Stoch Environ Res Risk Assess (2016) 30:1863–1873
DOI 10.1007/s00477-015-1151-0

ORIGINAL PAPER

An improved spectral turning-bands algorithm for simulating stationary vector Gaussian random fields

Xavier Emery[1,2] · **Daisy Arroyo**[1,2] · **Emilio Porcu**[3]

Published online: 9 September 2015

Abstract We propose a spectral turning-bands approach for the simulation of second-order stationary vector Gaussian random fields. The approach improves existing spectral methods through coupling with importance sampling techniques. A notable insight is that one can simulate any vector random field whose direct and cross-covariance functions are continuous and absolutely integrable, provided that one knows the analytical expression of their spectral densities, without the need for these spectral densities to have a bounded support. The simulation algorithm is computationally faster than circulant-embedding techniques, lends itself to parallel computing and has a low memory storage requirement. Numerical examples with varied spatial correlation structures are presented to demonstrate the accuracy and versatility of the proposal.

Keywords Matrix-valued covariance functions · Spectral density · Importance sampling · Matérn covariance · Compactly supported covariance

✉ Xavier Emery
xemery@ing.uchile.cl

Daisy Arroyo
darroyo@ing.uchile.cl

Emilio Porcu
emilio.porcu@usm.cl

[1] Department of Mining Engineering, University of Chile, Santiago, Chile

[2] Advanced Mining Technology Center, University of Chile, Santiago, Chile

[3] Department of Mathematics, Technical University Federico Santa Maria, Valparaiso, Chile

1 Introduction

The simulation of multivariate stationary Gaussian random fields with cross-correlated components is an active research topic in spatial statistics and has received an increasing attention in many disciplines of sciences and engineering where it is of interest to assess spatial uncertainty or to map heterogeneities at different spatial scales, such as geosciences, physics, cosmology, medicine, agriculture, soil, atmospheric and environmental sciences, to name a few.

Available algorithms for simulating Gaussian random fields have been extensively described in the literature, see (Chilès and Delfiner 2012; Christakos 1992; Lantuéjoul 2002; Schlather 2012) and references therein. Broadly speaking, they can be classified into two families: the former consists of algorithms that produce realizations of random fields with finite-dimensional distributions being exactly multivariate Gaussian, whereas the latter consists of algorithms that produce realizations of random fields having finite-dimensional distributions that are approximately Gaussian. Among exact algorithms, autoregressive, moving-average, circulant-embedding and discrete spectral simulation (Box and Jenkins 1976; Chan and Wood 1999; Chilès and Delfiner 1997, 2012; Dietrich and Newsam 1993; Mignolet and Spanos 1992; Pardo-Iguzquiza and Chica-Olmo 1993; Spanos and Mignolet 1992; Wood and Chan 1994) have been especially popular. However, these algorithms are restricted to the simulation of random fields at evenly spaced locations in low-dimensional spaces and cannot reproduce all the parametric families of covariance models available in the literature. The Cholesky decomposition of the covariance matrix (Alabert 1987; Davis 1987) and the sequential algorithms (Ripley 1987) produce perfect simulation of Gaussian random fields at non-evenly spaced locations, but computational costs become prohibitive when the number of locations exceeds a few thousands. These

costs can be decreased by using Markov Chain Monte Carlo methods such as the Gibbs sampler (Arroyo et al. 2012; Emery et al. 2014), but the number of target locations remains limited by computer storage capacity. On the other hand, the second family of simulation algorithms rely on the central limit theorem to produce approximately Gaussian random fields. The most notable are the Poisson dilution, tessellation, continuous spectral and turning bands algorithms (Chilès and Delfiner 2012; Christakos 1992; Emery 2008; Emery and Lantuéjoul 2006; Lantuéjoul 2002; Matheron 1973; Shinozuka 1971; Shinozuka and Jan 1972). All these algorithms allow simulating random fields at non-evenly spaced locations, but they have been notoriously limited in their ability to simulate from certain parametric models of covariance only.

To remove this limitation, our research presents a spectral turning-bands algorithm that improves the proposal in (Shinozuka 1971) and (Shinozuka and Jan 1972) by simulating from any multivariate covariance model having an available closed form for the associated matrix-valued spectral density. In particular, our main finding is obtained by coupling the idea in (Shinozuka 1971) with importance sampling approaches. The outline of the paper is the following: after a brief review on matrix-valued covariance functions and variograms in Sect. 2, we expose our proposal in Sect. 3. Section 4 is devoted to applications to synthetic case studies. Conclusions follow in Sect. 5. Some technical facts are reported in Appendix for a neater exposition.

2 Matrix-valued covariance functions and variograms

Throughout the paper, let us denote by $\mathbf{Y} = \{\mathbf{Y}(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^d\}$ a second-order stationary vector Gaussian random field with P cross-correlated components. Recall that a vector random field is Gaussian when any linear combination of its variables (associated with possibly different vector components and with different locations of $\mathbb{R}^d$) is a Gaussian random variable, and that it is second-order stationary when its first two moments (mean and covariance function) exist and are invariant under a translation in space. Assuming, with no loss of generality, that $\mathbf{Y}$ has zero mean, its finite-dimensional distributions are completely determined by the covariance function $\mathbf{C} : \mathbb{R}^d \to S_P^+$, with S_P^+ denoting the set of real-valued symmetric positive semi-definite matrices of size $P \times P$, such that

$$\mathbf{C}(\mathbf{h}) = \mathbb{E}\{\mathbf{Y}(\mathbf{x}+\mathbf{h}) \cdot \mathbf{Y}(\mathbf{x})^T\}, \qquad \mathbf{x}+\mathbf{h}, \mathbf{x} \in \mathbb{R}^d,$$

with $\mathbb{E}$ denoting stochastic expectation. The diagonal elements of $\mathbf{C}(\mathbf{h})$ are even functions called the auto or direct covariances of the components of $\mathbf{Y}$, while the off-diagonal elements are the cross-covariances between the components of $\mathbf{Y}$.

For the remainder of the paper, we work with covariances being continuous and absolutely integrable. For the scalar-valued case ($P = 1$) such covariances are described through Bochner's theorem (Bochner 1933) as being the Fourier transforms of uniquely determined bounded measures. Such a result has been extended to the matrix-valued case ($P > 1$) by (Cramér 1940):

$$\mathbf{C}(\mathbf{h}) = \int_{\mathbb{R}^d} \exp\{i\langle \mathbf{u}, \mathbf{h}\rangle\}\mathbf{f}(\mathbf{u})\mathrm{d}\mathbf{u}, \qquad \mathbf{h} \in \mathbb{R}^d, \tag{1}$$

with $i^2 = -1$ and $\langle \cdot, \cdot \rangle$ the usual scalar product in $\mathbb{R}^d$. The mapping $\mathbf{f} : \mathbb{R}^d \to H_P^+$, with H_P^+ denoting the set of complex-valued Hermitian positive semi-definite matrices of size $P \times P$, is called the matrix of spectral densities and the elements of $\mathbf{f}$ must be integrable on $\mathbb{R}^d$. Since $\mathbf{C}(\mathbf{h})$ is real-valued, one can rewrite Eq. (1) as follows:

$$\mathbf{C}(\mathbf{h}) = \int_{\mathbb{R}^d} [\cos(\langle \mathbf{u}, \mathbf{h}\rangle)\mathrm{Re}(\mathbf{f}(\mathbf{u})) - \sin(\langle \mathbf{u}, \mathbf{h}\rangle)\mathrm{Im}(\mathbf{f}(\mathbf{u}))]\,\mathrm{d}\mathbf{u}, \quad \mathbf{h} \in \mathbb{R}^d, \tag{2}$$

where Re and Im stand for the real and imaginary part, respectively.

Matrix-valued covariance functions have become ubiquitous and there has been a growing interest in proposing models for their construction. We refer the reader to the broad review in (Genton and Kleiber 2015) with the references therein, as well to the discussion by (Bevilacqua et al. 2015) for suggestions on how to choose a suitable multivariate model. Notable examples are the intrinsic correlation model (Journel and Huijbregts 1978), Markov-type models (Almeida and Journel 1994), linear coregionalization model (Wackernagel 2003), generalized linear coregionalization model (Marcotte 2012), bilinear coregionalization model (Grzebyk and Wackernagel 1994), kernel and covariance convolution models (Gaspari and Cohn 1999; Ver Hoef and Barry 1998), multivariate Matérn models (Gneiting et al. 2010), matrix-valued covariances with compact supports (Daley et al. 2015), and multivariate spectral densities based on Archimedean compositions (Porcu et al. 2012).

