation framework. In a spatio-temporal context, Allard et al. (2020) proposed two simulation methods for producing stationary space-time univariate random fields with a Gneiting-type covariance function. Allard et al. (2022) generalized this work to multivariate (stationary and space-time) random fields. However, despite all the above advances and to the best of our knowledge, simulating nonstationary spatio-temporal GRFs (univariate and multivariate) using the spectral simulation approach remains an open question.

In this context, this paper has a threefold aim: Firstly, it offers a literature review leading to a unified view of the spectral simulation approach whose initial description and subsequent evolutions are up to now fragmented across numerous contributions with not always consistent notations and parametrizations. Secondly, it provides a revisited perspective on this simulation approach, which will be regarded here as a Gaussian mixture in which each cosine wave relates to the spectral density of a Gaussian kernel. The mixture of those (Gaussian) cosine waves does not originate from the spectral density of the targeted covariance, but from its Gaussian mixture representation. For this reason, we refer to this simulation approach as a *Gaussian mixture simulation method*. Thirdly, it leverages the Gaussian mixture perspective to propose extensions covering new classes of covariance functions for nonstationary (univariate or multivariate) spatio-temporal GRFs, as well as simulation algorithms for those that are currently missing in the framework of spectral simulation.

For conciseness, we focus our exposition to Euclidean spaces. Other spaces such as spheres, spheres cross time and other manifolds are not considered here. Also, we neither address the

estimation of the parameters of the covariance functions at hand, nor the conditioning of the simulations to observed values. The discussion section will nevertheless touch upon these issues.

The remainder of the paper is structured as follows: Section 2 summarizes background material on stationary GRFs, including the original spectral simulation method. To make the link between former formulations of the spectral simulation approach and the proposed Gaussian mixture perspective, we explicitly state the relationship between the spectral density and the Gaussian mixture representation of a covariance function admitting such a representation. Next, Section 3 focuses on the simulation of stationary, univariate, spatial and space-time GRFs using the Gaussian mixture representation. Section 4 explains how to simulate stationary multivariate GRFs within the same framework. Then, Section 5 details the extension of the Gaussian mixture approach for the simulation of nonstationary GRFs, including nonstationary space-time and multivariate random fields. To make the paper self-sufficient, all simulation methods are translated into pseudo-code algorithms. An illustration is then provided in Section 6 for a bivariate nonstationary spatio-temporal example. Finally, Section 7 discusses the advantages and limitations of the Gaussian mixture simulation method, and makes suggestions for further research.

2 Reminders on stationary Gaussian random fields

In the rest of this paper, vectors and matrices will be denoted using bold fonts. In particular, $\mathbf{0}$ and $\mathbf{1}$ denote vectors or matrices of all-zeros and all-ones of appropriate dimension, respectively. All vectors are column vectors, $\boldsymbol{h}^t$ denotes the transpose vector of $\boldsymbol{h}$, $\boldsymbol{h}^t\boldsymbol{a}$ is the scalar product of the vectors $\boldsymbol{h}$ and $\boldsymbol{a}$ in $\mathbb{R}^d$, with $d \in \mathbb{N}^*$, and $||\boldsymbol{h}|| = \sqrt{\boldsymbol{h}^t\boldsymbol{h}}$ is the Euclidean norm of $\boldsymbol{h}$. $\mathcal{N}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ is the multivariate Gaussian random vector of dimension p with expectation $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$, and the determinant of $\boldsymbol{\Sigma}$ will be noted $|\boldsymbol{\Sigma}|$.

2.1 Stationary covariance functions

A random field $Z(\cdot)$ defined over $\mathbb{R}^d$ is an (uncountable) collection of random variables indexed over $\mathbb{R}^d$ by their coordinates $\boldsymbol{s} \in \mathbb{R}^d$, see e.g., Yaglom (1987) or Gneiting and Guttorp (2010) for a more precise definition. In the stochastic process literature, random fields are sometimes called *random functions* (Yaglom, 1987) and in the Machine Learning community they are usually referred to as *Gaussian processes* (Williams and Rasmussen, 2006). In geostatistics and spatial statistics, the term Random Field (RF) is generally used (Chilès and Delfiner, 2012) and we will stick to this usage. We shall assume that the RF is second order, i.e. that it has finite first and second moments. Without loss of generality, we shall further assume that the expectation of Z

is equal to 0 everywhere, because otherwise it could easily be subtracted from Z in order to get a centered random field.

The covariance of a centered second order random field $Z(\cdot)$ on $\mathbb{R}^d$ is the function $c : \mathbb{R}^d \times \mathbb{R}^d \to \mathbb{R}$, $c(\boldsymbol{s}_1, \boldsymbol{s}_2) = \mathbb{E}[Z(\boldsymbol{s}_1)Z(\boldsymbol{s}_2)]$. The random field is said to be second-order stationary (or simply, *stationary*) if c depends only on the separation distance, i.e. $c(\boldsymbol{s}_1, \boldsymbol{s}_2) := C(\boldsymbol{s}_2 - \boldsymbol{s}_1)$ for some function C, called the *covariance function*. Denoting $\boldsymbol{h} = \boldsymbol{s}_2 - \boldsymbol{s}_1$ the separation vector, the stationary covariance function $C : \mathbb{R}^d \to \mathbb{R}$ writes $C(\boldsymbol{h}) = \mathbb{E}[Z(\boldsymbol{s})Z(\boldsymbol{s} + \boldsymbol{h})]$. The covariance function of Z must be semi-definite positive on $\mathbb{R}^d$, i.e.

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j C(\boldsymbol{s}_i - \boldsymbol{s}_j) \geq 0,$$

for all $n \in \mathbb{N}^*$, $\forall (a_1, \ldots, a_n) \in \mathbb{R}^n$ and for all locations $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ in $\mathbb{R}^d$. A stationary covariance function is said to be isotropic if it depends only on the distance $||\boldsymbol{h}||$, and anisotropic otherwise.

2.2 Bochner's theorem and spectral densities

In this Section, we essentially follow Gneiting and Guttorp (2010). From Bochner's theorem (Bochner, 1933, 2005), there is a one-to-one correspondence between the family of stationary covariance functions on $\mathbb{R}^d$ (i.e. stationary, definite positive functions on $\mathbb{R}^d$) and the family of positive finite measures through Fourier transforms:

$$C(\boldsymbol{h}) = \int_{\mathbb{R}^d} \exp(i\boldsymbol{h}^t\boldsymbol{\omega}) d\mu(\boldsymbol{\omega}), \quad \forall \boldsymbol{h} \in \mathbb{R}^d. \tag{1}$$

The measure μ, referred to as the *spectral measure* associated to C, is positive and finite, i.e. $\int_{\mathbb{R}^d} d\mu(\boldsymbol{\omega}) < \infty$. In most cases it is continuous, in which case we write, with a slight abuse of notation, $d\mu(\boldsymbol{\omega}) = \mu(\boldsymbol{\omega})d\boldsymbol{\omega}$ and μ is called the *spectral density*. In this work, we shall only consider continuous measures, i.e. spectral densities.

Similarly to the Bochner theorem, Schoenberg (1938) established that all isotropic covariance functions, say $\phi_d(\boldsymbol{h})$, valid in $\mathbb{R}^d$, have an expression of the form

$$\phi_d(\boldsymbol{h}) = \Gamma(d/2) \int_{\mathbb{R}^+} \left(\frac{2}{r\omega}\right)^{(d-2)/2} J_{(d-2)/2}(r\omega) d\mu_{\text{iso}}(\omega), \quad r = \|\boldsymbol{h}\| \geq 0, \tag{2}$$

where Γ is the usual Gamma function, μ_{iso} is a probability measure on the positive half-line (hence integrating to 1) often referred to as the *radial spectral measure*, and J_k is the Bessel function of parameter k. In most cases of practical interest the measure is absolutely continuous with respect to the Lebesgue measure with $d\mu_{\text{iso}}(\omega) = \mu_{\text{iso}}(\omega)d\omega$. In this case, μ_{iso} is a probability density, called the *radial spectral density*.

The correspondence between the radial spectral measure and the spectral density is given by

$$\mu_{\text{iso}}(||\boldsymbol{\omega}||) = 2\frac{\pi^{d/2}}{\Gamma(d/2)}||\boldsymbol{\omega}||^{d-1}\mu(\boldsymbol{\omega}), \quad \boldsymbol{\omega} \in \mathbb{R}^d,$$

with the factor $2\pi^{d/2}/\Gamma(d/2)$ being the surface of the d-sphere of volume 1 entailed by the change from cartesian coordinates to polar coordinates. From (2) we thus see that all valid isotropic covariance functions over $\mathbb{R}^d$ can be seen as mixtures of generators $\Omega_d(x) = (2/x)^{(d-2)/2}\, J_{(d-2)/2}(x)$, $x \geq 0$. The Bessel functions J_d are non-monotonic functions with positive and negative values tending to 0 as $x \to \infty$. The amplitudes of the oscillations decrease as $(d-2)/2$ increases, and to the limit, as $d \to \infty$, $J_{(d-2)/2}$ is a non-negative function. This implies that isotropic functions that are covariance functions in $\mathbb{R}^d$ for any dimension d cannot be negative.

Example 1 (Gaussian covariance) *The Gaussian covariance function* $C_G(\boldsymbol{h}) = \exp\left(-(||\boldsymbol{h}||/a)^2\right)$ *with* $a > 0$*, whose support is* $\mathbb{R}^d$*, will be key to this work. Its spectral density is*

$$\mu_G(\boldsymbol{\omega}) = a^d \left(2\sqrt{\pi}\right)^{-d} \exp\left(-||a\boldsymbol{\omega}||^2/4\right), \tag{3}$$

which is a Gaussian density with mean 0 and variance $2/a^2$*.*

2.3 Anisotropies

Anisotropy can be modeled in many different ways (Allard et al., 2016), but the most common model is the geometric anisotropy and we will adopt it for the rest of the paper. A stationary covariance function with a geometric anisotropy is obtained by combining a stationary isotropic covariance function $C_{\text{iso}}(\boldsymbol{h})$ by a $d \times d$ symmetric positive definite matrix $\boldsymbol{\Sigma}$ such that $C(\boldsymbol{h}) = C_{\text{iso}}(\boldsymbol{\Sigma}^{-1/2}\boldsymbol{h})$, $\forall \boldsymbol{h} \in \mathbb{R}^d$ (Chilès and Delfiner, 2012; Allard et al., 2016). In $\mathbb{R}^2$, the matrix $\boldsymbol{\Sigma}^{-1/2}$ can be described as the combination of a scaling and a rotation with angle θ

$$\boldsymbol{\Sigma}^{-1/2} = \begin{pmatrix} r_1 & 0 \\ 0 & r_2 \end{pmatrix} \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix}, \tag{4}$$

where $r_1 > 0$ and $r_2 > 0$ are the scale parameters in the main anisotropy directions. Emery and Arroyo (2018) showed that the spectral density of C is $\mu(\boldsymbol{\omega}) = |\boldsymbol{\Sigma}|^{1/2}\mu_{\text{iso}}(\boldsymbol{\Sigma}^{1/2}\boldsymbol{\omega})$.

Example 2 (Gaussian spectral density) *Specifically, in the case of a Gaussian covariance, one gets:*

$$C_G(\boldsymbol{h}) = \exp\left(-\boldsymbol{h}^t\boldsymbol{\Sigma}^{-1}\boldsymbol{h}\right); \qquad \mu_G(\boldsymbol{\omega}) = \left(2\sqrt{\pi}\right)^{-d}|\boldsymbol{\Sigma}|^{1/2}|\exp\left(-\boldsymbol{\omega}^t\boldsymbol{\Sigma}\boldsymbol{\omega}/4\right), \quad \forall \boldsymbol{h}, \boldsymbol{\omega} \in \mathbb{R}^d.$$

2.4 Gaussian mixtures

A standard result on Bessel functions states that $\lim_{d\to\infty} \Omega_d\left((2d)^{1/2}x\right) = e^{-x^2}$ uniformly on $x \geq 0$. As a consequence, a mixture of squared exponential is a valid covariance function on $\mathbb{R}^d$, $\forall d \geq 1$. Schoenberg (1938) proved the necessary part and established that a function ϕ is a valid isotropic covariance function on $\mathbb{R}^d$, $\forall d \geq 1$ if and only if it is a mixture of isotropic Gaussian kernels

$$\phi_{\text{iso}}(r) = \int_{\mathbb{R}^+} \exp(-r^2\xi) dF(\xi), \qquad \forall r \geq 0, \tag{5}$$

where F is a probability measure on $[0, \infty)$. Focusing on the continuous case only, one usually writes $dF(\xi) = f(\xi)d\xi$, with f being a probability density.

Another way of seeing this is the following. Let $\varphi(x),\ x \geq 0$ be a continuous completely monotone function on $\mathbb{R}^+$, i.e. a function with $\varphi(0) = 1$ such that for any $n \in \mathbb{N}$, the derivative of order n, $\varphi^{(n)}$ exists and is finite with $(-1)^n\varphi^{(n)}(x) \geq 0$, $\forall x \geq 0$. One example of such a function is the exponential: $\exp(-\xi x)$ with $\xi > 0$. By Bernstein's theorem (see for example Feller (1966, p. 439)), a continuous completely monotone function can be written as a mixture of exponential functions. In other words, it is the Laplace transform of a nonnegative density, i.e. $\varphi(x) = \int_0^\infty \exp(-\xi x) f(\xi) d\xi$. Let us denote $\mathcal{C}_\infty$ the class of continuous isotropic covariance functions valid on $\mathbb{R}^d$, $\forall d \geq 1$, i.e. that can be expressed as in Eq. (5). Putting everything together, one thus have the following series of equality for covariance functions $\phi \in \mathcal{C}_\infty$:

$$\phi(\boldsymbol{h}) = \varphi(||\boldsymbol{h}||^2) = \int_0^\infty \exp(-||\boldsymbol{h}||^2\xi) f(\xi) d\xi, \qquad \forall \boldsymbol{h} \in \mathbb{R}^d,$$

where $f(\xi)$ is also referred to as the scale mixture. Gaussian mixtures are key for constructing large classes of spatio-temporal (Gneiting, 2002), multivariate (Gneiting et al., 2010; Genton and Kleiber, 2015), and multivariate spatio-temporal covariances (Bourotte et al., 2016; Allard et al., 2022). Gaussian mixtures will be at the core of the approaches and simulation algorithms presented in this work, and from now on they will be characterized by their scale mixture f and their covariance function $\phi \in \mathcal{C}_\infty$.

Note that for a covariance function $\phi \in \mathcal{C}_\infty$, there is a direct relationship between the scale mixture and the spectral density, as stated in Proposition 1.

Proposition 1 *Let $\phi \in \mathcal{C}_\infty$ be an isotropic covariance function on $\mathbb{R}^d$ and let f denote its scale mixture and μ its spectral density. Then,*

$$\mu(\boldsymbol{\omega}) = \left(2\sqrt{\pi}\right)^{-d} \int_0^{+\infty} \exp\left(-||\boldsymbol{\omega}||^2/4\xi\right) \xi^{-d/2} f(\xi)\, d\xi, \qquad \forall \boldsymbol{\omega} \in \mathbb{R}^d. \tag{6}$$

The proof is direct. For completeness, it is given in Appendix A. This proposition essentially states that whenever there is a Gaussian mixture representation of a covariance function (i.e., when $\phi \in \mathcal{C}_\infty$), its spectral density can be represented as a specific scale mixture of the Gaussian spectral density.

2.5 Simulation with spectral methods

Proposition 2 *Consider a stationary covariance function C and let μ denotes its spectral measure with $\int_{\mathbb{R}^d} d\mu(\boldsymbol{\omega}) = 1$. Let $\boldsymbol{\Omega}$ be a random variable drawn according to μ. Let Φ be a uniform random variable in $[0, 2\pi)$. Define*

$$Z(\boldsymbol{s}) = \sqrt{2}\cos\left(\boldsymbol{\Omega}^t\boldsymbol{s} + \Phi\right).$$

Then $\mathbb{E}[Z(\boldsymbol{s})] = 0$ and $\mathbb{C}\mathrm{ov}(Z(\boldsymbol{s}), Z(\boldsymbol{s}+\boldsymbol{h})) = \mathbb{E}[Z(\boldsymbol{s})Z(\boldsymbol{s}+\boldsymbol{h})] = C(\boldsymbol{h})$, $\forall \boldsymbol{s}, \boldsymbol{h} \in \mathbb{R}^d$.

Clearly the expectation is equal to zero since Φ is uniform on $[0, 2\pi)$. The proof for the covariance is obtained by computing $\mathbb{E}[Z(\boldsymbol{s})Z(\boldsymbol{s}+\boldsymbol{h})]$ and using product-to-sum trigonometric identities (Shinozuka, 1971), see also Rahimi and Recht (2007) for a Machine Learning perspective. The random field $Z(\boldsymbol{s})$ is not Gaussian but a Gaussian distribution can be approximated by simulating many independent samples $\{(\boldsymbol{\Omega}_l, \Phi_l)\}_{l=1,\dots,L}$ and considering the sum

$$\tilde{Z}_L(\boldsymbol{s}) = \sqrt{\frac{2}{L}} \sum_{l=1}^{L} \cos\left(\boldsymbol{\Omega}_l^t\boldsymbol{s} + \Phi_l\right).$$

Because of the central limit theorem, the finite-dimensional distribution of $\tilde{Z}_L$ tends to become multivariate Gaussian as L tends to infinity (Lantuéjoul, 2002, Chap. 15). Finally, the original spectral simulation algorithm is shown in Algorithm 1. In step 4, a multiplicative factor $\sqrt{-\ln U}$ is introduced, where U is an independent uniform variable on $(0, 1)$. It corresponds to the Box–Müller transformation (Box and Muller, 1958) which ensures that the marginal distributions of each term of the sum in step 4 is Gaussian. This additional step accelerates the convergence of the simulation algorithm, but it is not absolutely necessary. Sometimes, it will be omitted. We recall that, once the random elements $(\boldsymbol{\Omega}_l, \Phi_l, U_l)_{l=1,\dots,L}$ are stored, Step 6 can be performed separately from the previous steps, possibly using massive parallel computing.

Example 3 (Simulating from a Gaussian covariance) *The algorithm for simulating a Gaussian random field with Gaussian covariance thus follows by simply replacing Step 2 in Algorithm 1 by:*

$$2. \quad \text{Simulate } \boldsymbol{\Omega}_l \sim \sqrt{2}/a\mathcal{N}_d(0, \boldsymbol{I}_d). \tag{7}$$

Algorithm 1 Original spectral simulation algorithm

Require: Covariance function and its associated spectral density μ

Require: A set of points, $\mathcal{S} \in \mathbb{R}^d$

Require: A large number L

1: **for** $l = 1$ to L **do**

2: Simulate $\boldsymbol{\Omega}_l \sim \mu$

3: Simulate $\Phi_l \sim \mathcal{U}(0, 2\pi)$

4: Simulate $U_l \sim \mathcal{U}(0, 1)$

5: **end for**

6: For each $\boldsymbol{s} \in \mathcal{S}$ return

$$\tilde{Z}(\boldsymbol{s}) = \sqrt{\frac{2}{L}} \sum_{l=1}^{L} \sqrt{-\ln U_l}\, \cos\left(\boldsymbol{\Omega}_l^t \boldsymbol{s} + \Phi_l\right)$$

Using the original spectral algorithm requires to be able to simulate directly from the spectral density, a situation not always met, for instance for the spatio-temporal covariance functions of the Gneiting class (see Sect. 3.2). Approaches based on Gaussian mixtures offer additional flexibility, as shown in the following.

