SPARTAN GIBBS RANDOM FIELD MODELS FOR GEOSTATISTICAL APPLICATIONS

Abstract: The inverse problem of determining the spatial dependence of random fields from an experimental sample is a central issue in Geostatistics. We propose a computationally efficient approach based on Spartan Gibbs random fields. Their probability density function is determined by a small set of parameters, which can be estimated by enforcing sample-based constraints on the stochastic moments. The computational complexity of calculating the constraints increases linearly with the sample size. We investigate a specific Gibbs probability density with spatial dependence derived from generalized gradient and Laplacian operators, and we derive permissibility conditions for the model parameters. The optimal values for the parameters are determined by minimizing a normalized metric measuring the ÒdistanceÓbetween stochastic moments and the sample constraints. The computational complexity of the minimization depends on the number of the model parameters, but not on the sample size. We illustrate the method using simulated control samples with different types of spatial dependence. Spartan Gibbs random fields are useful for modeling large samples, and when reliable estimation of the variogram is not possible. Estimation of the field values at non-sampled positions, conditional simulations, as well as extensions to anisotropic spatial dependence, non-Gaussian probability densities, and off-lattice distributions are briefly discussed.

DIONISSIOS T. HRISTOPULOS, Pulp and Paper Research Institute of Canada
570 St.-John Blvd., Pointe-Claire, QC, H9R 3J9, Canada (dionisi@mailaps.org)

AMS Classification: 63M30, 62P12, 62P30, 62M40

Keywords: Inverse problem, spatial dependence, random fields, optimization, simulation

Abbreviated Title: Spartan Gibbs Random Fields in Geostatistics

Date Modified: 2/13/02 8:23 PM

1. Introduction

The purpose of this paper is to investigate geostatistical applications of Gibbs random fields that are ÒeconomicallyÓ determined by a small set of parameters, and thus we appropriately name them ÒSpartanÓ. In hydrology, oil reservoir forecasting, environmental engineering and geophysics random fields are used to model spatial heterogeneities and space-time variations [40,37,7,14,27]. They account for structural and dynamic variations as well as for uncertainties due to the limited precision of measurements. Multivariate normal (i.e., Gaussian) random fields are commonly used in practical applications. Distributions with non-Gaussian probability density functions (pdf) are often reduced to a Gaussian pdf by means of appropriate nonlinear transformations. The spatial dependence of a Gaussian random field $\mathrm{X}(\mathbf{s})$ is determined from the semi-variogram (one-half of the variogram) which is defined by means of the following stochastic expectation

$$\gamma_{\mathrm{x}}(\mathbf{r}) = \frac{1}{2} E\left[\,\mathrm{X}(\mathbf{s}) - \mathrm{X}(\mathbf{s}+\mathbf{r})\,\right]^2 . \qquad (1)$$

This is approximated by means of the empirical (experimental) semi-variogram, the values of which are estimated from an available sample $X(\mathbf{s})$ at a number of lag distances $\mathbf{r}_k$. Estimation of the experimental semi-variogram is a high-dimensional inverse problem. It is accomplished using the method of moments [40]

$$\hat{\gamma}_{\mathrm{x}}(\mathbf{r}_k) \equiv \frac{1}{2\; N(\mathbf{r}_k)} \sum_{i=1}^{N} \sum_{j=1}^{N} \left[X(\mathbf{s}_i) - X(\mathbf{s}_i + \mathbf{u}_{ij}) \right]^2, \;\; r_k - \delta r_k \leq \left\| \mathbf{u}_{ij} \right\| < r_k + \delta r_k. \qquad (2)$$

In the above N is the sample size, the interval $[r_k - \delta r_k, r_k + \delta r_k)$ defines the bin for the lag $\mathbf{r}_k$ and $N(\mathbf{r}_k)$ is the cardinality of the bin. The optimal variogram model is selected by model fits to the set $\{\mathbf{r}_k, \hat{\gamma}_{\mathrm{x}}(\mathbf{r}_k)\}$. Since estimation of the variogram involves all (or most) pairings of the sample positions, the computational complexity increases with the square of the sample size. Bin definition involves various assumptions, especially in the case of anisotropic dependence and irregularly spaced measurement locations. In addition, the estimates of the variogram involve considerable uncertainty [22,47], especially for the bins that contain only a small number of pairs. Robust variogram estimators [13,14] involve similar assumptions. The method of maximum likelihood (ML) [41,43,50] bypasses the empirical variogram calculation. The variogram parameters are evaluated by maximizing the likelihood of the data for a specific variogram model. However, in the ML approach the Cholesky decomposition of the covariance matrix, which is an $O(N^3)$ operation, is evaluated for every set of model parameters tried during the optimization search. For a discussion on ML’s merits and disadvantages see [51].

Hence, there is a need for computationally efficient models of spatial dependence, especially for studies that involve large samples. The models that we propose here are determined from a small number of parameters. Estimation of the empirical variogram is not required, thus reducing significantly the problem of dimensionality. The relative probabilities of different states (samples) can be efficiently calculated, which is favorable for Monte Carlo simulations. The paper is organized as follows: Section 2 provides general definitions used in the paper. In Section 3 we discuss the motivation for coarse-grained random fields and properties of coarse-graining kernels. In Section 4 the Gibbs random field model is introduced and its properties are studied. In Section 5 we propose a solution to the inverse problem of estimating the model parameters from available samples, based on a normalized distance metric. In Section 6 we present a numerical implementation of the estimation procedure and explain the results. In Section 7 we briefly discuss issues that deserve further investigation, and conclusions are given in Section 8.

2. Definitions

We use roman characters to denote a random field and italic to denote specific states. We consider weakly homogeneous random fields (i.e., fields with constant mean and strictly lag-dependent covariance). For multivariate normal distributions, weak homogeneity implies a homogeneous probability density. Let $\mathrm{X}(\mathbf{s})$ denote a random field in a continuum space $\Omega \subset R^d$ and $X(\mathbf{s})$ its specific states. The multivariate probability density is denoted by $f_{\mathrm{x}}[\mathrm{X}(\mathbf{s})]$, and the ensemble mean by $E[\mathrm{X}(\mathbf{s})]$. The zero-mean fluctuation is $\mathrm{x}(\mathbf{s}) \equiv \mathrm{X}(\mathbf{s}) - E[\mathrm{X}(\mathbf{s})]$, and the two-point covariance is given by $G_{\mathrm{x}}(\mathbf{s}, \mathbf{s}+\mathbf{r}) \equiv E[\,\mathrm{x}(\mathbf{s})\mathrm{x}(\mathbf{s}+\mathbf{r})\,]$.

DEFINITION 2.1. *The covariance spectral density,* $\tilde{G}_{\mathrm{x}}(\mathbf{k})$, *of the random field* $X(\mathbf{s})$ *is given by the following Fourier integral*

$$\tilde{G}_{\mathrm{x}}(\mathbf{k}) = \int \mathbf{dr}\, G_{\mathrm{x}}(\mathbf{r})\, e^{-i\,\mathbf{k}\cdot\mathbf{r}}. \qquad (3)$$

The inverse transform is then defined as follows

$$G_{\mathrm{x}}(\mathbf{r}) = \vartheta_d \int \mathbf{dk}\; \tilde{G}_{\mathrm{x}}(\mathbf{k})\, e^{i\,\mathbf{k}\cdot\mathbf{r}}, \qquad (4)$$

where the coefficient is given by $\vartheta_d = (2\pi)^{-d}$. ð

Similarly, let $\mathrm{X}(\mathbf{s}_i)$ be a field on a lattice $\Pi \subset R^d$. The lattice sites $\mathbf{s}_i$ $(i = 1,\ldots,N)$, are generated from unit vectors $\mathbf{e}_j$ and lattice steps a_j $(j = 1,\ldots,d)$ in d orthogonal directions. The X_i $(i = 1,\ldots,N)$ denote the value of a field state (sample) at $\mathbf{s}_i$, the vector $\mathbf{X} \equiv (X_1,\ldots,X_N)$, the entire state, and $f_{\mathrm{x}}(\mathbf{X})$ the lattice multivariate density.

DEFINITION 2.2. *A homogeneous random field is isotropic if its covariance depends only on the Euclidean distance* $r = \|\mathbf{r}\|$. *We define the correlation radius,* r_c, *by means of*

$$G_{\mathrm{x}}(r_c) \equiv 0.05\, G_{\mathrm{x}}(0)\ . \qquad (5)$$

3. Coarse-Graining Transformations

We introduce coarse-graining transformations and use them below to define the Gibbs probability densities. The symbol $\mathrm{X}_{\lambda}(\mathbf{s})$ denotes a coarse-grained average of the primary field X over length scales smaller than λ. Coarse-graining is useful for defining differentiable Gibbs random fields. The physical motivation is that experimental data are often coarse-grained, because limited resolution prohibits probing the variability at very small scales [6,20]. In mining engineering the commonplace Ònugget effectÓ[40] refers to the offset of the variogram at zero distance due to unresolved small-scale variability. Coarse-graining is also motivated by the representative elementary volume (REV) hypothesis, which underlies the continuum modeling of flow and transport in porous media [2,5,33].