Another tool that is commonly used for modeling the spatial correlation structure of a vector Gaussian random field is the (semi-)variogram, defined as

$$\boldsymbol{\Gamma}(\mathbf{h}) = \frac{1}{2}\mathbb{E}\{[\mathbf{Y}(\mathbf{x}+\mathbf{h}) - \mathbf{Y}(\mathbf{x})] \cdot [\mathbf{Y}(\mathbf{x}+\mathbf{h}) - \mathbf{Y}(\mathbf{x})]^T\}, \quad \mathbf{x}+\mathbf{h}, \mathbf{x} \in \mathbb{R}^d.$$

The diagonal elements of $\boldsymbol{\Gamma}(\mathbf{h})$ are the direct variograms of the components of $\mathbf{Y}$, while the off-diagonal elements are the cross-variograms between components of $\mathbf{Y}$. In the case

of second-order stationary random fields, there is a one-to-one relationship between direct variograms and direct covariances, but the cross-variograms only contain information on the even part of the cross-covariance functions (Wackernagel 2003):

$$\boldsymbol{\Gamma}(\mathbf{h}) = \mathbf{C}(\mathbf{0}) - \frac{1}{2}[\mathbf{C}(\mathbf{h}) + \mathbf{C}(-\mathbf{h})]. \tag{3}$$

3 Our proposal

Let us define a simulated vector random field $\mathbf{Y}_S$ as follows:

$$\begin{aligned}\forall \mathbf{x} \in \mathbb{R}^d, \mathbf{Y}_S(\mathbf{x}) = &\sum_{p=1}^{P} \boldsymbol{\alpha}_p(\mathbf{U}_p)\cos(\langle \mathbf{x}, \mathbf{U}_p\rangle + \phi_p) \\ &+ \sum_{p=1}^{P} \boldsymbol{\beta}_p(\mathbf{U}_p)\sin(\langle \mathbf{x}, \mathbf{U}_p\rangle + \phi_p),\end{aligned} \tag{4}$$

where $\{\mathbf{U}_p : p = 1, \ldots, P\}$ are mutually independent random vectors with probability density $g : \mathbb{R}^d \to \mathbb{R}_+$, $\{\phi_p : p = 1, \ldots, P\}$ are mutually independent random variables (phases) being uniformly distributed over the interval $[0, 2\pi)$, independent of the $\mathbf{U}_p$, and $\{\boldsymbol{\alpha}_p : p = 1, \ldots, P\}$ and $\{\boldsymbol{\beta}_p : p = 1, \ldots, P\}$ are deterministic vector-valued mappings with P components. It is straightforward to show that the random field so defined has a zero mean vector. The covariance between $\mathbf{Y}_S(\mathbf{x}+\mathbf{h})$ and $\mathbf{Y}_S(\mathbf{x})$ is then

$$\begin{aligned}&\mathbb{E}\{\mathbf{Y}_S(\mathbf{x}+\mathbf{h})\cdot\mathbf{Y}_S(\mathbf{x})^T\} \\ &= \frac{1}{2}\sum_{p=1}^{P}\Big[\mathbb{E}\{\boldsymbol{\alpha}_p(\mathbf{U}_p)\boldsymbol{\alpha}_p^T(\mathbf{U}_p)\} + \mathbb{E}\{\boldsymbol{\beta}_p(\mathbf{U}_p)\boldsymbol{\beta}_p^T(\mathbf{U}_p)\}\Big]\cos(\langle \mathbf{h}, \mathbf{U}_p\rangle) \\ &\quad -\frac{1}{2}\sum_{p=1}^{P}\Big[\mathbb{E}\{\boldsymbol{\alpha}_p(\mathbf{U}_p)\boldsymbol{\beta}_p^T(\mathbf{U}_p)\} - \mathbb{E}\{\boldsymbol{\beta}_p(\mathbf{U}_p)\boldsymbol{\alpha}_p^T(\mathbf{U}_p)\}\Big]\sin(\langle \mathbf{h}, \mathbf{U}_p\rangle).\end{aligned}$$

This can be rewritten as follows:

$$\begin{aligned}&\mathbb{E}\{\mathbf{Y}_S(\mathbf{x}+\mathbf{h})\cdot\mathbf{Y}_S(\mathbf{x})^T\} \\ &= \frac{1}{2}\int_{\mathbb{R}^d}\sum_{p=1}^{P}\Big[\boldsymbol{\alpha}_p(\mathbf{u})\boldsymbol{\alpha}_p^T(\mathbf{u}) + \boldsymbol{\beta}_p(\mathbf{u})\boldsymbol{\beta}_p^T(\mathbf{u})\Big]\cos(\langle \mathbf{h}, \mathbf{u}\rangle)g(\mathbf{u})\,\mathrm{d}\mathbf{u} \\ &\quad -\frac{1}{2}\int_{\mathbb{R}^d}\sum_{p=1}^{P}\Big[\boldsymbol{\alpha}_p(\mathbf{u})\boldsymbol{\beta}_p^T(\mathbf{u}) - \boldsymbol{\beta}_p(\mathbf{u})\boldsymbol{\alpha}_p^T(\mathbf{u})\Big]\sin(\langle \mathbf{h}, \mathbf{u}\rangle)g(\mathbf{u})\,\mathrm{d}\mathbf{u} \\ &= \int_{\mathbb{R}^d}\frac{\mathbf{A}(\mathbf{u})\mathbf{A}^T(\mathbf{u}) + \mathbf{B}(\mathbf{u})\mathbf{B}^T(\mathbf{u})}{2}\cos(\langle \mathbf{h}, \mathbf{u}\rangle)g(\mathbf{u})\,\mathrm{d}\mathbf{u} \\ &\quad -\int_{\mathbb{R}^d}\frac{\mathbf{A}(\mathbf{u})\mathbf{B}^T(\mathbf{u}) - \mathbf{B}(\mathbf{u})\mathbf{A}^T(\mathbf{u})}{2}\sin(\langle \mathbf{h}, \mathbf{u}\rangle)g(\mathbf{u})\,\mathrm{d}\mathbf{u},\end{aligned} \tag{5}$$

where $\mathbf{A}(\mathbf{u})$ and $\mathbf{B}(\mathbf{u})$ are the matrices whose p-th columns are $\boldsymbol{\alpha}_p(\mathbf{u})$ and $\boldsymbol{\beta}_p(\mathbf{u})$, respectively.

Equation (5) proves that the simulated vector random field $\mathbf{Y}_S$ is second-order stationary, insofar as the covariance between $\mathbf{Y}_S(\mathbf{x}+\mathbf{h})$ and $\mathbf{Y}_S(\mathbf{x})$ does not depend on $\mathbf{x}$, but only on $\mathbf{h}$. Moreover, this covariance is equal to $\mathbf{C}(\mathbf{h})$ in Eq. (2) when the following conditions are fulfilled:

$$\begin{aligned}\mathbb{E}\{\mathbf{Y}_S(\mathbf{x}+\mathbf{h})\cdot\mathbf{Y}_S(\mathbf{x})^T\} = &\sum_{p=1}^{P}\sum_{q=1}^{P}\mathbb{E}\{\boldsymbol{\alpha}_p(\mathbf{U}_p)\boldsymbol{\alpha}_q^T(\mathbf{U}_q)\cos(\langle \mathbf{x}+\mathbf{h}, \mathbf{U}_p\rangle + \phi_p)\times\cos(\langle \mathbf{x}, \mathbf{U}_q\rangle + \phi_q)\} \\ &+ \sum_{p=1}^{P}\sum_{q=1}^{P}\mathbb{E}\{\boldsymbol{\beta}_p(\mathbf{U}_p)\boldsymbol{\beta}_q^T(\mathbf{U}_q)\sin(\langle \mathbf{x}+\mathbf{h}, \mathbf{U}_p\rangle + \phi_p)\sin(\langle \mathbf{x}, \mathbf{U}_q\rangle + \phi_q)\} \\ &+ \sum_{p=1}^{P}\sum_{q=1}^{P}\mathbb{E}\{\boldsymbol{\alpha}_p(\mathbf{U}_p)\boldsymbol{\beta}_q^T(\mathbf{U}_q)\cos(\langle \mathbf{x}+\mathbf{h}, \mathbf{U}_p\rangle + \phi_p)\sin(\langle \mathbf{x}, \mathbf{U}_q\rangle + \phi_q)\} \\ &+ \sum_{p=1}^{P}\sum_{q=1}^{P}\mathbb{E}\{\boldsymbol{\beta}_p(\mathbf{U}_p)\boldsymbol{\alpha}_q^T(\mathbf{U}_q)\times\sin(\langle \mathbf{x}+\mathbf{h}, \mathbf{U}_p\rangle + \phi_p)\cos(\langle \mathbf{x}, \mathbf{U}_q\rangle + \phi_q)\}.\end{aligned}$$