3 Simulation of stationary univariate GRFs using Gaussian mixtures

3.1 Spatial Gaussian Random Fields

Since one can easily simulate a RF with Gaussian covariance as shown in Example 3, spectral simulation algorithms based on Gaussian mixtures can easily be designed. By construction, this approach is thus restricted to the simulation of covariance functions being element of $\mathcal{C}_\infty$. The general principle is that for each $l \in \{1, \ldots, L\}$, a scale is randomly drawn according to the scale mixture f. Then, a single cosine wave with a frequency drawn according to (7) is simulated, conditional on the random scale previously drawn. This leads to Algorithm 2, which will then be specified for two important examples of covariance functions: the Matérn and the Cauchy.

3.1.1 The Matérn covariance

The Matérn covariance function is a widely used covariance function, in particular because it is related to the Wittle-Matérn SPDE over $\mathbb{R}^d$ (Whittle, 1954; Lindgren et al., 2011; Carrizo-

Algorithm 2 Generic isotropic scale mixture simulation algorithm

Require: Covariance function in $\mathcal{C}_\infty$ and its associated scale mixture f

Require: A set of points, $\mathcal{S} \in \mathbb{R}^d$

Require: A large number L

1: **for** $l = 1$ to L **do**

2: Simulate $\xi_l \sim f$

3: Simulate $\boldsymbol{\Omega}_l \sim \sqrt{2\xi_l}\mathcal{N}_d(0, \boldsymbol{I}_d)$

4: Simulate $\Phi_l \sim \mathcal{U}(0, 2\pi)$

5: Simulate $U_l \sim \mathcal{U}(0, 1)$

6: **end for**

7: For each $\boldsymbol{s} \in \mathcal{S}$ return

$$\tilde{Z}(\boldsymbol{s}) = \sqrt{\frac{2}{L}} \sum_{l=1}^{L} \sqrt{-\ln U_l} \cos\left(\boldsymbol{\Omega}_l^t \boldsymbol{s} + \Phi_l\right)$$

Vergara et al., 2022),

$$\left(\kappa^2 - \Delta\right)^{\alpha/2} Z = \mathcal{W}, \qquad \alpha = \nu + d/2 > d/2, \tag{8}$$

where Δ is the Laplacian operator, $\kappa > 0$ and $\mathcal{W}$ is a white noise. The Matérn covariance is

$$C_{\mathcal{M}}(\boldsymbol{h}) = \frac{\sigma^2}{2^{\nu-1}\Gamma(\nu)} (\kappa||\boldsymbol{h}||)^\nu K_\nu(\kappa||\boldsymbol{h}||), \quad \boldsymbol{h} \in \mathbb{R}^d, \tag{9}$$

where K_ν is the Bessel function of second kind with parameter $\nu > 0$. Its spectral density is

$$\mu_{\mathcal{M}}(\boldsymbol{\omega}) = \frac{\Gamma(\nu + d/2)}{\kappa^d \pi^{d/2} \Gamma(\nu)} \frac{1}{(1 + ||\boldsymbol{\omega}||^2/\kappa^2)^{\nu+d/2}}, \quad \boldsymbol{\omega} \in \mathbb{R}^d. \tag{10}$$

The Matérn covariance function is also an Inverse Gamma mixture of Gaussian kernels (Tronarp et al., 2018). First recall that the probability density of an Inverse Gamma random variable $IG(\alpha, \beta)$ with shape and rate parameters (α, β) is

$$f_{IG}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-1-\alpha} e^{-\beta/x}, \quad x \geq 0.$$

Proposition 3 *Let $\xi \sim IG(\nu, \kappa^2/4)$ be an Inverse Gamma random variable with shape and rate parameters $(\nu, \kappa^2/4)$, whose density is denoted $f_{IG_{\nu,\kappa^2/2}}$. Then, the mixture of Gaussian covariances with scale parameter ξ is a Matérn covariance as in* (9)*. In other words, the scale mixture of a Matérn covariance is*

$$f_{\mathcal{M}}(\xi) = \left(\frac{\kappa^2}{4}\right)^\nu \frac{\xi^{-1-\nu}}{\Gamma(\nu)} e^{-\kappa^2/4\xi}.$$

The proof can be found in Bourotte et al. (2016, Theorem 1) and in Tronarp et al. (2018) with slightly different parametrization. The relationship (6) is easy to verify (see Appendix A). A simulation algorithm then follows directly by replacing Step 2 in Algorithm 2 by

$$2: \text{ Simulate } \quad \xi \sim IG(\nu, \kappa^2/4)$$

Remark: Notice that simulating $\xi \sim IG(\nu, \kappa^2/4)$ consists of simulating $S \sim G(\nu, \kappa^2/4)$ and setting $\xi = 1/S$.

3.1.2 The Cauchy covariance

As an alternative to the Matérn class, with a more slowly decaying behavior to 0 as $||\boldsymbol{h}|| \to \infty$, let us consider the Cauchy family, having expression

$$C_{\mathcal{C}}(\boldsymbol{h}) = \left(1 + a||\boldsymbol{h}||^2\right)^{-\nu}, \quad \boldsymbol{h} \in \mathbb{R}^d, \tag{11}$$

with $\nu, a > 0$. Here, ν characterizes the dependence at large distances. It was considered in Bourotte et al. (2016) because it allows for a straightforward Gaussian mixture model as shown below, whereas the spectral density of the Cauchy covariance function has no simple explicit expression, see also Appendix A. Notice that the covariance in (11) is a special case of the generalized Cauchy covariance function proposed in Gneiting and Schlather (2004), but no closed form expression for the scale mixture corresponding to the generalized Cauchy covariance is known.

Proposition 4 *Let $\xi \sim G(\nu, a)$ be a Gamma random variable with density $f_{\nu,a}$. Then, the mixture of Gaussian covariances with scale parameter ξ is a Cauchy covariance as in* (11).

Proof: The probability density of ξ is $f_{\mathcal{C}}(\xi) = a^{-\nu}\Gamma(\nu)^{-1}\xi^{\nu-1}e^{-\xi/a}$. Then, writing $r = ||\boldsymbol{h}||$ we get

$$\begin{aligned} C_{\mathcal{C}}(\boldsymbol{h}) &= a^{-\nu}\Gamma(\nu)^{-1}\int_0^\infty e^{-\xi r^2}\xi^{\nu-1}e^{-\xi/a}\,d\xi \\ &= a^{-\nu}\Gamma(\nu)^{-1}\int_0^\infty \exp\left(-\xi(r^2 + a^{-1})\right)\xi^{\nu-1}\,d\xi \\ &= a^{-\nu}\left(r^2 + a^{-1}\right)^{-\nu} = \left(1 + ar^2\right)^{-\nu}. \end{aligned}$$

□

In this case, Step 2. in Algorithm 2 becomes

$$2: \text{ Simulate } \quad \xi \sim G(\nu, a).$$

3.1.3 Simulation in the presence of anisotropy

In order to simulate a GRF with geometric anisotropies, the last step of the spectral simulation Algorithm 1 must be adapted according to

$$\tilde{Z}(\boldsymbol{s}) = \sqrt{2/L}\sum_{l=1}^{L}\sqrt{-\ln U_l}\,\cos\big(\boldsymbol{\Omega}_l^t\boldsymbol{\Sigma}^{-1/2}\boldsymbol{s}+\Phi_l\big) = \sqrt{2/L}\sum_{l=1}^{L}\sqrt{-\ln U_l}\,\cos\big((\boldsymbol{\Sigma}^{-1/2}\boldsymbol{\Omega}_l)^t\boldsymbol{s}+\Phi_l\big). \tag{12}$$

The last equality is useful for simulating anisotropic random fields with scale mixtures. Notice that the single scale ξ_l in Algorithm 2 acts as a multiplicative factor for the entire vector $\boldsymbol{\Omega}_l$. Using scale mixtures only, it is thus not possible to change the geometrical characteristics (anisotropy angle θ, range ratio a_1/a_2) of the anisotropy. By decomposing $\boldsymbol{\Omega}_l$ as the product of a scale ξ and a direction $\boldsymbol{V}_l$ and using (12), it is straightforward to generalize Algorithm 2 for anisotropic RFs. This is shown in Algorithm 3 below. In order to avoid an overparametrization, we recommend to either set $r_1 = 1$, or to set the scale parameter in f equal to 1. In the first case, when $d = 2$, the only two parameters of the matrix $\boldsymbol{\Sigma}^{-1/2}$ would thus be $r_2 > 0$ and $\theta \in [0, \pi)$.

Algorithm 3 Generic scale mixture simulation algorithm with anisotropy

Require: Covariance function in $\mathcal{C}_\infty$ and its associated scale mixture f

Require: anisotropy matrix $\boldsymbol{\Sigma}^{-1/2}$

Require: A set of points, $\mathcal{S} \in \mathbb{R}^d$

Require: A large number L

1: **for** $l = 1$ to L **do**
2: Simulate $\xi_l \sim f$
3: Simulate $\boldsymbol{V}_l \sim \mathcal{N}_d(0, \boldsymbol{I}_d)$
4: Compute $\boldsymbol{\Omega}_l = \sqrt{2\xi_l}\boldsymbol{\Sigma}^{-1/2}\,\boldsymbol{V}_l$
5: Simulate $\Phi_l \sim \mathcal{U}(0, 2\pi)$
6: Simulate $U_l \sim \mathcal{U}(0, 1)$
7: **end for**
8: For each $\boldsymbol{s} \in \mathcal{S}$ return

$$\tilde{Z}(\boldsymbol{s}) = \sqrt{\frac{2}{L}}\sum_{l=1}^{L}\sqrt{-\ln U_l}\,\cos\left(\boldsymbol{\Omega}_l^t\boldsymbol{s}+\Phi_l\right)$$

3.2 Spatio-temporal Gaussian Random Fields

Spatio-temporal random fields can be seen as special cases of random fields defined over $\mathbb{R}^d$ with $d \geq 2$ and one dimension being identified as time. Then, one could proceed with all the

theory seen above. However, it is our common daily experience that time is not a dimension directly commensurable to space. In particular, most phenomena follow physical dynamics that are clearly different over space and time. We will therefore separate space from time and write $(\boldsymbol{s}, t) \in \mathbb{R}^d \times \mathbb{R}$ for a space-time coordinate. A stationary spatio-temporal covariance function is a positive definite function C over $\mathbb{R}^d \times \mathbb{R}$ with $(\boldsymbol{h}, u) \mapsto C(\boldsymbol{h}, u) \in \mathbb{R}$. Setting $u = 0$ defines the purely spatial covariance function $C_S(\boldsymbol{h}) := C(\boldsymbol{h}, 0)$ whereas setting $\boldsymbol{h} = \mathbf{0}$ defines purely temporal covariance function $C_T(u) := C(\mathbf{0}, u)$.

3.2.1 Simple extensions of spatial and temporal models

A considerable literature has been devoted to the construction of valid (i.e. positive definite) spatio-temporal covariance functions using valid covariances C_S and C_T as building blocks. The simplest construction is to set $C_{\text{sep}}(\boldsymbol{h}, u) := C_S(\boldsymbol{h})C_T(u)$, for all $(\boldsymbol{h}, u) \in \mathbb{R}^d \times \mathbb{R}$. The covariance C_{sep} is said to be *separable* because under a Gaussian assumption on Z, the following conditional independence property holds: $Z(\boldsymbol{s}_1, t_2) \perp Z(\boldsymbol{s}_2, t_1) \mid Z(\boldsymbol{s}_1, t_1), \forall \boldsymbol{s}_1, \boldsymbol{s}_2, t_1, t_2$. In other words, given $Z(\boldsymbol{s}, t)$ the spatial field $Z(\cdot, t)$ behaves independently to the time process $Z(\boldsymbol{s}, \cdot)$, i.e. the processes $Z(\cdot, t)$ and $Z(\boldsymbol{s}, \cdot)$ *are separated* by $Z(\boldsymbol{s}, t)$. A spatio-temporal covariance function is said to be positively (resp. negatively) space-time nonseparable if the separability function $S(\boldsymbol{h}, u) = C(\boldsymbol{h}, u)C(\mathbf{0}, 0) - C_S(\boldsymbol{h})\, C_T(u) \geq 0$ (resp. ≤ 0) for all $(\boldsymbol{h}, u) \in \mathbb{R}^d \times \mathbb{R}$ (Rodrigues and Diggle, 2010; De Iaco and Posa, 2013). Its empirical version can be used as an exploratory tool to characterize the type of nonseparability in a dataset. Statistical tests for spatio-temporal separability have been proposed, see for example Li et al. (2007), Huang and Sun (2019) or Cappello et al. (2018). As exemplified in these references, separable covariances are very often overly simplistic models for environmental or climate applications. A wealth of more flexible models have thus been proposed, among which the product sum model (De Iaco et al., 2001), linear combinations or mixtures of separable components (Ma, 2003), the Gneiting class (Gneiting, 2002) and its many extensions (see e.g. Porcu et al. (2006, 2008); Emery and Porcu (2023)) are the most widely used. For a more extensive presentation of space-time covariance functions, we refer to the excellent overviews in Porcu et al. (2021) and Chen et al. (2021).

Simulating random fields characterized by the product-sum or linear combinations of separable components is a straightforward extension of the algorithms presented in the previous section. Algorithms for these models will not be further considered in this work, and we refer for example to Schlather et al. (2015) for a general exposition regarding these extensions.

3.2.2 The Gneiting class

This class of spatio-temporal covariance functions was originally proposed in Gneiting (2002) with a construction essentially based on Gaussian mixtures. For an easier exposition of the simulation algorithms presented below, we will slightly change the notations and consider the class of functions of the form

$$C(\boldsymbol{h}, u) = \frac{1}{(\gamma(u)+1)^{d/2}} \phi\left(\frac{||\boldsymbol{h}||}{\sqrt{\gamma(u)+1}}\right), \qquad (\boldsymbol{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \tag{13}$$

which is a second-order stationary spatio-temporal covariance function on $\mathbb{R}^d \times \mathbb{R}$ for any function $\phi \in \mathcal{C}_\infty$ if and only if the function $\gamma : \mathbb{R} \to [0, \infty)$ is a variogram on $\mathbb{R}$, i.e., a conditionally negative semidefinite function (Zastavnyi and Porcu, 2011). We refer to Chilès and Delfiner (2012, Chapter 2) and Dörr and Schlather (2023a) for a full characterization of variograms. For clarity of exposition, we first discuss the Gneiting class in an isotropic context. Spatial anisotropy can be accounted for by replacing the vector $\boldsymbol{h}$ with $\boldsymbol{\Sigma}^{-1/2}\boldsymbol{h}$.

The Gneiting class of spatio-temporal covariances thus involves two functions ϕ and γ: the former is associated with the spatial structure, since $C_S(\boldsymbol{h}) = \phi(||\boldsymbol{h}||)$, whereas the latter is associated with the temporal structure, with $C_T(u) = (\gamma(u)+1)^{-d/2}$. Notice that $C_T(u) \to 0$ as $|u| \to \infty$ if and only if $\gamma(u)$ is unbounded. Examples of such classes include the Gneiting-Matérn and Gneiting-Cauchy covariance functions, in the cases where $\phi(||\boldsymbol{h}||)$ is the Matérn or the Cauchy spatial covariance, respectively. Using direct computations of the separability function $S(\boldsymbol{h}, u)$, it is easy to show that covariances from the Gneiting class are uniformly positively nonseparable, owing to the fact that $\phi\left(||\boldsymbol{h}||/\sqrt{\gamma(u)+1}\right) \geq \phi(||\boldsymbol{h}||)$, $\forall (\boldsymbol{h}, u) \in \mathbb{R}^d \times \mathbb{R}$.

From a modeling point of view, the model in (13) lacks flexibility because the nonseparability is imposed. In order to get a separability that can vary from fully separable to fully nonseparable Gneiting (2002) proposed a construction that uses the fact that the product of two positive definite functions is positive definite. First, we need the following lemma.

Lemma 1 *Let γ be a variogram defined on $\mathbb{R}$. Then, the function $\gamma_b : \mathbb{R} \to \mathbb{R}$ defined by $\gamma_b(u) := (1 + \gamma(|u|))^b - 1$ with $b \in [0, 1]$ is a variogram.*

Proof: The cases $b \in \{0, 1\}$ are trivial. Using the fact that the function $x \mapsto x^b$, $x > 0$, is a complete Bernstein function for $b \in (0, 1)$, the proof of this Lemma is a direct consequence of Theorem 5 in Porcu and Schilling (2011). □

Using Lemma 1, and multiplying a Gneiting model as in (13) with γ_b by a purely temporal model of the form $(1 + \gamma(u))^{-\delta}$ with $\delta > 0$, a spatio-temporal covariance with a separability

parameter $b \in [0,1]$ thus writes

$$C(\boldsymbol{h},u) = \frac{1}{(\gamma(u)+1)^{\delta}} \frac{1}{(\gamma_b(u)+1)^{d/2}} \phi\left(\frac{||\boldsymbol{h}||}{\sqrt{\gamma_b(u)+1}}\right) = \frac{1}{(\gamma(u)+1)^{\tau}} \phi\left(\frac{||\boldsymbol{h}||}{(\gamma(u)+1)^{b/2}}\right) \quad (14)$$

with $\tau = bd/2+\delta > bd/2$. In Gneiting (2002) as well as in many applications, including Bourotte et al. (2016) and Allard et al. (2022), the variogram $\gamma(u)$ in the denominator of (14) is

$$\gamma(u) = a|u|^{\alpha}, \qquad u \in \mathbb{R}, \quad (15)$$

with $a > 0$, $\alpha \in (0,2]$. In particular $\phi(||\boldsymbol{h}||) := C_{\mathcal{F}}(\boldsymbol{h})$ provides the Gneiting-Matérn and Gneiting-Cauchy spatio-temporal covariance functions when $\mathcal{F}$ stands for $\mathcal{M}$ as in (9) for the Matérn and for $\mathcal{C}$ as in (11) for the Cauchy, respectively.

3.2.3 Simulation of GRFs with covariance functions in the Gneiting class

The simulation algorithm for RFs characterized by a Gneiting-type covariance function relies on the following proposition. Let W be a GRF on $\mathbb{R}$ characterized by its variogram γ. Consider the RF Z defined as follows for $(\boldsymbol{s},t) \in \mathbb{R}^d \times \mathbb{R}$:

$$Z(\boldsymbol{s},t) = \sqrt{2}\cos\left(\sqrt{2\xi}\,\boldsymbol{\Sigma}^{-1/2}\boldsymbol{V}^t\boldsymbol{s} + \frac{\|\boldsymbol{V}\|}{\sqrt{2}}W(t) + \Phi\right), \quad (16)$$

with $\boldsymbol{V} \sim \mathcal{N}_d(\boldsymbol{0},\boldsymbol{I}_d)$, $\Phi \sim \mathcal{U}(0,2\pi)$ and $\xi \sim f$ where f is the scale mixture characterizing the covariance function $\phi \in \mathcal{C}_{\infty}$ (see Sect. 2.4).