DEFINITION 3.1. *A linearly coarse-grained random field* $\mathrm{X}_\lambda(\mathbf{s})$ *in a continuum space* $\Omega \subset R^d$ *is defined by the convolution integral*

$$\mathrm{X}_\lambda(\mathbf{s}) = \int \mathbf{ds}'\, Q_\lambda(\mathbf{s},\mathbf{s}')\,\mathrm{X}(\mathbf{s}'), \tag{6}$$

where $Q_\lambda(\mathbf{s},\mathbf{s}')$ *is a smoothing kernel of radius* λ *and* $\mathrm{X}(\mathbf{s})$ *is the primary field. The kernel is assumed to be:*

i. *A real function, i.e.* $Q_\lambda(\mathbf{s},\mathbf{s}')$: $\Re^d \times \Re^d \to \Re$.

ii. *Translation invariant, i.e.* $Q_\lambda(\mathbf{s},\mathbf{s}') = Q_\lambda(\mathbf{s}-\mathbf{s}')$.

iii. *Differentiable, at least in the generalized sense* [21].

iv. *Integrable and normalized, i.e.* $\int \mathbf{ds}'\, Q_\lambda(\mathbf{s}-\mathbf{s}') = 1$.

v. *Isotropic, i.e.* $Q_\lambda(\mathbf{s}-\mathbf{s}') = Q_\lambda(\|\mathbf{s}-\mathbf{s}'\|)$.

vi. *Asymptotically declining to zero or equal to zero outside a finite radius. The coarse-graining scale* λ *is defined by* $Q_\lambda(\|\mathbf{s}-\mathbf{s}'\| = \lambda) = 0.05\, Q_\lambda(0)$.

vii. *Heaviest at zero-frequency, i.e.* $\left|\tilde{Q}_\lambda(\mathbf{k})\right|^2 \le \left|\tilde{Q}_\lambda(0)\right|^2, \quad \forall\, k \ge 0$.

The primary field is recovered if $Q_{\lambda}(\mathbf{s}-\mathbf{s}')=\delta(\mathbf{s}-\mathbf{s}')$. Condition (ii) ensures that $\mathrm{X}_{\lambda}(\mathbf{s})$ is homogeneous and condition (iii) that it is differentiable. Condition (iv) guarantees that the Fourier transform of the kernel exists and $\tilde{Q}_{\lambda}(0)=1$. It also implies conservation of the mean, $E[\mathrm{X}_{\lambda}(\mathbf{s})]=E[\mathrm{X}(\mathbf{s})]$. Condition (v) is not restrictive, because an anisotropic kernel can be reduced to an isotropic form by rescaling the coordinate axes. Condition (vii) ensures that the kernel has the least impact on large wavelengths. From the above definitions follow three useful properties of coarse-grained random fields, that is, the spectral density relation, the reduction of variance, and the extension of the correlation range. They are given by the following lemmas.

LEMMA 3.1. *The spectral density of the coarse-grained field is equal to the product of the spectral density of the primary and that of the kernel, i.e.,*

$$\tilde{G}_{\mathrm{x};\lambda}(\mathbf{k})=\left|\tilde{Q}_{\lambda}(\mathbf{k})\right|^{2}\tilde{G}_{\mathrm{x}}(\mathbf{k}). \tag{7}$$

Proof. Based on Equation (6) $G_{\mathrm{x};\lambda}(\mathbf{r})=\int \mathbf{dr}' \int \mathbf{dr}''\, Q(\mathbf{s}+\mathbf{r}-\mathbf{r}'')\, Q(\mathbf{s}-\mathbf{r}')\, G_{\mathrm{x}}(\mathbf{r}'-\mathbf{r}'')$. Equation (7) then follows by replacing the functions with their Fourier transforms and using the orthogonality relation $\int \mathbf{dr}' \exp(i(\mathbf{k}_1+\mathbf{k}_2)\mathbf{r}')=(2\pi)^{d}\,\delta(\mathbf{k}_1+\mathbf{k}_2)$.

LEMMA 3.2. *The variance of the primary provides an upper bound for the variance of the coarse-grained field.*

Proof. By expressing the ratio of the variances in terms of the spectral density integrals, using Equation (7) for the spectral density of the coarse-grained field, and the inequality of condition (vii) we obtain

$$\frac{\sigma^2_{\mathrm{x};\lambda}}{\sigma^2_{\mathrm{x}}} = \frac{\int \mathbf{dk} \left| \tilde{Q}_\lambda(\mathbf{k}) \right|^2 \tilde{G}_{\mathrm{x}}(\mathbf{k})}{\int \mathbf{dk}\, \tilde{G}_{\mathrm{x}}(\mathbf{k})} \leq 1. \tag{8}$$

LEMMA 3.3. *Let* $\zeta_{\mathrm{x}} \equiv \left[\int \mathbf{dr}\, G_{\mathrm{x}}(\mathbf{r}) / G_{\mathrm{x}}(0) \right]^{1/d}$ *be the correlation range of the primary field. The range of the primary is a lower bound on the range of the coarse-grained field, i.e.,* $\zeta_{\mathrm{x};\lambda} \geq \zeta_{\mathrm{x}}$.

Proof. The covariance integrals are expressed in terms of Fourier transforms. Lemma 3.3 follows from the condition (iv) of the Definition (3.1) and Lemma (3.2)

$$\frac{\zeta_{\mathrm{x};\lambda}}{\zeta_{\mathrm{x}}} = \left[\frac{\tilde{G}_{\mathrm{x}}(0) \left| \tilde{Q}_\lambda(0) \right|^2}{\sigma^2_{\mathrm{x};\lambda}} \frac{\sigma^2_{\mathrm{x}}}{\tilde{G}_{\mathrm{x}}(0)} \right]^{1/d} = \left(\frac{\sigma^2_{\mathrm{x}}}{\sigma^2_{\mathrm{x};\lambda}} \right)^{1/d}. \tag{9}$$

In the case of lattice random fields the coarse-graining equation (6) is replaced by the discrete summation

$$\mathrm{X}_\lambda(\mathbf{s}_i) = \sum_j Q_\lambda(\mathbf{s}_i, \mathbf{s}_j)\, X(\mathbf{s}_j),\ \mathbf{s}_j \in B_\lambda(\mathbf{s}_i). \tag{10}$$

where the neighborhood $B_\lambda(\mathbf{s}_i)$ of $\mathbf{s}_i$, is defined by: $\mathbf{s}_j \in B_\lambda(\mathbf{s}_i)$ if $\left| s_{i,k} - s_{j,k} \right| \le R_k(\lambda)$, $k = 1,...,d$. For lattice fields the step size a_i, $i = 1,...,d$, should satisfy $a_i \ge R_i(\lambda)$.

3.1 Kernel with sharp frequency cutoff

PROPOSITION 3.1. *The inverse Fourier transform of the step function*

$$\tilde{Q}_\lambda(\mathbf{k}) = \theta(k_{\max} - |\,\mathbf{k}\,|), \tag{11}$$

where $k_{\max} \sim 1/\lambda$ *is a kernel that satisfies the conditions of Definition 3.1.*

Proof: Conditions (i)-(ii), (iii)-(v) and (vii) follow trivially. The kernel in real space is

$$Q_\lambda(\mathbf{r}) = k_{\max}^d \sqrt{\vartheta_d}\, \frac{J_{d/2}(k_{\max}\, r)}{(k_{\max}\, r)^{d/2}}, \tag{12}$$

where $J_{d/2}(x)$ is the Bessel function of the first kind of order $d/2$. Its asymptotic behavior for $x >> 1$ is given by $J_{d/2-1}(x) \sim \cos\left[\, x - (\tfrac{d}{2} - 1)\tfrac{\pi}{2} - \tfrac{\pi}{4}\,\right]\sqrt{2}\big/\sqrt{\pi\, x}$. This suffices to prove the condition (vi). The derivative is calculated using $d\{x^{-\nu}\, J_\nu(x)\}\big/dx = -x^{-\nu}\, J_{\nu+1}(x)$ [52], which leads to

$$\frac{dQ_\lambda(\mathbf{r})}{dr} = -k_{\max}^{d+1} \sqrt{\vartheta_d}\, \frac{J_{d/2+1}(k_{\max}\, r)}{(k_{\max}\, r)^{d/2}}. \tag{13}$$

In view of the Bessel function dependence for $x << 1$, $J_{d/2-1}(x) \sim x^{d/2-1} / 2^{d/2-1}\,\Gamma(d/2)$, the first derivative vanishes at zero. The second derivative is evaluated using the relation $d\left\{J_{\nu+1}(x)/x^{\nu}\right\}/dx = -J_{\nu+2}(x)/x^{\nu} + J_{\nu+1}(x)/x^{\nu+1}$, which leads to

$$\frac{d^2 Q_\lambda(\mathbf{r})}{dr^2} = k_{\max}^{d+2}\,\sqrt{\vartheta_d}\,\left\{\frac{J_{d/2+2}(k_{\max}\,r)}{(k_{\max}\,r)^{d/2}} - \frac{J_{d/2+1}(k_{\max}\,r)}{(k_{\max}\,r)^{d/2+1}}\right\}. \tag{14}$$

Hence, the second derivative at zero is $d^2 Q_\lambda(\mathbf{r})/dr^2 = -k_{\max}^{d+2}\,\sqrt{\vartheta_d}\,\left\{\Gamma(d/2)\,2^{d/2+1}\right\}^{-1}$.