Using the product-to-sum trigonometric identities and accounting for the fact that the phases $\{\phi_p : p = 1, \ldots, P\}$ are independent and uniformly distributed in $[0, 2\pi)$, the only terms that do not vanish are found when $p = q$, so that the previous equation simplifies into

$$\begin{aligned}\frac{\mathbf{A}(\mathbf{u})\mathbf{A}^T(\mathbf{u}) + \mathbf{B}(\mathbf{u})\mathbf{B}^T(\mathbf{u})}{2}g(\mathbf{u}) &= \mathrm{Re}(\mathbf{f}(\mathbf{u})) \\ \frac{\mathbf{A}(\mathbf{u})\mathbf{B}^T(\mathbf{u}) - \mathbf{B}(\mathbf{u})\mathbf{A}^T(\mathbf{u})}{2}g(\mathbf{u}) &= \mathrm{Im}(\mathbf{f}(\mathbf{u})),\end{aligned}$$

or, equivalently,

$$[\mathbf{A}(\mathbf{u}) - i\mathbf{B}(\mathbf{u})][\mathbf{A}(\mathbf{u}) - i\mathbf{B}(\mathbf{u})]^* = \frac{2\mathbf{f}(\mathbf{u})}{g(\mathbf{u})}, \tag{6}$$

where the asterisk represents the conjugate transpose operator.

For $\mathbf{A}(\mathbf{u})$ and $\mathbf{B}(\mathbf{u})$ to remain finite for any $\mathbf{u} \in \mathbb{R}^d$, the support of g must contain the support of $\mathbf{f}$ (i.e., g is strictly positive when $\mathbf{f}$ is not zero). For instance, g can be a probability density with support on $\mathbb{R}^d$, while $\mathbf{A}(\mathbf{u})$ and $-\mathbf{B}(\mathbf{u})$ can be the real and imaginary parts of the Cholesky factor of $2\mathbf{f}(\mathbf{u})/g(\mathbf{u})$ (if $\mathbf{f}(\mathbf{u})$ is strictly definite positive for any $\mathbf{u}$) or, in its defect, of a square root of $2\mathbf{f}(\mathbf{u})/g(\mathbf{u})$.

The proposed approach Eq. (4) borrows from importance sampling, insofar as the random vectors $\{\mathbf{U}_p : p = 1, \ldots, P\}$ are simulated with a density g different from the target spectral density $\mathbf{f}$. This allows using the same set of cosine and sine basic functions to simulate all the components of the desired vector random field. A similar simulation algorithm has been proposed by (Shinozuka and Jan 1972) and (Mantoglou 1987), using a uniform density for g, thus applicable only to the simulation of random fields whose spectral densities have a bounded support, which is quite restrictive.

According to the previous statements, the simulated vector random field $\mathbf{Y}_S$ is second-order stationary with zero mean and covariance function $\mathbf{C}(\mathbf{h})$. Based on the central limit theorem, to obtain an approximately Gaussian random field, it suffices to sum up and properly normalize many of such independent random fields:

$$\begin{aligned}\forall \mathbf{x} \in \mathbb{R}^d, \mathbf{Y}_S(\mathbf{x}) = &\frac{1}{\sqrt{L}} \sum_{l=1}^{L} \sum_{p=1}^{P} \boldsymbol{\alpha}_p(\mathbf{U}_{l,p}) \cos(\langle \mathbf{x}, \mathbf{U}_{l,p} \rangle + \phi_{l,p}) \\ &+ \frac{1}{\sqrt{L}} \sum_{l=1}^{L} \sum_{p=1}^{P} \boldsymbol{\beta}_p(\mathbf{U}_{l,p}) \sin(\langle \mathbf{x}, \mathbf{U}_{l,p} \rangle + \phi_{l,p}),\end{aligned} \tag{7}$$

where L is a large integer, $\{\mathbf{U}_{l,p} : l = 1, \ldots, L; p = 1, \ldots, P\}$ are mutually independent and distributed as $\mathbf{U}_p$, and $\{\phi_{l,p} : l = 1, \ldots, L; p = 1, \ldots, P\}$ are mutually independent and distributed as ϕ_p. From Eq. (7), the target location $\mathbf{x}$ appears to be successively projected onto the random vectors $\mathbf{U}_{l,p}$, which makes our proposal a special case of the turning-bands algorithm (Matheron 1973), with cosines and sines as the basic functions used for the one-dimensional simulation.

Computing Eq. (7) for N target locations in $\mathbb{R}^d$ corresponds to generating LP uniform random variables in $[0, 2\pi)$, generating LP random vectors with probability density g, calculating LP Cholesky decompositions of complex-valued matrices of size $P \times P$, $2NLP$ scalar-vector multiplications, NLP scalar products and $N(LP + 1)$ summations. Let q be the computational cost of generating one uniform random variable in $[0, 2\pi)$, one random vector $\mathbf{u}$ with density g and calculating the matrix $\frac{2\mathbf{f}(\mathbf{u})}{g(\mathbf{u})}$. As the cost of a Cholesky factorization is $P^3/3$ floating point operations (flops) (Golub and Van 1989), the total cost of computing Eq. (7) is $LP(q + P^3/3) + N(2LP^2 + 2LPd + 1)$ flops. Accordingly, the fixed cost (irrespective of the number of locations targeted for simulation) is about $LP^4/3$, whereas the varying cost is about $2NLP^2$. Interestingly, this varying cost is directly proportional to the number N of target locations, which makes the proposal a competitive alternative to other simulation algorithms. As an example, the computational cost of the circulant-embedding algorithm (one of the fastest alternatives to date) is $O(N \log N)$ (Gneiting et al. 2006). Another advantage of the proposed algorithm over circulant-embedding techniques is the possibility to split the set of target locations into small subsets and to process these subsets consecutively and/or independently, allowing for a considerable reduction of memory storage requirements and for parallel computations.

4 Examples

In this section, the proposed algorithm is tested to simulate random fields on a regular two-dimensional grid ($d = 2$) with 500×500 nodes and a unit mesh. The simulation is performed by using $L = 500$ basic random fields in Eq. (7). The density g is chosen as the spectral density of an isotropic Matérn (Bessel-K) covariance model with scale parameter $a = 10$ and shape parameter $\nu = 0.5$:

$$M(\mathbf{h}, a, \nu) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\|\mathbf{h}\|}{a} \right)^{\nu} K_{\nu} \left(\frac{\|\mathbf{h}\|}{a} \right), \tag{8}$$

that is (Lantuéjoul 2002)

$$g(\mathbf{u}, a, \nu) = \frac{a^d \Gamma\left(\nu + \frac{d}{2}\right)}{\Gamma(\nu) \pi^{d/2}} \frac{1}{\left(1 + a^2 \|\mathbf{u}\|^2\right)^{\nu + d/2}}. \tag{9}$$

A random vector with such a density in $\mathbb{R}^d$ can be obtained by computing the ratio of a Gaussian random vector and the square root of a gamma random variable with shape parameter ν (Emery and Lantuéjoul 2006).

