Non-parametric Identification of Anisotropic (Elliptic) Correlations in Spatially Distributed Data Sets

Arsenia Chorti[§] and Dionissios T. Hristopulos[*]

Abstract

Random fields are useful models of spatially variable quantities, such as those occurring in environmental processes and medical imaging. The fluctuations obtained in most natural data sets are typically anisotropic. The parameters of anisotropy are often determined from the data by means of empirical methods or the computationally expensive method of maximum likelihood. In this paper we propose a systematic method for the identification of geometric (elliptic) anisotropy parameters of scalar fields. The proposed method is computationally efficient, non-parametric, non-iterative, and it applies to differentiable random fields with normal or lognormal probability density functions. Our approach uses sample based estimates of the random field spatial derivatives that we relate through closed form expressions to the anisotropy parameters. This paper focuses on two spatial dimensions. We investigate the performance of the method on synthetic samples with Gaussian and Matérn correlations, both on regular and irregular lattices. The systematic anisotropy detection provides an important pre-processing stage of the data. Knowledge of the anisotropy parameters, followed by suitable rotation and rescaling transformations restores isotropy thus allowing classical interpolation and signal processing methods to be applied.

Index Terms

Manuscript Received July 11, 2007, Latex

A. Chorti is with the Department of Mineral Resources Eng., Technical University of Crete, Chania 73100, Greece. [§]Current address: Department of Electronic and Electrical Eng., University College London. (e-mail: ersi.chorti@gmail.com)

[*]Corresponding author, D.T. Hristopulos is with the Department of Mineral Resources Eng., Technical University of Crete, Chania 73100, Greece (e-mail: dionisi@mred.tuc.gr, Tel:+30 2821037688)

Anisotropy, Spatial Random Fields, Geostatistics, Mapping, Parameter Inference, Image Processing

I. INTRODUCTION

The efficient stochastic modeling of multi-dimensional data is at the epicenter of geophysical, informational and environmental sciences as well as image processing and even transistor circuit design methodologies [1]. Spatial statistics provides important tools for modeling the underlying processes in a variety of applications, using classical correlation methods [2] or semivariograms [3] and spatial interpolation based on kriging [4], [5] and Wiener filtering [6], [7] methods. In recent years, there is also a growing interest in spatio-temporal variogram and covariance models [8]. Furthermore, over the past few years variogram methods have been introduced in image processing applications such as medical imaging [9], [10], [11], [12] and brain mapping [13], [14].

Most signal processing methods are based on the assumption that the Spatial Random Fields (SRFs) are statistically homogeneous and isotropic. Statistical homogeneity implies that the SRF correlation function depends only on the lag vector between two points, but not on their exact location [15]. Statistical isotropy implies that the correlation function depends purely on the Euclidean distance, but not on the direction of the lag vector, i.e., that the correlation lengths are equal along every direction [16]. The assumption of isotropy generally fails in natural SRFs. Spatial data often exhibit continuity properties that depend on the direction in space.

Two types of anisotropy are usually considered in geostatistical studies: (i) geometric anisotropy [17] (also known as elliptic or range anisotropy) implies that the correlation iso-level contours are ellipses (in 2D) or ellipsoids (in 3D), (ii) zonal anisotropy implies a direction dependent variogram sill. The focus of this paper is on geometric anisotropy in two spatial dimensions (called anisotropy hereon).

Determining the parameters of anisotropy allows a coordinate system transformation that renders the spatial dependence statistically isotropic in the new system (i.e. the iso-level contours become circles). The geometric anisotropy in a given 2D sampling coordinate system of axes x and y is fully characterized by the anisotropy ratio R and the orientation angle θ. R represents the ratio of the correlation lengths along the principal axes of anisotropy (i.e. it is the ratio of the semi-axes of the elliptical iso-level contours) and θ is the rotation angle between the principal coordinates system and the sampling system. The isotropic transformations in spatial data are depicted in Fig. 1.

The principal system of axes is first aligned with the sampling coordinate system by rotation. Following

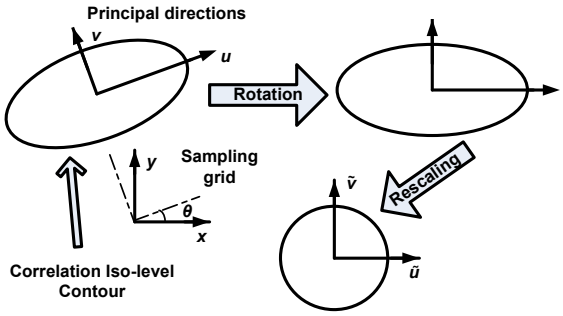

Fig. 1. Isotropic transformations in two dimensions.

that, the rotated principal coordinate system is rescaled so that we obtain an isotropic coordinate system in which iso-level correlation contours become circles. The coordinates in the rotated and rescaled isotropic principal coordinate system are expressed as follows, in relation to the initial sampling coordinate system:

$$\tilde{u} = x\cos\theta + y\sin\theta \tag{1}$$

$$\tilde{v} = (-x\sin\theta + y\cos\theta)\, R \tag{2}$$

Empirical estimates of geometric anisotropy are based on the visual inspection of geological maps or the estimation of the experimental directional variograms (structure functions). The latter are estimated along the principal directions, if these are known a priori (not a typical case), or arbitrarily in the North-South and East-West directions, or in multiple defined directions) [18], [17]. The estimation of directional variograms is an operation with numerical complexity $O(N^2)$ and requires making empirical choices pertaining to discretization (i.e., the number of radial distance and orientation angle classes). In addition, fitting of the directional variograms to a parametric model can be conducted by means of various methods, and the choice influences the anisotropy estimation.

An alternative approach for the evaluation of the anisotropy parameters of spatial data is based on maximum likelihood estimation (MLE) [19]–[21], which presumes a given parametric correlation model. The MLE seeks optimal parameters (in 2D these involve the variance, correlation length, anisotropic ratio, and orientation angle) for this model. This task is computationally greedy, since for every iteration of the optimization algorithm the covariance matrix of the data is calculated and inverted. This procedure has to be repeated for all possible correlation models before the ML estimator determines the most likely one. The number of iterations and the accuracy of the MLE estimates depend crucially on the initial guesses of the inferred parameters.

Another method for the estimation of geometric anisotropy is the Bayesian updating method presented in [22]. In this approach, the posterior distribution of the anisotropy parameters is sampled using (i) a prior distribution and (ii) a noniterative importance-sampling Monte Carlo method with an approximating importance sampling density. The prior distribution employed in [22] was based on information from earlier years. The importance sampling density, used to approximate the likelihood of the anisotropic model, employs an adaptive mixture of multivariate normal distributions.

All of the aforementioned methods are not easily automated and either presume prior knowledge or are computationally expensive. We propose a non-parametric and non-iterative approach for the systematic detection of anisotropic correlations in spatial data based on two fundamental assumptions: (i) the SRF samples are drawn from jointly normal or lognormal distributions and (ii) the SRF be statistically homogeneous and differentiable. The proposed method determines the parameters of geometric anisotropy by analyzing the expectations of sample based spatial derivatives along orthogonal directions and is of order $O(N)$ (for data on regular lattices) or at most $O(N \log_2 N)$ (for scattered data). The principal directions are not assumed to be known in advance, but they are determined self-consistently from the data. Once the anisotropy parameters are determined, it becomes possible to transform into a coordinate system where the correlation function is isotropic, as illustrated in Fig. 1. The systematic anisotropy detection can constitute a valuable pre-processing stage allowing for isotropic SRF signal processing stages to follow.

The paper is organized as follows: In Section II we introduce preliminary concepts and necessary notation for modeling the anisotropy of scalar SRFs. In Section III we present the Covariance Hessian Identity (CHI) for jointly normal and differentiable SRFs. This leads to a nonlinear system of equations, which can be used to infer the anisotropy parameters from the data. We also show that the anisotropy parameters of an SRF with lognormal probability distribution are the same as those of its logarithm (which is normally distributed). In Section IV we derive the closed form solution of the nonlinear system of equations for the anisotropy parameters in two dimensions. Next, we apply the method to simulated data on a square grid in Section V and to simulated data at scattered locations in Section VI. Finally, we present our conclusions in Section VII.

II. Preliminary Notions and Nomenclature

To illustrate preliminary notions, we begin by assuming $\mathbf{s} \in \mathbb{R}^d$, $\mathbf{r}_{ij} = \mathbf{s}_i - \mathbf{s}_j$ the distance vector and $X(\mathbf{s})$ a scalar SRF on the probability space (Ω, F, P)[1]. The SRF represents a physical variable of interest, i.e., temperature, and the events constitute the observables of the SRF.

A. Ensemble Moments of First and Second Order

In general, a state (realization) of the SRF can be decomposed into a deterministic trend $m_{\mathrm{x}}(\mathbf{s})$, a fluctuation $\chi_\lambda(\mathbf{s})$, and a noise term $\eta(\mathbf{s})$ according to [18]

$$X(\mathbf{s}) = m_{\mathrm{x}}(\mathbf{s}) + \chi_\lambda(\mathbf{s}) + \eta(\mathbf{s}). \tag{3}$$

The trend represents large-scale variations of the field, corresponding to the ensemble average [16],

$$m_{\mathrm{x}}(\mathbf{s}) = \langle X(\mathbf{s}) \rangle \tag{4}$$

with $\langle \cdot \rangle$ denoting statistical expectation over the ensemble of states. The fluctuation models faster variations, which may appear as quasi-random changes, and have a cut-off scale λ determined by the spatial resolution of the experiment. Any fluctuations that occur at smaller scales can not be resolved, and they are incorporated with various other random perturbations in the noise term, which is modeled as a zero-mean white Gaussian SRF. In the following, we assume that the trend is removed from the initial SRF, and we will use the symbol $X(\mathbf{s})$ for the residual, i.e., the fluctuation and the noise components.

The spatial dependence of the SRF fluctuations is determined by means of the covariance function or the structure function (semivariogram). In many cases, one can assume that the fluctuation is a Wide Sense Stationary (WSS) SRF, or a SRF with WSS increments [23]. For WSS SRFs the structure function (semivariogram) $\gamma_{\mathrm{x}}(\mathbf{r})$, defined as $2\gamma_{\mathrm{x}}(\mathbf{r}) = \langle X(\mathbf{s}) - X(\mathbf{s}+\mathbf{r}) \rangle^2$, contains the same information as the covariance $c_{\mathrm{x}}(\mathbf{r}) = \langle X(\mathbf{s})\, X(\mathbf{s}+\mathbf{r}) \rangle$. In practice the structure function is often estimated instead of the covariance [15]. The structure function has a sill that is equal to the variance σ_{x}^2 of the SRF, and the following relation holds $\gamma_{\mathrm{x}}(\mathbf{r}) = \sigma_{\mathrm{x}}^2 - c_{\mathrm{x}}(\mathbf{r})$ [23], [24].

An SRF $X(\mathbf{s})$ has second-order stationary increments if the SRF $\Psi(\mathbf{s}) = X(\mathbf{s}) - X(\mathbf{s_0})$ is WSS. Such an SRF is also called intrinsic [25]. The category of intrinsic SRF's includes fractional Brownian motion

[1] Ω denotes the sample space (ensemble) that includes all the possible states (realizations) of the SRF, F is the set of all the observable events, $F \subset \Omega$, and $P(F) \in [0, 1]$ is the probability associated with each event.