4. Gibbs Random Field Models

DEFINITION 4.1. *The multivariate pdf of Gibbs random fields is in general expressed as follows*

$$f_{\mathrm{x}}[X_\lambda(\mathbf{s})] = Z^{-1}\,\exp\left\{-H[X_\lambda(\mathbf{s})]\right\}. \tag{15}$$

The exponent $H[X_\lambda(\mathbf{s})]$ represents the value of the energy functional $H[\mathrm{X}_\lambda]$ for the state $X_\lambda(\mathbf{s})$ and the normalization constant $Z = \sum_{X_\lambda(\mathbf{s})} \exp\left\{-H[X_\lambda(\mathbf{s})]\right\}$ is the partition function.

Gibbs fields were introduced to model the equilibrium states of physical systems [24]. For these systems the functional $H[\mathrm{X}_\lambda]$ is the Hamiltonian that governs the distribution of energy among different states. We can in general think of $H[\mathrm{X}_\lambda]$ as an ÒenergyÓ functional, because states with higher values $H[X_\lambda(\mathbf{s})]$ have lower probability. If the random field has infinite or semi-infinite support the energy functional must obey $\lim_{|X_\lambda|\to\infty} H[X_\lambda] = \infty$. This forces the term that dominates $H[\mathrm{X}_\lambda]$ for large X_λ to be an even function of X_λ.

Gibbs random fields have also found applications in image analysis and robotic vision [23,53,46], where they are used to represent the spatial dependence of visual patterns. Gaussian random fields and maximum entropy probability densities [7,9,8] used in geostatistics can be viewed as special cases of Gibbs random fields. In these cases, $H[\mathrm{X}_\lambda]$ incorporates sample-based constraints such as the variogram. Non-Gaussian terms in the Òmaximum-entropy-priorÓ energy functional involve higher-order correlation functions, which are difficult to obtain from the data. Also, determining the parameters of high-order correlation functions is a multidimensional optimization problem, which to our knowledge has been solved only under restrictive conditions [36]. Here we focus on Gibbs probability densities that involve a small set of model parameters (in contrast to multidimensional constraints).

4.1 A class of Gibbs random fields with local coupling

DEFINITION 4.2. *Let* $B(\mathbf{s};\mathbf{R}_m)$ *define a local neighborhood of the point* $\mathbf{s}$, *such that* $\mathbf{s}+\mathbf{r} \in B(\mathbf{s};\mathbf{R}_m)$ *if* $|r_i| \le R_{m,i}$, $i=1,\ldots,d$ *where* $\mathbf{R}_m$ *is a vector of positive radii. We define a local-coupling energy functional as follows*

$$H[X_\lambda(\mathbf{s})] \equiv \int \mathbf{ds} \sum_{m=1}^{M} \alpha_m \sum_{\mathbf{r}_m \in B(\mathbf{s};\mathbf{R}_m)} L_m^{(2)}[X_\lambda(\mathbf{s}), X_\lambda(\mathbf{s}+\mathbf{r}_m)]. \qquad (16)$$

The corresponding Gibbs random field has a multivariate normal pdf if the $L_m^{(2)}$ *are quadratic functionals of the fluctuations. The energy coefficients* α_m *satisfy permissibility conditions specified by Bochner's theorem.*

In the case of lattice fields the integral is replaced by a summation. Specific examples of $L_m^{(2)}$ operators include quadratic combinations of various derivatives (for continuum random fields) or finite differences. For Gaussian random fields differentiability of the sample functions is equivalent to the existence of generalized derivatives for the appropriate covariance functions, e.g. [32].

DEFINITION 4.3. *The following energy functional where* $\mathrm{X}_\lambda(\mathbf{s})$ *is coarse-grained with the kernel of Eq. (11), is locally coupled in the sense of Definition 4.2:*

$$H[\mathrm{X}_\lambda] = \frac{1}{2}\int \mathbf{ds}\left\{ \alpha_0 \left[\mathrm{X}_\lambda(\mathbf{s}) - E[\mathrm{X}_\lambda(\mathbf{s})] \right]^2 + \alpha_1 \left[\nabla \mathrm{X}_\lambda(\mathbf{s}) \right]^2 + \alpha_2 \left[\nabla^2 \mathrm{X}_\lambda(\mathbf{s}) \right]^2 \right\}, \tag{17}$$

The energy functional is isotropic and couples the field values via the derivatives. For a specified cutoff $k_{\max}$ the multivariate density is determined from the energy coefficients α_0, α_1 and α_2. The first term ($\propto \alpha_0$) represents the energy cost of deviations from the mean state. The gradient ($\propto \alpha_1$) and curvature ($\propto \alpha_2$) terms represent energy costs related to spatial variations of the field. For $\alpha_1, \alpha_2 > 0$ the first term favors low-gradient and the second term low-curvature states.

PROPOSITION 4.1. *The energy functional of Equation (17) leads to a permissible random field density if the coefficients* $\alpha_0,\ \alpha_1,\ \alpha_2$ *satisfy conditions specified below. Then, the multivariate density is (i) even with respect to the fluctuation* $\mathrm{x}_\lambda \equiv \mathrm{X}_\lambda - E[\mathrm{X}_\lambda]$, *(ii) unimodal, and (iii) Gaussian.*

Proof: (i). All the terms in $H[X_\lambda]$ are even functions of X_λ. (ii). The mode is the highest-probability field state $X_\lambda^*(\mathbf{s})$. Since $H[X_\lambda]$ is convex (provided the coefficients satisfy the permissibility conditions), the mode(s) is determined from the roots of the variational derivative $\delta H[X_\lambda]/\delta X_\lambda$, i.e., from

$$\left.\frac{\delta H[X_\lambda]}{\delta X_\lambda}\right|_{X_\lambda = X_\lambda^{(0)}} = 0 \Rightarrow \alpha_0 \left[X_\lambda^{(0)}(\mathbf{s}) - E[X_\lambda^{(0)}(\mathbf{s})] \right] - \alpha_1 \nabla^2 X_\lambda^{(0)}(\mathbf{s}) + \alpha_2 \nabla^4 X_\lambda^{(0)}(\mathbf{s}) = 0\,, \qquad (18)$$

For a homogeneous random field the mode $X_\lambda^{(0)}(\mathbf{s})$ is spatially uniform, and thus the only solution is $X_\lambda^{(0)}(\mathbf{s}) = E\,[\mathrm{X}_\lambda^{(0)}(\mathbf{s})]$. (iii). The probability density is Gaussian since all the energy terms are quadratic functionals of the field. The permissibility conditions on the energy coefficients are determined by means of Bochner's theorem [4]:

BOCHNER'S THEOREM: A function $G_{\mathrm{x};\lambda}(\mathbf{r})$ is a permissible covariance only if it is positive semi-definite. This requires the spectral density $\tilde{G}_{\mathrm{x};\lambda}(\mathbf{k})$ to be (i) positive semi-definite (i.e., $\tilde{G}_{\mathrm{x};\lambda}(\mathbf{k}) \geq 0$ for all $\mathbf{k}$) and (ii) integrable. Condition (ii) ensures that the variance is finite. ð

PROPOSITION 4.2. *The energy functional of Equation (17) is permissible if the energy coefficients satisfy the following relations:*

$$\alpha_0 > 0,\ \alpha_1 \geq 0,\ \alpha_2 > 0, \tag{19}$$

or

$$\alpha_2 > 0,\ \alpha_1 < 0,\ \alpha_1^2 < 4\alpha_0\,\alpha_2. \tag{20}$$

The variance of the Gibbs field is then given by the following integral

$$\sigma^2_{\mathrm{x};\lambda} = \vartheta_d \int \mathbf{dk}\, \tilde{G}_{\mathrm{x}}(\mathbf{k}) = \frac{\vartheta_d\, A_d}{\alpha_2} \int_0^\infty dk \frac{k^{d-1}\, \left|\tilde{Q}_\lambda(\mathbf{k})\right|^2}{k^4 + \beta k^2 + \gamma}, \tag{21}$$

where $\beta = \alpha_1/\alpha_2$, $\gamma = \alpha_0/\alpha_2$, *and* $A_d = 2\pi^{d/2}/\Gamma(d/2)$ *is the surface area of the unit sphere in d dimensions.* .