The examples presented in the next subsections and their specific parameters have been taken from the literature or chosen arbitrarily, in order to illustrate the capabilities and versatility of the proposed algorithm.

4.1 Example 1: bivariate random field with Matérn covariance model

Consider a vector random field with $P = 2$ components and a bivariate Matérn covariance model (Gneiting et al. 2010)

$$\mathbf{C}(\mathbf{h}) = \begin{pmatrix} M(\mathbf{h}, a_1, \nu_1) & \rho_{12} M(\mathbf{h}, a_{12}, \nu_{12}) \\ \rho_{12} M(\mathbf{h}, a_{12}, \nu_{12}) & M(\mathbf{h}, a_2, \nu_2) \end{pmatrix}. \quad (10)$$

The two components have unit variances and collocated correlation coefficient ρ_{12}. This model is especially interesting as it allows for different levels of smoothness for the components of the associated vector Gaussian random field, associated with the shape parameters ν_1 and ν_2. Following (Gneiting et al. 2010), we here consider the case when $a_1 = 20$, $a_2 = 100$, $a_{12} = 100/3$, $\nu_1 = 1.5$, $\nu_2 = 0.5$, $\nu_{12} = 1.0$ and $\rho_{12} = 0.5$.

The steps for simulation are the following:

(i) For $p = 1, 2$ and $l = 1, \ldots, 500$

a. Simulation of a phase $\phi_{l,p}$ uniformly distributed in $[0, 2\pi)$.

b. Simulation of a random vector $\mathbf{U}_{l,p}$ in $\mathbb{R}^2$ with density $g(., a, \nu)$ Eq. (9).

c. Calculation of the spectral density matrix

$$\mathbf{f}(\mathbf{U}_{l,p}) = \begin{pmatrix} g(\mathbf{U}_{l,p}, a_1, \nu_1) & \rho_{12} g(\mathbf{U}_{l,p}, a_{12}, \nu_{12}) \\ \rho_{12} g(\mathbf{U}_{l,p}, a_{12}, \nu_{12}) & g(\mathbf{U}_{l,p}, a_2, \nu_2) \end{pmatrix}.$$

Given the values chosen for parameters a_1, a_2, a_{12}, ν_1, ν_2, ν_{12} and ρ_{12}, this matrix is positive semi-definite for all $\mathbf{U}_{l,p} \in \mathbb{R}^2$, which ensures the validity of the bivariate Matérn model under consideration.

d. Eigendecomposition and calculation of the principal square root of matrix $\frac{2\mathbf{f}(\mathbf{U}_{l,p})}{g(\mathbf{U}_{l,p}, a, \nu)}$.

e. Identification of matrices $\mathbf{A}(\mathbf{U}_{l,p})$ and $-\mathbf{B}(\mathbf{U}_{l,p})$ as the real and imaginary parts of this square root matrix Eq. (6).

f. Identification of vectors $\boldsymbol{\alpha}_p(\mathbf{U}_{l,p})$ and $\boldsymbol{\beta}_p(\mathbf{U}_{l,p})$ as the p-th columns of $\mathbf{A}(\mathbf{U}_{l,p})$ and $\mathbf{B}(\mathbf{U}_{l,p})$, respectively.

(ii) Calculation of the simulated random field $\mathbf{Y}_S$ at all the grid nodes, as per Eq. (7).

Figure 1 shows the map of one realization, while Fig. 2 compares the experimental variograms of one hundred realizations (calculated along the abscissa axis) with the theoretical Matérn variograms. The experimental variograms fluctuate around the theoretical model and the amplitude of the fluctuation increases with the lag separation distance, as expected by theory (Chilès and Delfiner 2012; Emery 2007). However, on average over the realizations, the experimental variograms almost match the theoretical model, corroborating that the spatial correlation structure of the desired vector random field is reproduced, see further discussion in Sect. 4.5.

4.2 Example 2: bivariate random field with compactly supported covariance functions

In this subsection, let us consider the following Wendland covariance model in the two-dimensional space (Daley et al. 2015):

$$W(\mathbf{h}, a) = \left(1 - \frac{\|\mathbf{h}\|}{a}\right)_+^5 \left(1 + 5\frac{\|\mathbf{h}\|}{a}\right), \quad (11)$$

where $(\cdot)_+$ denotes positive part and a identifies the compact support of the covariance.

Following (Daley et al. 2015), one can define a bivariate random field ($P = 2$) with covariance functions of the form:

$$\mathbf{C}(\mathbf{h}) = \begin{pmatrix} W(\mathbf{h}, a_{11}) & \rho_{12} W(\mathbf{h}, a_{12}) \\ \rho_{12} W(\mathbf{h}, a_{12}) & W(\mathbf{h}, a_{22}) \end{pmatrix}. \quad (12)$$

The spectral density of $W(\mathbf{h}, a)$ is (Zastavnyi 1991)

$$f(\mathbf{u}, a) = \frac{a^2}{120} {}_1F_2\left(\frac{5}{2}; \frac{9}{2}, 5, -\frac{a^2 \|\mathbf{u}\|^2}{4}\right), \quad (13)$$

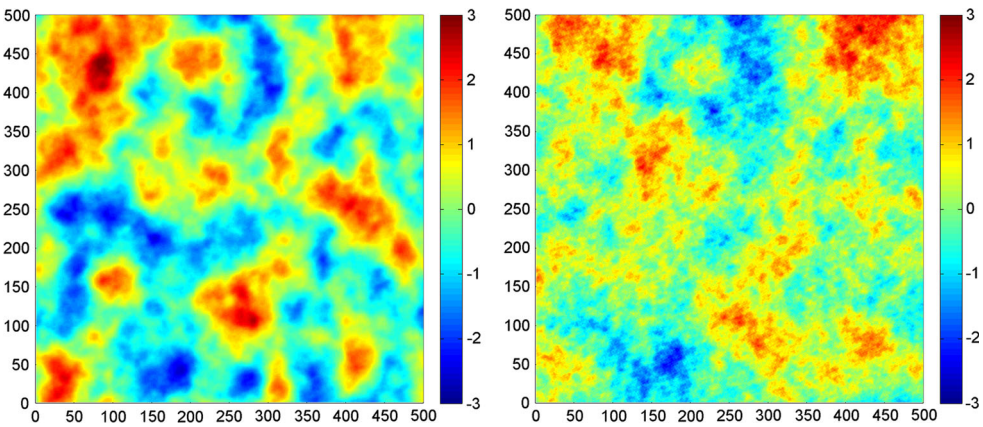

Fig. 1 Realization of a vector Gaussian random field with bivariate Matérn covariance structure (first component on the *left*, second component on the *right*), with scale parameters $a_1 = 20, a_2 = 100$ and $a_{12} = 100/3$, and shape parameters $\nu_1 = 1.5, \nu_2 = 0.5$ and $\nu_{12} = 1$