Proposition 5 (Allard et al. (2020)) *The random field Z defined in* (16) *is second-order stationary in $\mathbb{R}^d \times \mathbb{R}$, with zero mean and Gneiting-type covariance function as given in* (13).

The proof of this proposition is given in Allard et al. (2020, Appendix B). The construction in (16) is a particular case of a substitution random field, obtained by combining a directing function with stationary increments in $\mathbb{R}^d \times \mathbb{R}$ and a stationary coding process in $\mathbb{R}$, as proposed in Lantuéjoul (2002, Chapter 17), a construction that generalizes the subordination approach introduced by Feller (1966, Chapter 10). We will therefore refer to this approach as the *substitution approach*.

The random field Z defined in (16) is centered and second-order stationary. Its covariance function belongs to the Gneiting class (13), i.e. it is nonseparable. In order to introduce a separability parameter, we have seen that it is necessary to multiply Z by a temporal process Z_T with covariance function $C_T(u) = (1+\gamma(u))^{-\delta}$ with $\delta > 0$. To obtain a random field whose finite-dimensional distributions are approximately Gaussian, one thus simulates a large number of independent random fields $(Z_l, Z_{T,l})$, $l = 1,\ldots,L$ and properly renormalize the sum of their

products. Algorithm 4 below is the spatio-temporal generalization of Algorithm 3 that includes a possible spatial geometric anistropy $\boldsymbol{\Sigma}^{-1/2}$. Notice that here we do not introduce the Box-Müller factor because we now consider the product of the cosine waves with the temporal process Z_T, which is Gaussian. This algorithm is generic in the sense that it is valid for any covariance function $\phi \in \mathcal{C}_\infty$ characterized by its scale mixture f, and for any variogram γ. For the specific case of Gneiting-Matérn and Gneiting-Cauchy models, we refer to Sect. 3.1.

Algorithm 4 Simulation algorithm for Gneiting-type RFs

Require: Spatial covariance function in $\mathcal{C}_\infty$ ans its associated scale mixture f

Require: Variogram γ

Require: Parameters $b \in [0,1]$ and $\delta > 0$

Require: Spatial anisotropy $\boldsymbol{\Sigma}^{-1/2}$

Require: A set of points, $\mathcal{S} \in \mathbb{R}^d \times \mathbb{R}$

Require: A large number L

1: **for** $l = 1$ to L **do**

2: Simulate a temporal RF $Z_{T,l}$ with covariance function $C_T(\cdot) = (1 + \gamma(\cdot))^{-\delta}$

3: Simulate an intrinsic temporal RF W_l with Gaussian increments and variogram $\gamma_b(\cdot) = (1 + \gamma(\cdot))^b - 1$

4: Simulate $\xi_l \sim f$

5: Simulate $\boldsymbol{V}_l \sim \mathcal{N}_d(0, \boldsymbol{I}_d)$

6: set $\boldsymbol{\Omega}_l = \sqrt{2\xi_l}\boldsymbol{\Sigma}^{-1/2}\boldsymbol{V}_l$

7: Simulate $\Phi_l \sim \mathcal{U}(0, 2\pi)$

8: **end for**

9: For each $(\boldsymbol{s}, t) \in \mathcal{S}$ return

$$\tilde{Z}_L(\boldsymbol{s}, t) = \sqrt{\frac{2}{L}} \sum_{l=1}^{L} Z_{T,l}(t) \cos\left(\boldsymbol{\Omega}_l^t \boldsymbol{s} + \frac{\|\boldsymbol{V}_l\|}{\sqrt{2}} W_l(t) + \Phi_l\right)$$

The simulation of the temporal process Z_T can be performed using the Cholesky decomposition of the Gram matrix associated to C_T. The simulation of the intrinsic random field W_l at Step 3 can be done by the same method applied to the increment $\tilde{W}(t) = W(t) - W(0)$, by setting $W(0) = 0$ and using the nonstationary covariance $\mathbb{C}\mathrm{ov}(\tilde{W}(t_1), \tilde{W}(t_2)) = \gamma_b(t_1) + \gamma_b(t_2) - \gamma_b(t_1 - t_2)$. This is possible as long as the number of time steps considered for the simulation is not too large (less than a few tens of thousands). In both cases, the covariance matrix involved in the simulation of the temporal process needs to be inverted only once.

4 Extension to spatio-temporal multivariate GRFs

We now consider spatio-temporal multivariate GRFs. We will restrict ourselves to models defined by Gaussian mixtures, which are essentially extensions to a multivariate context of the Gneiting construction seen in the previous section. Purely spatial multivariate GRFs will be considered at the end of Section 5, as a special case of spatio-temporal multivariate models where there is only one time coordinate. For other stationary multivariate models, we refer the reader to Genton and Kleiber (2015) for a full review.

4.1 Preliminary material

4.1.1 Separability, proportional models and linear models of coregionalization

Consider a multivariate space-time random field $\boldsymbol{Z}(\boldsymbol{s},t) = [Z_i(\boldsymbol{s},t)]_{i=1}^p$ defined on $\mathbb{R}^d \times \mathbb{R}$, where p is the number of random field components, each being real-valued. In this section, it will also be assumed that the multivariate random field $\boldsymbol{Z}(\boldsymbol{s},t)$ is second-order stationary, so that its covariance functions exist and depend only on the space-time lag $(\boldsymbol{h},u) \in \mathbb{R}^d \times \mathbb{R}$ (Chilès and Delfiner, 2012):

$$\mathrm{Cov}(Z_i(\boldsymbol{s},t), Z_j(\boldsymbol{s}+\boldsymbol{h}, t+u)) = C_{ij}(\boldsymbol{h},u), \tag{17}$$

for any pair $i,j = 1,\ldots,p$, $\forall(\boldsymbol{s},\boldsymbol{s}+\boldsymbol{h}) \in \mathbb{R}^d \times \mathbb{R}^d$ and $\forall(t,t+u) \in \mathbb{R}\times\mathbb{R}$. The functions C_{ij} are called direct covariance functions when $i=j$ and cross-covariance functions otherwise. The matrix-valued covariance function $(\boldsymbol{h},u) \mapsto \boldsymbol{C}(\boldsymbol{h},u) = [C_{ij}(\boldsymbol{h},u)]_{i,j=1}^p$, is positive semidefinite on $\mathbb{R}^d \times \mathbb{R}$, that is, for any finite collection of space-time coordinates $(\boldsymbol{s}_k,t_k)_{k=1}^N$, the matrix $\left[[C_{ij}(\boldsymbol{s}_l-\boldsymbol{s}_k, t_l-t_k)]_{i,j=1}^p\right]_{k,l=1}^N$ is positive semidefinite (Wackernagel, 2003).

The concept of separability seen in Section 3.2.2 between the space and time dimensions can take a slightly different meaning in the multivariate context. According to Genton and Kleiber (2015), a multivariate cross-covariance matrix is separable if

$$C_{ij}(\boldsymbol{s}_l-\boldsymbol{s}_k, t_l-t_k) = \sigma_{ij}\rho(\boldsymbol{s}_l-\boldsymbol{s}_k, t_l-t_k) \tag{18}$$

for all $i,j = 1,\ldots,p$ where ρ is a valid spatio-temporal correlation function and where $\sigma_{ij} = C_{ij}(\boldsymbol{0},0)$ is the covariance between variables i and j and $\boldsymbol{\sigma} = [\sigma_{ij}]_{i,j=1}^p$ is a covariance matrix. In the geostasticics literature, separable cross-covariances as in (18) are referred to as proportional models (Chilès and Delfiner, 2012, Chap. 5), and also sometimes as intrinsic correlation models (Wackernagel, 2003). To make a clear distinction between the (usual) spatio-temporal separability and the multivariate separability in (18), we will opt for the geostatistical usage and refer to the latter as a proportional model. The simulation for the proportional model is straightforward.

1. Using Algorithm 4, simulate p i.i.d spatio-temporal random fields with the same spatio-temporal correlation function ρ, denoted $\boldsymbol{Y}(\boldsymbol{s},t) = (Y_1(\boldsymbol{s},t),\dots,Y_p(\boldsymbol{s},t))^t$.

2. At each space-time location $(\boldsymbol{s},t)$, compute $\boldsymbol{Z}(\boldsymbol{s},t) = \boldsymbol{L}\boldsymbol{Y}(\boldsymbol{s},t)$, where $\boldsymbol{L}$ is the Cholesky factorization of the covariance matrix $\boldsymbol{\sigma}$, i.e. $\boldsymbol{L}\boldsymbol{L}^t = \boldsymbol{\sigma}$.

A nonproportional model can be obtained as a linear combination of $1 \leq r \leq p$ proportional models with different correlation functions. The resulting cross covariance function takes the form

$$C_{ij}(\boldsymbol{h},u) = \sum_{k=1}^{r} \sigma_{k,ij}\rho_k(\boldsymbol{h},u), \qquad (\boldsymbol{h},u) \in \mathbb{R}^d \times \mathbb{R}, \tag{19}$$

where ρ_k are valid spatio-temporal correlation functions (some being possibly purely spatial and/or temporal), and $C_{k,ij}$ are r covariance matrices. The model in (19) is referred to as the Linear Model of Coregionalization (LMC), see for example De Iaco et al. (2013) for an application to air pollution in Lombardy. Note also that the product sum model introduced in De Iaco et al. (2001) being the sum of one spatial covariance, one temporal covariance and a separable spatio-temporal covariance model can be seen as a special case of (19). Since a LMC is the sum of r proportional models, simulations from a spatio-temporal LMC is a straightforward extension of the approach described above for proportional models. Other models are possible such as spatio-temporal extensions of convolution methods (Paciorek and Schervish, 2006; Ver Hoef et al., 2004) or latent dimensions (Apanasovich and Genton, 2010; Porcu and Zastavnyi, 2011), but they are not often used for modeling spatio-temporal processes.

In Section 4.2, we will present a valid and flexible parametric class of matrix-valued space–time covariance functions that is inspired by the Gneiting construction using Gaussian mixtures. To our knowledge, this class, presented in Allard et al. (2022), is the most general class built on Gaussian mixtures allowing for different range and smoothness parameters in space and in time for each variable. It is at the same time nonproportional and space-time nonseparable, but it contains those more simple models as special cases. It is therefore referred to as being *fully nonseparable*.

4.1.2 Pseudo-variograms

A key building block of fully nonseparable models are pseudo-variograms, a rather unusual tool. For completeness, we recall here the definition and the main properties of pseudo-variograms. An in-depth presentation is available in Dörr and Schlather (2023a). Let $W_i(\cdot)_{i=1,\dots,p}$ be p RFs on $\mathbb{R}$ (more general spaces could be considered but are not necessary here). Then, the function

$$\gamma_{ij}(u) = 0.5\mathbb{V}\mathrm{ar}\big(W_i(t) - W_j(t+u)\big), \quad u \in \mathbb{R},$$

provided it exists for all $t, u \in \mathbb{R}$ and for all $i, j = 1, \ldots, p$, is called a pseudo-variogram. For any $i = 1, \ldots, p$, the function γ_{ii}, is a usual variogram, i.e., a conditionally negative semidefinite function (Chilès and Delfiner, 2012). For $i \neq j$, the pseudo-variogram has nonnegative entries and is not necessarily an even function. The pseudo-variogram must be unbounded for the direct and cross-covariances $C_{ij}(\boldsymbol{h}, u)$ to vanish as $|u| \to \infty$. Necessary and sufficient conditions for a matrix-valued function to be a pseudo-variogram have been provided in Dörr and Schlather (2023a) and applications to covariance functions, including spatio-temporal covariances, is available in Dörr and Schlather (2023b). Despite these recent results, building valid unbounded matrix-valued pseudo-variograms with different diagonal entries (direct variograms) is still an open question. As a simple model for this, Allard et al. (2022) proposed a constructive model where $W_i(t) = V_0(t) + V_i(t), t \in \mathbb{R}, i = 1, \ldots, p$, where $(V_i)_{i=1}^p$ are second-order stationary random fields on $\mathbb{R}$ with matrix-valued covariance function $\boldsymbol{R}(u) = [R_{ij}(u)]_{i,j=1}^p$ and Y_0 is an independent intrinsic random field with unbounded variogram γ_0. The resulting matrix-valued pseudo-variogram is thus

$$\boldsymbol{\gamma}(u) = \gamma_0(u)\mathbf{1} + \boldsymbol{R}^0 - \boldsymbol{R}(u), \qquad u \in \mathbb{R} \tag{20}$$

where $\boldsymbol{R}^0$ is the matrix with entries $R^0_{ij} = \big(R_{ii}(0) + R_{jj}(0)\big)/2$.

4.2 A fully nonseparable spatio-temporal multivariate model

The following Theorem extends the construction in Allard et al. (2022) to a separability parameter $b \in [0, 1]$.

Theorem 1 *Let $\boldsymbol{\sigma} = [\sigma_{ij}]_{i,j=1}^p$ be a positive semidefinite matrix, $u \mapsto \boldsymbol{\gamma}(u) = [\gamma_{ij}(u)]_{i,j=1}^p$ be a matrix-valued pseudo-variogram on $\mathbb{R}$, and $t \mapsto \boldsymbol{\phi}(t) = [\phi_{ij}(t)]_{i,j=1}^p$ be a matrix of completely monotone functions on $[0, \infty)$ such that $\phi_{ij}(\boldsymbol{h}) = \int_0^\infty e^{-\xi||\boldsymbol{h}||^2}(f_{ii}(\xi)f_{jj}(\xi))^{1/2}d\xi$, where f_{ii} and f_{jj} are probability density functions on $(0, \infty)$. Then, the matrix-valued function $\mathbf{C} : (\boldsymbol{h}, u) \mapsto [C_{ij}(\boldsymbol{h}, u)]_{i,j=1}^p$ with*

$$C_{ij}(\boldsymbol{h}, u) = \frac{\sigma_{ij}}{(\gamma_{ij}(u) + 1)^\tau}\phi_{ij}\left(\frac{\boldsymbol{\Sigma}^{-1/2}\boldsymbol{h}}{(\gamma_{ij}(u) + 1)^{b/2}}\right), \quad (\boldsymbol{h}, u) \in \mathbb{R}^d \times \mathbb{R}, \tag{21}$$

is positive semidefinite in $\mathbb{R}^d \times \mathbb{R}$, with $\tau = bd/2 + \delta > bd/2$ and $b \in [0, 1]$ and where $\boldsymbol{\Sigma}^{-1/2}$ is an asnitropy matrix, as defined in Section 2.3.

The proof for $b = 1$ and $\delta = 0$ is constructive and can be found in Allard et al. (2022, Appendix B). The general case with $b \in (0, 1)$ and $\delta > 0$ is then easily proved in a way similar to the univariate case shown in Sect. 3.2.2, see also Proposition 1 in Dörr and Schlather (2023a). In

this construction, all direct spatio-temporal covariances are positively nonseparable. Notice that in this construction the anisotropy matrix is unique, and hence identical for all the p components of the multivariate random field. A more general construction allowing for different anisotropies for each component will be presented in Section 5.3. As was already the case in Section 3.1.3, the scale parameters are present in the anisotropy matrix and in the mixture f. Care must be taken in order to avoid overparametrization.

The following two examples generalize the models considered in Bourotte et al. (2016) (see also Allard et al. (2022) and Dörr and Schlather (2023a)) with the introduction of a separability parameter $b \in [0,1]$, and where, instead of a univariate variogram γ, a matrix-valued pseudo-variogram $\boldsymbol{\gamma}$ in the denominators allows for different temporal covariance functions for each variable.

Example 4 (Gneiting-Matérn covariance) *Let $f_{ii} : \xi \mapsto \Gamma(\nu_{ii})^{-1}(\kappa_{ii}^2/4)^{\nu_{ii}}\xi^{-\nu_{ii}-1}\exp(-\kappa_{ii}^2/(4\xi))$ be the probability density of an inverse gamma distribution on $(0,\infty)$ with shape parameter $\nu_{ii} > 0$ and scale parameter $\kappa_{ii}^2/4 > 0$. Based on Proposition 3, all direct and cross-covariances belong to the Gneiting-Matérn family:*

$$C_{ij}(\boldsymbol{h},u) = \frac{\Gamma(\nu_{ij})}{\sqrt{\Gamma(\nu_{ii})\,\Gamma(\nu_{jj})}}\frac{\kappa_{ii}^{\nu_{ii}}\,\kappa_{jj}^{\nu_{jj}}}{\kappa_{ij}^{2\nu_{ij}}}\frac{\sigma_{ij}}{(\gamma_{ij}(u)+1)^{\tau}}C_{\mathcal{M}}\left(\frac{\boldsymbol{\Sigma}^{-1/2}\boldsymbol{h}}{(\gamma_{ij}(u)+1)^{b/2}};\kappa_{ij},\nu_{ij}\right), \tag{22}$$

where $C_{\mathcal{M}}$ is the Matérn covariance defined in (9) with $2\nu_{ij} = \nu_{ii}+\nu_{jj}$, $2\kappa_{ij}^2 = \kappa_{ii}^2+\kappa_{jj}^2$, and where b,τ are as in Theorem 1.

Example 5 (Gneiting-Cauchy covariance) *Let $f_{ii} : \xi \mapsto \Gamma(\nu_{ii})^{-1}\exp(-\xi/a_{ii})\,a_{ii}^{-\nu_{ii}}\,\xi^{\nu_{ii}-1}$ be the Gamma probability density on $(0,\infty)$ with shape parameter $\nu_{ii} > 0$ and scale parameter $a_{ii} > 0$. Based on Proposition 4, the direct and cross-covariances belong to the Gneiting-Cauchy family:*

$$C_{ij}(\boldsymbol{h},u) = \frac{\Gamma(\nu_{ij})}{\sqrt{\Gamma(\nu_{ii})\,\Gamma(\nu_{jj})}}\frac{a_{ij}^{\nu_{ij}}}{a_{ii}^{\nu_{ii}/2}a_{jj}^{\nu_{jj}/2}}\frac{\sigma_{ij}}{(\gamma_{ij}(u)+1)^{\tau}}C_{\mathcal{C}}\left(\frac{\boldsymbol{\Sigma}^{-1/2}\boldsymbol{h}}{(\gamma_{ij}(u)+1)^{b/2}};a_{ij},\nu_{ij}\right), \tag{23}$$

where $C_{\mathcal{C}}$ is the Cauchy covariance defined in (11) with $2\nu_{ij} = \nu_{ii}+\nu_{jj}$, $2a_{ij}^{-1} = a_{ii}^{-1}+a_{jj}^{-1}$, and where b,τ are as in Theorem 1.

The advantage of the construction in the examples above is that it is easy to understand and that it is straightforward to simulate realizations of the multivariate random field $\boldsymbol{Z}$, as will be shown in Section 4.3. However, an implicit assumption is that $\phi_{ij}(t) = \int_0^\infty e^{-rt^2}f_{ij}(r)dr$, with f_{ij} verifying $f_{ij} = (f_{ii}f_{jj})^{1/2}$. This is a parsimonious parametrization, but also a restrictive condition, which has been alleviated in Allard et al. (2022, Section 4.1) to propose a very general

Gneiting-Matérn class that allows for both positively or negatively space–time nonseparability. This is a very original feature, differing from all other Gneiting-type spatio-temporal models of our knowledge. However, no scale mixture representation is known for this new class, and hence the simulation makes use of importance sampling and the square root of a large matrix whose entries are given by the spectral density of the Matérn covariance. For this reason, this model is not presented further in this work. Interested readers will find all details as well as a spectral simulation algorithm based on spetral densities in Allard et al. (2022).