Proof: The energy functional is expressed in the Fourier domain as follows

$$H[x_\lambda(\mathbf{s})] = \frac{\vartheta_d}{2} \int \mathbf{dk}\, \tilde{x}_\lambda(\mathbf{k})\, \tilde{x}_\lambda(-\mathbf{k})\, \tilde{G}^{-1}_{\mathrm{x};\lambda}(\mathbf{k}),$$

where $\tilde{x}_\lambda(\mathbf{k})$ is the Fourier transform of the coarse-grained fluctuations and $\tilde{G}^{-1}_{\mathrm{x};\lambda}(\mathbf{k})$is the inverse of the following spectral density function

$$\tilde{G}_{x;\lambda}(\mathbf{k}) = \frac{\left|\tilde{Q}_\lambda(\mathbf{k})\right|^2}{\alpha_2 k^4 + \alpha_1 k^2 + \alpha_0}. \tag{22}$$

The condition (i) of Bochner's theorem is fulfilled when the inequalities (19) or (20) hold: If the inequalities (19) are satisfied the roots of the polynomial $p_4(k) \equiv \alpha_2 k^4 + \alpha_1 k^2 + \alpha_0$ are imaginary (for $\alpha_1 = 0$), complex (for $\alpha_1 > 0$, $\alpha_1^2 < 4\alpha_0\,\alpha_2$), or real and negative (for $\alpha_1 > 0$, $\alpha_1^2 \geq 4\alpha_0\,\alpha_2$). If the inequalities (20) are satisfied the roots of $p_4(k)$ are complex. In all cases, $p_4(k)$ is strictly positive for $k \geq 0$. If $d \leq 3$ the condition (ii) of Bochner's theorem is satisfied for any bounded $\tilde{Q}_\lambda(\mathbf{k})$, and if $\tilde{Q}_\lambda(\mathbf{k}) = \theta(k_{\max} - |\mathbf{k}|)$ it is satisfied for all d.

REMARK 4.1. The curvature term in the energy functional adds stiffness to high-frequency fluctuations and is thus responsible for the finite variance, even if $\alpha_1 < 0$.

REMARK 4.2. In the coefficient space $(\alpha_2, \beta, \gamma)$ the permissibility conditions are $\alpha_2 > 0,\ \beta \geq 0,\ \gamma > 0$ or $\alpha_2 > 0,\ \beta < 0,\ \beta^2 < 4\gamma$. Hence, the branch $\beta < 0$ of the parabola $\beta^2 = 4\gamma$ is the boundary that separates the half-plane $\beta \in R$ and $\gamma \in R^+$ into permissible and non-permissible domains. At the boundary the stochastic moments of the field diverge, because the poles of the spectral density approach the real line.

4.2 Physical parameterization

The energy functional of Equation (17) follows the general formulation of Equation (16). However, the physical meaning of the coefficients α_i ($i = 0,1,2$), is not clear. The following formulation uses a more intuitive set of parameters:

DEFINITION 4.4. *Let us define the scale parameter* η_0, *the dimensionless coefficient* η_1, *and the correlation length* ξ *such that* $\alpha_0 = (\eta_0\,\xi^d)^{-1}$, $\alpha_1 = (\eta_1/\eta_0)\xi^{2-d}$ *and* $\alpha_2 = \xi^{4-d}/\eta_0$. *The Gibbs random field defined in (4.1) is then expressed as follows*

$$H[\mathrm{X}_\lambda] = \frac{1}{2\,\eta_0\,\xi^d}\int \mathbf{ds}\Big\{ \big[\mathrm{X}_\lambda(\mathbf{s}) - E[\mathrm{X}_\lambda(\mathbf{s})]\big]^2 + \eta_1\,\xi^2\,\big[\nabla \mathrm{X}_\lambda(\mathbf{s})\big]^2 + \xi^4\big[\nabla^2 \mathrm{X}_\lambda(\mathbf{s})\big]^2 \Big\}. \quad (23)$$

The scale parameter η_0 determines the total variance of the fluctuations, the coefficient η_1 the shape of the covariance spectral density, and ξ the correlation radius.

REMARK 4.3. The spectral density is expressed in terms of η_0, η_1, ξ as follows

$$\tilde{G}_{\mathrm{x};\lambda}(\mathbf{k}) = \frac{\left|\tilde{Q}_\lambda(\mathbf{k})\right|^2 \eta_0\,\xi^d}{1+\eta_1\,(k\xi)^2 + (k\xi)^4}. \quad (24)$$

The permissibility conditions for the new parameters become:

$\eta_0 > 0,\ \xi > 0$ and $\left\{ \text{(i) } \eta_1 \geq 0, \text{ or (ii) } \eta_1 < 0,\ \eta_1^2 < 4 \right\}$.

4.3 The Covariance function

The covariance function is the simplest measure of spatial dependence and is used below to enforce the sample constraints. The frequency dependence of the primary covariance spectral density is shown in Figure 1 for $\eta_1 > 0$ and in Figure 2 for $\eta_1 < 0$. The spectral density declines fast toward zero, and its tail beyond $k_{\max} = 2\pi/a$ where a is the spatial resolution (taken as the length unit) can be ignored. Hence, the covariance of the primary is a good approximation of the coarse-grained covariance. The maximum of the spectral density for $\eta_1 > 0$ is at $k_m = 0$. For $\eta_1 < 0$ the maximum occurs at $k_m = \xi^{-1}\sqrt{|\eta_1|/2}$ suggesting oscillatory behavior of the covariance. The value of the spectral peak is $\tilde{G}_{\mathrm{x}}(k_m) = 4\eta_0\, \xi^d / (4 - \eta_1^2)$. The curvature at the peak, $K_{\mathrm{G}}(k_m) \equiv \left[d^2 \tilde{G}_{\mathrm{x}}(k) / dk \right]_{k=k_m}$, is $K_{\mathrm{G}}(k_m) = -64\, \eta_0\, \xi^d \left|\eta_1\right| / (4 - \eta_1^2)^2$. The width of the peak, i.e. $\left| \tilde{G}_{\mathrm{x}}(k_m) / K_{\mathrm{G}_{\mathrm{x}}}(k_m) \right| = \left| (4 - \eta_1^2) / 16\eta_1 \right|$, tends to zero as the permissibility boundary is approached. In real space the covariance is given by the following integral, e.g. [54]

$$G_{\mathrm{x}}(\mathbf{r}) = \left(\frac{\eta_0\, \xi^d \sqrt{\vartheta_d}}{r^{d/2-1}}\right) \int_0^\infty dk\, \frac{k^{d/2}\, J_{d/2-1}(k\,r)}{1 + \eta_1 (k\xi)^2 + (k\xi)^4}. \qquad (25)$$

The covariance functions for three ξ>0, which are plotted in Figure 5, show the increase of the correlation radius with ξ. The plots of the covariance function for η_1<0 in Figure

4 exhibit a characteristic negative "hole" the depth of which increases with $|\eta_1|$.

REMARK 4.4. For $\eta_1 = 2$, the spectral density of the primary field is $\tilde{G}_{\mathrm{x}}(\mathbf{k}) = \eta_0\, \xi^d / (1+(k\xi)^2)^2$. In $d=3$ this is equivalent to the spectral density of an exponential covariance $G_{\mathrm{x}}(\mathbf{r}) = \sigma_{\mathrm{x}}^2 \exp(-r/\xi)$ with variance given by $\sigma_x^2 = \eta_0 / 8\pi^{5/2}$.

4.4. A lattice Gibbs random field model

The Gibbs random field model is extended on the lattice as follows:

DEFINITION 4.5. *On a hypercubic lattice generated by the unit vectors* $\hat{\mathbf{e}}_i \equiv (0,...,1_i,..,0)$ *and lattice steps* a_i *(*$i=1,...,d$*), we define probability density functions by means of the following quadratic energy functional*

$$
\begin{aligned}
H[\mathrm{X}_\lambda] = \frac{1}{2\,\eta_0\,\xi^d} \sum_n \Bigg\{ & \left[\mathrm{X}_\lambda(\mathbf{s}_n) - E[\mathrm{X}_\lambda(\mathbf{s}_n)]\right]^2 + \eta_1\,\xi^2 \sum_{i=1}^{d} \frac{\left[\mathrm{X}_\lambda(\mathbf{s}_n + a_i\,\hat{\mathbf{e}}_i) - \mathrm{X}_\lambda(\mathbf{s}_n)\right]^2}{a_i^{\,2}} \\
& + \xi^4 \sum_{i=1}^{d} \left[\frac{\mathrm{X}_\lambda(\mathbf{s}_n + a_i\,\hat{\mathbf{e}}_i) - 2\mathrm{X}_\lambda(\mathbf{s}_n) + \mathrm{X}_\lambda(\mathbf{s}_n - a_i\,\hat{\mathbf{e}}_i)}{a_i^{\,2}}\right]^2 \Bigg\}
\end{aligned} \quad .(26)
$$

REMARK 4.5. The probability density function (26) is anisotropic, even if $a_i \equiv a$ ($i=1,...,d$), due to the anisotropy of the lattice. The coarse-graining scale λ denotes the vector $(\lambda_1,...,\lambda_d)$.