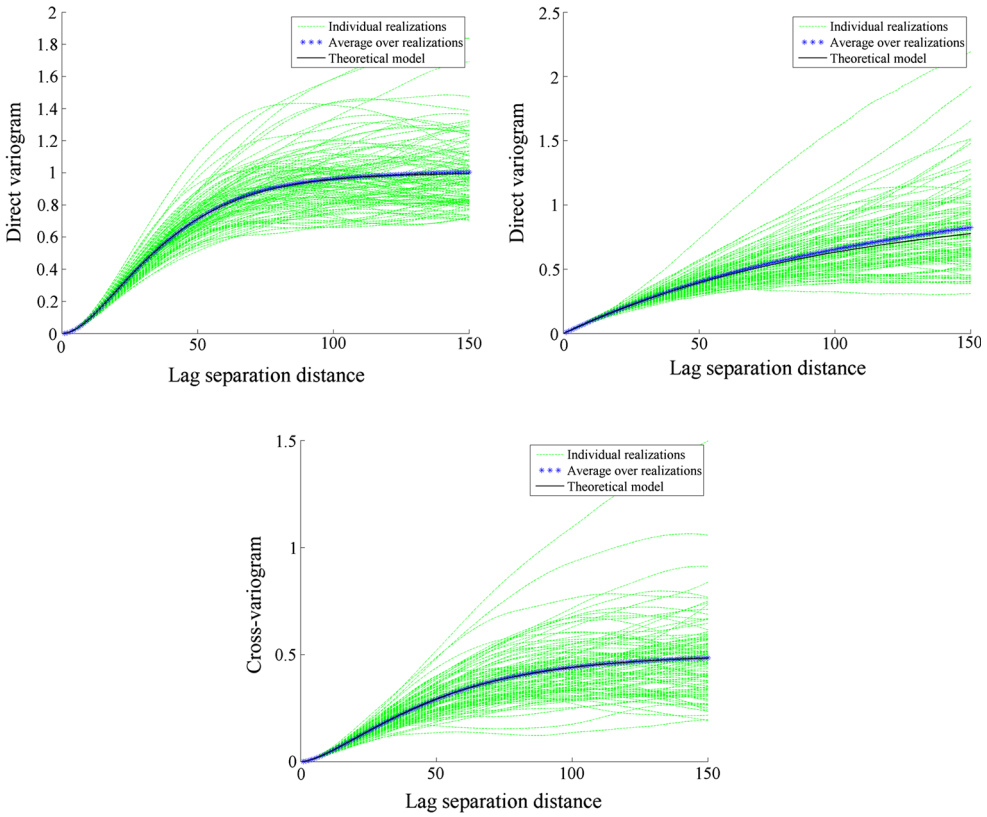

Fig. 2 Experimental variograms for 100 realizations (*green dashed lines*), average of experimental variograms (*blue stars*) and theoretical variograms (*black solid lines*) (bivariate Matérn model). From *left* to *right* and *top* to *bottom* variograms for first component, variograms for second component, and cross-variograms

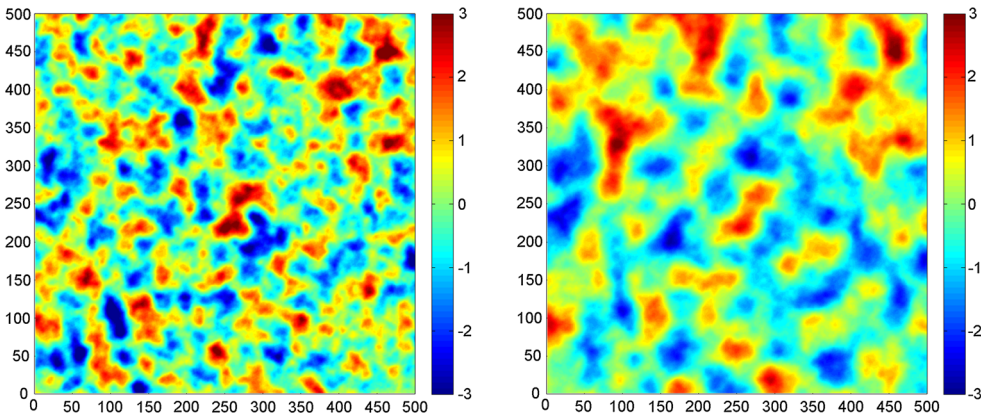

Fig. 3 Realization of a vector Gaussian random field with compactly supported covariances (Eq. (12)) with $a_{11} = 50, a_{22} = 100$, $a_{12} = 70$ and $\rho_{12} = 0.5$. *Left* first component; *right* second component

where ${}_1F_2$ is a generalized hypergeometric function, which can be calculated by means of Bessel functions of the first kind and Struve functions (http://functions.wolfram.com/HypergeometricFunctions/hypergeometric1F2/03/03/11/15/0004/, accessed on April 27, 2015).

Figure 3 shows a map of one realization obtained with $a_{11} = 50, a_{12} = 70, a_{22} = 100$ and $\rho_{12} = 0.5$, while Fig. 4 displays the experimental direct and cross-variograms calculated over one hundred realizations, which fluctuate around the theoretical model (Eq. (12)).

4.3 Example 3: random field regularized at different supports

Let us consider a scalar random field Y_0 with a Matérn covariance model ($a_0 = 100$, $\nu_0 = 0.6$) in $\mathbb{R}^d$ and its regularization Y_r over a ball with radius r:

$$Y_r = Y_0 * \breve{\omega}_r,$$

where $*$ is the convolution operator, ω_r is the indicator function of the ball of radius r centered at the origin, and

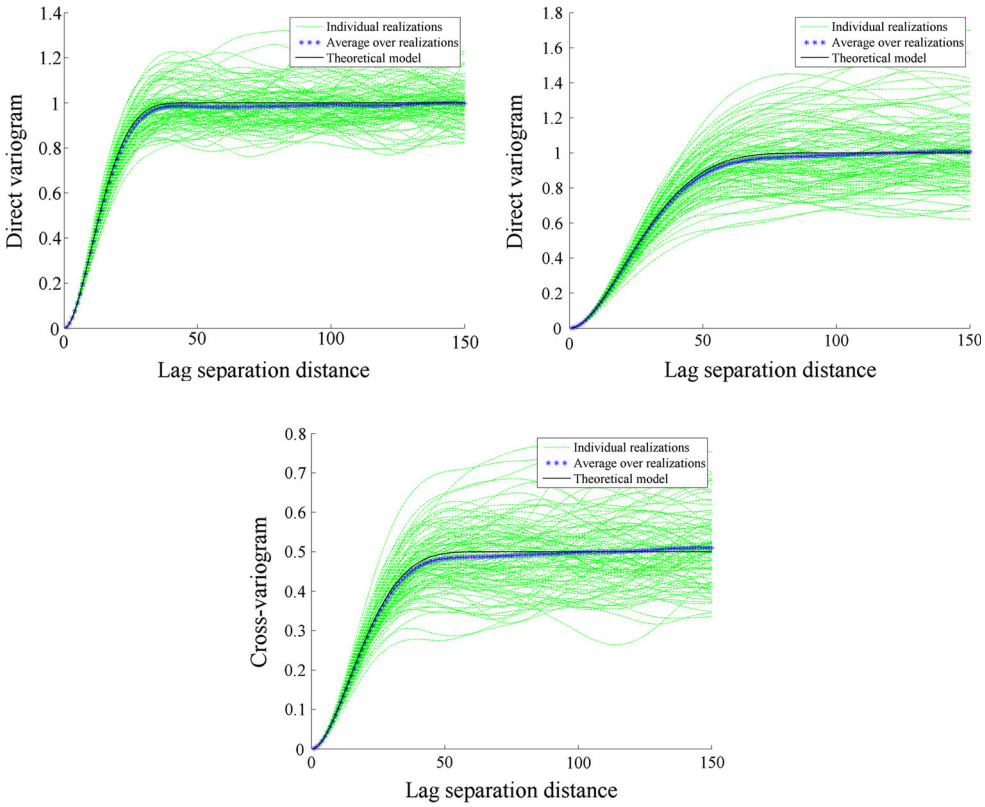

Fig. 4 Experimental variograms for 100 realizations (*green dashed lines*), average of experimental variograms (*blue stars*) and theoretical variogram models (*black solid lines*) (compactly supported model). From *left* to *right* and *top* to *bottom* variograms for first component, variograms for second component and cross-variograms

$\breve{\omega}_r(\mathbf{t}) = \omega_r(-\mathbf{t})$. The spectral density f_{rr} of Y_r can be derived from the spectral density f_{00} of Y_0 (Proof in Appendix)

$$f_{rr}(\mathbf{u}) = f_{00}(\mathbf{u})\xi^2(\mathbf{u}, r), \tag{14}$$

where

$$\xi(\mathbf{u}, r) = \frac{2^{d/2}\Gamma\left(\frac{d}{2}+1\right)J_{d/2}(\,r\|\mathbf{u}\|\,)}{(r\|\mathbf{u}\|)^{d/2}} \tag{15}$$

with $J_{d/2}$ being the Bessel function of the first kind with index d / 2.