[26]. In this case, the translation invariance of the covariance is lost. The structure function is still purely a function of the space lag, but it increases without bound. One can investigate anisotropy in the intrinsic SRFs by focusing on their increments, which are WSS.

The fluctuation SRF is isotropic if the covariance function depends purely on the magnitude of the distance vector, i.e., $c_{\mathrm{x}}(\mathbf{r}) = c_{\mathrm{x}}(\|\mathbf{r}\|)$, where $\|\mathbf{r}\|$ is the Euclidean norm of the vector $\mathbf{r}$. An anisotropic covariance function can be expressed as $c_{\mathrm{x}}(\mathbf{r}) = \Phi(\mathbf{r}^{\mathrm{T}}\mathbf{A}\mathbf{r})$, where $\mathbf{A}$ is a non-diagonal matrix. In the system of the principal axes $c'_{\mathrm{x}}(\mathbf{r}') = \Phi'(\mathbf{r}'^{\mathrm{T}}\mathbf{A}'\mathbf{r}')$, where $\mathbf{A}'$ is diagonal. The local rate of change of the covariance function along the principal directions is determined by respective correlation lengths ξ_j, $j = 1 \ldots, d$. The anisotropy parameters of the SRF are a set of orientation (rotation) angles, $\theta_j, j = 1, \ldots, d_m$ that determine the orientation of the principal axes in d dimensions, and the aspect vector $\mathbf{R}_1 = (1,\ R_{2(1)}, \ldots, R_{d(1)})$ which represents the ratios of the principal correlation lengths ξ_j, $j = 1, \ldots, d$ over ξ_1 (chosen arbitrarily). In $d = 2$ one rotation angle is sufficient ($d_m = 1$), while in $d = 3$ three Euler angles are necessary ($d_m = 3$).

B. On SRF Differentiability

The method that we propose for anisotropy detection relies on the estimation of expected values of the partial SRF derivatives. Let $\vec{\mathbf{e}}_i$ denote the unit vector in the spatial direction $i\,(i = 1, \ldots, d)$, and $\partial_i X(\mathbf{s}) = \lim_{\alpha \to 0} \left[X(\mathbf{s} + \alpha\vec{\mathbf{e}}_i) - X(\mathbf{s})\right] / \alpha$ denote the usual definition of the partial derivative (for a given state with index ω). The SRF X is *mean square differentiable* in $\mathcal{D} \subset \mathbb{R}^d$ if for every $\mathbf{s} \in \mathcal{D}$ and sequence $\{\mathbf{s}_n\}$ such that $\|\mathbf{s}_n - \mathbf{s}\| \to 0$ as $n \to \infty$, $\left\langle \left|\partial_i X(\mathbf{s}) - \partial_i X(\mathbf{s}_n)\right|^2 \right\rangle \to 0$, $\forall i = 1, \ldots, d$. In particular, if the partial derivatives $\partial^2 c_{\mathrm{x}}(\mathbf{r}) / \partial r_i^2$ exist for all orthogonal directions $i = 1, \ldots, d$ at $\mathbf{r} = \mathbf{0}$ (where $\mathbf{0}$ is the zero vector in d dimensions) the SRF is differentiable in the mean square sense for every $\mathbf{s} \in \mathbb{R}^d$. This paper focuses on multinormal and jointly lognormal SRFs. In these cases, mean square differentiability implies that the derivatives of the sample states exist almost surely [15].

Differentiable covariance models include the Gaussian and the Matérn class [27], [28], [29]. The latter involves a parameter ν that controls the SRF smoothness. A Matérn SRF is $\lceil \nu \rceil - 1$ times mean square differentiable, where $\lceil \nu \rceil$ is the lowest integer that is equal or higher than ν. The recently proposed Spartan SRFs [30]–[32], provide another class of differentiable random fields with controlled smoothness. Their differentiability is due to a band-limited spectrum that eliminates high frequency fluctuations [32].

For non-differentiable SRFs, one can distinguish three types based on the origin of the non-differentiability:

(a) Due to a discontinuity of the covariance at the origin: Additive white Gaussian noise (AWGN), which is present in most spatial data sets, creates a covariance discontinuity at zero lag (in geostatistics this is known as the nugget effect). This effect is manifested by a jump in the experimental structure function at zero lag. The robustness of the CHI method to noise is investigated in Section V. *(b) Due to discontinuity of the covariance derivatives at the origin:* Many WSS covariance models used in geostatistics are non-differentiable (e.g., the exponential, spherical, and logistic models). This effect is due to the change in the slope of the covariance at zero lag, and it is milder than that induced by noise. Selection of a non-differentiable model as optimal is based on parametric fits with the experimental structure function and does not actually imply that the SRF from which the data are sampled is inherently non-differentiable. Non-differentiability due to model selection is practically not a problem, since to date flexible differentiable models (as mentioned above) are available. *(c) Due to non-stationarity:* Non-stationary SRF's with WSS increments, such as the fractional Brownian motion (fBm), have non-differentiable covariance functions. In this case, one can apply the CHI to the SRF increments in orthogonal directions, which are differentiable[2].

III. The Covariance Hessian Identity

A. Background

The method that we propose is based on the *Covariance Hessian Identity* introduced by Swerling [33]. The CHI links the second order derivatives of the covariance function (in directions i and j), evaluated at zero distance, to the mean value of the product of the first order field derivatives (in the respective directions). We show that the covariance derivatives are explicitly expressed in terms of the correlation lengths along the principal axes and the orientation angles. On the other hand, the mean value of the field first order derivative products can be estimated by means of suitable sample averages assuming ergodic conditions. In this manner, the anisotropy parameters are expressed in terms of equations that involve sample averages. We solve the nonlinear equations explicitly in 2D, obtaining closed-form solutions for the anisotropy aspect ratio and the orientation angle. Unless otherwise noted we use the Einstein convention, which implies a summation over repeated indices on the same side of an equation.

[2]The CHI method focuses on short-range anisotropic correlations, while for fBm long-range properties are often the focus.

If the covariance of the SRF $X(\mathbf{s})$ is differentiable, the *Covariance Hessian Matrix* $\mathbf{H}$ (CHM) is defined as follows

$$H_{ij}(\mathbf{r}) = -\frac{\partial^2 c_{\mathrm{x}}(\mathbf{r})}{\partial r_i \, \partial r_j}, \quad i,j = 1,\ldots,d. \tag{5}$$

For normal, stationary SRFs the existence and finiteness of the second order derivatives of the covariance function at zero lag ensures the existence of the first order field derivatives [34]. We define the tensor that represents the *Gradient Kronecker Product* (GKP), $\mathbb{X}$, with elements X_{ij}, as follows,

$$X_{ij} \doteq \nabla X(\mathbf{s}) \otimes \nabla X^T(\mathbf{s}) = \partial_i X(\mathbf{s}) \, \partial_j X(\mathbf{s}). \tag{6}$$

We also define the expectation, $\mathbf{Q}$, of the GKP, $\mathbb{X}$, as

$$Q_{ij} \doteq \langle X_{ij} \rangle = \Big\langle \partial_i X(\mathbf{s}) \, \partial_j X(\mathbf{s}) \Big\rangle. \tag{7}$$

Swerling has proved in [33] the following equation, which we call the *Covariance Hessian Identity* (CHI):

$$\mathbf{Q} = \mathbf{H}(\mathbf{r})|_{\mathbf{r}=\mathbf{0}} \, . \tag{8}$$

B. CHI Equations

Let us assume that the covariance function is anisotropic. In the following, we derive a relation between the CHM in the original coordinate system, and the second-order derivative of the covariance function expressed in an isotropic system of dimensionless coordinates.

Theorem 1: *Let us assume that $X(\mathbf{s})$ is a differentiable, statistically homogeneous SRF with a correlation function that exhibits geometric (elliptic) anisotropy. The CHM $H^*_{ij}(\mathbf{0})$ in the principal system is a diagonal tensor given by*

$$H^*_{ij}(\mathbf{h})\big|_{\mathbf{h}=\mathbf{0}} = -\frac{\delta_{ij}}{d\,\xi_i \xi_j} \triangle \tilde{c}_{\mathrm{x}}(0), \tag{9}$$

where $\triangle \tilde{c}_{\mathrm{x}}(0) = \sum_{i=1}^{d} \partial^2 \tilde{c}_{\mathrm{x}}(0) \big/ \partial r_i^2$ is the Laplacian of the reduced isotropic covariance function $\tilde{c}_{\mathrm{x}}(\mathbf{h})$ evaluated at zero lag. The function $\tilde{c}_{\mathrm{x}}(\mathbf{h})$ is obtained by rotation and rescaling of the axes. The vector $\mathbf{h}$ is the lag in the rotated and rescaled system. The $\xi_i, \xi_j \, (i,j = 1,\ldots,d)$ are the correlation lengths in the respective principal directions and δ_{ij} is the Kronecker delta, defined by $\delta_{ij} = 1$, if $i = j$ and $\delta_{ij} = 0$ for $i \neq j$.

Proof: It is assumed that the covariance function in the initial coordinate system is $c_{\mathrm{x}}(\mathbf{r})$, where $\mathbf{r}$ is the lag vector in the initial system. Let the lag vector be $\mathbf{r}'$ in the principal system, which is obtained from

the initial by a rotation; in this system the covariance function becomes $c'_{\rm x}(\mathbf{r}')$, so that $c'_{\rm x}(\mathbf{r}') = c_{\rm x}(\mathbf{r})$. The rotation matrix is $\mathbf{U}(\boldsymbol{\theta})$, where $\boldsymbol{\theta}$ is the vector of the rotation angles in d dimensions. The elements of the rotation matrix are $U_{ij} = \partial r'_i/\partial r_j$. Using the chain rule of differentiation on $c_{\rm x}(\mathbf{r})$ we find

$$\frac{\partial c_{\rm x}(\mathbf{r})}{\partial r_i} = \frac{\partial c'_{\rm x}(\mathbf{r}')}{\partial r'_k}\frac{\partial r'_k}{\partial r_i} = \frac{\partial c'_{\rm x}(\mathbf{r}')}{\partial r'_k}U_{ki}. \tag{10}$$

An additional differentiation leads to

$$\frac{\partial^2 c_{\rm x}(\mathbf{r})}{\partial r_i\,\partial r_j} = \frac{\partial^2 c'_{\rm x}(\mathbf{r}')}{\partial r'_k\,\partial r'_l}\,\frac{\partial r'_l}{\partial r_j}\,U_{ki} = \frac{\partial^2 c'_{\rm x}(\mathbf{r}')}{\partial r'_k\,\partial r'_l}\,U_{ki}\,U_{lj}. \tag{11}$$

We define the reduced isotropic covariance $\tilde{c}_{\rm x}(\mathbf{h}) = c'_{\rm x}(\mathbf{r}')$, where $\mathbf{h}$ is the dimensionless lag vector in the rescaled coordinate system defined by $\mathbf{h} = (r'_1/\xi_1, \ldots, r'_d/\xi_d)$. The function $\tilde{c}_{\rm x}(\mathbf{h})$ is isotropic with respect to $\mathbf{h}$. In the principal coordinate system, the CHM is given by

$$\frac{\partial^2 c'_{\rm x}(\mathbf{r}')}{\partial r'_l\,\partial r'_m} \;=\; \frac{\partial}{\partial r'_m}\left[\frac{\partial \tilde{c}_{\rm x}(\mathbf{h})}{\partial h_l}\frac{\partial h_l}{\partial r'_l}\right] = \frac{\partial h_m}{\partial r'_m}\frac{\partial}{\partial h_m}\left[\frac{\partial \tilde{c}_{\rm x}(\mathbf{h})}{\partial h_l}\frac{1}{\xi_l}\right] = \frac{1}{\xi_m\xi_l}\frac{\partial^2 \tilde{c}_{\rm x}(\mathbf{h})}{\partial h_l \partial h_m}. \tag{12}$$

In order to evaluate the CHM at zero lag in the rescaled isotropic system, i.e.,

$$\tilde{H}_{ij}(\mathbf{h}) \doteq \frac{\partial^2 \tilde{c}_{\rm x}(\mathbf{h})}{\partial h_i \partial h_j} \tag{13}$$

one can use the spectral representation of the covariance. A more intuitive approach, which is valid in any d, is as follows: due to the isotropy of $\tilde{c}_{\rm x}(\mathbf{h})$, $\tilde{H}_{ij}(0)$ is a diagonal matrix with equal elements in all directions. Hence, it follows that $\tilde{H}_{ij}(0) = \delta_{ij}\,Tr(\tilde{\mathbf{H}})/d$, where Tr denotes the trace of a matrix. Based on (13), it follows that $Tr(\tilde{\mathbf{H}}) = \sum_{i=1}^{d} \tilde{H}_{ij}(0) \equiv \triangle\,\tilde{c}_{\rm x}(0)$. Thus, the CHM in the principal system is indeed given by (9).