4.5. The lattice covariance function

For the lattice pdf given by Equation (26), the covariance spectral density is estimated by approximating the summation over the lattice sites in the energy functional with an integral. This leads to

$$\tilde{G}_{x;(a_1,\ldots,a_d)}(\mathbf{k}) = \frac{\eta_0\, \xi^d}{1+\eta_1 \sum_{i=1}^{d} \frac{4\xi^2}{a_i^2} \sin^2(\frac{a_i k_i}{2}) + \sum_{i=1}^{d} \frac{16\xi^4}{a_i^4} \sin^4(\frac{a_i k_i}{2})}\,. \tag{27}$$

PROPOSITION 4.3. *The energy functional of Equation* (26) *is permissible if the model parameters satisfy the following relations:*

$$\eta_0 > 0,\ \xi > 0,\ \eta_1 \geq 0\,, \tag{28}$$

or

$$\eta_0 > 0,\ \xi > 0,\ \eta_1 < 0\,,\ \frac{2+d}{\sqrt{2}\,d} > \left|\eta_1\right|. \tag{29}$$

Proof: If the inequalities (28) hold the spectral density is always positive. If $\eta_1 < 0$ Bochner©s theorem requires that $1+\eta_1 \sum_{i=1}^{d} \frac{4\xi^2}{a_i^2} \sin^2(\frac{a_i k_i}{2}) + \sum_{i=1}^{d} \frac{16\xi^4}{a_i^4} \sin^4(\frac{a_i k_i}{2}) > 0$ be positive. Using the definition $\psi_i \equiv \frac{4\xi^2}{a_i^2} \sin^2(\frac{a_i k_i}{2})$, the above inequality is expressed as $1+\sum_{i=1}^{d} \psi_i^2 > \left|\eta_1\right| \sum_{i=1}^{d} \psi_i$. This is satisfied for any ψ_i if $\min\left\{ \varphi(\psi_1,\ldots,\psi_d) \equiv \frac{1+\sum_{i=1}^{d} \psi_i^2}{\sum_{i=1}^{d} \psi_i},\ 0 \leq \psi_i \leq \frac{4\xi^2}{a_i^2} \right\} > \left|\eta_1\right|$. The minimum of the function $\varphi(\psi_1,...,\psi_d)$ is obtained for $\psi_1 = ... = \psi_d = 1/\sqrt{2}$, which leads to the inequality (29).

In Figure 5 we compare the lattice spectral density in the direction $\hat{\mathbf{e}} \equiv (1/\sqrt{2}, 1/\sqrt{2})$ with its continuum counterpart for $\eta_1 = 0.2$. The two functions are almost identical for small k. For larger k the lattice density declines slower, because the sine functions are bounded. Plots of the lattice and continuum spectral densities for a negative shape coefficient ($\eta_1 = -1.2$) are shown in Figure 6. The peak of the lattice density is considerably more pronounced, reflecting the fact that η_1 is closer to the permissibility boundary in the lattice case ($\eta_{1,\min} = -\sqrt{2}$), than in the continuum case ($\eta_{1,\min} = -2$).

5. Estimation of the Model Parameters

Determining the model parameters requires information (observable quantities) that can be reliably estimated from data. In many geostatistical applications it is difficult to estimate reliably the long-range properties [22] of the variogram. Then it makes sense to focus on the short-range moments of the random field and formulate appropriate sample-based constraints.

DEFINITION 5.1. *Let us define the observable energy terms* S_0, S_1 *and* S_2 *as follows:*

$$\mathrm{S}_0 \equiv [\mathrm{X}_\lambda(\mathbf{s})]^2 - \left\{ E[\mathrm{X}_\lambda(\mathbf{s})] \right\}^2. \tag{30}$$

$$\mathrm{S}_1 \equiv \sum_{i=1}^{d} \left[\frac{\mathrm{X}_\lambda(\mathbf{s}+a_i\,\hat{\mathbf{e}}_i) - \mathrm{X}_\lambda(\mathbf{s}-a_i\,\hat{\mathbf{e}}_i)}{2a_i} \right]^2 , \tag{31}$$

$$\mathrm{S}_2 \equiv \sum_{i,j=1}^{d} \Delta_2^{(i)}[\mathrm{X}_\lambda(\mathbf{s})]\,\Delta_2^{(j)}[\mathrm{X}_\lambda(\mathbf{s})]^2 , \tag{32}$$

where $\Delta_2^{(i)}$ *is the centered second difference operator in the direction* $\hat{\mathbf{e}}_i$, *i.e.,*

$$\Delta_2^{(i)}[\mathrm{X}_\lambda(\mathbf{s})] \equiv \frac{\mathrm{X}_\lambda(\mathbf{s}+a_i\,\hat{\mathbf{e}}_i) + \mathrm{X}_\lambda(\mathbf{s}-a_i\,\hat{\mathbf{e}}_i) - 2\,\mathrm{X}_\lambda(\mathbf{s})}{a_i^2}. \tag{33}$$

These random field functionals are used below to enforce sample-based constraints on the Gibbs probability density.

REMARK 5.1. The energy terms S_0, S_1, and S_2 are motivated from the couplings in the energy exponent of the Gibbs pdf. However, there is no a priori reason for choosing these in favor of some other local interaction measures. The fields S_1 and S_2 are defined in terms of centered differences in the domain interior, but forward differences are used at the domain boundaries.

DEFINITION 5.2. *The stochastic expectations* $E[\mathrm{S}_m]$ *(*$m = 0,1,2$*), are given by the following correlations:*

$$E[\mathrm{S}_0] = E\left\{[\mathrm{X}_\lambda(\mathbf{s})]^2\right\} - E\left\{\mathrm{X}_\lambda(\mathbf{s})\right\}^2 = \sigma^2_{\mathrm{x};\lambda}. \tag{34}$$

$$E[\mathrm{S}_1] = \sum_{i=1}^{d} \frac{2\sigma^2_{\mathrm{x};\lambda} - G_{\mathrm{x};\lambda}(2\,a_i\,\hat{\mathbf{e}}_i) - G_{\mathrm{x};\lambda}(-2\,a_i\,\hat{\mathbf{e}}_i)}{4a_i^2}. \tag{35}$$

$$E[\mathrm{S}_2] = \sum_{i,j=1}^{d} E\left\{\Delta_2^{(i)}[\mathrm{X}_\lambda(\mathbf{s})]\,\Delta_2^{(j)}[\mathrm{X}_\lambda(\mathbf{s})]\right\}. \tag{36}$$

LEMMA 5.1. *The expectation* $E[\mathrm{S}_2]$ *is expressed as a double sum that is symmetric over the signs of the arguments (as denoted below by the subscript s)*

$$\begin{aligned} E\left[\mathrm{S}_2\right] = \sum_{i,j=1}^{d} \frac{1}{a_i^2\,a_j^2} \Big\{ & 2G_{\mathrm{x};\lambda}(0) - 2\,G_{\mathrm{x};\lambda}(a_i\,\hat{\mathbf{e}}_i) - 2\,G_{\mathrm{x};\lambda}(a_j\,\hat{\mathbf{e}}_j) + \\ & G_{\mathrm{x};\lambda}(a_i\,\hat{\mathbf{e}}_i - a_j\,\hat{\mathbf{e}}_j) + G_{\mathrm{x};\lambda}(a_i\,\hat{\mathbf{e}}_i + a_j\,\hat{\mathbf{e}}_j)\Big\}_s \end{aligned} \tag{37}$$

Proof. It follows directly from Eq. (33) that

$$\begin{aligned} E[\mathrm{S}_2] &= \sum_{i,j=1}^{d} \sum_{\kappa,\mu=0,\pm1} g_\kappa\, g_\mu \frac{E[\mathrm{X}_\lambda(\mathbf{s} + u_{i\kappa}\,\hat{\mathbf{e}}_i)\,\mathrm{X}_\lambda(\mathbf{s} - u_{j\mu}\,\hat{\mathbf{e}}_j)]}{a_i^2\,a_j^2} \\ &= \sum_{i,j=1}^{d} \sum_{\kappa,\mu=0,\pm1} \frac{g_\kappa\, g_\mu}{a_i^2\,a_j^2} G_{\mathrm{x};\lambda}(u_{i\kappa}\,\hat{\mathbf{e}}_i - u_{j\mu}\,\hat{\mathbf{e}}_j) \end{aligned},$$

where $u_{i,-1} = -a_i$, $u_{i,0} = 0$, $u_{i,1} = a_i$, and $g_{-1} = g_1 = 1$, $g_0 = -2$. Expansion of the double sums proves that $E\left[\mathrm{S}_2\right]$ is given by Equation (37) above.

LEMMA 5.2. *The energy moments are given by the following frequency integrals*

$$E[\mathrm{S}_m] = \vartheta_d\, \eta_0\, \xi^d \int_0^\infty dk \frac{k^{d-1} \left| \tilde{Q}_\lambda(k) \right|^2}{1 + \eta_1 (k\xi)^2 + (k\xi)^4} \left\{ \int d\Omega_d \sum_{i,j=1}^{d} A_{ij}^{(m)}(\mathbf{k}) \right\} \qquad (m = 0, 1, 2). \quad (38)$$

$\int d\Omega_d$ *denotes the integral over the surface of the unit sphere in d dimensions, and the functions* $A_{ij}^{(m)}(\mathbf{k})$ *are as follows:*

$$A_{ij}^{(0)}(\mathbf{k}) = \frac{\delta_{ij}}{d}, \quad (39)$$

$$A_{ij}^{(1)}(\mathbf{k}) = \delta_{ij} \left[\frac{1 - \cos(2\, a_i\, k_i)}{2 a_i^2} \right], \quad (40)$$

$$A_{ij}^{(2)}(\mathbf{k}) = \frac{1}{a_i^2\, a_j^2} \Big[4 - 4\cos(a_i\, k_i) - 4\cos(a_j\, k_j) + 2\cos(a_i\, k_i - a_j\, k_j) + 2\cos(a_i\, k_i + a_j\, k_j) \Big]. \quad (41)$$

Proof: The Equations (38) follow from the Equations (34)-(36) using the Fourier transform of the covariance function. The Equations (40) and (41) for the tensor functions $A_{ij}^{(m)}(\mathbf{k})$ ($m = 1, 2$) follow from the shifting theorem of Fourier transforms.