Similarly, the density f_{0r} associated with the cross-covariance between Y_0 and Y_r is

$$f_{0r}(\mathbf{u}) = f_{00}(\mathbf{u})\xi(\mathbf{u}, r). \tag{16}$$

Likewise, the density associated with the cross-covariance between the two regularized random fields Y_r and $Y_{r'}$ is

$$f_{rr'}(\mathbf{u}) = f_{00}(\mathbf{u})\xi(\mathbf{u}, r)\xi(\mathbf{u}, r') \tag{17}$$

The above formulae have been used for simulating the vector random field

$$\mathbf{Y} = [Y_0, Y_3, Y_6, Y_9, Y_{12}, Y_{15}, Y_{18}, Y_{21}, Y_{24}]^T$$

on the previously defined two-dimensional grid. A map of one realization is displayed in Fig. 5, where one observes a progressive smoothing when the radius of the regularization ball increases, in agreement with the change-of-support theory (Chilès and Delfiner 2012). Note that the standard approach for change of support, which consists in simulating Y_0 and regularizing it over balls of radii 3, 6, 9, 12, 15, 18, 21 and 24, would be approximate since, in practice, the balls need to be discretized into a finite set of point-support locations. In contrast, our algorithm directly simulates the random field Y_0 and its regularizations over different supports without the need for any discretization.

4.4 Example 4: bivariate random field with an odd cross-covariance function

In the examples shown so far, the cross-covariances are even functions, so that the spectral densities are real-valued. Accordingly, matrix $\mathbf{B}$ in Eq. (6) is zero and the sine terms vanish from Eq. (7). The following gives an example of random fields with an odd cross-covariance function, for which the spectral density is purely imaginary.

Consider the following bilinear coregionalization model (De Iaco et al. 2003; Grzebyk and Wackernagel 1994):

$$\mathbf{C}(\mathbf{h}) = M(\mathbf{h}, a_0, \nu_0)[B_1 \cos\langle \mathbf{u}_0, \mathbf{h}\rangle - B_2 \sin\langle \mathbf{u}_0, \mathbf{h}\rangle]$$

with $\mathbf{u}_0$ fixed, and

$$B_1 = \begin{pmatrix} b_1 & 0 \\ 0 & b_1 \end{pmatrix} \quad \text{and} \quad B_2 = \begin{pmatrix} 0 & b_2 \\ -b_2 & 0 \end{pmatrix}.$$

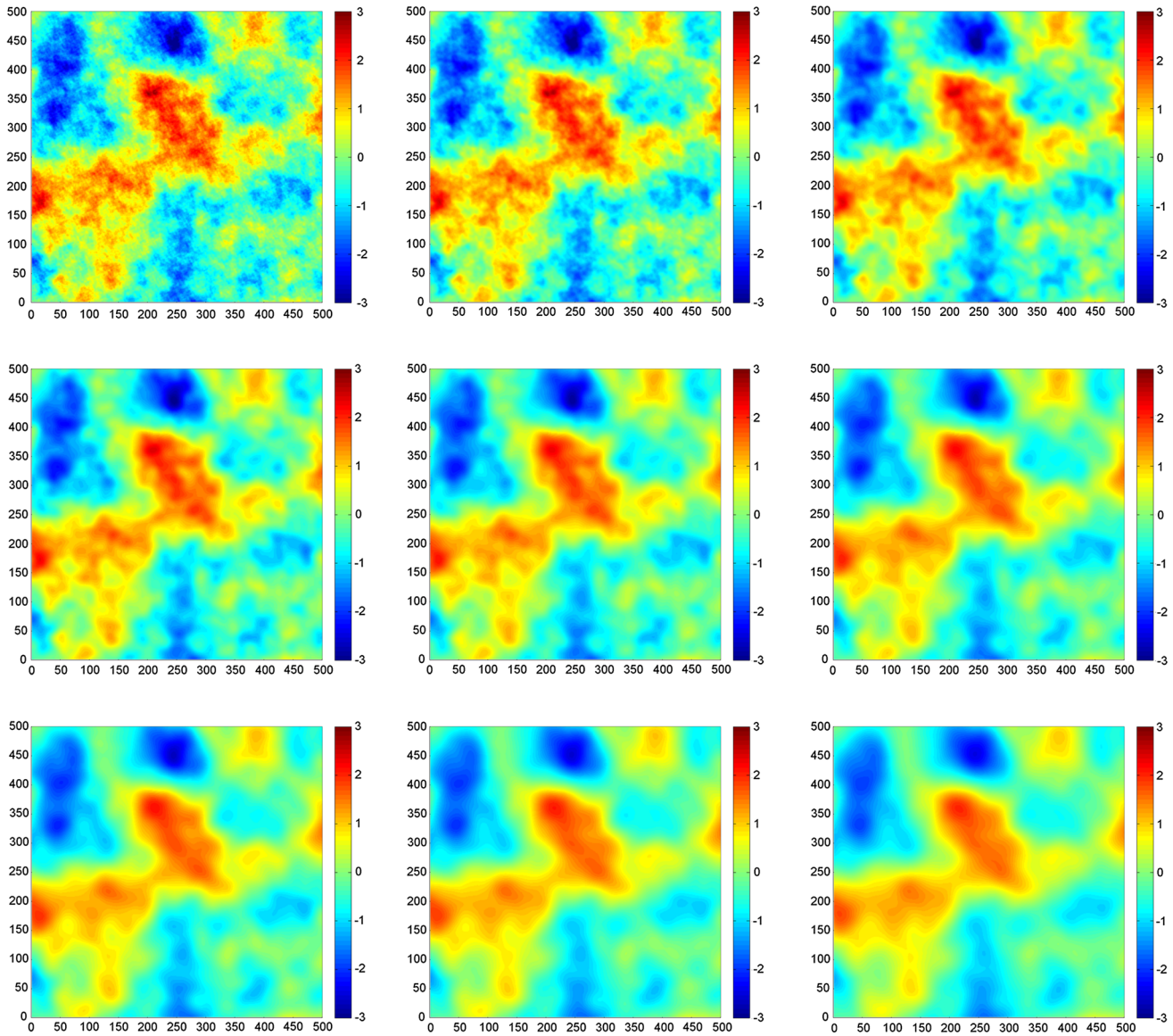

Fig. 5 Realization of a vector Gaussian random field with Matérn covariance model (*top left*) and its regularizations (from *left* to *right* and *top* to *bottom*, starting from the top medium, $r = 3, 6, 9, 12, 15, 18, 21$ and 24)

The corresponding spectral density matrix is

$$\mathbf{f}(\mathbf{u}) = \begin{pmatrix} f_{11}(\mathbf{u}) & f_{12}(\mathbf{u}) \\ f_{21}(\mathbf{u}) & f_{22}(\mathbf{u}) \end{pmatrix}$$

with

$$\begin{aligned} f_{11}(\mathbf{u}) &= f_{22}(\mathbf{u}) \\ &= b_1 \left[\frac{g(\mathbf{u}-\mathbf{u}_0, a_0, \nu_0) + g(\mathbf{u}+\mathbf{u}_0, a_0, \nu_0)}{2} \right], \\ f_{12}(\mathbf{u}) &= -f_{21}(\mathbf{u}) \\ &= -i\, b_2 \left[\frac{g(\mathbf{u}-\mathbf{u}_0, a_0, \nu_0) - g(\mathbf{u}+\mathbf{u}_0, a_0, \nu_0)}{2} \right]. \end{aligned}$$

Figure 6 shows the map of one realization for $a_0 = 100, \nu_0 = 0.6, b_1 = 1, b_2 = 0.7$ and $\mathbf{u}_0 = 0.04$, while Fig. 7 compares the experimental covariances (calculated along the abscissa axis) of one hundred realizations with the theoretical covariance. On average over the realizations, the experimental covariances almost match the theoretical model. The variograms have not been calculated in this example, insofar as the cross-variograms only contain information on the even part of the cross-covariance, which is zero (Eq. (3)).