Equation (9) provides a closed-form expression for the CHM in the principal system, which can be used to recast the CHI equations in a form that is explicit in the anisotropy parameters. Assuming that the correlation lengths are non-zero in all directions, it follows from (9) that

$$\frac{\xi_i}{\xi_j} = \sqrt{\frac{H^*_{jj}(\mathbf{0})}{H^*_{ii}(\mathbf{0})}} \quad i,j = 1,\ldots,d. \tag{14}$$

Thus, one can define $d-1$ aspect ratios for the correlation lengths as follows

$$R_{i(1)} \doteq \frac{\xi_1}{\xi_i} = \sqrt{\frac{H^*_{ii}(\mathbf{0})}{H^*_{11}(\mathbf{0})}} \quad i = 1,...,d. \tag{15}$$

The vector of correlation aspect ratios, anchored to the first principal direction, $\mathbf{R}_1 = (1, R_{2(1)}, \ldots, R_{d(1)})$ has $d-1$ independent components. Anchoring the aspect ratios on the length ξ_1 is arbitrary. Based on

(9) and (11), the CHM in the original system is given by

$$H_{ij}(\mathbf{0}) \;=\; U_{li}(\boldsymbol{\theta})\, U_{lj}(\boldsymbol{\theta}) H^{*}_{ll}(\mathbf{0}) = -\,\mathbf{U_{li}}(\boldsymbol{\theta})\,\mathbf{U_{lj}}(\boldsymbol{\theta})\,\frac{\mathbf{R^{2}_{l(1)}}}{\mathbf{d}\,\boldsymbol{\xi}_{\mathbf{1}}^{\mathbf{2}}} \triangle\, \tilde{\mathbf{c}}_{\mathrm{x}}(\mathbf{0}). \tag{16}$$

Finally, in light of equation (8), the system of CHI equations is expressed as follows

$$Q_{ij} = -\frac{R^{2}_{l(1)}}{\xi_1^{2} d}\, \triangle\, \tilde{c}_{\mathrm{x}}(0)\, U_{li}(\boldsymbol{\theta})\, U_{lj}(\boldsymbol{\theta}) \quad i,j = 1,\ldots,d. \tag{17}$$

The expressions (17) form a system of nonlinear equations that involve the components of the rotation matrix $\mathbf{U}$ and the aspect ratio vector $\mathbf{R}_1$ as unknown variables, which should be determined from the Q_{ij}. Taking into account the symmetry of the matrix Q_{ij}, the number of independent equations is $d(d+1)/2$. The correlation length ξ_1 and the Laplacian of the covariance $\triangle\, \tilde{c}_{\mathrm{x}}(0)$ are also unknown at this stage. However, as shown below, they are not required for determining the anisotropy parameters. Instead, they can be estimated at a later stage following the isotropic transformation.

C. Variance of the GKP Tensor

Let us define the shorthand notation for the partial field derivatives $\partial_i X(\mathbf{s}) = \partial X(\mathbf{s})/\partial s_i$, and for the GKP $X_{ij} = \partial_i X(\mathbf{s})\, \partial_j X(\mathbf{s})$. The CHI gives the mean value of the GKP, according to (8). The width of the GKP distribution is measured by the square root of $\mathrm{Var}(X_{ij})$. The latter can be evaluated explicitly if $X(\mathbf{s})$ is a Gaussian SRF.

Theorem 2 *For a Gaussian, statistically homogeneous SRF $X(\mathbf{s})$, the variance of the GKP tensor is given by the following expression*[3]:

$$\mathrm{Var}(X_{ij}) = H_{ii}(\mathbf{0}) H_{jj}(\mathbf{0}) + H^{2}_{ij}(\mathbf{0}), \quad i,j = 1,\ldots,d. \tag{18}$$

Proof: By definition

$$\mathrm{Var}(X_{ij}) \;=\; \left\langle X^{2}_{ij} \right\rangle - \left\langle X_{ij} \right\rangle^{2} = \left\langle [\partial_i X(\mathbf{s})]^{2} [\partial_j X(\mathbf{s})]^{2} \right\rangle - \left\langle \partial_i X(\mathbf{s})\, \partial_j X(\mathbf{s}) \right\rangle^{2} \tag{19}$$

Note that $\langle \partial_i X(\mathbf{s})\, \partial_j X(\mathbf{s}) \rangle = H_{ij}(\mathbf{0})$. Since $\partial_i X(\mathbf{s})$, $i = 1,\ldots,d$ are zero-mean, normally distributed SRFs, we can apply the product-of-pairs decomposition property of central Gaussian moments to expand the term $\left\langle [\partial_i X(\mathbf{s})]^{2}\, [\partial_j\, X(\mathbf{s})]^{2} \right\rangle$ as follows:

$$\left\langle [\partial_i X(\mathbf{s})]^{2}\, [\partial_j X(\mathbf{s})]^{2} \right\rangle \;=\; 2 \left\langle \partial_i X(\mathbf{s})\, \partial_j X(\mathbf{s}) \right\rangle^{2} + \left\langle [\partial_i X(\mathbf{s})]^{2} \right\rangle \left\langle [\partial_j X(\mathbf{s})]^{2} \right\rangle \tag{20}$$

[3]Summation is not implied over the indices i and j here.

The first term on the right hand side is equal to $2\,H_{ij}^2(\mathbf{0})$, and cancels $\langle X_{ij}\rangle^2$ thus completing the proof.

For normal, stationary SRF's, the GKP is a tensorial SRF, and the normalized diagonal elements $X_{ii}/\sqrt{2}H_{ii}(\mathbf{0})$ follow the χ^2 distribution with one degree of freedom. Based on the above, the coefficient of variation (COV) for the non-zero GKP elements becomes

$$\text{COV}(X_{ij}) = \sqrt{\frac{H_{ii}(\mathbf{0})H_{jj}(\mathbf{0})}{H_{ij}^2(\mathbf{0})} + 1}. \tag{21}$$

The $\text{COV}(X_{ij})$ is always greater than unity, indicating high variability of the GKP.

D. Anisotropy of Lognormal SRF

Let us consider a jointly lognormal WSS SRF $X(\mathbf{s})$. [4] Then, the SRF $Y(\mathbf{s}) = \log X(\mathbf{s})$ follows the joint normal (multinormal) distribution. The mean and the covariance functions of the primary and the logarithmic SRF are related by means of the equations:

$$m_{\text{x}} = e^{m_{\text{y}}+c_{\text{y}}(0)/2}, \tag{22}$$

$$c_{\text{x}}(\mathbf{r}) = m_{\text{x}}^2\ \left[e^{c_{\text{y}}(\mathbf{r})} - 1\right] \tag{23}$$

where

$$m_{\text{y}} = \langle Y(\mathbf{s})\rangle, \quad m_{\text{x}} = \langle X(\mathbf{s})\rangle = \langle e^{Y(\mathbf{s})}\rangle. \tag{24}$$

Equations (22) and (23) follow both from the *cumulant expansion* [35], [36]: Any random function $Z(\mathbf{s})$ can be expressed as follows:

$$\langle e^{Z(\mathbf{s})}\rangle = \exp\left[\sum\nolimits_{m=1}^{\infty} C_m[Z(\mathbf{s})]\Big/ m!\right], \tag{25}$$

where $C_m[Z(\mathbf{s})]$ is the cumulant of order m of the SRF $Z(\mathbf{s})$. If $Z(\mathbf{s})$ is a Gaussian SRF, only the first two cumulants are non-zero, i.e. $C_1[Z(\mathbf{s})] = \langle Z(\mathbf{s})\rangle$, $C_2[Z(\mathbf{s})] = \langle Z^2(\mathbf{s})\rangle - \langle Z(\mathbf{s})\rangle^2$. More specifically, for m_{x} it holds that $Z(\mathbf{s}) = Y(\mathbf{s})$, and thus $C_1[Y(\mathbf{s})] = m_{\text{y}}$, $C_2[Y(\mathbf{s})] = c_{\text{y}}(0)$, which in view of (25) leads to (22). The definition of the covariance function is $c_{\text{x}}(\mathbf{r}) = \langle X(\mathbf{s})\,X(\mathbf{s}+\mathbf{r})\rangle - m_{\text{x}}^2$, or equivalently $c_{\text{x}}(\mathbf{r}) = \langle e^{Y(\mathbf{s})+Y(\mathbf{s}+\mathbf{r})}\rangle - m_{\text{x}}^2$. In this case $Z(\mathbf{s}) = Y(\mathbf{s}) + Y(\mathbf{s}+\mathbf{r})$ and the respective cumulants are $C_1[Z(\mathbf{s})] = 2m_{\text{y}}$, and $C_2[Z(\mathbf{s})] = 2c_{\text{y}}(0) + 2c_{\text{y}}(\mathbf{r})$, thus leading to (23).

[4] In this section, we consider that $m_{\text{x}} \neq 0$, since the skewness of the lognormal distribution does not confer any advantages on centered fluctuations.