The Fourier representation leads to explicit expressions for the energy moments in terms of the model parameters. Hence, it is thus more suitable for numerical calculations than the real-space representation.

DEFINITION 5.3. *The expectations* $E[\mathrm{S}_i]$ $(i = 0,1,2)$ *are approximated by the following sample-based constraints (spatial moments),* $\bar{\mathbf{S}} = (\overline{S_0}, \overline{S_1}, \overline{S_2})$

$$\overline{S_0} = \frac{1}{N}\sum_{l=1}^{N}\left\{ X_\lambda(\mathbf{s}_l) - \left[\frac{1}{N}\sum_{l=1}^{N} X_\lambda(\mathbf{s}_l)\right]\right\}^2, \tag{42}$$

$$\overline{S_1} = \frac{1}{N}\sum_{l=1}^{N}\left\{ \sum_{j=1}^{d}\left[\frac{X_\lambda(\mathbf{s}_l + a_j\,\hat{\mathbf{e}}_j) - X_\lambda(\mathbf{s}_l - a_j\,\hat{\mathbf{e}}_j)}{2a_j}\right]^2 \right\}, \tag{43}$$

$$\overline{S_2} = \frac{1}{N}\sum_{l=1}^{N}\left\{ \sum_{j=1}^{d}\left[\frac{X_\lambda(\mathbf{s}_l + a_j\hat{\mathbf{e}}_j) - 2X_\lambda(\mathbf{s}_l) + X_\lambda(\mathbf{s}_l - a_j\hat{\mathbf{e}}_j)}{a_j^2}\right]^2 \right\}. \tag{44}$$

PROPOSITION 5.1. *The optimal parameters* η_1^* *and* ξ^* *of the Gibbs random field are estimated by minimizing the distance metric given by the following functional*

$$\Phi_s[\mathrm{X}_\lambda(\mathbf{s})] \equiv \left| 1 - \sqrt{\frac{\overline{S_1}}{\overline{S_0}}\frac{E[\mathrm{S}_0]}{E[\mathrm{S}_1]}} \right|^2 + \left| 1 - \sqrt{\frac{\overline{S_2}}{\overline{S_1}}\frac{E[\mathrm{S}_1]}{E[\mathrm{S}_2]}} \right|^2. \tag{45}$$

Then, the optimal scaling parameter is given by

$$\eta_0^* = \frac{\overline{S_0}}{E\left[\mathrm{S}_0(\eta_0 = 1, \eta_1^*, \xi^*)\right]}. \tag{46}$$

Comment: The metric $\Phi_s[\mathrm{X}_\lambda(\mathbf{s})]$ is expressed in terms of dimensionless combinations of the energy moments that do not involve the scaling coefficient η_0. This is possible because η_0 multiplies all the energy terms. The application of the distance metric is based on two assumptions: Firstly, the spatial averages $\overline{S_m}$ $(m = 0,1,2)$, should be accurate and precise estimators of the stochastic expectations $E'[\mathrm{S}_m]$ of the underlying random field (where $E'[\cdot]$ denotes the expectation with respect to the unknown probability density, which is approximated by the Gibbs model). This assumes that ergodic conditions are satisfied, e.g., [1,32]. Secondly, the stochastic expectations $E[\mathrm{S}_m]$ of the Gibbs random field should accurately approximate the expectations $E'[\mathrm{S}_m]$ of the underlying field.

REMARK 5.2. The above choice for the distance metric is not the only one possible; in principle any two terms based on dimensional combinations of the energy moments can be used. In general, for an n-component constraint vector **S** the following normalized distance functional can be defined

$$\Phi_s[\mathrm{X}_\lambda(\mathbf{s})] \equiv \sum_{i=1}^{n-1} \left| 1 - \sqrt{\frac{\overline{S_1}}{\overline{S_0}} \frac{E\left[\mathrm{S}_0\right]}{E\left[\mathrm{S}_i\right]}} \right|^2. \tag{47}$$

REMARK 5.3. We choose to minimize the distance functional (45) instead of solving the nonlinear system of equations $\overline{S_m} = E\left[\mathrm{S}_m\right]$ $(m = 0, 1, 2)$, because the existence and uniqueness of a solution for any set of experimentally realizable $\overline{S_m}$ is not obvious. We attempted to solve this system of equations using a globally convergent Newton method [48] and more powerful, trust-region methods [11,12]. The algorithms were eventually trapped inside local minima that gave poor matches with the sample-based constraints.

6. Numerical Implementation

In this section we describe a numerical approach for estimating the model parameters of the Spartan Gibbs random field defined by Equation (23). We consider a simulated control sample on a square lattice with $a_1 = a_2 = a$ and L nodes per side. The kernel of Equation (11) with frequency cutoff $k_{\max} \equiv 2\pi/a$ is used. First, we derive one-dimensional integral expressions for the energy moments based on Equation (38).

6.1 Energy moments in two dimensions

LEMMA 6.1. *Let the frequency functions* $B^{(m)}(k)$ *denote the surface integrals* $B^{(m)}(k) \equiv \int_0^{2\pi} d\phi \sum_{i,j=1}^{d} A_{ij}^{(m)}(\mathbf{k})$. *The functions* $B^{(m)}(k)$ *are then given by*

$$B^{(0)}(k) = 2\pi , \tag{48}$$

$$B^{(1)}(k) = \frac{2\pi}{a^2}\left[\, 1 - J_0(2ak) \,\right], \tag{49}$$

$$B^{(2)}(k) = \frac{2\pi}{a^4}\left[\, 20 - 32\, J_0(a\,k) + 8\, J_0(\sqrt{2}\,a\,k) + 4 J_0(a\,k)\,\right]. \qquad (50)$$

Proof: The $B^{(0)}(k)$ is obvious. The functions $A_{ij}^{(m)}(\mathbf{k})$ ($m = 1,2;\, i,j = 1,2$) are given by

$$A_{11}^{(1)}(\mathbf{k}) = \frac{1-\cos(2\,a\,k\cos\phi)}{2a^2};\; A_{22}^{(1)}(\mathbf{k}) = \frac{1-\cos(2\,a\,k\sin\phi)}{2a^2};\; A_{12}^{(1)}(\mathbf{k}) = A_{21}^{(1)}(\mathbf{k}) = 0\,,$$

$$A_{11}^{(4)}(\mathbf{k}) = \frac{2\,[3+\cos(2\,a\,k\cos\phi) - 4\cos(a\,k\cos\phi)]}{a^4},$$

$$\begin{aligned} A_{12}^{(4)}(\mathbf{k}) = A_{21}^{(4)}(\mathbf{k}) = \frac{2}{a^4}\{&2 - 2\cos(a\,k\cos\phi) - 2\cos(a\,k\sin\phi) \\ &+\cos[a\,k(\sin\phi - \cos\phi)] + \cos[(a\,k(\sin\phi+\cos\phi)]\}\end{aligned},$$

$$A_{22}^{(4)}(\mathbf{k}) = \frac{2\,[3+\cos(2\,a\,k\sin\phi) - 4\cos(a\,k\sin\phi)]}{a^4}.$$

The integrals in the $B^{(m)}(k)$ involve the following integrations [29, pp. 441-442]

$$\int_0^{2\pi} d\phi\, \cos(2\,a\,k\cos\phi) \;=\; \int_0^{2\pi} d\phi\, \cos(2\,a\,k\sin\phi) \;=\; 2\pi\, J_0(2\,a\,k),$$

$$\int_0^{2\pi} d\phi\, \cos[\,a\,k(\cos\phi \pm \sin\phi)\,] = \int_0^{2\pi} d\phi\, \cos[\sqrt{2}\,a\,k\cos(\phi \pm \frac{\pi}{4})] \;=\; 2\pi\, J_0(\sqrt{2}\,a\,k).$$

The Equations (48)-(50) follow from these results. An expansion of the right side of Equations (49) and (50) in a Taylor series verifies that the low frequency behavior is $B^{(1)}(k) \propto k^2$ and $B^{(2)}(k) \propto k^4$, in agreement with the continuum limits for the squares of the gradient and the Laplacian.

6.2. Computational scheme

We simulate data by numerically generating control random field samples. The simulations are based on the Fourier Filtering Method [42] on a $L \times L$ lattice with L ranging from 100 to 800. Three models of covariance dependence commonly used in geostatistics are investigated: Exponential, $G_{\mathrm{x;E}}(r) = \sigma_x^2 \exp(-r/\xi_E)$, hole-type, $G_{\mathrm{x;H}}(\mathbf{r}) = \sigma_x^2 \prod_{i=1}^{2} \sin c(r_i / \sqrt{2}\,\xi_H)$, and Gaussian $G_{\mathrm{x;G}}(r) = \sigma_x^2 \exp(-r^2/\xi_G^2)$. The lengths ξ_E and ξ_G are related to the correlation radius as follows: $\xi_E = -r_c / \log(0.05)$, and $\xi_G = r_c / \sqrt{-\log(0.05)}$. Since the hole-type covariance is anisotropic we extend the definition (5) of the correlation radius to $G_{\mathrm{x}}(r_c/\sqrt{2}, r_c/\sqrt{2}) \equiv 0.05\, G_{\mathrm{x}}(0)$, from which we obtain $\xi_H \cong r_c / 3.60$. We use the correlation radius $r_c = 5$ for all the control samples.