4.3 Simulation of multivariate GRFs with a fully nonseparable model

The simulation algorithm for the model (21) was given in Allard et al. (2022) for a RF characterized by an isotropic covariance and space-time full nonseparability (i.e., $b = 1$). For the sake of completeness, we provide a more general version below. Algorithm 5 allows for spatial anisotropy through the anisotropy matrix $\boldsymbol{\Sigma}^{-1/2}$, as discussed after Theorem 1, and also allows for any degree of nonseparability through the parameter $b \in [0, 1]$ ($b = 0$ corresponds to a separable model while $b = 1$ corresponds to a fully nonseparable model, as in Allard et al. (2022, Algorithm 1)). This algorithm deserves a few comments:

1. The direct spatial covariance functions are fully characterized by their scale mixtures f_{ii}. Random scales for each variable are not directly simulated. Instead, a common scale ξ_l is simulated for each cosine wave according to an instrumental scale mixture f whose support must contain the supports of $f_{11}, \ldots, f_{pp}$. Then, an importance ratio $\sqrt{f_{ii}(\xi_l)/f(\xi_l)}$ is computed and weights the cosine waves differently for each variable $i = 1, \ldots, p$. It is this construction that leads to the cross-scale mixture $f_{ij} = \sqrt{f_{ii} f_{jj}}$.

2. Theoretically, any instrumental scale mixture f whose support contains the support of $f_{11}, \ldots, f_{pp}$ is valid. In practice however, for numerical reasons, one should choose a scale mixture f that does not differ too much from the above ones. One possibility is to chose $i \in \{1, \ldots, p\}$ and to set $f = f_{ii}$.

3. Compared to Algorithm 4, the temporal processes $\boldsymbol{W}_l$ are now multivariate. Direct computation of $C_T(t_1, t_2) = \mathbb{E}[\tilde{Z}_{L,i}(\boldsymbol{s}, t_1)\tilde{Z}_{L,j}(\boldsymbol{s}, t_2) \mid \xi]$, with use of the product-to-sum trigonometric identities, shows that $C_T(t_1, t_2)$ is proportional to $\mathbb{E}[\cos(||\boldsymbol{V}||\big(W_i(t_1) - W_j(t_2)\big)/\sqrt{2})]$, with $W_i(t_1) - W_j(t_2)$ being a Gaussian random variable with mean 0 and variance equal to $2\gamma_{ij}(t_1 - t_2)$, i.e. twice the pseudo variogram computed at $t_1 - t_2$. A simulation algorithm for those is given in Allard et al. (2022, Algorithm 2), which is applicable as long as the number of target time coordinates is not too large (in practice, up

to a few thousands). For the rare cases with larger sets of time coordinates, other simulations techniques are available (circular embedding methods, spectral simulation techniques, Gibbs sampling, etc.). The temporal processes $\boldsymbol{Z}_{T,l}$ are easily simulated using Cholesky factorization under the same conditions.

Algorithm 5 Simulation algorithm for multivariate spatio-temporal RFs as in (21).

Require: Covariance function in $\mathcal{C}_\infty$ characterized by its scale mixture f.

Require: Admissible pseudo variogram $\boldsymbol{\gamma}$

Require: Parameters $b \in [0,1]$ and $\delta > 0$

Require: A spatial anisotropy $\boldsymbol{\Sigma}^{-1/2}$

Require: A probability density function f_1, with support equal to $(0, \infty)$

Require: A covariance matrix $\boldsymbol{\sigma} = \boldsymbol{L}\boldsymbol{L}^t$

Require: A set of points, $\mathcal{S} \in \mathbb{R}^d \times \mathbb{R}$

Require: A large number L

1: **for** $l = 1$ to L **do**

2: Simulate a p-variate temporal RF $\boldsymbol{Z}_{T,l}$ with matrix-valued covariance function $\boldsymbol{C}_T(\cdot) = (1 + \boldsymbol{\gamma}(\cdot))^{-\delta}$

3: Simulate a p-variate temporal RF $\boldsymbol{W}_l = [W_{l,i}]_{i=1}^p$ with Gaussian direct and cross-increments, with 0 mean and pseudo-variogram $\boldsymbol{\gamma}_b(\cdot) = (1 + \boldsymbol{\gamma}(\cdot))^b - 1$ following Allard et al. (2022, Algorithm 2)

4: Simulate $\xi_l \sim f$

5: Simulate $\boldsymbol{V}_l \sim \mathcal{N}_d(0, \boldsymbol{I}_d)$

6: set $\boldsymbol{\Omega}_l = \sqrt{2\xi_l}\boldsymbol{\Sigma}^{-1/2}\boldsymbol{V}_l$

7: Simulate $\Phi_l \sim \mathcal{U}(0, 2\pi)$

8: Simulate $\boldsymbol{A}_l \sim \mathcal{N}_p(0, \boldsymbol{\sigma})$

9: **end for**

10: For each $(\boldsymbol{s}, t) \in \mathcal{S}$ return

$$\tilde{Z}_{L,i}(\boldsymbol{s}, t) = \sqrt{\frac{2}{L}} \sum_{l=1}^{L} Z_{T,l,i}(t) \sqrt{\frac{f_{ii}(\xi_l)}{f_1(\xi_l)}} A_{l,i} \cos\left(\boldsymbol{\Omega}_l^t \boldsymbol{s} + \frac{\|\boldsymbol{V}_l\|}{\sqrt{2}} W_{l,i}(t) + \Phi_l\right), \qquad i = 1, \ldots, p$$

5 Extensions to Gaussian Random Fields with nonstationary covariances

We now show how the methods presented above can be extended to GRFs with a nonstationary covariance. In Section 5.1, we first recall some nonstationary models and their simulation algorithms, mostly originating from Paciorek and Schervish (2006) and Emery and Arroyo (2018). For completeness and for an easier reading of the rest of this section, these models are presented using our own notations and with a perspective based on Gaussian mixtures instead of spectral densities. This change of perspective paves the way to the generalization to a much larger class of covariance functions. For all covariance functions in $\mathcal{C}_\infty$ characterized by a Gaussian mixture belonging to the exponential family of distributions, we establish in Theorem 2 a very general nonstationary construction with closed-form formulae for the parameters of the nonstationary covariance function. This result leads to Algorithm 6. This generic algorithm is then exemplified for the Matérn and Cauchy covariance functions. Building on this, we then propose in Section 5.3 a spatio-temporal extension which provides new nonstationary spatio-temporal models, univariate and multivariate, along with their simulation Algorithm 7.

5.1 Background material on nonstationary covariances

Let us assume that the parameters of the spectral density μ vary over $\mathbb{R}^d$ and let us denote $\mu_{\boldsymbol{s}}$ the spectral density at location $\boldsymbol{s} \in \mathbb{R}^d$. Define

$$Z(\boldsymbol{s}) = \sqrt{2\frac{\mu_{\boldsymbol{s}}(\boldsymbol{\Omega})}{\mu_0(\boldsymbol{\Omega})}}\cos(\boldsymbol{\Omega}^t\boldsymbol{s} + \Phi),$$

where $\boldsymbol{\Omega}$ is distributed according to the spectral density μ_0 and Φ is a uniform random variable over $[0, 2\pi)$. Then, straightforward trigonometric manipulations show that the covariance between $Z(\boldsymbol{s}_1)$ and $Z(\boldsymbol{s}_2)$ is equal to the Fourier transform of the geometric average of the spectral densities associated with locations $\boldsymbol{s}_1$ and $\boldsymbol{s}_2$:

$$\mathbb{C}\mathrm{ov}\left(Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2)\right) = \int_{\mathbb{R}^d} \cos\left(\boldsymbol{\omega}^t(\boldsymbol{s}_1 - \boldsymbol{s}_2)\right) \sqrt{\mu_{\boldsymbol{s}_1}(\boldsymbol{\omega})\mu_{\boldsymbol{s}_2}(\boldsymbol{\omega})}\, d\boldsymbol{\omega}. \tag{24}$$

This very general construction is at the core of the nonstationary extensions of the spectral algorithms presented in Emery and Arroyo (2018). Since in this work we focus on Gaussian mixture simulation approaches, we must first consider nonstationary Gaussian covariance functions.

Example 6 (Nonstationary Gaussian covariance) *Let us consider the family of Gaussian spectral densities*

$$\mu^G_{\boldsymbol{\Sigma}_{\boldsymbol{s}}}(\boldsymbol{\omega}) = \left(2\sqrt{\pi}\right)^{-d} |\boldsymbol{\Sigma}_{\boldsymbol{s}}|^{1/2} \exp\left(-\boldsymbol{\omega}^t\boldsymbol{\Sigma}_{\boldsymbol{s}}\boldsymbol{\omega}/4\right), \qquad \forall \boldsymbol{s} \in \mathbb{R}^d,$$

where $\boldsymbol{\Sigma}_{\boldsymbol{s}}^{-1/2}$ *is the anisotropy as defined in* (4)*, with parameters* $r_1(\boldsymbol{s}), r_2(\boldsymbol{s})$ *and* $\theta(\boldsymbol{s})$*, all depending on* $\boldsymbol{s} \in \mathbb{R}^d$*. Emery and Arroyo (2018) established that the application of* (24) *to this family of Gaussian densities leads to*

$$\mathbb{C}\mathrm{ov}\left(Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2)\right) = |\boldsymbol{\Sigma}_{\boldsymbol{s}_1}|^{1/4} |\boldsymbol{\Sigma}_{\boldsymbol{s}_2}|^{1/4} \left| \frac{\boldsymbol{\Sigma}_{\boldsymbol{s}_1} + \boldsymbol{\Sigma}_{\boldsymbol{s}_2}}{2} \right|^{-1/2} \exp\left\{ -(\boldsymbol{s}_1 - \boldsymbol{s}_2)^t \left(\frac{\boldsymbol{\Sigma}_{\boldsymbol{s}_1} + \boldsymbol{\Sigma}_{\boldsymbol{s}_2}}{2} \right)^{-1} (\boldsymbol{s}_1 - \boldsymbol{s}_2) \right\}. \tag{25}$$

This family of nonstationary covariance models was already introduced in Paciorek and Schervish (2006) using nonstationary kernel convolutions of white noise. Using the Gaussian mixture representation the same authors further proved that, for a covariance function $\phi(\boldsymbol{h}) \in \mathcal{C}_\infty$, the function defined by

$$\phi_{NS}(\boldsymbol{s}_1, \boldsymbol{s}_2) = |\boldsymbol{\Sigma}_{\boldsymbol{s}_1}|^{1/4} |\boldsymbol{\Sigma}_{\boldsymbol{s}_2}|^{1/4} |\boldsymbol{\Sigma}_{\boldsymbol{s}_1, \boldsymbol{s}_2}|^{-1/2} \phi\Big((\boldsymbol{s}_1 - \boldsymbol{s}_2)^t \boldsymbol{\Sigma}_{\boldsymbol{s}_1, \boldsymbol{s}_2}^{-1/2} \Big), \tag{26}$$

is a covariance function with a nonstationary anisotropy on $\mathbb{R}^d$, $d \geq 1$, with $\boldsymbol{\Sigma}_{\boldsymbol{s}_1, \boldsymbol{s}_2} = \big(\boldsymbol{\Sigma}_{\boldsymbol{s}_1} + \boldsymbol{\Sigma}_{\boldsymbol{s}_2}\big)/2$. Notice that in a nonstationary context, it is more convenient to let all three parameters of the matrix $\boldsymbol{\Sigma}_{\boldsymbol{s}}^{-1/2}$: $(r_1(\boldsymbol{s}), r_2(\boldsymbol{s}), \theta(\boldsymbol{s}))$ vary in space. In order to avoid overparametrization, the scale parameter in ϕ is thus preferably set to 1.

5.2 A general result for spatial nonstationary GRFs

The relationship (24) can readily be used to simulate nonstationary covariance functions depending only on scale parameters, such as the exponential or the Gaussian covariance, using a reference spectral density μ_0 whose support is $\mathbb{R}^d$ and an importance sampling approach, see Emery and Arroyo (2018). However, for Matérn and Cauchy covariance functions for example, there is also a shape parameter that can vary in space. For the Matérn covariance function, Emery and Arroyo (2018) proposed an extra randomization step, where a random shape, say ξ, is drawn according to the exponential distribution. The nonstationarity of the parameters of the Gaussian mixture is then accounted for using importance weights. Notice that the choice of an exponential density as instrumental density is arbitrary. Other choices are possible.

This approach is now generalized to any mixture belonging to the exponential family of distributions, thus encompassing all cases considered in Stein (2005). The exponential family is a parametric set of families whose probability distribution functions can be expressed in the form

$$f(\xi; \boldsymbol{\theta}) = h(\boldsymbol{\theta}) \exp\left(-\boldsymbol{\ell}(\boldsymbol{\theta})^t \boldsymbol{T}(\xi) \right), \tag{27}$$

where $\boldsymbol{T}(\xi)$ is a vector of functions of ξ, $\boldsymbol{\ell}(\boldsymbol{\theta})$ is a vector of functions of the parameters and $\boldsymbol{\theta}$ is a vector of p parameters. This class includes in particular the Gamma and Inverse Gamma scale

mixtures which define the Cauchy and Matérn covariances, respectively. It also includes the Gaussian, Inverse Gaussian and Beta (with fixed support) distributions, among many others.

Theorem 2 *Let $\phi(\cdot, \boldsymbol{\theta}) \in \mathcal{C}_\infty$ be a stationary, isotropic covariance function defined by its scale mixture $f(\cdot; \boldsymbol{\theta})$ belonging to the exponential family of distributions as in* (27) *and let us denote $f_1(\cdot) = f(\cdot\,; \mathbf{1})$. Assume that the parameters $\boldsymbol{\theta}_{\boldsymbol{s}}$ and the anisotropy $\boldsymbol{\Sigma}_{\boldsymbol{s}}^{-1/2}$ can depend on $\boldsymbol{s} \in \mathbb{R}^d$. Let $\xi, \boldsymbol{\Omega}, \Phi$ be random variables distributed according to $f_1, \sqrt{2\xi}\mathcal{N}_d(\mathbf{0}, \boldsymbol{I}_d)$ and $\mathcal{U}(0, 2\pi)$, respectively. Then, the covariance function of*

$$Z(\boldsymbol{s}) = \sqrt{2\frac{f(\xi; \boldsymbol{\theta}_{\boldsymbol{s}})}{f_1(\xi)}} \sqrt{\frac{\mu^G_{\boldsymbol{\Sigma}_{\boldsymbol{s}}}(\boldsymbol{\Omega})}{\mu^G_{\boldsymbol{I}_d}(\boldsymbol{\Omega})}} \cos\big(\boldsymbol{\Omega}^t \boldsymbol{s} + \Phi\big),$$

is equal to

$$\mathbb{C}\mathrm{ov}\left(Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2)\right) = |\boldsymbol{\Sigma}_{\boldsymbol{s}_1}|^{1/4} |\boldsymbol{\Sigma}_{\boldsymbol{s}_2}|^{1/4} |\boldsymbol{\Sigma}_{\boldsymbol{s}_1, \boldsymbol{s}_2}|^{-1/2} \phi(\boldsymbol{\Sigma}^{-1/2}_{\boldsymbol{s}_1, \boldsymbol{s}_2}(\boldsymbol{s}_1 - \boldsymbol{s}_2); \boldsymbol{\theta}_{\boldsymbol{s}_1, \boldsymbol{s}_2}), \tag{28}$$

where $\phi(\cdot; \boldsymbol{\theta}_{\boldsymbol{s}_1, \boldsymbol{s}_2})$ is the covariance function defined by the scale mixture $f(\cdot, \boldsymbol{\theta}_{\boldsymbol{s}_1, \boldsymbol{s}_2})$ with $\boldsymbol{\theta}_{\boldsymbol{s}_1, \boldsymbol{s}_2}$ being such that

$$\boldsymbol{\ell}(\boldsymbol{\theta}_{\boldsymbol{s}_1, \boldsymbol{s}_2}) = \frac{\boldsymbol{\ell}(\boldsymbol{\theta}_{\boldsymbol{s}_1}) + \boldsymbol{\ell}(\boldsymbol{\theta}_{\boldsymbol{s}_2})}{2}. \tag{29}$$

The proof is given in Appendix B. This Theorem will be illustrated with the Matérn and Cauchy covariance functions in the next Section. It leads to the following very general algorithm:

Algorithm 6 Generic nonstationary scale mixture simulation algorithm

Require: A scale mixture pdf, $f(\cdot;\boldsymbol{\theta})$, belonging to the exponential family

Require: A set of points, $\mathcal{S} \in \mathbb{R}^d$

Require: Varying parameters $\boldsymbol{\theta}_{\boldsymbol{s}}$ and anisotropy matrices $\boldsymbol{\Sigma}_{\boldsymbol{s}}^{-1/2}$, $\forall \boldsymbol{s} \in \mathcal{S}$

Require: A large number L

1: Set $f_1 := f(\cdot;\boldsymbol{\theta})$ with $\boldsymbol{\theta} = \mathbf{1}$
2: **for** $l = 1$ to L **do**
3: Simulate $\xi_l \sim f_1$
4: Simulate $\boldsymbol{V}_l \sim \mathcal{N}_d(0, \boldsymbol{I}_d)$
5: Compute $\boldsymbol{\Omega}_l = \sqrt{2\xi_l}\,\boldsymbol{V}_l$
6: Simulate $\Phi_l \sim \mathcal{U}(0, 2\pi)$
7: **end for**
8: For each $\boldsymbol{s} \in \mathcal{S}$ return

$$\tilde{Z}(\boldsymbol{s}) = \sqrt{\frac{2}{L}} \sum_{l=1}^{L} \sqrt{\frac{f(\xi_l;\boldsymbol{\theta}_{\boldsymbol{s}})}{f_1(\xi_l)}} \sqrt{\frac{\mu^G_{\boldsymbol{\Sigma}_{\boldsymbol{s}}}(\boldsymbol{\Omega}_l)}{\mu^G_{\boldsymbol{I}_d}(\boldsymbol{\Omega}_l)}} \cos\left(\boldsymbol{\Omega}_l^t \boldsymbol{s} + \Phi_l\right).$$