Theorem 3: *Let $Y(\mathbf{s})$ be a lognormal, WSS and differentiable SRF, and $X(\mathbf{s}) = \exp[Y(\mathbf{s})]$ denote its lognormal counterpart. The CHMs of the two fields are related by means of the equation $\mathbf{H}_{\mathrm{x}}(\mathbf{0}) = \rho\,\mathbf{H}_{\mathrm{y}}(\mathbf{0})$, where $\rho \in \mathbb{R}_+$.*

Proof: For a WSS, lognormal SRF, $m_{\mathrm{x}} \in \mathbb{R}_+$ is constant. Hence, if the covariance function $c_{\mathrm{y}}(\mathbf{r})$ is differentiable, so is the covariance $c_{\mathrm{x}}(\mathbf{r})$. It follows from (23) that

$$\frac{\partial^2 c_{\mathrm{x}}(\mathbf{r})}{\partial r_i \partial r_j} \;=\; m_{\mathrm{x}}^2\,\frac{\partial}{\partial r_i}\left[e^{c_{\mathrm{y}}(\mathbf{r})}\,\frac{\partial c_{\mathrm{y}}(\mathbf{r})}{\partial r_j}\right] = m_{\mathrm{x}}^2\,e^{c_{\mathrm{y}}(\mathbf{r})}\left[\frac{\partial c_{\mathrm{y}}(\mathbf{r})}{\partial r_i}\frac{\partial c_{\mathrm{y}}(\mathbf{r})}{\partial r_j} + \frac{\partial^2 c_{\mathrm{y}}(\mathbf{r})}{\partial r_i \partial r_j}\right] \tag{26}$$

Since $c_{\mathrm{y}}(\mathbf{r})$ is differentiable, its first derivative vanishes at zero lag, where the maximum is attained $\partial c_{\mathrm{y}}(\mathbf{r})/\partial r_j|_{\mathbf{r}=\mathbf{0}} = 0$. Hence, from (26) one obtains the equation

$$\mathbf{H}_{\mathrm{x}}(\mathbf{0}) = m_{\mathrm{x}}^2\,e^{c_{\mathrm{y}}(\mathbf{0})}\,\mathbf{H}_{\mathrm{y}}(\mathbf{0}). \tag{27}$$

Equation (27) proves the theorem using $\rho = m_{\mathrm{x}}^2\,\exp(\sigma_{\mathrm{y}}^2)$, where $\sigma_{\mathrm{y}}^2 = c_{\mathrm{y}}(\mathbf{0})$.

Theorem 4: *The aspect ratio vector $\mathbf{R}_{\mathrm{x}(1)}$ is identical with the vector $\mathbf{R}_{\mathrm{y}(1)}$.*

Proof: The theorem follows directly from (15) and (27).

Theorems 3 and 4 ensure that the anisotropy parameters of the SRF $X(\mathbf{s})$ are identical to those of the SRF $Y(\mathbf{s})$. This implies that if the initial SRF $X(\mathbf{s})$ follows the lognormal distribution, its anisotropy parameters can be determined from those of the normal SRF $Y(\mathbf{s})$.

IV. Estimation of Anisotropy Parameters

First we consider the estimation of the anisotropy parameters in any dimension $d \geq 2$. This requires solving the nonlinear equations (17), which is hindered by two factors: First, the right hand side of (17) involves the parameters ξ_1 and $\triangle\,\tilde{c}_{\mathrm{x}}(0)$, which are not known in advance. Secondly, the expected GKP $\mathbf{Q}$ is estimated from the sample by means of $\hat{\mathbf{Q}}$, which is most likely contaminated by noise and other estimation errors (e.g., discretization effects).

A. Estimation of Anisotropy Parameters in Any Dimension

We assume without loss of generality that $\hat{Q}_{11}$ is the GKP element with the maximum value. Then, we cast the system of equations in terms of ratios of the CHM elements, $H_{ij}(\mathbf{0})/H_{11}(\mathbf{0})$ – in which the dependence on ξ_1 and $\triangle\,\tilde{c}_{\mathrm{x}}(0)$ is eliminated – and the respective sample slope tensor ratios $\hat{Q}_{ij}/\hat{Q}_{11}$.

To account for possible sampling or modeling errors, we allow for the possibility of an imperfect match between the CHM ratios and the respective sample slope tensor ratios by introducing the residual

$$\hat{Q}_{ij}/\hat{Q}_{11} - H_{ij}(\mathbf{0})/H_{11}(\mathbf{0}) = \varepsilon_{ij}(\boldsymbol{\theta}; \mathbf{R}_1). \tag{28}$$

The anisotropy parameters can then be determined by minimizing numerically the following nonlinear functional $\Lambda(\boldsymbol{\theta}; \mathbf{R}_1) = \sum_{i=2}^{d} \sum_{j \le i}^{d} \varepsilon_{ij}^2(\boldsymbol{\theta}; \mathbf{R}_1)$, i.e.,

$$\{\hat{\boldsymbol{\theta}}; \hat{\mathbf{R}}_1\} = \underset{\boldsymbol{\theta}; \mathbf{R}_1}{\arg\min}\, \Lambda(\boldsymbol{\theta}; \mathbf{R}_1) \tag{29}$$

The $\Lambda(\boldsymbol{\theta}; \mathbf{R}_1)$ is non-negative, and its global minimum (zero), is obtained when the residuals vanish. This numerical approach is general and can be useful if an explicit solution for the anisotropy parameters is not available (e.g., for $d > 2$.)

An alternative numerical approach is based on the diagonalization of the CHM, where the latter is estimated from the sample by means of $\hat{\mathbf{Q}}$. Then, the $d-1$ square roots of the eigenvalue ratios (e.g., with respect to the first eigenvalue) correspond the anisotropic ratios, while the orientation angles can be obtained from the elements of the diagonalizing transformation matrix.

B. Estimation of Anisotropy Parameters in Two Dimensions

Based on (17), we derive expressions for the CHI equations that are valid in two dimensions. We then solve the equations explicitly to obtain estimates of the anisotropy parameters. The rotation matrix $\mathbf{U}$ is given by $U_{11} = U_{22} = \cos\theta$ and $U_{12} = -U_{21} = \sin\theta$. The anisotropy ratio vector is $\mathbf{R}_1 = (1,\ R_{2(1)})$. The equations (17) are then expressed as follows

$$Q_{11} = \frac{\sigma_{\mathrm{x}}^2 \zeta^2}{\xi_1^2} \left(\cos^2\theta + R_{2(1)}^2 \sin^2\theta \right) \tag{30}$$

$$Q_{22} = \frac{\sigma_{\mathrm{x}}^2 \zeta^2}{\xi_1^2} \left(R_{2(1)}^2 \cos^2\theta + \sin^2\theta \right) \tag{31}$$

$$Q_{12} = Q_{21} = \frac{\sigma_{\mathrm{x}}^2 \zeta^2}{\xi_1^2} \left[\sin\theta \cos\theta (1 - R_{2(1)}^2) \right] \tag{32}$$

where $\zeta = \frac{1}{2} \triangle \tilde{c}_{\mathrm{x}}(0)$ is an unknown parameter. The orientation of the principal axes is determined by θ and the aspect ratio $R_{2(1)}$. By dividing both sides of (31)-(32) by the terms on the respective sides of (30) we obtain the following set of equations

$$q_{\mathrm{diag}} \doteq \frac{Q_{22}}{Q_{11}} = \frac{R_{2(1)}^2 + \tan^2\theta}{1 + R_{2(1)}^2 \tan^2\theta}, \tag{33}$$

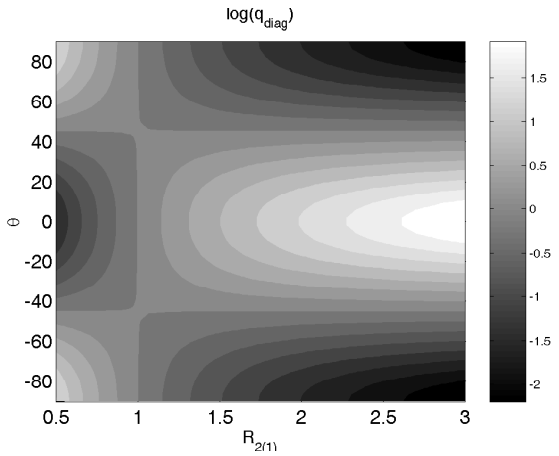

Fig. 2. Dependence of the ratio $q_{\rm diag}$ on $R_{2(1)}$ and θ. The color scale corresponds to $\log(q_{\rm diag})$ for enhanced level resolution.

$$q_{\rm off} \doteq \frac{Q_{12}}{Q_{11}} = \frac{\tan\theta(1-R_{2(1)}^2)}{1+R_{2(1)}^2\tan^2\theta}, \tag{34}$$

where $\theta \in [-\pi/2, \pi/2]$ and $R_{2(1)} \in [0, \infty)$.

It is straightforward to show that (33) and (34) are invariant under the transformation $\tan\theta \rightarrow -1/\tan\theta$ and $R_{2(1)} \rightarrow 1/R_{2(1)}$. Equivalently, this means that $q_{\rm diag}$ and $q_{\rm off}$ are invariant under the transformations $\theta \rightarrow \theta \pm \pi/2$ and $R_{2(1)} \rightarrow 1/R_{2(1)}$. Hence, the non-degenerate parameter space is defined by $\theta \in [-\pi/2, \pi/2)$ and $R_{2(1)} \in [1, \infty)$. The degeneracy reflects the fact that an elliptical curve (i.e., an isolevel covariance contour) is equivalent to an ellipse rotated by 90^o, with the orientation of the minor and major axes interchanged. The dependence of $q_{\rm diag}$ and $q_{\rm off}$ on the anisotropy parameters is shown in Figs. 2 and 3. Note that $q_{\rm diag}$ is symmetrical, while $q_{\rm off}$ is antisymmetrical around $\theta = 0$.

The anisotropy parameters are the roots of the nonlinear system of equations (33) and (34), and they can be expressed explicitly as follows

$$\theta = \frac{1}{2}\tan^{-1}\left(\frac{2q_{\rm off}}{1-q_{\rm diag}}\right), \tag{35}$$

$$R_{2(1)} = \sqrt{1+\frac{1-q_{\rm diag}}{q_{\rm diag}-(1+q_{\rm diag})\cos^2\theta}}. \tag{36}$$

The solution for $R_{2(1)}$ is valid for all $q_{\rm diag}$ and $q_{\rm off}$. The solution for θ is well defined provided that $q_{\rm diag} \neq 1$ and $q_{\rm off} \neq 0$. For $q_{\rm diag} = 1$ and $q_{\rm off} = 0$ one obtains $R_{2(1)} = 1$ but θ is indeterminate. This result is meaningful since the case $q_{\rm diag} = 1$, $q_{\rm off} = 0$ corresponds to isotropic dependence, and thus no preferred angle. This is obvious from equations (33) and (34), which are satisfied by any θ if $R_{2(1)} = 1$.

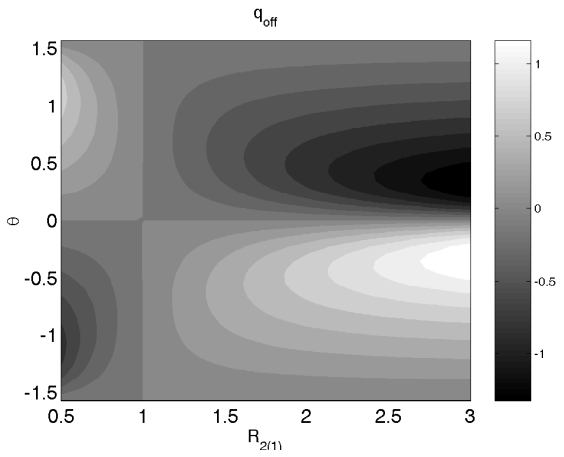

Fig. 3. Dependence of the ratio $q_{\rm off}$ on $R_{2(1)}$ and θ.