For the optimization we use the Nelder-Mead simplex search algorithm [48], that does not require gradient evaluations. The optimization is terminated when both the model parameters (η_0, η_1, ξ) and the distance metric change between consecutive steps less than specified tolerances, for which we use the value 10^{-6}. The initial guesses for the model parameters are $\eta_1^{(0)} = -1$ and $\xi^{(0)} = 3$ for all the runs. We calculate the frequency integrals of the energy moments with an adaptive-recursive Newton-Cotes 8-panel algorithm [18]. We impose a cutoff, $\eta_{1,co} = -1.99$, on the shape parameter to avoid proximity to the permissibility boundary and numerical instabilities. The cutoff is implemented by assigning a large value (e.g., 10^6) to the distance metric if η_1 becomes smaller than the cutoff. The numerical investigations were conducted on a desktop computer with 800 MHz processor and 256MB RAM.

6.3 Results

We minimize the normalized distance metric (45) using the scheme outlined above. In Table 1 we list the minimum values of the distance metric, the number of iterations of the Nelder-Mead algorithm, the total CPU time (including the optimization and the calculation of the sample moments), and the estimated Gibbs model parameters. The same set of random numbers was used to generate all three control samples. The energy moments of the optimal Gibbs random field are an excellent approximation of the sample with the exponential dependence, as evidenced by the extremely low value of the distance metric. Estimation of the optimal model parameters requires less CPU time if the control sample has exponential dependence than if it has Gaussian dependence, even though more Nelder-Mead iterations are needed in the former case. Increased computational effort is required for calculating the moment integrals for negative η_1. This slow-down affects the optimization time for the Gaussian sample, the shape coefficient of which is attracted towards the permissibility boundary. Since η_1 is positive for the sample with exponential dependence, less time per iteration is required in spite of the higher number of iterations. The value of the distance metric is also low for the samples with Gaussian and hole-type dependence. However, this is achieved at the expense of a η_1^* very close to the permissible boundary, especially for the hole-type covariance. The negative η_1^* implies that the covariance of the optimal Gibbs random field has a negative hole, a feature absent from the covariance of the control sample with Gaussian dependence. However, the constraints we imposed do not consider the entire covariance function, so this deviation should not be surprising. The proximity of η_1^* to the permissibility boundary may indicate that (i) ergodic conditions are not satisfied or (ii) the specific Spartan Gibbs energy functional or smoothing kernel permits an accurate matching of the constraints only by means of an almost singular spectral density. The issue (i) is discussed below. Competition between the smoothness of the solution and its ability to match the data is typical of many inverse problems [48, p. 796].

In Table 2 we list the three energy moments for each of the control samples and their respective values for simulated states of the optimal Gibbs random field with the parameters from Table 1. There is good agreement in general, except in the case of hole-type dependence, which presents a 14% difference between the standard deviation of the control sample and the Gibbs state. This is most likely due to ergodic fluctuations that are stronger for hole-type dependence due to the non-monotonic decline, than for the other two covariance models. To test this hypothesis we derive explicit expressions for the variance of the Spartan Gibbs random field in two dimensions. We evaluate the spectral density integral by approximating the upper limit, $k_{\max}\,\xi$, with infinity. This leads to the following

$$\sigma_{\mathrm{x}}^{2}=\begin{cases}\frac{\eta_{0}}{4\pi\sqrt{\eta_{1}^{2}-4}}\ln\left(\frac{\eta_{1}+\sqrt{\eta_{1}^{2}-4}}{\eta_{1}-\sqrt{\eta_{1}^{2}-4}}\right), & |\eta_{1}|>2\\ \frac{\eta_{0}}{2\pi\sqrt{4-\eta_{1}^{2}}}\left(\frac{\pi}{2}-\tan^{-1}\frac{\eta_{1}}{\sqrt{4-\eta_{1}^{2}}}\right), & |\eta_{1}|<2\,.\\ \frac{\eta_{0}}{2\pi\,\eta_{1}}, & |\eta_{1}|=2\end{cases} \qquad (51)$$

Using the optimal parameters for hole-type dependence, i.e. $\eta_1^* = -1.99$ and $\eta_0^* = 0.38$ in the second of the Equations (51), we obtain $\sigma_{\mathrm{x}}^2 = 0.92$, in excellent agreement with the variance of the control sample. In further support of the ergodic effect, we repeat the calculations for a larger system size, *L*=200. The results are listed in Table 3 and Table 4. It is seen from Table 3 that η_1^* increases slightly to —1.98, and from Table 4 that the sample variance of the Gibbs state, i.e. 0.90, is closer to the value for the control sample, i.e., 1.00. We have also checked that the trends shown in Tables 1-4 persist if we generate 100 samples for each type of dependence and calculate the average values of the parameters over the population.

Next, we performed the optimization for control samples on lattices with size that varies between 100 and 800 nodes per side. The CPU time is plotted versus the number of Nelder-Mead iterations in Figure 7. The graphs show that the CPU time in general increases with the number of iterations, but there is also significant scattering. As pointed out above, this is due to the dependence of the computational cost per iteration on the shape parameter. The optimization is subject to this slow-down more for control samples with Gaussian and hole-type dependence than for exponential dependence. The dependence of the CPU time on the number of iterations implies that there is significant dependence on the initial values of the model parameters. Finally, in Figure 8 we plot the

CPU time versus the sample size. The graphs show that there is no systematic dependence on the sample size. The CPU time quoted includes calculation of the constraints, which is an $O(L^2)$ operation. This explains the slight increase in the CPU time for the control samples with exponential and Gaussian dependence. In the case of hole-type dependence, most of the CPU time is spent on the optimization that dominates any size-related increase of the time required for constraint calculation.

7. Future Directions

We presented a general framework for investigating spatial dependence based on the Spartan Gibbs probability densities given by the Equation (16). The detailed analysis focused on determining the parameters for a specific model from this class given by Equation (23). Other issues of relevance to geostatistical applications that require further study include estimation of the field at non-sampled locations, simulation, anisotropic statistical dependence, non-Gaussian probability densities, and off-lattice distributions.

Once the model parameters have been determined, estimation of the field at a non-sampled location $\mathbf{s}_m$ is straightforward: the optimal value $\mathrm{X}_\lambda^*(\mathbf{s}_m)$ is obtained by maximizing the conditional probability density $f_{\mathrm{x}}[\mathrm{X}_\lambda(\mathbf{s}_m)|\mathbf{X}_\lambda(S)]$, where $\mathbf{X}_\lambda(S)$ is the vector representing the field sample at the locations in the observation set S. Since $f_{\mathrm{x}}[\mathrm{X}_\lambda(\mathbf{s}_m)|\mathbf{X}_\lambda(S)] = f_{\mathrm{x}}[\mathrm{X}_\lambda(\mathbf{s}_m) \cap \mathbf{X}_\lambda(S)] / f_{\mathrm{x}}[\mathbf{X}_\lambda(S)]$, this is equivalent to maximizing the joint density $f_{\mathrm{x}}[\mathrm{X}_\lambda(\mathbf{s}_m) \cap \mathbf{X}_\lambda(S)]$, and therefore to solving the linear equation

$$\left. \frac{\partial H[X_\lambda(\mathbf{s})]}{\partial \mathrm{X}_\lambda(\mathbf{s}_m)} \right|_{\mathrm{X}_\lambda^*(\mathbf{s}_m)} = 0. \qquad (52)$$

Unconditional simulation can be conducted with any of the standard techniques for multivariate normal distributions [7,14,22,27,32,42]. Due to the localized nature of the coupling it is computationally efficient to calculate Monte Carlo transition probabilities between different states [30,45], without calculating the inverse of very large covariance matrices. For the same reason, conditional simulations can also be performed efficiently with the simulated annealing method [48,17,26,28] or with the recently proposed energy-landscape-paving global optimization approach [31]. This approach is a low-temperature Monte Carlo method with memory that modifies the statistical weight of each state according to the number of times it has been visited. The memory enables the algorithm to escape from local minima and to explore efficiently the parameter space. The method has not yet been tested in geostatistical simulations, but it outperformed simulated annealing in two protein folding problems.