Example 7 (Nonstationary Matérn covariance) *The scale mixture for the Matérn covariance is* $f_{\mathcal{M}}(\xi;(\nu,\kappa)) = \left(\kappa^2/4\right)^{\nu} \Gamma(\nu)^{-1} \xi^{-1-\nu} e^{-\kappa^2/4\xi}$ *(see Proposition 3). It belongs to the exponential family as in* (27) *with* $\boldsymbol{\ell}(\boldsymbol{\theta}) = (\nu+1, \kappa^2)^t$, $\boldsymbol{T}(\xi) = (\ln\xi, 1/4\xi)^t$ *and* $h(\boldsymbol{\theta}) = \left(\kappa^2/4\right)^{\nu} \Gamma(\nu)^{-1}$. *Application of Theorem 2 leads to*

$$\boldsymbol{\ell}(\boldsymbol{\theta}_{\boldsymbol{s}_1,\boldsymbol{s}_2}) = \left((\nu_{\boldsymbol{s}_1} + \nu_{\boldsymbol{s}_2})/2 + 1, (\kappa^2_{\boldsymbol{s}_1} + \kappa^2_{\boldsymbol{s}_2})/2\right)^t$$

and

$$h(\boldsymbol{\theta}_{\boldsymbol{s}_1,\boldsymbol{s}_2}) = \frac{1}{\Gamma\left((\nu_{\boldsymbol{s}_1} + \nu_{\boldsymbol{s}_2})/2\right)} \left(\frac{(\kappa^2_{\boldsymbol{s}_1} + \kappa^2_{\boldsymbol{s}_2})/2}{4}\right)^{(\nu_{\boldsymbol{s}_1}+\nu_{\boldsymbol{s}_2})/2}.$$

As seen in Section 3.1.3, when modeling anisotropies, the ranges are sometimes parameterized in $\boldsymbol{\Sigma}_{\boldsymbol{s}_1}$, *while setting* $\kappa^2_{\boldsymbol{s}} = 1$. *In this case, the above expressions simplify to*

$$\boldsymbol{\ell}(\boldsymbol{\theta}_{\boldsymbol{s}_1,\boldsymbol{s}_2}) = \left((\nu_{\boldsymbol{s}_1} + \nu_{\boldsymbol{s}_2})/2 + 1, 1\right)^t, \quad h(\boldsymbol{\theta}_{\boldsymbol{s}_1,\boldsymbol{s}_2}) = \frac{1}{\Gamma\left((\nu_{\boldsymbol{s}_1} + \nu_{\boldsymbol{s}_2})/2\right)} \left(\frac{1}{4}\right)^{(\nu_{\boldsymbol{s}_1}+\nu_{\boldsymbol{s}_2})/2}.$$

574 *Hence* $\nu_{\boldsymbol{s}_1,\boldsymbol{s}_2} = (\nu_{\boldsymbol{s}_1} + \nu_{\boldsymbol{s}_2})/2$ *and* $\kappa_{\boldsymbol{s}_1,\boldsymbol{s}_2} = 1$, *thus leading to the nonstationary Matérn covariance*
575 *already obtained in Emery and Arroyo (2018, Eq. (27)),*

$$C(\boldsymbol{s}_1, \boldsymbol{s}_2) = \frac{\Gamma\left((\nu_{\boldsymbol{s}_1} + \nu_{\boldsymbol{s}_2})/2\right)}{\sqrt{\Gamma(\nu_{\boldsymbol{s}_1})\Gamma(\nu_{\boldsymbol{s}_2})}} C_{\mathcal{M}}\left(\boldsymbol{\Sigma}^{-1/2}_{\boldsymbol{s}_1,\boldsymbol{s}_2}(\boldsymbol{s}_1 - \boldsymbol{s}_2); (\nu_{\boldsymbol{s}_1} + \nu_{\boldsymbol{s}_2})/2\right). \tag{30}$$

Example 8 (Nonstationary Cauchy covariance) *The scale mixture for the Cauchy covariance is* $f_{\mathcal{C}}(\xi;(\nu,a)) = a^{-\nu}\Gamma(\nu)^{-1}\xi^{\nu-1}e^{-\xi/a}$, *and thus* $\boldsymbol{\ell}(\boldsymbol{\theta}) = (1-\nu, 1/a)^t$, $\boldsymbol{T}(\xi) = (\ln\xi, \xi)^t$ *and* $h(\boldsymbol{\theta}) = a^{-\nu}\Gamma(\nu)^{-1}$. *Application of Theorem 2 leads to*

$$\boldsymbol{\ell}(\boldsymbol{\theta}_{\boldsymbol{s}_1,\boldsymbol{s}_2}) = \left(1-(\nu_{\boldsymbol{s}_1}+\nu_{\boldsymbol{s}_2})/2, (a_{\boldsymbol{s}_1}^{-1}+a_{\boldsymbol{s}_2}^{-1})/2\right)^t$$

and

$$h(\boldsymbol{\theta}_{\boldsymbol{s}_1,\boldsymbol{s}_2}) = \frac{1}{\Gamma\left((\nu_{\boldsymbol{s}_1}+\nu_{\boldsymbol{s}_2})/2\right)}\left(\frac{2a_{\boldsymbol{s}_1}a_{\boldsymbol{s}_2}}{a_{\boldsymbol{s}_1}+a_{\boldsymbol{s}_2}}\right)^{-(\nu_{\boldsymbol{s}_1}+\nu_{\boldsymbol{s}_2})/2}.$$

When setting $a_{\boldsymbol{s}} = 1$, *the above expressions simplify, with* $\nu_{\boldsymbol{s}_1,\boldsymbol{s}_2} = (\nu_{\boldsymbol{s}_1}+\nu_{\boldsymbol{s}_2})/2$ *and* $a_{\boldsymbol{s}_1,\boldsymbol{s}_2} = 1$, *thus leading to the same right-hand side as in* (30) *for* $C_{\mathcal{C}}$.

5.3 Nonstationary multivariate spatio-temporal random fields

Capitalizing on all results seen above, we are now able to extend the parsimonious multivariate spatio-temporal model of Theorem 1 to be nonstationary in space and time. In this section, for ease of notation, we will denote $\boldsymbol{x} = (\boldsymbol{s}, t) \in \mathbb{R}^d \times \mathbb{R}$ a space-time coordinate.

Theorem 3 *Let* $\boldsymbol{\sigma}_{\boldsymbol{x}} = [\sigma_{ij,\boldsymbol{x}}]_{i,j=1}^p$ *be a collection of positive semidefinite matrices, defined at all space-time locations* $\boldsymbol{x} \in \mathbb{R}^d \times \mathbb{R}$, *and let* $\boldsymbol{L}_{\boldsymbol{x}}$ *be such that* $\boldsymbol{\sigma}_{\boldsymbol{x}} = \boldsymbol{L}_{\boldsymbol{x}}\boldsymbol{L}_{\boldsymbol{x}}^t$. *Let* $u \mapsto \boldsymbol{\gamma}(u) = [\gamma_{ij}(u)]_{i,j=1}^p$ *be a matrix-valued pseudo-variogram on* $\mathbb{R}$. *Define*

$$\phi(r;\boldsymbol{\theta}_{i,j,\boldsymbol{x}_1\boldsymbol{x}_2}) = \int_0^\infty e^{-\xi r^2} f(r;\boldsymbol{\theta}_{i,\boldsymbol{x}_1}) f(r;\boldsymbol{\theta}_{j,\boldsymbol{x}_2}) d\xi$$

where $f(\cdot;\boldsymbol{\theta}_{i,\boldsymbol{x}})$ *is a probability density function on* $(0,\infty)$, *varying with* $\boldsymbol{x}$ *and* $i = 1,\ldots,p$, *that belong to a family of scale mixtures associated to* $\phi \in \mathcal{C}_\infty$. *Then, the matrix-valued function* $\mathbf{C} : (\boldsymbol{x}_1, \boldsymbol{x}_2) \mapsto [C_{ij}(\boldsymbol{x}_1,\boldsymbol{x}_2)]_{i,j=1}^p$ *with*

$$C_{ij}(\boldsymbol{x}_1,\boldsymbol{x}_2) = C_{ij}(\boldsymbol{s}_1,\boldsymbol{s}_2;t_1,t_2) = \frac{\sigma_{ij,\boldsymbol{x}_1\boldsymbol{x}_2}}{(\gamma_{ij}(t_1-t_2)+1)^{d/2}}\phi\left(\frac{\boldsymbol{s}_1-\boldsymbol{s}_2}{(\gamma_{ij}(t_1-t_2)+1)^{1/2}};\boldsymbol{\theta}_{i,j,\boldsymbol{x}_1,\boldsymbol{x}_2}\right), \quad (31)$$

is positive semidefinite on $\mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}$, *where* $\boldsymbol{\sigma}_{\boldsymbol{x}_1\boldsymbol{x}_2} = \boldsymbol{L}_{\boldsymbol{x}_1}\boldsymbol{L}_{\boldsymbol{x}_2}^t$, *and where* $\boldsymbol{\theta}_{i,j,\boldsymbol{x}_1,\boldsymbol{x}_2}$ *is defined as in Eq.* (29) *with* $\boldsymbol{\ell}(\boldsymbol{\theta}_{i,j,\boldsymbol{x}_1,\boldsymbol{x}_2}) = 0.5\left(\boldsymbol{\ell}(\boldsymbol{\theta}_{i,\boldsymbol{x}_1})\right) + \boldsymbol{\ell}(\boldsymbol{\theta}_{j,\boldsymbol{x}_2}))$.

Moreover, in the case of anisotropies $\boldsymbol{\Sigma}_{ii,\boldsymbol{x}}^{-1/2}$ *defined for all* $\boldsymbol{x} \in \mathbb{R}^d \times \mathbb{R}$ *and all* $i = 1,\ldots,p$, *the following nonstationary matrix-valued spatio-temporal function is a valid covariance:*

$$C_{ij}(\boldsymbol{s}_1,\boldsymbol{s}_2;t_1,t_2) = |\boldsymbol{\Sigma}_{ii,\boldsymbol{x}_1}|^{1/4}|\boldsymbol{\Sigma}_{jj,\boldsymbol{x}_2}|^{1/4}\frac{\sigma_{ij,\boldsymbol{x}_1\boldsymbol{x}_2}}{|\boldsymbol{\Lambda}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2}|^{1/2}}\phi\left(\boldsymbol{\Lambda}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2}^{-1/2}(\boldsymbol{s}_1-\boldsymbol{s}_2);\boldsymbol{\theta}_{i,j,\boldsymbol{x}_1,\boldsymbol{x}_2}\right) \quad (32)$$

where $\boldsymbol{\Lambda}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2} = (\boldsymbol{\Sigma}_{ii,\boldsymbol{x}_1}+\boldsymbol{\Sigma}_{jj,\boldsymbol{x}_2})/2 + \gamma_{ij}(t_1-t_2)\boldsymbol{I}_d$.

The proof, given in the appendix C, is based on Theorem 1 and 2 and the simulation procedure is detailed in Algorithm 7. Note that the square root matrix $\boldsymbol{\Lambda}^{-1/2}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2}$ is uniquely defined because $\boldsymbol{\Lambda}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2}$ is positive definite.

This Theorem deserves some comments. Firstly, we emphasize that all parameters of ϕ (parameters $\boldsymbol{\theta}_{\boldsymbol{x}}$ and anisotropy matrices $\boldsymbol{\Sigma}_{\boldsymbol{x}}$), as well as the correlation matrices $\boldsymbol{\sigma}_{\boldsymbol{x}}$ are nonstationary in space and in time.

Secondly, we emphasize that in the anisotropic case the temporal covariance varies in space. Setting $\boldsymbol{s}_1 = \boldsymbol{s}_2 = \boldsymbol{s}$ in Eq. (32), the temporal covariance at $\boldsymbol{s} \in \mathbb{R}^d$ is

$$C_{ij}(\boldsymbol{s},\boldsymbol{s},t_1,t_2) = |\boldsymbol{\Sigma}_{ii,\boldsymbol{s},t_1}|^{1/4}|\boldsymbol{\Sigma}_{jj,\boldsymbol{s},t_2}|^{1/4}\frac{\sigma_{ij,\boldsymbol{s},t_1,t_2}}{|(\boldsymbol{\Sigma}_{ii,\boldsymbol{s},t_1}+\boldsymbol{\Sigma}_{jj,\boldsymbol{s},t_2})/2+\gamma_{ij}(t_1-t_2)\boldsymbol{I}_d|^{1/2}}. \tag{33}$$

This shows that having different anisotropies in each component of a multivariate GRF has a profound impact even on the temporal covariances.

Furthermore, in the stationary case with a unique anisotropy for all components of the multivariate GRF, $\boldsymbol{\Sigma}_{ii,\boldsymbol{x}} = \boldsymbol{\Sigma}$, $\forall i = 1,\dots,p$ and $\forall \boldsymbol{x} \in \mathbb{R}^d \times \mathbb{R}$, Eq. (32) simplifies to

$$C_{ij}(\boldsymbol{s},\boldsymbol{s},t_1,t_2) = \frac{|\boldsymbol{\Sigma}|^{1/2}\sigma_{ij,\boldsymbol{s},t_1,t_2}}{|\boldsymbol{\Sigma}+\gamma_{ij}(t_1-t_2)\boldsymbol{I}_d|^{1/2}}\phi\left((\boldsymbol{\Sigma}+\gamma_{ij}(t_1-t_2)\boldsymbol{I}_d)^{-1/2}(\boldsymbol{s}_1-\boldsymbol{s}_2);\boldsymbol{\theta}_{i,j,\boldsymbol{x}_1,\boldsymbol{x}_2}\right). \tag{34}$$

This last result shows an important difference from the construction obtained in Theorem 1. However, in the isotropic case with $\boldsymbol{\Sigma} = \boldsymbol{I}_d$, we have $\boldsymbol{\Lambda}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2} = (1+\gamma_{ij}(t_1-t_2))\boldsymbol{I}_d$, and we get the simpler expression (31).

Thirdly, it must be noticed that in the construction of Theorem 3 it was not possible to include a separability parameter b, as was done in the stationary setting. In order to get some flexibility, it is however possible to multiply the covariance function in Eq. (32) by a temporal covariance parameterized by the same pseudo-variogram, i.e. $\boldsymbol{C}_T(t_1,t_2) = \left[(1+\gamma_{ij}(t_1-t_2))^{-\delta}\right]_{i,j=1,p}$ with $\delta > 0$, which is similar to what was proposed in Section 3.2.2 for the univariate case. Algorithm 7 includes this possibility.

Fourthly, for ease of exposition, Theorem 3 is stated for a stationary matrix-valued variogram that depends only on $t_1 - t_2$. However, it can easily be extended to the more general case where $\boldsymbol{\gamma}(t_1,t_2) = \gamma_0(t_1-t_2)\mathbf{1}+\boldsymbol{R}^0_{t_1,t_2}-\boldsymbol{R}(t_1,t_2)$, where $\boldsymbol{R}$ is a valid multivariate nonstationary temporal covariance function, as in Kleiber and Nychka (2012) or Genton and Kleiber (2015) for example. $\boldsymbol{R}^0_{t_1,t_2}$ is the matrix with entries $R^0_{ij,t_1,t_2} = \big(R_{ii}(t_1,t_1)+R_{ii}(t_2,t_2)+R_{jj}(t_1,t_1)+R_{jj}(t_2,t_2)\big)/4$. In that case the simulation step 4 in Algorithm 7 remains unchanged.

As in the stationary case in Section 4.2, we provide an example of the parsimonious nonstationary Gneiting-Matérn model, which follows directly from Theorem 3 and Proposition 3. Although not explicitly given here, a nonstationary version of the Gneiting-Cauchy model from 5 can be constructed in a similar manner.

Algorithm 7 Simulation algorithm for nonstationary multivariate spatio-temporal RF.

Require: A family of scale mixtures, $f(\cdot;\boldsymbol{\theta})$, belonging to the exponential family

Require: A set of points, $\mathcal{S} \in \mathbb{R}^d \times \mathbb{R}$

Require: Varying parameters $\boldsymbol{\theta}_{i,\boldsymbol{x}}$ and anisotropy matrices $\boldsymbol{\Sigma}_{ii,\boldsymbol{x}}^{-1/2} \ \ \forall \boldsymbol{x} \in \mathcal{S},\ \forall i \in \{1,\ldots,p\}$

Require: Varying covariance matrices $\boldsymbol{\sigma}_{\boldsymbol{x}} = \boldsymbol{L}_{\boldsymbol{x}}\boldsymbol{L}_{\boldsymbol{x}}^t$

Require: Admissible pseudo variogram $\boldsymbol{\gamma}$

Require: Parameter $\delta > 0$

Require: A large number L

1: Set $f_1 := f(\boldsymbol{\theta})$ for $\boldsymbol{\theta} = \mathbf{1}$

2: **for** $l = 1$ to L **do**

3: Simulate a p-variate temporal RF $\boldsymbol{Z}_{T,l}$ with matrix-valued covariance function $\boldsymbol{C}_T(\cdot) = (1 + \boldsymbol{\gamma}(\cdot))^{-\delta}$

4: Simulate a p-variate temporal RF $\boldsymbol{W}_l = [W_{l,i}]_{i=1}^p$ with Gaussian direct and cross-increments, with 0 mean and pseudo-variogram $\boldsymbol{\gamma}$

5: Simulate $\xi_l \sim f_1$

6: Simulate $\boldsymbol{V}_l \sim \mathcal{N}_d(0, \boldsymbol{I}_d)$

7: set $\boldsymbol{\Omega}_l = \sqrt{2\xi_l}\boldsymbol{V}_l$

8: Simulate $\Phi_l \sim \mathcal{U}(0, 2\pi)$

9: Simulate $\boldsymbol{A}_l \sim \mathcal{N}_p(0, \boldsymbol{I}_p)$

10: **end for**

11: For each $\boldsymbol{x} = (\boldsymbol{s}, t) \in \mathcal{S}$, and for $i = 1, \ldots, p$ return

$$\tilde{Z}_{L,i}(\boldsymbol{s},t) = \sqrt{\frac{2}{L}} \sum_{l=1}^{L} Z_{T,l,i}(t) \sqrt{\frac{f(\xi_l;\boldsymbol{\theta}_{i,\boldsymbol{x}})}{f_1(\xi_l)}} \sqrt{\frac{\mu^G_{\boldsymbol{\Sigma}_{ii,\boldsymbol{x}}}(\sqrt{2}\boldsymbol{V}_l)}{\mu^G_{\boldsymbol{I}_d}(\sqrt{2}\boldsymbol{V}_l)}} (\boldsymbol{L}_{\boldsymbol{x}}\boldsymbol{A}_l)_i \cos\left(\boldsymbol{\Omega}_l^t\boldsymbol{s} + \frac{\|\boldsymbol{V}_l\|}{\sqrt{2}} W_i(t) + \Phi_l\right)$$