V. Numerical Results on Regular Supports

Below we focus on the estimation of the anisotropy parameters from data distributed on rectangular grids with dimensions (number of nodes per side) $N = L_1 \times L_2$. In practice, the true (unknown) values of $q_{\rm diag}$ and $q_{\rm off}$ are approximated by the sample estimates $\hat{q}_{\rm diag}$ and $\hat{q}_{\rm off}$. In view of (6) and (7), the latter require estimates of the GKP $\hat{\mathbb{X}}$, which involves the calculation of numerical derivatives. The estimates $\hat{\theta}$ and $\hat{R}_{2(1)}$ for the anisotropy parameters are then obtained from (35) and (36) by replacing $q_{\rm diag}$ and $q_{\rm off}$ with their sample estimates. Using either centered differences or Savitzky Golay (SG) differentiation filters [37], [38] to estimate the partial derivatives, the algorithmic complexity of the method is $O(N)$.

We use the discrete spectral method (Fourier filtering method) [39], [40] to generate anisotropic random fields with Gaussian, and Matérn $(\nu = 2, 3, 5)$ covariance models. On regular grids this method takes advantage of the computational efficiency of the Fast Fourier Transform (FFT). Fig. 4 shows realizations of these fields with anisotropic ratio $R_{2(1)} = 2$ and rotation angle $\theta = 20^o$ on a regular square lattice with $L = 512$ nodes per side and the smallest correlation length, $\xi_{\rm min} = \min(\xi_1, \xi_2)$, given by $\xi_{\rm min} = \{4, 2, 1.5, 1\}$ respectively. The derivations of the anisotropic power spectral density for SRFs with Gaussian and Matérn correlations are given in the Appendix. Unless otherwise noted, in this section we use square grids with $L = 512$ sites per side and the covariance models with the correlation lengths specified above. In addition, the mean estimates and the errors presented are based on an ensemble of 100 independent realizations from a multivariate normal distribution with the specified covariance model.

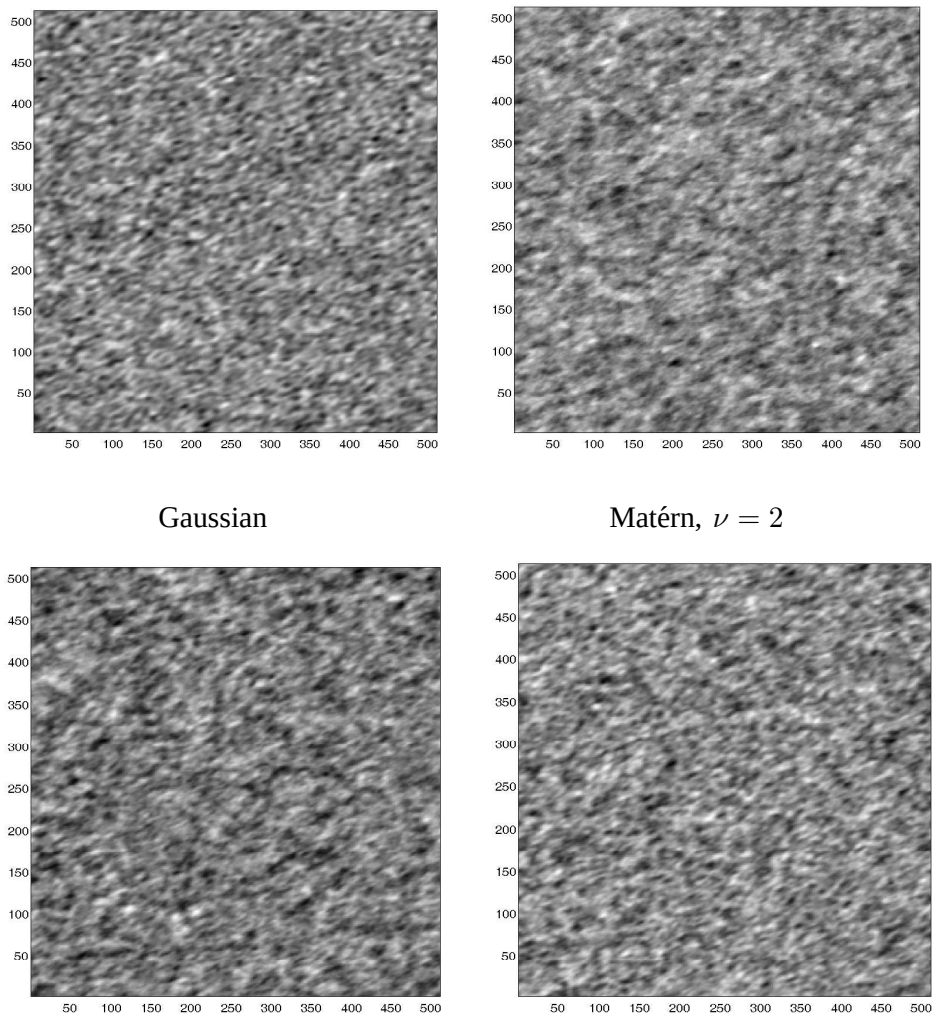

Fig. 4. Realizations of SRF with correlations lengths $\xi_{\min} = \{4, 2, 1.5, 1\}$ for respectively the Gaussian, Matérn $\nu = 2$, Matérn $\nu = 3$, and Matérn $\nu = 5$ covariance models and anisotropy parameters $R_{2(1)} = 2$ and $\theta = 20^o$.

We estimate the anisotropy parameters for SRF realizations that correspond to the following cases: $R_{2(1)} = \{0.3, 0.5, 0.7, 1, 1.5, 2, 3, 5\}$, $\theta = 20^o$ and $R_{2(1)} = 2$, $\theta = \{-45, -30, -15, 0.1, 15, 30, 45\}^o$. The mean values of the estimates are presented in Tables I and II for $\hat{R}_{2(1)}$ and $\hat{\theta}$ respectively. First, the angle $\hat{\theta}$ is estimated from (35) using the sample based estimates $\hat{q}_{\text{diag}}$, $\hat{q}_{\text{off}}$; the angle $\hat{\theta}$ is used in (36) to estimate the anisotropic ratio. In both cases, the errors are reduced as the degree of differentiability

TABLE I

TRUE (FIRST COLUMN) AND ESTIMATED VALUES OF $R_{2(1)}$ FOR VARIOUS COVARIANCE MODELS

$R_{2(1)}$	Gaussian	Matérn $\nu = 2$	Matérn $\nu = 3$	Matérn $\nu = 5$
0.3	0.3104	0.3259	0.3161	0.3136
0.5	0.5135	0.5290	0.5195	0.5172
0.7	0.7125	0.7242	0.7172	0.7156
1	0.9994	0.9991	0.9991	0.9993
1.5	1.4693	1.4416	1.4576	1.4618
2	1.9434	1.8851	1.9195	1.9289
3	2.8920	2.7612	2.8394	2.8625
5	4.7624	4.4200	4.6220	4.6916

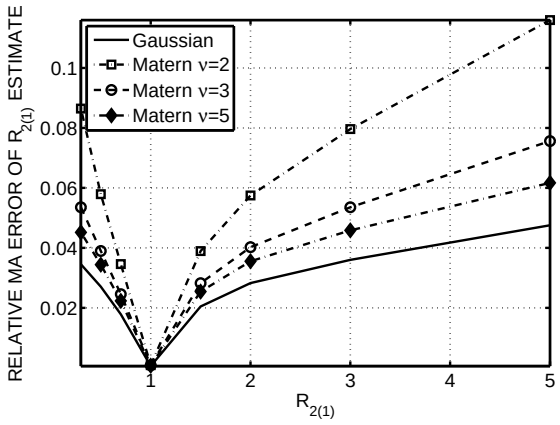

Fig. 5. Relative mean absolute error of $\hat{R}_{2(1)}$ versus the true $R_{2(1)}$ for SRFs with Gaussian and Matérn ($\nu = 2, 3, 5$) correlations ($\theta = 20^o$, $\xi_{\min} = 4, 2, 1.5, 1$ respectively) on a square 512×512 grid.

increases. In Figs. 5 and 6 we illustrate the relative Mean Absolute (RMA) error for $\hat{R}_{2(1)}$ and the MA error for $\hat{\theta}$. Smoother covariances lead to smaller errors. The RMA in all the cases studied is less than 10% for $R_{2(1)} \in [0.3, 3]$. The MA error of $\hat{\theta}$ is low (less than 2^o for all cases studied); it appears to be almost symmetrical around $\theta = 0^o$, where, as well as at $\theta = \pm 45^o$, it reaches a minimum. In the case of Gaussian correlation, the MA error is lower than 1^o.

The curves in Fig. 5 are asymmetric around $R_{2(1)} = 1$, due to $\theta \neq 45^o$. Note that $R_{2(1)}$ is un-

TABLE II

TRUE (FIRST COLUMN) AND ESTIMATED VALUES OF θ FOR VARIOUS COVARIANCE MODELS

θ in ^{o}	Gaussian	Matérn $\nu = 2$	Matérn $\nu = 3$	Matérn $\nu = 5$
-45	-45.0002	-44.9999	-45.0001	-44.9999
-30	-30.9210	-31.6937	-31.2442	-31.1469
-15	-15.9282	-16.7110	-16.2460	-16.1569
0.1	0.1576	0.1911	0.1837	0.1693
10	10.7704	11.3942	11.0419	10.9560
20	21.1279	22.0469	21.5233	21.4004
30	30.9658	31.7536	31.3953	31.1985
45	45.0000	45.0001	45.0000	44.9996

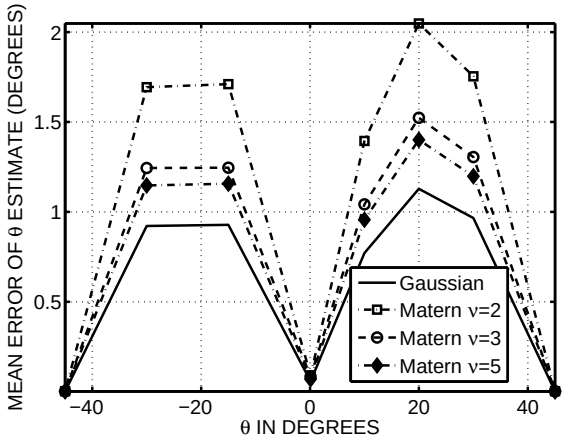

Fig. 6. Mean absolute error of $\hat{\theta}$ versus the true θ for SRFs with Gaussian and Matérn ($\nu = 2, 3, 5$) correlations ($R_{2(1)} = 2$, $\xi_{\min} = 4, 2, 1.5, 1$ respectively) on a square 512×512 grid.

derestimated (overestimated) as its true value increases above (drops below) unity (cf. Table I). This "compression" of the estimated $R_{2(1)}$ is due to the limited grid resolution and sample size. This effect highlights the contradicting requirements for keeping the grid size large as compared to the correlation lengths in order to ensure ergodicity [23], while keeping all correlation lengths larger than the grid step in order to resolve the correlation. These two requirements can be summarized as $\frac{\min(\xi_i,\xi_j)}{\max(a_i,a_j)} \gg 1$

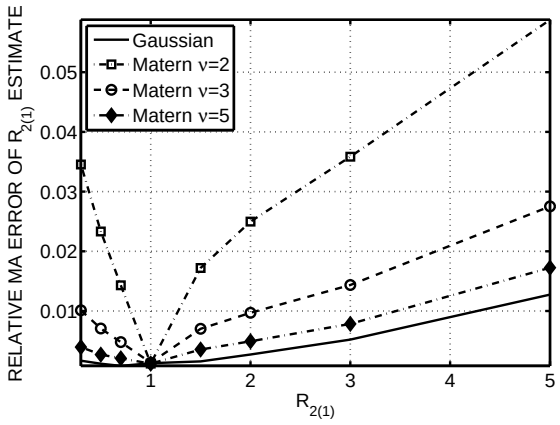

Fig. 7. Relative mean absolute error of $\hat{R}_{2(1)}$ versus the true $R_{2(1)}$ on a 512×512 square grid. SG filtering (p=5, W=7) is used.

and $\frac{\min(L_i,L_j)}{\max(\xi_i,\xi_j)} \gg 1$. When $R_{2(1)} \gg 1$ or $R_{2(1)} \ll 1$, it is not possible to simultaneously satisfy both requirements, unless the sample size is substantially increased.