4.5 Discussion: fluctuation between experimental covariance or variogram and theoretical model

The experimental covariance or variogram of a random field simulated in a domain of finite size does not perfectly match the underlying theoretical covariance or variogram

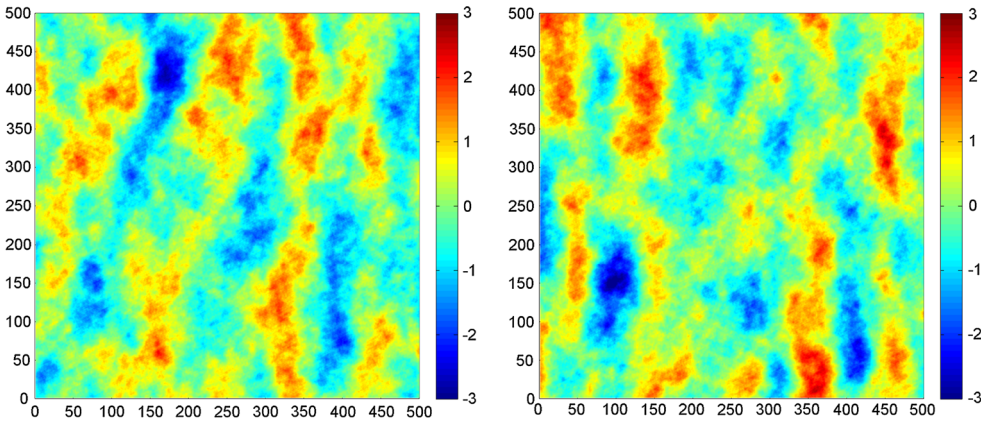

Fig. 6 Realization of a vector Gaussian random field with a correlation structure given by a bilinear coregionalization model

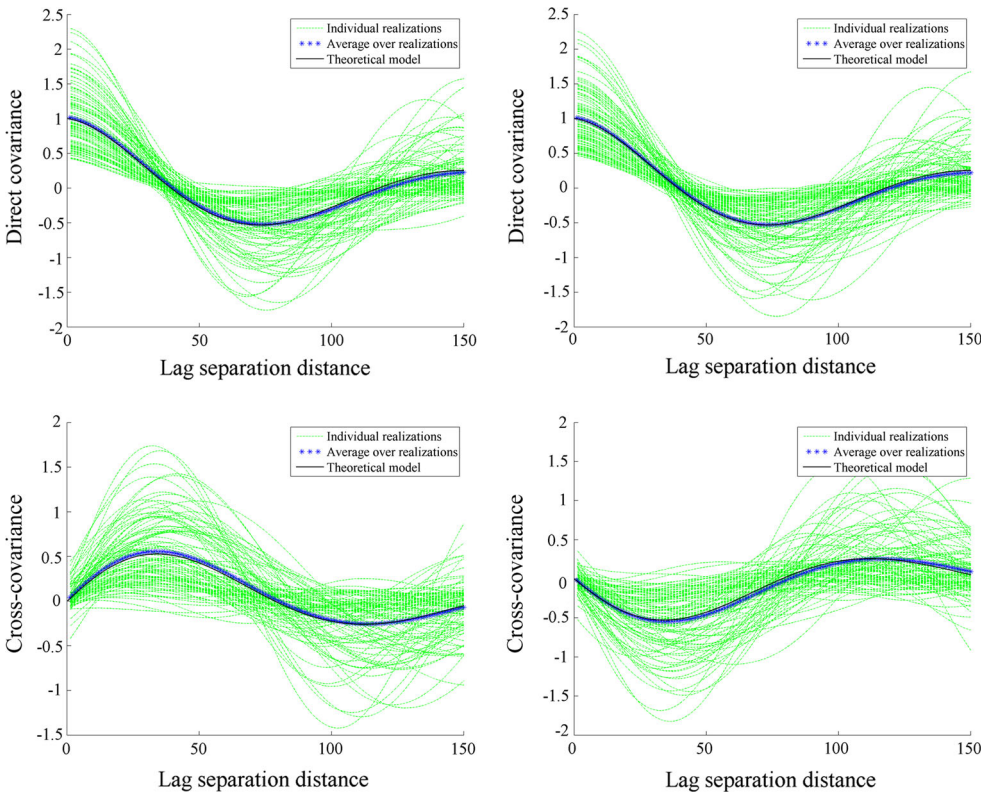

Fig. 7 Experimental covariances for 100 realizations (*green dashed lines*), average of experimental covariances (*blue stars*) and theoretical variogram models (*black solid lines*) (bilinear coregionalization model). From *left* to *right* and *top* to *bottom* covariances for first component, covariances for second component, and cross-covariances

model. The deviation between the experimental covariance or variogram and the model is called a fluctuation. It can be shown (Chilès and Delfiner 2012; Emery 2007) that, provided that the simulation algorithm is correct, the fluctuation at a given lag separation distance is a zero-mean random variable whose variance depends on the fourth-order moments of the underlying random field and on the size and shape of the simulation domain. Accordingly, when averaging the experimental covariances or variograms of a set of realizations drawn independently, the fluctuation variance should tend to zero as the number of realizations becomes very large. This fact agrees with the examples shown in the previous subsections, as the average experimental covariance or variogram is seen to get closer to the theoretical model when the number of realizations increases (Table 1). Such a result corroborates that the theoretical covariance of the simulated random field, as calculated in Eq. (5), matches the desired covariance model, i.e., that the proposed spectral-turning bands algorithm generates realizations that reproduce the spatial correlation structure of the target random field.

Table 1 Mean absolute deviation between (i) the theoretical variogram or covariance model and (ii) the average experimental variogram or covariance of the first R realizations, calculated for lag distances 1 to 150, as per Figs. 2, 4 and 7

Model	Number of realizations		
	$R = 1$	$R = 10$	$R = 100$
Bivariate Matérn			
Direct variogram 1	0.0557	0.0200	0.0038
Direct variogram 2	0.1769	0.0671	0.0168
Cross-variogram	0.0826	0.0160	0.0009
Compactly supported			
Direct variogram 1	0.1867	0.0205	0.0117
Direct variogram 2	0.0727	0.0276	0.0112
Cross-variogram	0.1271	0.0139	0.0068
Bilinear coregionalization			
Direct covariance 1	0.2136	0.0934	0.0267
Direct covariance 2	0.2646	0.0900	0.0182
Cross-covariance 1 and 2	0.1856	0.0444	0.0190
Cross-covariance 2 and 1	0.1673	0.0724	0.0242

5 Conclusions

We proposed a simulation algorithm having several nice features: (i) accuracy: the simulated random fields have the desired covariance structure; (ii) speed: the algorithm works very fast and lends itself to parallel computing; (iii) memory storage requirements: the entire field can be simulated progressively, with no need to be stored in random access memory; (iv) versatility: one can simulate any vector random field possessing continuous and absolutely integrable covariance functions, with any number of components (P), in any workspace dimension (d), for any number (N) of the target locations and any configuration of these locations (evenly spaced or not). The simulation at non-evenly spaced locations is a frequent problem in geostatistical applications, where it is of interest to condition the realizations to data known at scattered locations in space (Chilès and Delfiner 2012). The applicability of the algorithm has been demonstrated through several examples, where the vector components exhibit different spatial behaviors, such as their smoothness or their correlation ranges, which go well beyond the linear coregionalization model classically used in geostatistics.

Acknowledgments The authors acknowledge the funding by the Chilean Commission for Scientific and Technological Research, through Projects CONICYT / FONDECYT / REGULAR / No. 1130085, CONICYT / FONDECYT / POSTDOCTORADO / No. 3140568 and CONICYT / FONDECYT / REGULAR / No. 1130647, respectively.

Appendix: proof of equations (14), (16) and (17)

Let Y_0 be a scalar random field. The random field regularized by a sampling function $\omega_r(\cdot)$ being the indicator function of the ball of $\mathbb{R}^d$ with arbitrary radius r is defined as (Chilès and Delfiner 2012)

$$Y_r(\mathbf{x}) = \int_{\mathbb{R}^d} Y_0(\mathbf{x}+\mathbf{t})\omega_r(\mathbf{t})\,\mathrm{d}\mathbf{t}.$$

If Y_0 is a second-order stationary random field with covariance $C_0(\mathbf{h})$, then Y_r is also a second-order stationary random field, therefore one can define the following covariances that depend only on $\mathbf{h}$.