Example 9 (Nonstationary Gneiting-Matérn multivariate covariance) *Let*

$$\forall \boldsymbol{x} = (\boldsymbol{s}, t) \in \mathbb{R}^d \times \mathbb{R}, \quad f_{ii,\boldsymbol{x}} : \xi \mapsto \Gamma(\nu_{ii,\boldsymbol{x}})^{-1} (\kappa_{ii,\boldsymbol{x}}/2)^{2\nu_{ii,\boldsymbol{x}}} \xi^{-1-\nu_{ii,\boldsymbol{x}}} \exp(-\kappa_{ii,\boldsymbol{x}}^2/(4\xi))$$

be the probability density of an Inverse Gamma distribution on $(0, \infty)$ *with shape parameter* $\nu_{ii,\boldsymbol{x}} > 0$ *and scale parameter* $\kappa_{ii,\boldsymbol{x}}^2/4 > 0$ *and let* $\boldsymbol{\Sigma}_{ii,\boldsymbol{x}}^{-1/2}$ *be anisotropy matrices defined for all* $\boldsymbol{x} \in \mathbb{R}^d \times \mathbb{R}$ *and all* $i = 1, \ldots, p$*. Then the nonstationary Gneiting-Matérn multivariate spatio-temporal cross-covariance is:*

$$\begin{aligned} C_{ij}(\boldsymbol{s}_1, \boldsymbol{s}_2; t_1, t_2) = \frac{|\boldsymbol{\Sigma}_{ii,\boldsymbol{x}_1}|^{1/4} |\boldsymbol{\Sigma}_{jj,\boldsymbol{x}_2}|^{1/4}}{|\boldsymbol{\Lambda}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2}|^{1/2}} \frac{\Gamma(\nu_{ij,\boldsymbol{x}_1\boldsymbol{x}_2})}{\sqrt{\Gamma(\nu_{ii,\boldsymbol{x}_1\boldsymbol{x}_1})\,\Gamma(\nu_{jj,\boldsymbol{x}_2\boldsymbol{x}_2})}} \frac{\kappa_{ii,\boldsymbol{x}_1\boldsymbol{x}_1}^{\nu_{ii,\boldsymbol{x}_1\boldsymbol{x}_1}} \kappa_{jj,\boldsymbol{x}_2\boldsymbol{x}_2}^{\nu_{jj,\boldsymbol{x}_2\boldsymbol{x}_2}}}{\kappa_{ij,\boldsymbol{x}_1\boldsymbol{x}_2}^{2\nu_{ij,\boldsymbol{x}_1\boldsymbol{x}_2}}} \\ \times \sigma_{ij,\boldsymbol{x}_1\boldsymbol{x}_2}\; C_{\mathcal{M}}\left(\boldsymbol{\Lambda}_{ij,\boldsymbol{x}_1\boldsymbol{x}_2}^{-1/2}(\boldsymbol{s}_1 - \boldsymbol{s}_2); \kappa_{ij,\boldsymbol{x}_1\boldsymbol{x}_2}, \nu_{ij,\boldsymbol{x}_1\boldsymbol{x}_2}\right), \end{aligned} \tag{35}$$

where $C_{\mathcal{M}}(\cdot; \kappa_{ij,\boldsymbol{x}_1\boldsymbol{x}_2}, \nu_{ij,\boldsymbol{x}_1\boldsymbol{x}_2})$ *is the Matérn covariance with parameters* $2\nu_{ij,\boldsymbol{x}_1\boldsymbol{x}_2} = \nu_{ii,\boldsymbol{x}_1} + \nu_{jj,\boldsymbol{x}_2}$ *and* $2\kappa_{ij,\boldsymbol{x}_1\boldsymbol{x}_2}^2 = \kappa_{ii,\boldsymbol{x}_1}^2 + \kappa_{jj,\boldsymbol{x}_2}^2$*, and with* $\boldsymbol{\Lambda}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2}$ *being defined as in Theorem 3.*

5.4 A note on some special cases

Spatial multivariate nonstationary GRFs can be seen as a restriction of the corresponding spatio-temporal GRFs to $\mathbb{R}^d \times \{0\}$. Valid models are thus easily obtained as a special case of the spatio-temporal models presented above, by setting $t_1 = t_2$ in (32). Regarding the simulation, steps 3 and 4 in Algorithm 7 get simplified with $\boldsymbol{Z}_{T,l}(0) = \boldsymbol{1}$ and $\boldsymbol{W}_l = \boldsymbol{0}$, respectively. In a spatial context, our approach encompasses the one in Emery and Arroyo (2018) in several ways. Firstly, as already pointed out, we rely on Gaussian mixtures instead of spectral measures for computing the importance weights. Our approach is thus valid for all covariance function $\phi \in \mathcal{C}_\infty$. Secondly, instead of being (arbitrarily) set to exponential, the proposal density is set to f_1, which is the Gaussian mixture corresponding the covariance to be simulated with $\boldsymbol{\theta} = \boldsymbol{1}$, i.e., with all parameters set to 1. With these choices, the algorithm is more consistent and is likely to converge more rapidly to the Gaussian distribution. Notice that the correlation matrix can also vary in space, thanks to the use of its Cholesky decomposition.

A stationary anisotropic multivariate space-time Gneiting model with a different anisotropy matrix $\boldsymbol{\Sigma}_i$ for each variable $i = 1, \ldots, p$ can be obtained by setting $\boldsymbol{\theta}_{i,j,\boldsymbol{x}_1,\boldsymbol{x}_2} = \boldsymbol{\theta}_{i,j}$, $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^d \times \mathbb{R}$. This construction generalizes the one in Theorem 1 which instead has a unique anisotropy for all variables.

Finally, univariate nonstationary GRFs and simulation algorithm of those are easily obtained by setting $p = 1$. In the univariate case, there is only one (usual) time variogram γ, which can be any unbounded variogram.

5.5 Practical considerations

As a rule of thumb, several thousands cosine waves are usually recommended. In the illustration below, we used $L = 50,000$. Fewer waves can be used when the covariance function is smooth near the origin. Generally speaking, more numerous waves are required when the covariance function is not smooth (e.g. $\nu < 1$ for a Matérn covariance) and in presence of a strong non-stationarity, see also Lantuéjoul (2002, Chap. 15), Emery and Lantuéjoul (2006) and Schlather (2012) for guidelines on the number of cosine waves. From a theoretical point of view, the instrumental density can be any density whose support includes the support of all scale mixtures $f(\cdot;\boldsymbol{\theta})$. For simplicity, we have chosen$f_1 = f(\cdot;\boldsymbol{\theta})$ for $\boldsymbol{\theta} = \mathbf{1}$ in Algorithms 6 and 7, but in some cases this choice may not be optimal. Choosing a density such that for each $i = 1, \ldots, p$, the expectation $E_{f_1}\left[(f_{ii,\boldsymbol{x}}(\xi)/f_1(\xi))^{1/2}\right]$ is as close as possible to 1 for all $\boldsymbol{x} \in \mathcal{S}$, is likely to accelerate the convergence of the sum at the last step of the algorithm. Setting each parameter to some central tendency statistics over the domain of simulation, such as its mean or median, is a reasonable first order approximation.

6 Illustration

In this section, we present an illustration for the model defined in Theorem 3 and simulated based on Algorithm 7, because it is the most general construction, which covers all the other models and algorithms seen throughout this work.

It is illustrated with a bivariate, spatio-temporal, nonstationary RF $\boldsymbol{Z}(\cdot,\cdot) = (Z_1(\cdot,\cdot), Z_2(\cdot,\cdot))$, over the domain $[0,10]^2 \times [1,4]$, with a spatial resolution equal to 0.05 and temporal resolution equal to 1. The total number of grid points is thus $4 \times 201^2 = 161\,604$. We will consider a multivariate nonstationary Gneiting-Matérn model. The pseudo-variogram $\boldsymbol{\gamma}(u) = \gamma_0(u)+\boldsymbol{R}(0)-\boldsymbol{R}(u)$ is $\gamma_0(u) = |u|/2$ and $R_{ij}(u) = A_iA_j\exp(-u^2/2)$ with $A_1 = 0.2$ and $A_2 = 0.5$.

Z_1 is stationary and isotropic in space with $\nu_{11,\boldsymbol{x}} = 1$, $\kappa_{11,\boldsymbol{x}} = 1$ and $\boldsymbol{\Sigma}_{11,\boldsymbol{x}} = \boldsymbol{I}_2$. In contrast, Z_2 is nonstationary. We set $\kappa_{22,\boldsymbol{x}} = 1$ and $\nu_{22,\boldsymbol{x}} = 1 + 0.03(t-1)(s_1+s_2)$, where s_1 and s_2 are the spatial coordinates, so that the regularity of Z_2 increases from $\nu_{22} = 1$ in the bottom left corner at all time (and everywhere at $t = 1$) up to $\nu_{22} = 2.8$ in the upper right corner at $t = 4$. The anisotropy matrix of Z_2 varies with $\boldsymbol{x}$ in such a way that the anisotropy ratio decreases from 0.9 to 0.5 from the bottom left corner to the upper right corner. The anisotropy angle rotates with time according to $\theta(t) = 2\pi(t-1)/5$. We thus set

$$\boldsymbol{\Sigma}_2^{-1/2}(\boldsymbol{s},t) = \begin{pmatrix} 1 & 0 \\ 0 & 0.9-(s_1+s_2)/40 \end{pmatrix} \begin{pmatrix} \cos(2\pi(t-1)/5) & \sin(2\pi(t-1)/5) \\ -\sin(2\pi(t-1)/5) & \cos(2\pi(t-1)/5) \end{pmatrix}.$$

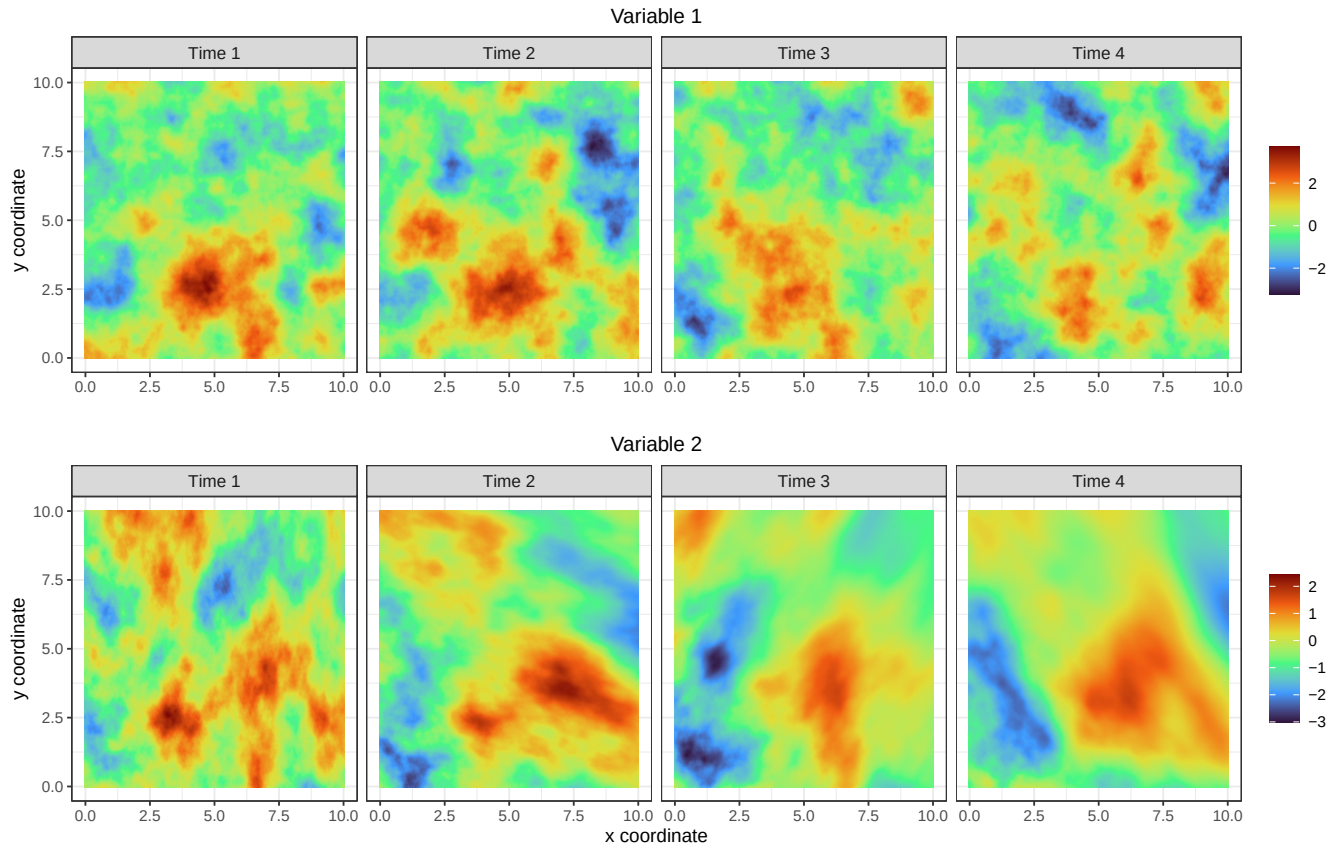

Figure 1: Realization of a bivariate random field with nonstationary Gneiting-Matérn covariance over four time steps, where the first variable is isotropic and stationary and the second is anisotropic and nonstationary.

Finally, in the correlation matrix $\boldsymbol{\sigma}$ we set $\sigma_{12,\boldsymbol{x}} = 9s_1/100$, so that the strength of the correlation between Z_1 and Z_2 increases from weak to strong as we move from left to right.

Figure 1 shows one realization of the space-time bivariate random field with nonstationary Gneiting-Matérn covariance simulated according to Algorithm 7 with $L = 50,000$. A visual assessment shows that, as prescribed, the first variable is stationary and isotropic while the second variable is nonstationary with the expected properties (e.g., the regularity increases from the bottom left to the upper right corner, the correlation increases from left to right, and the anisotropy rotates with time). This demonstrates the flexibility of the model in simulating complex spatio-temporal and inter-variable dependencies. To assess quantitatively if the simulated random field properly reproduces the theoretical covariance, 64 spatial locations are randomly drawn from the spatial domain, identical at all time steps. Then, 1000 random simulations of the multivariate spatio-temporal RF with the previous parametrization are run. Finally, for

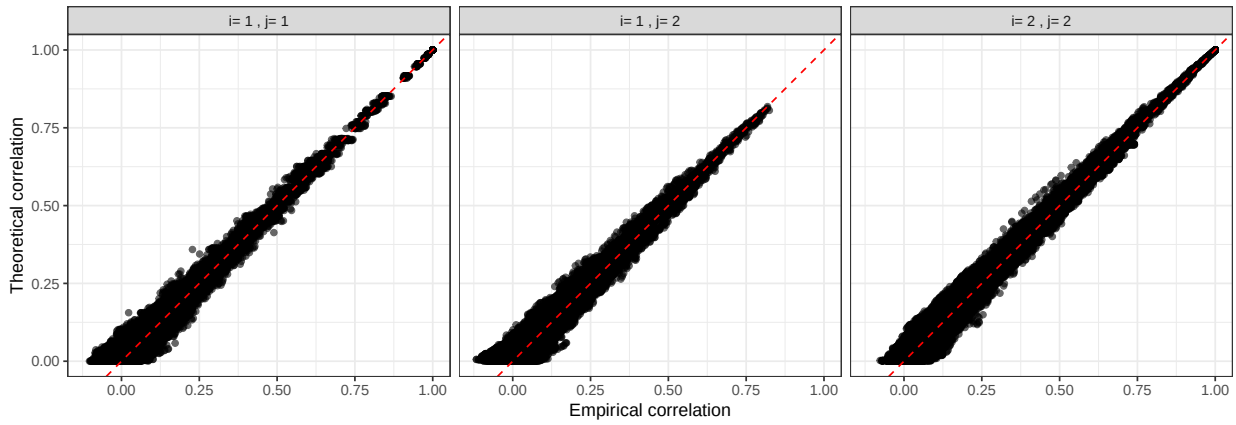

Figure 2: Empirical (based on 1000 simulations) and theoretical correlations between pairs of points at 64 random spatial locations and all time steps. Each panel corresponds to one pair of variables.

each variable-space-time pair, the empirical correlation is computed and compared with the theoretical one. The results are shown in Figure 2, where we show the correlations between all spatio-temporal pairs, for each pair of variables. The empirical correlations almost perfectly agree with the theoretical ones, which demonstrates that the simulation algorithm 7 with $L = 50,000$ converges to the theoretical nonstationary Gneiting-Matérn covariance in equation (35).

7 Discussion

Since its early foundation in Shinozuka (1971), the spectral approach has been one of the main techniques used for simulating Gaussian Random Fields, see for example Lantuéjoul (2002), Schlather et al. (2015) and references therein. In this study, instead of using the spectral density, we have adopted a Gaussian mixture representation of the covariance function grounded on the Schoenberg theorem (Schoenberg, 1938; Schlather, 2010). This change of perspective offers a unified view on earlier algorithms developed for the simulation of GRFs (Emery et al., 2016; Emery and Arroyo, 2018; Allard et al., 2020), which in our opinion did not sufficiently emphasized that they belonged to the Gaussian mixture representation. This new perspective also enables the design of simulation algorithms for covariance functions that explicitly or implicitly derive from a Gaussian mixture representation, which is the case for many covariance functions in various contexts: nonstationary spatial (Paciorek and Schervish, 2006), univariate space-time (Gneiting, 2002), multivariate (Genton and Kleiber, 2015) and multivariate space-

time (Allard et al., 2022). Finally, and importantly, the extension of Gaussian mixture-based simulation algorithms to nonstationary GRFs paved the way to a general theorem allowing for the construction of a new and wide class of nonstationary covariance functions, as detailed in Theorems 2 and 3 in Section 5. This new class of nonstationary covariance functions is expected to power GRF modeling, in particular in multivariate and space-time contexts.

The avenues for future research opened by the present work are diverse. From a methodological perspective, rather straightforward extensions touch on simulating GRFs on non-Euclidean domains, which includes the simulation of random fields on spheres cross time with covariance models similar to the Gneiting class, but involving Stieltjes functions instead of Bernstein functions, see for example Alegría et al. (2019), White and Porcu (2017), Lantuéjoul et al. (2019) or Alegría et al. (2020). Simulations on general manifolds (Pereira et al., 2022) or networks (Bolin et al., 2024) could also be explored.

Focusing on the space-time context, the substitution approach introduced in Section 3.2 offers numerous advantages. It is very general since it relies on the simulation of a (possibly multivariate) one dimensional process with Gaussian increments. In cases with less than 10^4 time steps (i.e., in most cases), this can easily be performed using a simple Cholesky decomposition. It is thus very general and does not necessitate to know the spectral density of the associated temporal variogram, see Allard et al. (2020) for an in-depth discussion about the differences between the spectral and the substitution approaches to simulate space-time GRFs. A possible extension of the present work would be to design even more general models of space-time GRFs, for instance enabling nonstationary dynamics in Lagrangian spatio-temporal covariance functions (Salvanña and Genton, 2020; Salvaña et al., 2023).

From a computational perspective, an appealing feature of all the simulation algorithms presented in this work is that the simulation of the stochastic representation (the L sets of random variables associated for instance with scales or phases along with the L temporal processes in a spatio-temporal context) is separated from the projection onto the set of points $(\boldsymbol{s}, t)$ on which the simulation is represented. When simulating on large space-time domains, the projection step is usually the most time consuming, with a computation time proportional to L and to the number of grid points. Since this step consists in a simple loop on all grid points in the simulation algorithms, it can be massively parallelized. In a nonstationary context, pre-computing and storing some of the values appearing in the importance weights can also accelerate the computations.