The estimates $\hat{R}_{2(1)}$ and $\hat{\theta}$ are quite accurate on the grid. It is also observed that the accuracy increases for smoother fields that admit higher order derivatives. If SG filtering (with polynomial order $p = 5$ and window size $2W + 1$, $W = 7$) is employed the error is considerably reduced as shown in Fig. 7. The above numerical results demonstrate the efficiency of the proposed method for anisotropy parameters identification in the case of differentiable normal fields on a regular grid. The average CPU time for the estimation of the anisotropic parameters on 512×512 square grids is $t_{\rm CPU} \approx 0.06$ sec using centered differences and $t_{\rm CPU} \approx 2.7$ sec using SG differentiation.[5]

In Fig. 8 we investigate the impact of the grid size on $\hat{R}_{2(1)}$. We use Gaussian, and Matérn ($\nu = 2, 3, 5$) anisotropic covariance models with $R_{2(1)} = 2$ on square grids of length $L = 2^p$, $p = 6, \ldots, 11$ per side. The RMA of $\hat{R}_{2(1)}$ declines asymptotically as L increases and the ergodic condition tends to be satisfied.

Finally, we investigate the effect of zero-mean AWGN on an SRF field. We characterize the intensity of the noise by the signal to noise ratio (SNR), i.e., the ratio of the SRF variance to the noise variance. We generate SRF realizations with $\xi_{\min} = 4$, to which we add zero-mean AWGN so that $\mathrm{SNR} =$

[5]The CPU times reported are averages over 100 realizations. They were obtained on a desktop with an Intel 6700 2.66 GHz processor, running Matlab®7.3 under the Windows XP OS (service pack 2). The simulations use less than 600 MB from the 3.24 GB RAM.

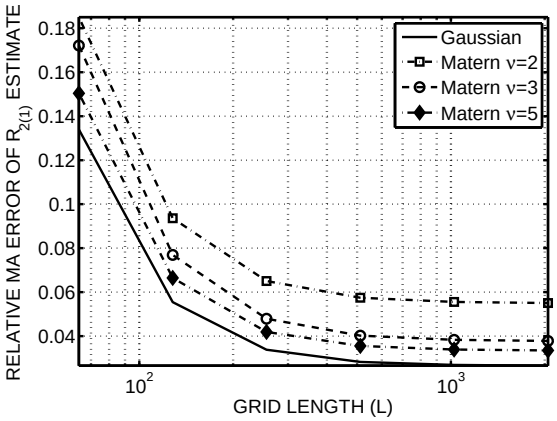

Fig. 8. Relative mean absolute error of $\hat{R}_{2(1)}$ versus the grid length per side L for SRFs with Gaussian and Matérn ($\nu = 2, 3, 5$) correlations ($\theta = 20^o$, $\xi_{\min} = 4, 2, 1.5, 1$ respectively). The scale of the horizontal axis is logarithmic.

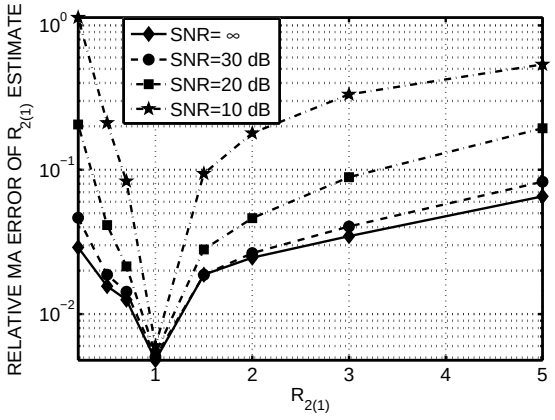

Fig. 9. Relative mean absolute error of $\hat{R}_{2(1)}$ versus the true $R_{2(1)}$ for different levels of zero-mean AWGN added to an SRF with Gaussian correlations ($\xi_{\min} = 10$, $\theta = 20^o$) on a 512×512 square grid. The SG differentiator (p=5, W=11) is used. The scale of the vertical axis is logarithmic.

$\{\inf, 30, 20, 10\}$ dB. We use SG filtering (with polynomial order $p = 5$ and window size $2W+1$, $W = 11$) to reduce the impact of noise on the derivatives estimation. The results are presented in Fig. 9. For $\mathrm{SNR} > 30$ dB, the anisotropy parameters are accurately estimated (RMA less than 10%) in the entire range of $\hat{R}_{2(1)}$ considered, while for $30 > \mathrm{SNR} > 20$ dB the RMA of the extreme anisotropic ratios does not exceed 20%. For lower SNR values the range of accurately estimated anisotropic ratios is progressively reduced.

VI. NUMERICAL RESULTS ON IRREGULAR SUPPORTS

In a number of applications, particularly in the field of geostatistics, the SRF is sampled on an irregular grid. As a result, the evaluation of the field's partial derivatives is not straightforward. In such cases, the results presented in this section indicate that it is preferable to interpolate the data on a regular lattice before estimating the anisotropy parameters. The choice of the above interpolator relies on the following desirable characteristics: (i) the interpolator should be kept simple in the framework of a pre-processing stage and (ii) the induced inaccuracies in estimating field partial derivatives should not introduce "too large" deviations in anisotropy parameters estimation. Another option that we investigate is the direct application (i.e., omitting grid interpolation) of SG filters on scattered data (henceforth, scattered SG).

In the following, we compare the performance of simple interpolators, namely linear, cubic spline, nearest neighbors and the biharmonic spline interpolator [41] followed by a differentiator (either employing centered differences or an SG filter) to the evaluation of the partial derivatives using scattered SG filters. The biharmonic spline interpolator relies on the minimum curvature principle, while the interpolated field is a linear combination of Green's functions centered on each data point. The weights of the relevant Green's functions are obtained by solving a linear system of equations, $O(N)$. The algorithm is unstable in the case of very proximate sampling points. This problem can be overcome by introducing a critical distance (typically a percentile of the point-to-point distance distribution), below which sampling points are considered identical and the respective field values are averaged. The numerical complexity increases to $O(N \log_2 N)$, while the use of scattered SG filters requires $O(N^2)$ operations. If the numerical complexity is not an issue, more elaborate interpolators based on kriging or Wiener filtering [27] could be employed.

We use the Fourier filtering method to generate SRFs with Gaussian correlations for the following combinations of anisotropy parameters: (i) $R_{2(1)} = \{0.3, 0.5, 0.7, 1, 1.5, 2, 3\}$ and $\theta = 20^o$ and (ii) $R_{2(1)} = 2$ and $\theta = \{-45, -30, -15, 0.1, 15, 30, 45\}^o$ on square lattices of size $L \times L$, where $L = 256$. These fields are then randomly sampled at $N_s = 4096$ points and interpolated over a regular lattice of size $L_r \times L_r$, where $L_r = 64$ using linear, cubic spline, nearest neighbor and biharmonic spline interpolation. The relative RMA error in estimating $R_{2(1)}$ and the MA error in estimating θ using a centered differences differentiator are presented in Figs. 10 and 11 respectively. The plots confirm that the biharmonic spline interpolator outperforms the other three interpolators. The RMA error of $\hat{R}_{2(1)}$ is higher than the respective

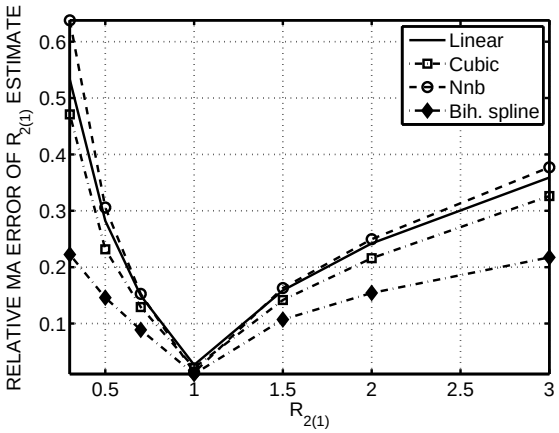

Fig. 10. Relative mean absolute error of $\hat{R}_{2(1)}$ versus the true $R_{2(1)}$ using four different interpolation algorithms (linear, cubic, nearest neighbor and biharmonic spline). The samples contain 4096 values of an SRF with Gaussian correlations ($\xi_{\min} = 8$, $\theta = 20^o$), randomly selected from a 256×256 grid.

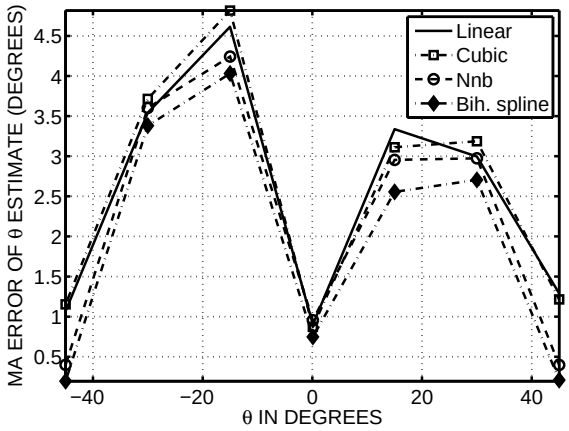

Fig. 11. Mean absolute error of $\hat{\theta}$ (in degrees) versus the true θ using four different interpolation algorithms (linear, cubic, nearest neighbor and biharmonic spline). The samples contain 4096 values of an SRF with Gaussian correlations ($\xi_{\min} = 8$, $R_{2(1)} = 2$), randomly selected from a 256×256 grid.

values on the grid (cf. Fig. 5). This loss in accuracy is inevitable due to the irregular sampling point distribution and, most significantly, the considerable loss of information due to sampling only $\approx 6\%$ of the field values. The estimation of the rotation angle remains quite satisfactory. We also compare the $\hat{R}_{2(1)}$ estimates for Gaussian and Matérn SRFs using the biharmonic spline interpolator. The RMA error of $\hat{R}_{2(1)}$ is shown in Fig. 12. In agreement with the regular grid case, the estimates are more accurate for SRFs that admit higher order derivatives.