The covariance between $Y_0(\mathbf{x}+\mathbf{h})$ and $Y_r(\mathbf{x})$ is

$$\begin{aligned} C_{0r}(\mathbf{h}) &= \int_{\mathbb{R}^d} \mathbb{E}\{Y_0(\mathbf{x}+\mathbf{h})\cdot Y_0(\mathbf{x}+\mathbf{t})\}\omega_r(\mathbf{t})\,\mathrm{d}\mathbf{t} \\ &= \int_{\mathbb{R}^d} C_0(\mathbf{h}-\mathbf{t})\omega_r(\mathbf{t})\,\mathrm{d}\mathbf{t}, \end{aligned}$$

that is

$$C_{0r} = C_0 * \omega_r. \tag{18}$$

Similarly, the covariance between $Y_r(\mathbf{x}+\mathbf{h})$ and $Y_{r'}(\mathbf{x})$ is (Chilès and Delfiner 2012)

$$C_{rr'} = C_0 * (\breve{\omega}_r * \omega_{r'}) \tag{19}$$

with $\breve{\omega}(\mathbf{t}) = \omega(-\mathbf{t})$.

Since the Fourier transformation exchanges convolution and multiplication, Eqs. (18) and (19) in terms of Fourier transforms follow respectively, as:

$$\begin{aligned} f_{0r} &= f_{00}\cdot\xi_{\omega_r} \\ f_{rr'} &= f_{00}\cdot\xi_{\breve{\omega}_r}\cdot\xi_{\omega_{r'}}, \end{aligned}$$

where f_{00} is the spectral density of Y_0, and ξ_{ω_r}, $\xi_{\breve{\omega}_r}$ and $\xi_{\omega_{r'}}$ are the Fourier transforms of ω_r, $\breve{\omega}_r$ and $\omega_{r'}$, respectively. As ω_r is the indicator of a ball with radius r, one has $\xi_{\omega_r} = \xi_{\breve{\omega}_r} = \xi(\cdot, r)$ (Eq. (15)) (Gradshtein and Ryzhik 1965); likewise, $\xi_{\omega_{r'}} = \xi(\cdot, r')$.

References

Alabert F (1987) The practice of fast conditional simulations through the LU decomposition of the covariance matrix. Math Geol 19(5):369–386

Almeida AS, Journel AG (1994) Joint simulation of multiple variables with a Markov-type coregionalization model. Math Geol 26(5):565–588

Arroyo D, Emery X, Peláez M (2012) An enhanced Gibbs sampler algorithm for non-conditional simulation of Gaussian random vectors. Comput Geosci 46:138–148

Bevilacqua M, Hering AS, Porcu E (2015) On the flexibility of multivariate covariance models. Stat Sci 30(2):167–169

Bochner S (1933) Monotone Funktionen Stieljessche integrale and harmonische analyse. Math Ann 108:378–410
Box GEP, Jenkins GM (1976) Time series analysis: forescasting and control, revised edn. Holden-Day, Oakland
Chan G, Wood A (1999) Simulation of stationary Gaussian vector fields. Stat Comput 9(4):265–268
Chilès JP, Delfiner P (1997) Discret exact simulation by the Fourier method. In: Baafi EY, Schofield NA (eds) Geostatistics Wollongong '96, vol 1. Kluwer, Dordrecht, pp 258–269
Chilès JP, Delfiner P (2012) Geostatistics: Modeling Spatial Uncertainty, 2nd edn. Wiley, New York
Christakos G (1992) Random field models in earth sciences. Academic Press, San Diego
Cramér H (1940) On the theory of Stationary random functions. Ann Math 41:215–230
Daley DJ, Porcu E, Bevilacqua M (2015) Classes of compactly supported covariance functions for multivariate random fields. Stoch Environ Res Risk Assess 29(4):1249–1263
Davis MW (1987) Production of conditional simulations via the LU triangular decomposition of the covariance matrix. Math Geol 19(2):91–98
De Iaco S, Palma M, Posa D (2003) Covariance functions and models for complex-valued random fields. Stoch Environ Res Risk Assess 17(3):145–156
Dietrich CR, Newsam GN (1993) A fast and exact method for multidimensional Gaussian stochastic simulations. Water Resour Res 29(8):2861–2869
Emery X (2007) Reducing fluctuations in the sample variogram. Stoch Environ Res Risk Assess 21(4):391–403
Emery X (2008) A turning bands program for conditional Co-simulation of cross-correlated random fields. Comput Geosci 34(12):1850–1862
Emery X, Arroyo D, Peláez M (2014) Simulating large Gaussian random vectors subject to inequality constraints by Gibbs sampling. Math Geosci 46(3):265–283
Emery X, Lantuéjoul C (2006) TBSIM: a computer program for conditional simulation of three-dimensional Gaussian random fields via the turning bands method. Comput Geosci 32(10):1615–1628
Gaspari G, Cohn SE (1999) Construction of correlation functions in two and three dimensions. Q J R Meteorol Soc 125:723–757
Genton M, Kleiber W (2015) Cross-covariance functions for multivariate geostatistics. (With discussion). Stat Sci 30(2):147–163
Gneiting T, Kleiber W, Schlather M (2010) Matérn cross-covariance functions for multivariate random fields. J Am Stat Assoc 105(491):1167–1177
Gneiting T, Ševčíhová H, Percival DB, Schlather M, Jiang Y (2006) Fast and exact simulation of large Gaussian lattice systems in $\mathbb{R}^2$: exploring the limits. J Comput Graph Stat 15(3):483–501
Golub G, Van Loan C (1989) Matrix computations. John Hopkins Press, Baltimore
Gradshtein IS, Ryzhik IM (1965) Table of integrals, series and products, 4th edn. Academic Press, New York
Grzebyk M, Wackernagel H (1994) Multivariate analysis and spatial/temporal scales: Real and complex Models. In Proceedings of XVIIth international biometric conference, Hamilton, Ontario vol 1 pp 19–33
Journel AG, Huijbregts CJ (1978) Mining geostatistics. Academic Press, London
Lantuéjoul C (2002) Geostatistical simulation: models and algorithms. Springer-Verlag, Berlin
Mantoglou A (1987) Digital simulation of multivariate two- and three-dimensional stochastic processes with a spectral turning bands method. Math Geol 19(2):129–149
Marcotte D (2012) Revisiting the linear model of coregionalization. In: Abrahamsen P, Hauge R, and Kolbjørnsen O, (eds). Proceedings of IX international geostatistics congress, Springer-Verlag, Oslo, 11–15 June 2012, pp 67–78
Matheron G (1973) The intrinsic random functions and their applications. Adv Appl Probab 5:439–468
Mignolet MP, Spanos PD (1992) Simulation of homogeneous two-dimensional random fields: part I-AR and ARMA models. J Appl Mech 59(2):260–269
Pardo-Iguzquiza E, Chica-Olmo M (1993) The fourier integral method: an efficient spectral method for simulation of random fields. Math Geol 25(2):177–217
Porcu E, Gregori P, Mateu J, Ostoja-Starzewski M (2012) New classes of spectral densities for lattice processes and random fields built from simple univariate margins. Stoch Environ Res Risk Assess 26(4):479–492
Ripley BD (1987) Stochastic Simulation. Wiley, New York
Schlather M (2012) Construction of covariance functions and unconditional simulation of random fields. Advances and Challenges in Space-Time modelling of natural events. Springer-Verlag, Lecture Notes in Statistics, pp 25–54
Shinozuka M (1971) Simulation of multivariate and multidimensional random processes. J Acoust Soc Am 49(1B):357–367
Shinozuka M, Jan CM (1972) Digital simulation of random processes and its applications. J Sound Vib 25(1):111–128
Spanos PD, Mignolet MP (1992) Simulation of homogeneous two-dimensional random fields: part II-MA and ARMA models. J Appl Mech 59(2):270–277
Ver Hoef JM, Barry RP (1998) Constructing and fitting models for cokriging and multivariate spatial prediction. J Stat Plan Infer 69:275–294
Wackernagel H (2003) Multivariate geostatistics: an introduction with applications. Springer-Verlag, Berlin
Wood ATA, Chan G (1994) Simulation of stationary Gaussian processes in $[0,1]^d$. J Comput Graph Stat 3(4):409–432
Zastavnyi VP (1991) Positive-definite functions depending on the norm. Solution of the Schoenberg problem. Institute of Applied Mathematics and Mechanics, Ukranian Academy of Sciences, Donetsk