From an applied perspective, the use of GRFs in environmental modeling often requires the conditioning to observed values, i.e., enforcing the random fields to honor a given set of observations. The question of conditioning the simulations to a set of N conditioning points

can be addressed using the so called "conditioning by kriging technique" (Chilès and Delfiner, 2012, Chap. 7), but since kriging has a complexity in the order of $\mathcal{O}(N^3)$ this approach is limited to conditioning sets of moderate size, and some form of approximation is needed when tens of thousands conditioning data are available. As an alternative to the classical approach of conditioning locally, Marcotte and Allard (2018) explored a half-tapering approach that was already suggested in Stein (2013). The idea is to produce unconditional simulations with an efficient algorithm such as those presented in this study, and to use a tapered covariance $C \times C_{\text{Taper}}$ for the conditioning step. The tapering function C_{Taper} is a correlation function with compact support, such as the Wendland functions (Wendland, 1995), and is identically equal to zero outside a particular range, say r. This ensures that the tapered covariance is definite positive since it is the product of two definite positive functions, and that by construction the kriging matrix involved in the conditioning step has a high proportion of zero elements when the range r is small compared to pairwise distances between data locations. In a spatial and stationary context, Marcotte and Allard (2018) established theoretically and numerically that under a tail condition, the half-tapering approach is asymptotically equivalent to the full kriging method. Extensions to the multivariate, spatio-temporal and nonstationary contexts remain to be developed, and would help broadening the field of applications of the unconditional simulation techniques designed in this study.

Finally, to help disseminating the simulation algorithms presented in this work as well as the above propositions of extension, we believe that an easy-to-use software should be implemented and made accessible to practitioners.

Code availability

An R code for simulating non-stationary multivariate space-time random fields is available at https://github.com/sobakrim/Sim-NSMuST. This repository contains also the code for reproducing the illustration in Section 6.

Acknowledgments

The authors acknowledge the financial support of the Chair Geolearning, funded by Andra, BNP Paribas, CCR and the SCOR Foundation for science. Said Obakrim acknowledges the funding of the Swiss National Science Foundation through project number 200021_204130, entitled "Deep-time synthetic data cubes to enable long-term hydrological modelling".

A Proof of Proposition 1

On the one hand, by Bochner's theorem, $\phi(\boldsymbol{h}) = \int_{\mathbb{R}^d} \cos(\boldsymbol{\omega}^t \boldsymbol{h}) \mu(\boldsymbol{\omega})\, d\boldsymbol{\omega}$. On the other hand, the mixture representation gives, $\forall \boldsymbol{h} \in \mathbb{R}^d$,

$$\begin{aligned}
\phi(\boldsymbol{h}) &= \int_0^{+\infty} \exp\left(-\xi ||\boldsymbol{h}||^2\right) f(\xi)\, d\xi \\
&= \left(2\sqrt{\pi}\right)^{-d} \int_0^{+\infty} \int_{\mathbb{R}^d} \cos(\boldsymbol{\omega}^t \boldsymbol{h}) \xi^{-d/2} \exp\left(-||\boldsymbol{\omega}||^2/4\xi\right) f(\xi)\, d\boldsymbol{\omega}\, d\xi \\
&= \int_{\mathbb{R}^d} \cos(\boldsymbol{\omega}^t \boldsymbol{h}) \left(2\sqrt{\pi}\right)^{-d} \int_0^{+\infty} \xi^{-d/2} \exp\left(-||\boldsymbol{\omega}||^2/4\xi\right) f(\xi)\, d\xi\, d\boldsymbol{\omega}.
\end{aligned}$$

The first equality uses Bochner's representation of the gaussian covariance and the second equality follows from (3). Inversion of integrals is permitted because $\phi(\boldsymbol{h})$ is finite. Then,

$$\mu(\boldsymbol{\omega}) = \left(2\sqrt{\pi}\right)^{-d} \int_0^{+\infty} \xi^{-d/2} \exp\left(-||\boldsymbol{\omega}||^2/4\xi\right) f(\xi)\, d\xi$$

follows by the uniqueness of the spectral density of ϕ. □

Examples

1. For the Matérn covariance with shape and scale parameters (ν, κ), ξ is distributed as an inverse Gamma distribution with shape and scale parameters $(\nu, \kappa^2/4)$:

$$f_{\mathcal{M}}(\xi) = \left(\frac{\kappa^2}{4}\right)^{\nu} \frac{\xi^{-1-\nu}}{\Gamma(\nu)} \exp\left(-\kappa^2/4\xi\right).$$

   Hence

$$\begin{aligned}
\mu_{\mathcal{M}}(\boldsymbol{\omega}) &= \left(2\sqrt{\pi}\right)^{-d} \left(\frac{\kappa^2}{4}\right)^{\nu} \frac{1}{\Gamma(\nu)} \int_0^{+\infty} \xi^{-1-\nu-d/2} \exp\left(-(||\boldsymbol{\omega}||^2 + \kappa^2)/4\xi\right) d\xi \\
&= \left(2\sqrt{\pi}\right)^{-d} \left(\frac{\kappa^2}{4}\right)^{\nu} 4^{\nu+d/2} \frac{\Gamma(\nu + d/2)}{\Gamma(\nu)} \left[||\boldsymbol{\omega}||^2 + \kappa^2\right]^{-\nu-d/2}.
\end{aligned}$$

   The second identity comes from the fact that we recognize in the integral (up to the normalizing constant) the Inverse Gamma probability density function with shape and scale parameters $(\nu + 1 + d/2, (||\boldsymbol{h}||^2 + \kappa^2)/4)$. After rearrangement and simplifications, we get the expression in (10).

2. For the Cauchy covariance with shape and scale parameters $(\nu, 1/a)$, ξ is distributed as a Gamma distribution with same shape and scale parameters:

$$f_{\mathcal{C}}(\xi) = a^{-\nu} \Gamma(\nu)^{-1} \xi^{\nu-1} e^{-\xi/a}.$$

   We thus get

$$\mu_{\mathcal{C}}(\boldsymbol{\omega}) = \left(2\sqrt{\pi}\right)^{-d} a^{-\nu} \Gamma(\nu) \int_0^{+\infty} \xi^{-1-\nu-d/2} \exp\left(-||\boldsymbol{\omega}||^2/4\xi\right) \exp\left(-\xi/a\right) d\xi. \tag{36}$$

This last expression cannot be further simplified, in accordance to the fact that, to the best of our knowledge, there is no closed form expression for the spectral density of the Cauchy covariance function. Nonetheless, the relationship (36) seems original, and perhaps easier to use than Theorem 1 in Bevilacqua and Faouzi (2019).

B Proof of Theorem 2

With the usual product-to-sum trigonometric identity and after some straightforward manipulation we get

$$\mathbb{C}\mathrm{ov}\left(Z(\boldsymbol{s}_1), Z(\boldsymbol{s}_2)\right) = \frac{|\boldsymbol{\Sigma}_{\boldsymbol{s}_1}|^{1/4}|\boldsymbol{\Sigma}_{\boldsymbol{s}_2}|^{1/4}}{|\boldsymbol{\Sigma}_{\boldsymbol{s}_1,\boldsymbol{s}_2}|^{1/2}} \int_0^{+\infty} \frac{\left(f(\xi;\boldsymbol{\theta}_{\boldsymbol{s}_1})f(\xi;\boldsymbol{\theta}_{\boldsymbol{s}_2})\right)^{1/2}}{f_1(\xi)} C_G\Big(\sqrt{\xi}\,\boldsymbol{\Sigma}_{\boldsymbol{s}_1,\boldsymbol{s}_2}^{-1/2}(\boldsymbol{s}_1-\boldsymbol{s}_2)\Big) f_1(\xi)d\xi. \tag{37}$$

Because f belongs to the exponential family, it holds that

$$\begin{aligned}\left(f(\xi;\boldsymbol{\theta}_{\boldsymbol{s}_1})f(\xi;\boldsymbol{\theta}_{\boldsymbol{s}_2})\right)^{1/2} &= \left(h(\boldsymbol{\theta}_{\boldsymbol{s}_1})h(\boldsymbol{\theta}_{\boldsymbol{s}_2})\right)^{1/2} \exp\left(-\frac{1}{2}(\boldsymbol{\ell}(\boldsymbol{\theta}_{\boldsymbol{s}_1}) + \boldsymbol{\ell}(\boldsymbol{\theta}_{\boldsymbol{s}_2}))^t \boldsymbol{T}(\xi)\right) \\ &= \frac{\left(h(\boldsymbol{\theta}_{\boldsymbol{s}_1})h(\boldsymbol{\theta}_{\boldsymbol{s}_2})\right)^{1/2}}{h(\boldsymbol{\theta}_{\boldsymbol{s}_1,\boldsymbol{s}_2})} f(\xi;\boldsymbol{\theta}_{\boldsymbol{s}_1,\boldsymbol{s}_2}),\end{aligned}$$

where $f(\xi;\boldsymbol{\theta}_{\boldsymbol{s}_1,\boldsymbol{s}_2})$ is the (exponential family) scale mixture defined by

$$f(\xi;\boldsymbol{\theta}_{\boldsymbol{s}_1,\boldsymbol{s}_2}) = h(\boldsymbol{\theta}_{\boldsymbol{s}_1,\boldsymbol{s}_2}) \exp\left(-\boldsymbol{\ell}(\boldsymbol{\theta}_{\boldsymbol{s}_1,\boldsymbol{s}_2})^t \boldsymbol{T}(\xi)\right)$$

with $\boldsymbol{\ell}(\boldsymbol{\theta}_{\boldsymbol{s}_1,\boldsymbol{s}_2}) = (\boldsymbol{\ell}(\boldsymbol{\theta}_{\boldsymbol{s}_1}) + \boldsymbol{\ell}(\boldsymbol{\theta}_{\boldsymbol{s}_2})/2$. □

C Proof of Theorem 3

Consider a space-time p-variate random field $\boldsymbol{Z}(\cdot,\cdot) = [Z_i(\cdot,\cdot)]_{i=1}^p$ defined as

$$Z_i(\boldsymbol{s},t) = \sqrt{2\frac{f(R;\boldsymbol{\theta}_{i,\boldsymbol{x}})}{f_1(R)}} \sqrt{\frac{\mu^G_{\boldsymbol{\Sigma}_{ii,\boldsymbol{x}}}(\sqrt{2}\boldsymbol{V})}{\mu^G_{\boldsymbol{I}_d}(\sqrt{2}\boldsymbol{V})}} (\boldsymbol{L}_{\boldsymbol{x}}\boldsymbol{A})_i \cos\left(\boldsymbol{\Omega}^t\boldsymbol{s} + \frac{\|\boldsymbol{V}\|}{\sqrt{2}} W_i(t) + \Phi\right), \qquad i = 1,\ldots,p, \tag{38}$$

where $R \sim f_1$, $\boldsymbol{V} \sim \mathcal{N}(0,\boldsymbol{I}_d)$, $\boldsymbol{A} \sim \mathcal{N}(0,\boldsymbol{I}_p)$ and $\Phi \sim (0,2\pi)$, all independent to each other and $\boldsymbol{\Omega} = R\boldsymbol{V}$. The matrix $\boldsymbol{L}_{\boldsymbol{x}}$ is the Cholesky factorization of $\boldsymbol{\sigma}_{\boldsymbol{x}}$ with $\boldsymbol{L}_{\boldsymbol{x}}\boldsymbol{L}_{\boldsymbol{x}}^t = \boldsymbol{\sigma}_{\boldsymbol{x}}$. Recall that $\boldsymbol{W}(\cdot) = [W_i(\cdot)]_{i=1}^p$ are p RFs on $\mathbb{R}$ whose direct and cross-increments are Gaussian and characterized by the matrix-valued pseudo-variogram $\boldsymbol{\gamma}$, i.e, $W_i(t) - W_j(t+u) = \sqrt{2\gamma_{ij}(u)}Y$, with $Y \sim \mathcal{N}(0,1)$. The process $\boldsymbol{W}(\cdot)$ is independent to all random variables. Clearly, all random fields $Z_i(\cdot,\cdot)$ have zero mean.

Write $Y_i(\boldsymbol{x}) = \sqrt{2f(R;\boldsymbol{\theta}_{i,\boldsymbol{x}})/f_1(R)}\sqrt{\mu^G_{\boldsymbol{\Sigma}_{ii,\boldsymbol{x}}}(\sqrt{2}\boldsymbol{V})/\mu^G_{\boldsymbol{I}_d}(\sqrt{2}\boldsymbol{V})}\cos\left(\boldsymbol{\Omega}^t\boldsymbol{s} + \frac{\|\boldsymbol{V}\|}{\sqrt{2}}W_i(t) + \Phi\right)$, $i = 1,\ldots,p$. Because of the independence assumption, we have

$$\mathbb{E}\left[\boldsymbol{Z}(\boldsymbol{s}_1,t_1)\boldsymbol{Z}^t(\boldsymbol{s}_2,t_2)\right] = \mathbb{E}\left[\boldsymbol{Z}_T(t_1)\boldsymbol{Z}_T^t(t_2)\right]\mathbb{E}\left[\boldsymbol{L}_{\boldsymbol{x}_1}\boldsymbol{A}\boldsymbol{A}^t\boldsymbol{L}^t_{\boldsymbol{x}_2}\right]\mathbb{E}\left[\boldsymbol{Y}(\boldsymbol{s}_1,t_1)\boldsymbol{Y}^t(\boldsymbol{s}_2,t_2)\right].$$

The first expectation is equal to $\boldsymbol{C}_T(t_2-t_1)$. Since $\mathbb{E}[\boldsymbol{A}\boldsymbol{A}^t] = \boldsymbol{I}_p$, the second expectation is equal to $\boldsymbol{\sigma}_{\boldsymbol{x}_1,\boldsymbol{x}_2}$. Let us now consider the third expectation:

$$\begin{aligned}\mathbb{E}\left[\boldsymbol{Y}(\boldsymbol{s}_1,t_1)\boldsymbol{Y}^t(\boldsymbol{s}_2,t_2)\right]_{ij} = 2\mathbb{E}\Bigg[&\frac{\sqrt{f(R;\boldsymbol{\theta}_{i,\boldsymbol{x}_1})f(R;\boldsymbol{\theta}_{j,\boldsymbol{x}_2})}}{f_1(R)}\frac{\sqrt{\mu^G_{\boldsymbol{\Sigma}_{ii,\boldsymbol{x}_1}}(\sqrt{2}\boldsymbol{V})\mu^G_{\boldsymbol{\Sigma}_{jj,\boldsymbol{x}_2}}(\sqrt{2}\boldsymbol{V})}}{\mu^G_{\boldsymbol{I}_d}(\sqrt{2}\boldsymbol{V})}\\ &\times\cos\left(\sqrt{2R}\,\boldsymbol{V}^t\boldsymbol{s}_1 + \frac{\|\boldsymbol{V}\|}{\sqrt{2}}W_i(t_1) + \Phi\right)\\ &\times\cos\left(\sqrt{2R}\,\boldsymbol{V}^t\boldsymbol{s}_2 + \frac{\|\boldsymbol{V}\|}{\sqrt{2}}W_j(t_2) + \Phi\right)\Bigg].\end{aligned} \tag{39}$$

Using trigonometric identities and the fact that Φ is uniformly distributed on $(0,2\pi)$ and independent on R, Ω, and $\boldsymbol{W}(\cdot)$, this expression simplifies to

$$\begin{aligned}\mathbb{E}\left[\boldsymbol{Y}(\boldsymbol{s}_1,t_1)\boldsymbol{Y}^t(\boldsymbol{s}_2,t_2)\right]_{ij} = \mathbb{E}\Bigg[&\frac{\sqrt{f(R;\boldsymbol{\theta}_{i,\boldsymbol{x}_1})f(R;\boldsymbol{\theta}_{j,\boldsymbol{x}_2})}}{f_1(R)}\frac{\sqrt{\mu^G_{\boldsymbol{\Sigma}_{ii,\boldsymbol{x}_1}}(\sqrt{2}\boldsymbol{V})\mu^G_{\boldsymbol{\Sigma}_{jj,\boldsymbol{x}_2}}(\sqrt{2}\boldsymbol{V})}}{\mu^G_{\boldsymbol{I}_d}(\sqrt{2}\boldsymbol{V})}\\ &\times\cos\left(\sqrt{2R}\,\boldsymbol{V}^t(\boldsymbol{s}_1-\boldsymbol{s}_2) + \frac{\|\boldsymbol{V}\|}{\sqrt{2}}\left(W_i(t_1)-W_j(t_2)\right)\right)\Bigg].\end{aligned} \tag{40}$$

Using again a trigonometric identity, we have

$$\begin{aligned}&\cos\left(\sqrt{2R}\,\boldsymbol{V}^t(\boldsymbol{s}_1-\boldsymbol{s}_2) + \frac{\|\boldsymbol{V}\|}{\sqrt{2}}\left(W_i(t_1)-W_j(t_2)\right)\right)\\ =\;&\cos\left(\sqrt{2R}\,\boldsymbol{V}^t(\boldsymbol{s}_1-\boldsymbol{s}_2)\right)\cos\left(\|\boldsymbol{V}\|\sqrt{\gamma_{ij}(t_1-t_2)}Y\right)\\ &+\sin\left(\sqrt{2R}\,\boldsymbol{V}^t(\boldsymbol{s}_1-\boldsymbol{s}_2)\right)\sin\left(\|\boldsymbol{V}\|\sqrt{\gamma_{ij}(t_1-t_2)}Y\right).\end{aligned} \tag{41}$$

Using symmetry and considering that the sine function is odd, the expectation of the second term of (41) is equal to 0. Let us denote $g(x) = (2\pi)^{-1/2}\exp(-x^2/2)$ the $(0,1)$ Gaussian pdf. Using the spectral density of Gaussian covariances as in Example 6, one gets that

$$\frac{\sqrt{\mu^G_{\boldsymbol{\Sigma}_{ii,\boldsymbol{x}_1}}(\sqrt{2}\boldsymbol{V})\mu^G_{\boldsymbol{\Sigma}_{jj,\boldsymbol{x}_2}}(\sqrt{2}\boldsymbol{V})}}{\mu^G_{\boldsymbol{I}_d}(\sqrt{2}\boldsymbol{V})} = |\boldsymbol{\Sigma}_{ii,\boldsymbol{x}_1}|^{1/4}|\boldsymbol{\Sigma}_{jj,\boldsymbol{x}_2}|^{1/4}e^{-\boldsymbol{V}^t\boldsymbol{\Sigma}_{i,j,\boldsymbol{x}_1,\boldsymbol{x}_2}\boldsymbol{V}/2}/e^{-\|\boldsymbol{V}\|^2/2},$$

with $\boldsymbol{\Sigma}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2} = (\boldsymbol{\Sigma}_{ii,\boldsymbol{x}_1} + \boldsymbol{\Sigma}_{jj,\boldsymbol{x}_2})/2$. Taking the expectation over $\boldsymbol{V}$, R, and Y, we obtain

$$\begin{aligned}\mathbb{E}\left[Y_i(\boldsymbol{s}_1,t_1)Y_j(\boldsymbol{s}_2,t_2)\right] = &\frac{|\boldsymbol{\Sigma}_{ii,\boldsymbol{x}_1}|^{1/4}|\boldsymbol{\Sigma}_{jj,\boldsymbol{x}_2}|^{1/4}}{(2\pi)^{d/2}}\int_0^\infty\int_{\mathbb{R}^d}\sqrt{f(r;\boldsymbol{\theta}_{i,\boldsymbol{x}_1})f(r;\boldsymbol{\theta}_{j,\boldsymbol{x}_2})}e^{-\boldsymbol{v}^t\boldsymbol{\Sigma}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2}\boldsymbol{v}/2}\\ &\times\cos\left(\sqrt{2r}\,\boldsymbol{v}^t(\boldsymbol{s}_1-\boldsymbol{s}_2)\right)\int_{\mathbb{R}}\cos\left(\|\boldsymbol{v}\|\sqrt{\gamma_{ij}(t_1-t_2)}y\right)g(y)dyd\boldsymbol{v}\,dr.\end{aligned} \tag{42}$$