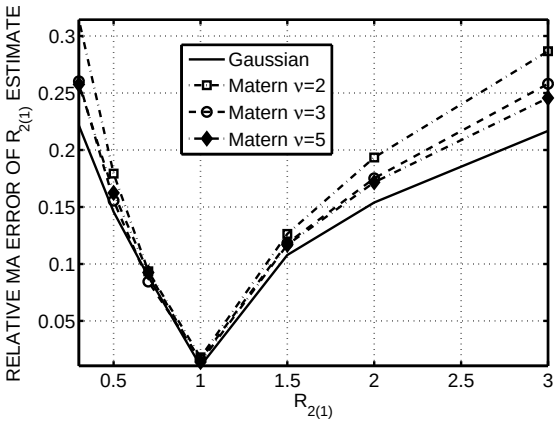

Fig. 12. Relative mean absolute error of $\hat{R}_{2(1)}$ versus the true $R_{2(1)}$ for SRFs with Gaussian and Matérn ($\nu = 2, 3, 5$) correlations ($\theta = 20^o$, $\xi_{\min} = 8, 4, 3, 2$ respectively). The samples contain 4096 values randomly selected from a 256×256 grid and interpolated on a 64×64 grid using the biharmonic spline.

Next, to improve the estimates we use SG smoothing in two ways: First, the scattered data are interpolated on the grid (as described above) and then SG filters ($p = 5$, $W = 7$) are used to estimate the derivatives. The algorithmic complexity of this method is also $O(N \log_2 N)$. Secondly, the filter is imposed directly on the scattered data (without interpolation). In this case, a 2D window that depends on the sampling point distribution is specified. We consider a third degree polynomial ($p = 3$), and around each point, a neighborhood that includes the nearest 48 neighbors. This increases the algorithmic complexity to $O(N^2)$. The results obtained are compared in Figs. 13-14. The computational cost of the operations required to estimate the anisotropic parameters plotted above is relatively small. The CPU time $t_{\rm CPU}$ required for the estimation of the anisotropy parameters per sample is: (i) $t_{\rm CPU} \approx 0.15$ sec for interpolation (linear, cubic, and nearest neighbor methods) followed by SG differentiation; (ii) $t_{\rm CPU} \approx 14$ sec if the biharmonic spline interpolator is used; (iii) $t_{\rm CPU} \approx 23$ sec for direct application of the SG to the scattered data.

VII. CONCLUSIONS

In this paper we present a non-parametric, systematic approach for the identification of elliptic anisotropy parameters of statistically homogeneous, differentiable, jointly normal or lognormal SRFs, focusing in the 2D case. The method is based on the Covariance Hessian Identity (CHI). It involves the evaluation of field partial derivatives in orthogonal directions, and their relation through closed form expressions, to

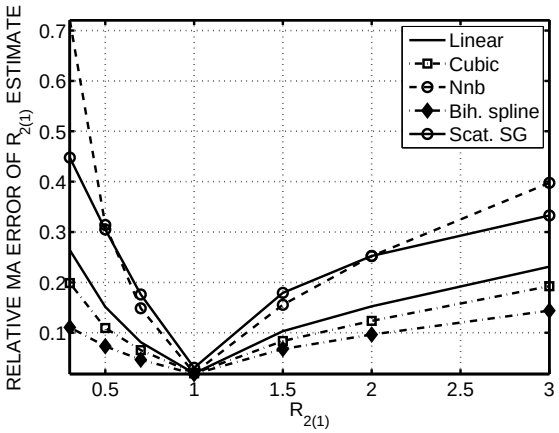

Fig. 13. Relative MA error of $\hat{R}_{2(1)}$ for irregularly sampled SRFs with Gaussian correlations using SG differentiation for SRFs with Gaussian correlations ($\theta = 20^o$, $\xi_{\min} = 4$). The samples contain 4096 values randomly selected from a 256×256 grid. For four of the plots the samples are first interpolated on a 64×64 grid using different interpolators (nearest neighbor, linear, cubic, biharmonic spline) and the sample derivatives are based on SG filtering. For the fifth plot (o) the scattered SG method is used without interpolation.

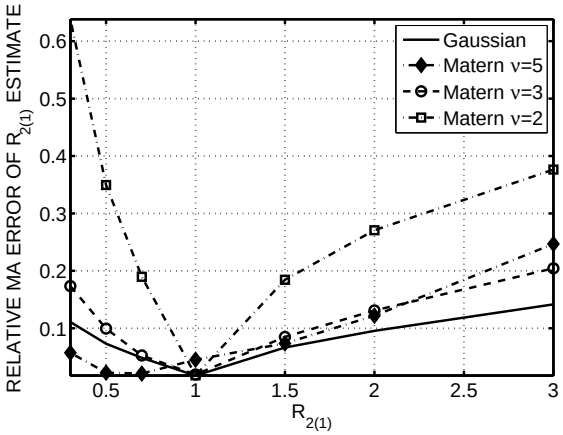

Fig. 14. Relative MA error of $\hat{R}_{2(1)}$ for irregularly sampled SRFs with Gaussian and Matérn ($\nu = 2, 3, 5$) correlations. The samples contain 4096 values randomly selected from a 256×256 grid. The samples are first interpolated on a 64×64 grid using the biharmonic spline interpolator and the sample derivatives are estimated based on SG filtering (p=5, W=7).

the anisotropy ratio and the rotation angle. For anisotropic ratios in the range $[1/3, 3]$, it is the only robust and computationally efficient non-parametric method available for anisotropy identification in spatial data. On regular grids, the method's algorithmic complexity is linear to the sample size, $O(N)$. The accuracy of the method is very good provided that ergodic conditions are respected. The accuracy can be further

improved using Savitzky Golay differentiation, at a small computational cost.

The CHI approach is tested on simulated data on regular grids. In general, the accuracy of the anisotropy estimates increases with the smoothness of the SRF realizations. For SRFs with Gaussian correlations, the RMA error in the anisotropy ratio estimation can be lower than 10% for moderate grid sizes (i.e. 128×128) even for $R_{2(1)}$ significantly different from unity. Use of the centered differences approximation for the derivatives allows very fast estimates of the anisotropy. If the data are contaminated by uncorrelated noise, or if higher accuracy is in general desired, Savitzky Golay differentiators can improve the estimation of the anisotropy parameters. The impact of zero-mean AWGN components is found negligible for $\mathrm{SNR} > 20$, dB but it becomes significant for lower SNRs.

The method is extended for application on scattered data. Two approaches are investigated. In the first approach, a simple interpolator is used to obtain field estimates on a regular grid. The derivatives of the interpolated field are then evaluated by means of either centered differences or SG filters. The best accuracy in the $R_{2(1)}$ estimates is obtained with the biharmonic spline interpolator. However, the numerical complexity increases to $O(N\log_2 N)$. Other interpolators (linear, cubic, nearest neighbor) are significantly faster but they give reduced accuracy. In the second approach, the SG filter is directly applied to the scattered data. This approach is considerably slower, but the obtained MA errors in $\hat{\theta}$ are more uniform and their maximum is reduced. Overall, the estimation accuracy depends on the sampling density and configuration. The accuracy reduction is due to the information loss resulting from the disorder and the rarefication of the sampled pattern. In all cases, the method's accuracy improves as the ratios of the domain to the correlation length, and of the correlation length over the lattice step (or the average minimum distance between sampling points) increase.

Future research should address the application of more elaborate denoising techniques to lattice data. Finally, in the case of non-differentiable SRFs it would be worth investigating the development of a differentiability test, and the application of a smoothing kernel to restore differentiability with minimal impact on the anisotropic structure of the correlations at larger scale.

VIII. Acknowledgments

This work is supported by the European Union in the framework of the INTAMAP "Interoperability and Automated Mapping" specific targeted research project (Contract no. 033811) under FP6 (Information Society Technologies for Environmental Risk Management).

APPENDIX

To use the spectral method for anisotropic SRF simulation, closed form expressions for the power spectral densities are needed. The dimensionless lags h_1 and h_2 in the transformed isotropic coordinate system can be expressed in relation to the sampling axes system r_1, r_2 as follows:

$$h_1 = \frac{\cos\theta}{\xi_1} r_1 + \frac{\sin\theta}{\xi_1} r_2, \tag{37}$$

$$h_2 = -\frac{\sin\theta}{\xi_2} r_1 + \frac{\cos\theta}{\xi_2} r_2. \tag{38}$$

The Euclidean distance $||\mathbf{h}||$ is thus expressed in terms of the original distance $\mathbf{r} = (r_1, r_2)$ as follows:

$$||\mathbf{h}||^2 = A r_1^2 + A_2 r_2^2 + A_{12} r_1 r_2 \tag{39}$$

where, (40)

$$A_1 = \xi_1^{-2}(\cos^2\theta + R^{-2}\sin^2\theta) \tag{41}$$

$$A_2 = \xi_1^{-2}(\sin^2\theta + R^{-2}\cos^2\theta) \tag{42}$$

$$A_{12} = -2\xi_1^{-2}\sin\theta\cos\theta(R^{-2} - 1). \tag{43}$$

The Gaussian, $c_G(r_1, r_2)$ and the Matérn, $c_M(r_1, r_2)$ correlation functions are expressed as follows:

$$c_G(r_1, r_2) = \exp(-||\mathbf{h}||^2) = \exp\left[-(A_1 r_1^2 + A_2 r_2^2 + A_{12} r_1 r_2)\right], \tag{44}$$

$$\begin{aligned} c_M(r_1, r_2) &= c_0\, 2^{1-\nu}\, \Gamma^{-1}(\nu)\, ||\mathbf{h}||^{\nu}\, K_\nu(||\mathbf{h}||) = c_0 2^{1-\nu}\, \Gamma^{-1}(\nu) \\ &\times \sqrt{\left(A_1\, r_1^2 + A_2\, r_2^2 + A_{12}\, r_1\, r_2\right)^{\nu}} K_\nu\left(\sqrt{A_1 r_1^2 + A_2 r_2^2 + A_{12} r_1 r_2}\right), \end{aligned} \tag{45}$$

where $K_\nu(\cdot)$ is the modified Bessel function of the second kind and order ν and $\Gamma(\cdot)$ is the Gamma function.