Since $\int_{\mathbb{R}} \cos\left(\|\boldsymbol{v}\|\sqrt{\gamma_{ij}(t_1-t_2)}y\right)\, g(y)dy$ is the Fourier transform of the standard Gaussian probability density g computed at $\|\boldsymbol{v}\|\sqrt{\gamma_{ij}(t_1-t_2)}$, we have

$$\begin{aligned}\mathbb{E}\left[Y_i(\boldsymbol{s}_1,t_1)Y_j(\boldsymbol{s}_2,t_2)\right] = \frac{|\boldsymbol{\Sigma}_{ii,\boldsymbol{x}_1}|^{1/4}|\boldsymbol{\Sigma}_{jj,\boldsymbol{x}_2}|^{1/4}}{(2\pi)^{d/2}} \int_0^\infty \int_{\mathbb{R}^d} \sqrt{f(r;\boldsymbol{\theta}_{i,\boldsymbol{x}_1})f(r;\boldsymbol{\theta}_{j,\boldsymbol{x}_2})}e^{-\boldsymbol{v}^t\boldsymbol{\Sigma}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2}\boldsymbol{v}/2} \\ \times \cos\left(\sqrt{2r}\,\boldsymbol{v}^t(\boldsymbol{s}_1-\boldsymbol{s}_2)\right)\exp\left(-\frac{\|\boldsymbol{v}\|^2}{2}\gamma_{ij}(t_1-t_2)\right)\, d\boldsymbol{v}\, dr.\end{aligned} \tag{43}$$

In view of grouping all arguments of the exponential, we define

$$\boldsymbol{\Lambda}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2} = \boldsymbol{\Sigma}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2} + \gamma_{ij}(t_1-t_2)\boldsymbol{I}_d.$$

We thus get

$$\begin{aligned}\mathbb{E}\left[Y_i(\boldsymbol{s}_1,t_1)Y_j(\boldsymbol{s}_2,t_2)\right] = \frac{|\boldsymbol{\Sigma}_{ii,\boldsymbol{x}_1}|^{1/4}|\boldsymbol{\Sigma}_{jj,\boldsymbol{x}_2}|^{1/4}}{(2\pi)^{d/2}} \int_0^\infty \int_{\mathbb{R}^d} \sqrt{f(r;\boldsymbol{\theta}_{i,\boldsymbol{x}_1})f(r;\boldsymbol{\theta}_{j,\boldsymbol{x}_2})} \\ \times \cos\left(\sqrt{2r}\,\boldsymbol{v}^t(\boldsymbol{s}_1-\boldsymbol{s}_2)\right)e^{-\boldsymbol{v}^t\boldsymbol{\Lambda}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2}\boldsymbol{v}/2}\, d\boldsymbol{v}\, dr.\end{aligned} \tag{44}$$

Up to a multiplicative factor, the integral $\int_{\mathbb{R}^d} \cos\left(\sqrt{2r}\,\boldsymbol{v}^t(\boldsymbol{s}_1-\boldsymbol{s}_2)\right)e^{-\boldsymbol{v}^t\boldsymbol{\Lambda}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2}\boldsymbol{v}/2}\, d\boldsymbol{v}$ is the Fourier transform of a multivariate standard Gaussian probability density. Therefore,

$$\begin{aligned}\mathbb{E}\left[Y_i(\boldsymbol{s}_1,t_1)Y_j(\boldsymbol{s}_2,t_2)\right] &= |\boldsymbol{\Sigma}_{ii,\boldsymbol{x}_1}|^{1/4}|\boldsymbol{\Sigma}_{jj,\boldsymbol{x}_2}|^{1/4} \int_0^\infty \sqrt{f(r;\boldsymbol{\theta}_{i,\boldsymbol{x}_1})f(r;\boldsymbol{\theta}_{j,\boldsymbol{x}_2})} \\ &\quad \times \frac{1}{|\boldsymbol{\Lambda}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2}|^{1/2}}\exp\left(-r(\boldsymbol{s}_1-\boldsymbol{s}_2)^t\boldsymbol{\Lambda}^{-1}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2}(\boldsymbol{s}_1-\boldsymbol{s}_2)\right)\, dr \\ &= |\boldsymbol{\Sigma}_{ii,\boldsymbol{x}_1}|^{1/4}|\boldsymbol{\Sigma}_{jj,\boldsymbol{x}_2}|^{1/4} \\ &\quad \times \frac{1}{|\boldsymbol{\Lambda}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2}|^{1/2}}\phi\left(\boldsymbol{\Lambda}^{-1/2}_{ij,\boldsymbol{x}_1,\boldsymbol{x}_2}(\boldsymbol{s}_1-\boldsymbol{s}_2);\boldsymbol{\theta}_{i,j,\boldsymbol{x}_1,\boldsymbol{x}_2}\right),\end{aligned} \tag{45}$$

where $\boldsymbol{\theta}_{i,j,\boldsymbol{x}_1,\boldsymbol{x}_2}$ is defined as in Eq. (29). The final result is then easily obtained.

□

References

Adler, R. J. (2010). *The geometry of random fields*. SIAM.

Adler, R. J. and Taylor, J. E. (2009). *Random fields and geometry*. Springer Science & Business Media.

Alegría, A., Emery, X., and Lantuéjoul, C. (2020). The turning arcs: a computationally efficient algorithm to simulate isotropic vector-valued gaussian random fields on the d-sphere. *Statistics and Computing*, 30:1403–1418.

Alegría, A., Porcu, E., Furrer, R., and Mateu, J. (2019). Covariance functions for multivariate gaussian fields evolving temporally over planet earth. *Stochastic Environmental Research and Risk Assessment*, 33:1593–1608.

Allard, D., Clarotto, L., and Emery, X. (2022). Fully nonseparable gneiting covariance functions for multivariate space–time data. *Spatial Statistics*, 52:100706.

Allard, D., Emery, X., Lacaux, C., and Lantuéjoul, C. (2020). Simulating space-time random fields with nonseparable gneiting-type covariance functions. *Statistics and Computing*, 30(5):1479–1495.

Allard, D., Senoussi, R., and Porcu, E. (2016). Anisotropy models for spatial data. *Mathematical Geosciences*, 48:305–328.

Apanasovich, T. V. and Genton, M. G. (2010). Cross-covariance functions for multivariate random fields based on latent dimensions. *Biometrika*, 97(1):15–30.

Arroyo, D. and Emery, X. (2017). Spectral simulation of vector random fields with stationary gaussian increments in d-dimensional euclidean spaces. *Stochastic Environmental Research and Risk Assessment*, 31(7):1583–1592.

Bagchi, P. and Dette, H. (2020). A test for separability in covariance operators of random surfaces. *The Annals of Statistics*, 48(4):2303–2322.

Benoit, L., Allard, D., and Mariethoz, G. (2018). Stochastic rainfall modeling at sub-kilometer scale. *Water Resources Research*, 54(6):4108–4130.

Bevilacqua, M. and Faouzi, T. (2019). Estimation and prediction of Gaussian processes using generalized Cauchy covariance model under fixed domain asymptotics. *Electronic Journal of Statistics*, 13(2):3025 – 3048.

Bochner, S. (1933). Monotone funktionen, stieltjessche integrale und harmonische analyse. *Mathematische Annalen*, 108(1):378–410.

Bochner, S. (2005). *Harmonic analysis and the theory of probability*. Courier Corporation.

Bolin, D., Simas, A. B., and Wallin, J. (2024). Gaussian whittle–matérn fields on metric graphs. *Bernoulli*, 30(2):1611–1639.

Bolin, D. and Wallin, J. (2020). Multivariate type G Matérn stochastic partial differential equation random fields. *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 82(1):215 – 239.

Bourotte, M., Allard, D., and Porcu, E. (2016). A flexible class of non-separable cross-covariance functions for multivariate space–time data. *Spatial Statistics*, 18:125–146.

Box, G. E. and Muller, M. E. (1958). A note on the generation of random normal deviates. *The Annals of Mathematical Statistics*, 29(2):610–611.

Cappello, C., De Iaco, S., and Posa, D. (2018). Testing the type of non-separability and some classes of space-time covariance function models. *Stochastic Environmental Research and Risk Assessment*, 32:17–35.

Carrizo-Vergara, R., Allard, D., and Desassis, N. (2022). A general framework for spde-based stationary random fields. *Bernoulli*, 28(1):1–32.

Chan, G. and Wood, A. T. (1999). Simulation of stationary gaussian vector fields. *Statistics and computing*, 9(4):265–268.

Chen, W., Genton, M. G., and Sun, Y. (2021). Space-time covariance structures and models. *Annual Review of Statistics and Its Application*, 8:191–215.

Chilès, J.-P. and Delfiner, P. (2012). *Geostatistics: Modeling Spatial Uncertainty, Second Edition*. John Wiley & Sons.

Christakos, G. (2013). *Random field models in earth sciences*. Elsevier.

Clarotto, L., Allard, D., Romary, T., and Desassis, N. (2024). The spde approach for spatio-temporal datasets with advection and diffusion. *Spatial Statistics*, 62:100847.

De Iaco, S., Myers, D., Palma, M., and Posa, D. (2013). Using simultaneous diagonalization to identify a space–time linear coregionalization model. *Mathematical Geosciences*, 45:69–86.

De Iaco, S., Myers, D. E., and Posa, D. (2001). Space–time analysis using a general product–sum model. *Statistics & Probability Letters*, 52(1):21–28.

De Iaco, S. and Posa, D. (2013). Positive and negative non-separability for space–time covariance models. *Journal of Statistical Planning and Inference*, 143(2):378–391.

Dörr, C. and Schlather, M. (2023a). Characterization theorems for pseudo cross-variograms. *Journal of Applied Probability*, 60(4):1219–1231.

Dörr, C. and Schlather, M. (2023b). Covariance models for multivariate random fields resulting from pseudo cross-variograms. *Journal of Multivariate Analysis*, 197:105199.

Emery, X. and Arroyo, D. (2018). On a continuous spectral algorithm for simulating non-stationary gaussian random fields. *Stochastic Environmental Research and Risk Assessment*, 32:905–919.

Emery, X., Arroyo, D., and Porcu, E. (2016). An improved spectral turning-bands algorithm for simulating stationary vector Gaussian random fields. *Stochastic Environmental Research and Risk Assessment*, 30(7):1863–1873.

Emery, X. and Lantuéjoul, C. (2006). Tbsim: A computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. *Computers & Geosciences*, 32(10):1615–1628.

Emery, X. and Porcu, E. (2023). Extending the gneiting class for modeling spatially isotropic and temporally symmetric vector random fields. *Journal of Mathematical Analysis and Applications*, 525(2):127194.

Feller, W. (1966). *An Introduction to Probability Theory and its Applications, Vol. II.* John Wiley & Sons.

Genton, M. G. and Kleiber, W. (2015). Cross-covariance functions for multivariate geostatistics. *Statistical Science*, 30(2):147–163.

Gneiting, T. (2002). Nonseparable, stationary covariance functions for space–time data. *Journal of the American Statistical Association*, 97(458):590–600.

Gneiting, T. and Guttorp, P. (2010). Continuous parameter spatio-temporal processes. *Handbook of Spatial Statistics*, 97:427–436.

Gneiting, T., Kleiber, W., and Schlather, M. (2010). Matérn cross-covariance functions for multivariate random fields. *Journal of the American Statistical Association*, 105(491):1167–1177.

Gneiting, T. and Schlather, M. (2004). Stochastic models that separate fractal dimension and the hurst effect. *SIAM review*, 46(2):269–282.

Huang, H. and Sun, Y. (2019). Visualization and assessment of spatio-temporal covariance properties. *Spatial Statistics*, 34:100272.

Jurek, M. and Katzfuss, M. (2022). Hierarchical sparse Cholesky decomposition with applications to high-dimensional spatio-temporal filtering. *Statistics and Computing*, 32(1):15.

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

Lantuéjoul, C., Freulon, X., and Renard, D. (2019). Spectral simulation of isotropic gaussian random fields on a sphere. *Mathematical Geosciences*, 51(8):999–1020.

Lantuéjoul, C. (2002). *Geostatistical Simulation: Models and Algorithms*. Springer.

Li, B., Genton, M. G., and Sherman, M. (2007). A nonparametric assessment of properties of space–time covariance functions. *Journal of the American Statistical Association*, 102(478):736–744.

Linde, N., Renard, P., Mukerji, T., and Caers, J. (2015). Geological realism in hydrogeological and geophysical inverse modeling: A review. *Advances in Water Resources*, 86:86–101.

Lindgren, F., Bolin, D., and Rue, H. (2022). The SPDE approach for Gaussian and non-Gaussian fields: 10 years and still running. *Spatial Statistics*, 50:100599.

Lindgren, F., Rue, H., and Lindström, J. (2011). An explicit link between gaussian fields and gaussian markov random fields: the stochastic partial differential equation approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 73(4):423–498.

Ma, C. (2003). Families of spatio-temporal stationary covariance models. *Journal of Statistical Planning and Inference*, 116(2):489–501.

Marcotte, D. and Allard, D. (2018). Half-tapering strategy for conditional simulation with large datasets. *Stochastic environmental research and risk assessment*, 32:279–294.

Matheron, G. (1973). The intrinsic random functions and their applications. *Advances in applied probability*, 5(3):439–468.

Obakrim, S., Benoit, L., and Allard, D. (2025). A multivariate and space-time stochastic weather generator using a latent gaussian framework. *Stochastic Environmental Research and Risk Assessment*, hal.science/hal-04715860/.

Paciorek, C. J. and Schervish, M. J. (2006). Spatial modelling using a new class of nonstationary covariance functions. *Environmetrics*, 17.

Peleg, N., Fatichi, S., Paschalis, A., Molnar, P., and Burlando, P. (2017). An advanced stochastic weather generator for simulating 2-D high-resolution climate variables. *Journal of Advances in Modeling Earth Systems*, 9(3):1595–1627.

Pereira, M., Desassis, N., and Allard, D. (2022). Geostatistics for large datasets on riemannian manifolds: A matrix-free approach. *Journal of Data Science*, 20(4):512–532.

Porcu, E., Furrer, R., and Nychka, D. (2021). 30 years of space–time covariance functions. *Wiley Interdisciplinary Reviews: Computational Statistics*, 13(2):e1512.

Porcu, E., Gregori, P., and Mateu, J. (2006). Nonseparable stationary anisotropic space–time covariance functions. *Stochastic Environmental Research and Risk Assessment*, 21(2):113–122.

Porcu, E., Mateu, J., and Saura, F. (2008). New classes of covariance and spectral density functions for spatio-temporal modelling. *Stochastic Environmental Research and Risk Assessment*, 22(Suppl 1):65–79.

Porcu, E. and Schilling, R. L. (2011). From schoenberg to pick–nevanlinna: Toward a complete picture of the variogram class. *Bernoulli*, 17(1):441–455.

Porcu, E. and Zastavnyi, V. (2011). Characterization theorems for some classes of covariance functions associated to vector valued random fields. *Journal of Multivariate Analysis*, 102(9):1293–1301.

Rahimi, A. and Recht, B. (2007). Random features for large-scale kernel machines. *Advances in neural information processing systems*, 20.

Rodrigues, A. and Diggle, P. (2010). A class of convolution based models for spatio-temporal processes with non-separable covariance structure. *Scandinavian Journal of Statistics*, 37(4):553–567.

Romary, T. (2013). Incomplete cholesky decomposition for the kriging of large datasets. *Spatial Statistics*, 5:85–99.

Salvaña, M. L. O., Lenzi, A., and Genton, M. G. (2023). Spatio-temporal cross-covariance functions under the lagrangian framework with multiple advections. *Journal of the American Statistical Association*, 118(544):2746–2761.

Salvanña, M. L. O. and Genton, M. G. (2020). Nonstationary cross-covariance functions for multivariate spatio-temporal random fields. *Spatial Statistics*, 37:100411.

Schlather, M. (2010). Some covariance models based on normal scale mixtures. *Bernoulli*, pages 780–797.

Schlather, M. (2012). Construction of covariance functions and unconditional simulation of random fields. In *Advances and challenges in space-time modelling of natural events*, pages 25–54. Springer.

Schlather, M., Malinowski, A., Menck, P. J., Oesting, M., Strokorb, K., et al. (2015). Analysis, simulation and prediction of multivariate random fields with package RandomFields. *Journal of Statistical Software*, 63(8):1–25.

Schoenberg, I. J. (1938). Metric spaces and completely monotone functions. *Annals of Mathematics*, 39(4):811–841.

Shinozuka, M. (1971). Simulation of multivariate and multidimensional random processes. *The Journal of the Acoustical Society of America*, 49(1B):357–368.

Stein, A., van der Meer, F. D., and Gorte, B. (1999). *Spatial statistics for remote sensing*, volume 1. Springer Science & Business Media.

Stein, M. L. (2005). Nonstationary spatial covariance functions. Technical report, University of Chicago.

Stein, M. L. (2013). Statistical properties of covariance tapers. *Journal of Computational and Graphical Statistics*, 22(4):866–885.

Tronarp, F., Karvonen, T., and Särkkä, S. (2018). Mixture representation of the matérn class with applications in state space approximations and bayesian quadrature. In *2018 IEEE 28th International Workshop on Machine Learning for Signal Processing (MLSP)*, pages 1–6. IEEE.

Ver Hoef, J. M., Cressie, N., and Barry, R. P. (2004). Flexible spatial models for kriging and cokriging using moving averages and the fast fourier transform (fft). *Journal of Computational and Graphical Statistics*, 13(2):265–282.

Wackernagel, H. (2003). *Multivariate Geostatistics: an Introduction with Applications*. Springer Science & Business Media.

Wendland, H. (1995). Piecewise polynomial, positive definite and compactly supported radial functions of minimal degree. *Advances in computational Mathematics*, 4:389–396.

White, P. and Porcu, E. (2017). Towards a complete picture of covariance functions on spheres cross time. *Biometrika*, 103(1):1–23.

Whittle, P. (1954). On stationary processes in the plane. *Biometrika*, 41(3-4):434–449.

Williams, C. K. and Rasmussen, C. E. (2006). *Gaussian processes for machine learning*, volume 2. MIT press Cambridge, MA.

Wood, A. T. and Chan, G. (1994). Simulation of stationary Gaussian processes in $[0, 1]^d$. *Journal of Computational and Graphical Statistics*, 3(4):409–432.

Yaglom, A. M. (1987). *Correlation Theory of Stationary and Related Random Functions, Volume I: Basic Results*, volume 1. Springer.

Zakeri, F. and Mariethoz, G. (2021). A review of geostatistical simulation models applied to satellite remote sensing: Methods and applications. *Remote Sensing of Environment*, 259:112381.

Zastavnyi, V. P. and Porcu, E. (2011). Characterization theorems for the gneiting class of space-time covariances. *Bernoulli*, 17(1):456–465.