The respective power spectral densities $C(k_1, k_2)$ can be evaluated from the Fourier transform of the correlation function $c(r_1, r_2)$ expressed as in (44) or (45) according to the correlation type:

$$C(k_1, k_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} c(r_1, r_2) e^{-jk_1 r_1} e^{-jk_2 r_2} \mathrm{d}r_1 \mathrm{d}r_2. \tag{46}$$

For the evaluation of (46), we use the inverse of the transformations (37) and (38):

$$r_1 = h_1\xi_1\cos\theta - h_2\xi_2\sin\theta = \xi_1(h_1\cos\theta - h_2 R\sin\theta) \tag{47}$$

$$r_2 = h_1\xi_1\sin\theta + h_2\xi_2\cos\theta = \xi_1(h_1\sin\theta + h_2 R\cos\theta). \tag{48}$$

The Jacobian of this transformation is

$$|\mathbb{J}_1| = \left| \frac{\partial(r_1, r_2)}{\partial(h_1, h_2)} \right| = R\xi_1^2. \tag{49}$$

In light of the above, the spectrum is expressed in terms of the correlation function $\tilde{c}(h_1, h_2)$ in the isotropic system as follows:

$$C(k_1, k_2) = |\mathbb{J}_1| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{c}(h_1, h_2)\, e^{-(jh_1\tilde{k}_1 + jh_2\tilde{k}_2)} \mathrm{d}h_1 \mathrm{d}h_2 \tag{50}$$

where $\tilde{\mathbf{k}} = (\tilde{k}_1, \tilde{k}_2)$ is the wave-vector in the isotropic system, and

$$\tilde{k}_1 = \xi_1(k_1 \cos\theta + k_2 \sin\theta), \tag{51}$$

$$\tilde{k}_2 = \xi_1 R(-k_1 \sin\theta + k_2 \cos\theta). \tag{52}$$

A. Gaussian Correlation

In the case of a Gaussian correlation function, equation (50) becomes:

$$C(k_1, k_2) = |\mathbb{J}_1| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(h_1^2 + h_2^2 + jh_1\tilde{k}_1 + jh_2\tilde{k}_2)} \mathrm{d}h_1 \mathrm{d}h_2 = \pi\xi_1^2 R e^{-0.25(\tilde{k}_1^2 + \tilde{k}_2^2)} \Rightarrow \tag{53}$$

$$C(k_1, k_2) = \pi\xi_1^2 R e^{-0.25(G_1 k_1^2 + G_2 k_2^2 + G_{12} k_1 k_2)}, \tag{54}$$

where

$$G_1 = \xi_1^2(\cos^2\theta + R^2 \sin^2\theta) \tag{55}$$

$$G_2 = \xi_1^2(\sin^2\theta + R^2 \cos^2\theta) \tag{56}$$

$$G_{12} = 2\xi_1^2 \sin\theta \cos\theta (1 - R^2) \tag{57}$$

B. Matérn Correlation

In the case of the Matérn correlation, equation (50) becomes:

$$C(k_1, k_2) = c_0' \, |\mathbb{J}_1| \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} ||\mathbf{h}||^{\nu} K_{\nu}(||\mathbf{h}||) \exp(-jh_1\tilde{k}_1 - jh_2\tilde{k}_2) \, \mathrm{d}h_1 \mathrm{d}h_2, \tag{58}$$

where $c_0' = c_0 2^{1-\nu} \Gamma^{-1}(\nu)$ and $\mathbb{J}_1$ is given by (49). Changing to polar coordinates $\mathbf{h} = (h, \phi)$ where $h = ||\mathbf{h}||$, the h_1 and h_2 are expressed as $h_1 = h\cos\phi$ and $h_2 = h\sin\phi$. Finally, using the Jacobian of the transformation $|\mathbb{J}_2| = h$, it follows that

$$C(k_1, k_2) = c_0' \, |\mathbb{J}_1| \int_0^{\infty} \int_0^{2\pi} h^{\nu+1} K_{\nu}(h)\, e^{-jh(\cos\phi\tilde{k}_1 + \sin\phi\tilde{k}_2)} \, \mathrm{d}\phi \mathrm{d}h. \tag{59}$$

The angle integral leads to $\int_0^{2\pi} \mathrm{d}\phi\, e^{-jh(\cos\phi\tilde{k}_1+\sin\phi\tilde{k}_2)} = 2\pi\, J_0\big(h\sqrt{\tilde{k}_1^2+\tilde{k}_2^2}\big)$, where J_0 is the Bessel function of the first kind and order zero. Hence, the $C(k_1,k_2)$ is given by the following *Hankel transform*

$$C(k_1,k_2) \;=\; 2\pi\, c_0'\,|\mathbb{J}_1| \int_0^{\infty} h^{\nu+1}\, K_\nu(h)\, J_0\big(h\sqrt{\tilde{k}_1^2+\tilde{k}_2^2}\big)\mathrm{d}h. \tag{60}$$

This Hankel transform is evaluated explicitly using [42]

$$2\pi \int_0^{\infty} h^{\nu+1}\, K_\nu(h) J_0(h\beta)\, \Gamma(\nu+1)\, 2^{\nu+2} \tfrac{1}{\pi\,(1+\beta^2)^{\nu+1}}. \tag{61}$$

Hence, we finally obtain

$$C(k_1,k_2) = \frac{C_0(\nu)}{(1+G_1k_1^2+G_2k_2^2+G_{12}k_1k_2)^{\nu+1}}, \tag{62}$$

where $C_0(\nu)$ is a function of ν and the variance of the SRF.

REFERENCES

[1] J. Xiong, V. Zolotov, and L. He, "Robust extraction of spatial correlation," *IEEE Journal on Technology in Computer Aided Design*, vol. 26, no. 4, pp. 619–631, 2007.

[2] D. Williams and D. Johnson, "Robust estimation of structured covariance matrices," *IEEE Transactions on Signal Processing*, vol. 41, no. 9, pp. 2891–2906, 1993.

[3] J. Carr and F. de Miranda, "The semivariogram in comparison to the co-occurrence matrix for classification of image texture," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 36, no. 6, pp. 1945–1952, 1998.

[4] L. Sanchez-Brea and E. Bernabeu, "Uncertainty estimation by convolution using spatial statistics," *IEEE Transactions on Image Processing*, vol. 15, no. 10, pp. 3131–3137, 2006.

[5] R. Parrott, M. Stytz, P. Amburn, and D. Robinson, "Towards statistically optimal interpolation for 3D medical imaging," *IEEE Engineering in Medicine and Biology Magazine*, vol. 12, no. 3, pp. 49–59, 1993.

[6] B. Girod, "Motion-compensating prediction with fractional-pel accuracy," *IEEE Transactions on Communications*, vol. 41, no. 4, pp. 604–612, 1993.

[7] M. Unser and T. Blu, "Generalized smoothing splines and the optimal discretization of the Wiener filter," *IEEE Transactions on Signal Processing*, vol. 53, no. 6, pp. 2146–2159, 2005.

[8] C. Ma, "Linear combinations of space-time covariance functions and variograms," *IEEE Transactions on Signal Processing*, vol. 53, no. 3, pp. 857–864, 2005.

[9] R. Wilson and C. T. Li, "A class of discrete multiresolution random fields and its application to image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 42–56, Jan. 2002.

[10] J. Ruiz-Alzola *et al.*, "Geostatistical medical image registration," in *Proc. of 6th Intern. Conf. Medical Image Comput. Computer-Assisted Intervention (MICCAI'03)*, 2003, pp. 894–901.

[11] A. Kourgli and A. Belhadj-Aissa, "Texture primitives description and segmentation using variography and mathematical morphology," *2004 IEEE International Conference on Systems, Man and Cybernetics*, vol. 7, pp. 6360–6365, Oct. 2004.

[12] F. Gui and L. Wei, *Application of variogram function in image analysis*. IEEE, 31 Aug. - 4 Sep. 2004, vol. 2, ch. 2004 7th International Conference on Signal Processing, pp. 1099–1102.

[13] J. Cao and K. J. Worsley, "Applications of random fields in human brain mapping," in *Spatial Statistics: Methodological Aspects and Applications*, M. Moore, Ed. Berlin: Springer Lecture Notes in Statistics, 2001, vol. 159, pp. 169–182.

[14] A. Leow *et al.*, "Brain structural mapping using a novel hybrid implicit/explicit framework based on the level-set method," *NeuroImage*, vol. 24, no. 3, pp. 910–927, 2004.

[15] A. M. Yaglom, *Correlation Theory of Stationary and Related Random Functions I*. New York: Springer Verlag, 1987.

[16] M. Kanevski and M. Maignan, *Analysis and Modelling of Spatial Environmental Data*. EPFL Press, 2004.

[17] H. Wackernagel, *Multivariate Geostatistics*. Berlin: Springer Verlag, 1997.

[18] P. Goovaerts, *Geostatistics for Natural Resources Evaluation*. Oxford University Press, 1997.

[19] P. K. Kitanidis, "Statistical estimation of polynomial generalized covariance functions and hydrologic applications," *Water Resources Res.*, vol. 19, no. 2, pp. 909–921, 1983.

[20] ——, "Parametric estimation of covariances of regionalized variables," *Water Resources Res.*, vol. 23, no. 4, pp. 671–680, 1987.

[21] E. Pardo-Igúzquiza, "Maximum likelihhod estimation of spatial covariance parameters," *Mathematical Geology*, vol. 30, no. 1, pp. 95–108, 1998.

[22] M. Ecker and A. Gelfand, "Bayesian modeling and inference for geometrically anisotropic spatial data," *Mathematical Geology*, vol. 32, no. 1, pp. 67–82, 1999.

[23] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. New York: McGraw-Hill, 2002.

[24] D. Jupp, A. Strahler, and C. Woodcock, "Autocorrelation and regularization in digital images. I. Basic theory," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 4, pp. 463–473, 1988.

[25] M. L. Stein, "Local stationarity and simulation of self-affine intrinsic random functions," *IEEE Transactions on Information Theory*, vol. 47, no. 4, pp. 1385–1390, 2001.

[26] B. Mandelbrot, "Fractional Brownian motions, fractional noises and applications," *SIAM Review*, vol. 10, no. 4, pp. 422–437, 1968.

[27] M. L. Stein, *Interpolation of Spatial Data: Some Theory for Kriging*. New York: Springer, 1999.

[28] S. Ramani and M. Unser, "Matérn b-splines and the optimal reconstruction of signals," *IEEE Signal Processing Letters*, vol. 13, no. 7, pp. 437–440, 2006.

[29] S. Ramani, D. V. D. Ville, and M. Unser, "Non-ideal sampling and adapted reconstruction using the stochastic Matérn model," in *Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on*, vol. 2, May 2006, pp. 73–76.

[30] D. Hristopulos, "Spartan Gibbs random field models for geostatistical applications," *SIAM Journal of Scientific Computing*, vol. 24, no. 6, pp. 2125–2162, 2003.

[31] S. N. Elogne, D. Hristopulos, and E. Varouchakis, "An application of Spartan spatial random fields in environmental mapping: focus on automatic mapping capabilities," *Stochastic Environmental Research and Risk Assessment*, 2007, to appear, DOI:10.1007/s00477-007-0167-5.

[32] D. Hristopulos and S. Elogne, "Analytic properties and covariance functions of a new class of generalized Gibbs random fields," *IEEE Transactions on Information Theory*, vol. 53, no. 12, pp. 4467–4679, 2007.

[33] P. Swerling, "Statistical properties of the contours of random surfaces," *IRE Transactions on Information Theory*, pp. 315–321, July 1962.

[34] H. Cramér and M. R. Leadbetter, *Stationary and Related Stochastic Processes*. New York: John Wiley and Sons, 1967.

[35] H. F. R.J. Creswick and C. Poole, *Introduction to Renormalization Group Methods in Physics*. Wiley, 1991.

[36] D. Hristopulos, "Approximate methods for explicit calculations of non-gaussian moments," *Stochastic Environmental Research and Risk Assessment*, vol. 20, no. 4, pp. 278–290, 2006.

[37] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedure," *Anal. Chem*, vol. 36, pp. 1627–1639, 1964.

[38] P. H. J. B. J. Luo, K. Ying, "Properties of Savitzky Golay digital differentiators," *Digital Signal Processing*, vol. 15, pp. 122–136, 2005.

[39] E. Padro-Igúzquiza and M. Chica-Olmo, "The Fourier integral method: an efficient spectral method for simulation of random fields," *Mathematical Geology*, vol. 25, no. 2, pp. 177–217, 1993.

[40] C. Lantuéjoul, *Geostatistical Simulation: Models and Algorithms*. New York: Springer, 2002.

[41] D. Sandwell, "Interpolation of GEOS-3 and SEASAT altimeter data," *Geophysical Research Letters*, vol. 2, pp. 139–142, 1987.

[42] R. D. Lord, "The use of the Hankel transform in statistics: I. general theory and examples," *Biometrika*, vol. 41, no. 1/2, pp. 44–55, 1954.