The definition of the Gibbs random fields can be extended to include anisotropic correlations by introducing direction-dependent correlation parameters. In the case of global statistical anisotropy (that is for uniform orientation of the principal anisotropy axes in the sampled domain), estimation of the model parameters can be reduced to an equivalent isotropic problem. A method for identifying the direction of the principal axes and the ratios of the correlation parameters is presented in [34]. Non-Gaussian probability densities can be obtained by adding higher-order terms in the energy functional. Non-Gaussian probability densities, such as the one based on the Landau model, e.g. [25], can even exhibit a transition between regimes of exponential and power-law dependence. Approximations must be used in the calculation of the energy moments, because general, closed-form expressions for the covariance spectral density of non-Gaussian fields do not exist. Nonetheless, approximations can be obtained with various methods of many-body theories, e.g. [49]. One approach is based on variational techniques [19,3,35], which expand the non-Gaussian probability density around an optimal Gaussian. The variational spectral density can then be used as a zero-point approximation for a diagrammatic perturbation expansion of the spectral density in terms of self-energy operators [44,36]. Samples with off-lattice distributions can also be modeled using Equation (16). The main difference between regularly spaced samples and off-lattice distributions is the estimation of the energy moments. A different approach, which was not pursued here, is to estimate the model parameters by means of the ML method. This approach does not require estimating the energy moments, and unlike the standard application of the ML method that involves the entire covariance matrix, it does not suffer from the dimensionality problem.

8. Conclusions

We investigated Gibbs random fields with Spartan parameter spaces as models of spatial dependence. We defined a normalized distance between experimental moments (constraints) and the respective stochastic moments of the Gibbs probability density model, and we proposed solving the inverse problem of determining the model parameters by minimizing the distance functional. We calculated the parameters for a specific Gibbs random field model with local coupling, using simulated control samples to provide the constraints. We showed that the parameter estimation problem is solved very efficiently numerically, and that the computational complexity of the optimization is independent of the sample size. Some generalizations of the method and applications to estimation and conditional simulation were briefly discussed. The proposed approach is well suited for modeling the spatial dependence of large samples, especially if the full covariance function can not be reliably estimated from the sample.

REFERENCES

1. R. J. ADLER, *The Geometry of Random Fields*, John Wiley and Sons, New York, 1981.

2. G. ALLAIRE, *One-Phase Newtonian Flow*, in *Homogenization and Porous Media*, U. Hornung, ed., Springer Verlag, New York, NY, 1996, pp. 45-70.

3. M. BARTHELEMY, H. ORLAND AND G. ZERAH, *Propagation in random media: Calculation of the effective dispersive permittivity by use of the replica method*, Phys. Rev. E, 52(1) (1995), pp. 1123-1127.

4. S. BOCHNER, *Lectures on Fourier Integrals*, Princeton University Press, Princeton, NJ, 1959.

5. BOURGEAT, *One-Phase Newtonian Flow*, in *Homogenization and Porous Media*, U. Hornung, ed., Springer Verlag, New York, 1996, pp. 95-125.

6. P. M. CHAIKIN, and T. C. LUBENSKY, *Principles of Condensed Matter Physics*, Cambridge University Press, Cambridge, U.K, 1995.

7. G. CHRISTAKOS, *Random Field Models in Earth Sciences*, Academic Press, San Diego, CA, 1992.

8. G. CHRISTAKOS, *Modern Spatiotemporal Geostatistics*, Oxford University Press, New York, 2000.

9. G. CHRISTAKOS and D. T. HRISTOPULOS, *Spatiotemporal Environmental Health Modelling: A Tractatus Stochasticus*, Kluwer Academic Publishers, Boston, MA, 1998.

10. G. CHRISTAKOS, D. T. HRISTOPULOS and P. BOGAERT, *On the Physical Geometry Concept at the Basis of Space/Time Geostatistical Hydrology*, Adv. Water Resources, 23(8) (2000), pp. 799-810.

11. T. F. COLEMAN AND Y. LI, *On the convergence of reflective Newton methods for large-scale nonlinear minimization subject to bounds*, Mathem. Programming, 67(2) (1994), pp. 189-224.

12. T. F. COLEMAN AND Y. LI, *An interior, trust-region approach for nonlinear minimization subject to bounds*, SIAM J. Optimization, 6(4) (1996), pp. 1040-1058.

13. N. CRESSIE AND D. M. HAWKINS, *Robust estimation of the variogram I*, Math. Geology, 12 (1980), pp. 115-125.

14. N. CRESSIE, *Statistics for Spatial Data*, John Wiley, 2nd Edition, New York, 1993.

15. J. H. CUSHMAN, *On unifying the concepts of scale, instrumentation and stochastics in the development of multiphase transport theory*, Water Resour. Res., 20(11) (1984), pp. 1668-1676.

16. J. H. CUSHMAN, *On measurement, scale and scaling*, Water Resour. Res. 22(2) (1986), pp. 129-134.

17. C. V. DEUTSCH AND P. COCKERHAM, *Practical considerations in the application of simulated annealing to stochastic simulation*, Math. Geol., 26 (1994), pp. 67-82.

18. J. R. DORMAND AND P. J. PRINCE, *A family of embedded Runge-Kutta formulae*, J. Comp. Appl. Math., 6 (1980), pp. 19-26.

19. P. FEYNMAN, *Statistical Mechanics*, Benjamin/Cummings, Reading, MA, 1982.

20. J. B. FOURNIER, A. ADJARI AND L. PELITI, *Effective-area elasticity and tension of micromanipulated membranes*, Phys. Rev. Let., 86(21) (2001), pp. 4970-4974.

21. M. GEL'FAND AND G. E. SHILOV, *Generalized Functions Vol. I*, Academic Press, New York, 1964.

22. L. W. GELHAR, *Stochastic Subsurface Hydrology*, Prentice Hall, Englewood Cliffs, NJ, 1993.

23. S. GEMAN AND D. GEMAN, *Stochastic Relaxation, Gibbs distribution and Bayesian reconstruction of images*, IEEE Trans. Patt. Anal. Mach. Int., 6 (1984), pp. 721-741.

24. W. GIBBS, *Elementary Principles in Statistical Mechanics, Vol. II of Collected Works*, Longmans Green and Company, New York, 1928.

25. N. GOLDENFELD, *Lectures on Phase Transitions and the Renormalization Group*, Addison Wesley, Reading, MA, 1992.

26. P. GOOVAERTS, *Accounting for estimation optimality criteria in simulated annealing*, Math. Geol., 30 (1998), pp. 511-534.

27. P. GOOVAERTS, Geostatistics for Natural Resources Evaluation, Oxford Univ. Press, New York, 1997.

28. P. GOOVAERTS, *Combining minimum error variance and spatial variability in the modeling of petrophysical properties*, in Proceedings of the 5th Annual Conference of the International Association for Mathematical Geology, S. J. Lippard, A. Naess and R. Sinding-Larsen, eds., Trodheim, Norway, 1999, Vol. 1, p. 57-62.

29. S. GRADSHTEYN AND I. M. RYZHIK, *Table of Integrals, Series and Products*, Academic Press, 5th edition, London, England, 1994, pp. 72, 80.

30. J. M. HAMMERSLEY AND D. C. HANDSCOMB, *Monte Carlo Methods*, Methuen, London, 1964.

31. U. H. E. HANSMANN AND L. T. WILLE, *Global optimization by energy landscape paving*, Phys. Rev. Let., 88(6) (2002), pp. 068105-(1-3).

32. D. B. HERNÇNDEZ, *Lectures on Probability and Second Order Random Fields,* World Scientific, River Edge, NY, 1995.

33. U. HORNUNG, *Introduction*, in *Homogenization and Porous Media*, U. Hornung, ed., Springer Verlag, New York, 1996.

34. D. T. HRISTOPULOS, *New anisotropic covariance models and estimation of anisotropic parameters based on the covariance tensor identity*, Stoch. Environmental Res. and Risk Assessment, 16(1) (2002), pp. .

35. D. T. HRISTOPULOS AND G. CHRISTAKOS, *A variational calculation of the effective fluid permeability of heterogeneous media*, *Phys. Rev. E,* 55(6) (1997), pp. 7288-7298.

36. D. T. HRISTOPULOS AND G. CHRISTAKOS, *Practical calculation of non-Gaussian multivariate moments in spatiotemporal Bayesian maximum entropy analysis*, Math. Geol., 33(5) (2001), pp. 543-568.

37. E. H. ISAAKS AND R. M. SRIVASTAVA, *An Introduction to Applied Geostatistics*, Oxford University Press, New York, 1989.

38. M. B. ISICHENKO, *Percolation, statistical topography, and transport in porous media*, Rev. Mod. Phys., 64(4) (1992), pp. 961-1043.

39. T. JAYNES, *Information theory and statistical mechanics*, Phys. Rev. 106 (1957), pp. 620-630.

40. A. G. JOURNEL AND C. J. HUIJBREGTS, *Mining Geostatistics*, Academic Press, London, England, 1978.

41. P. K. KITANIDIS, *Statistical estimation of polynomial generalized covariance functions and hydrologic applications*, Water Resour. Research, 19 (1983), pp. 909-921.

42. H. A. MAKSE, S. HAVLIN, M. SCHWARTZ AND H. E. STANLEY, *Method for generating long-range correlations for large systems*, Phys. Rev. E 53(5) (1996), pp. 5445-5449.

43. K. V. MARDIA AND R. J. MARSHALL, *Maximum likelihood estimation of models for residual covariance in spatial regression*, Biometrika 71 (1984), pp. 289-295.

44. R. D. MATTUCK, *A Guide to Feynman Diagrams in the Many-Body Problem*, McGraw Hill, New York, 1976.

45. N. METROPOLIS, A. W. ROSENBLUTH, M. N. ROSENBLUTH, A. N. TELLER AND E. TELLER, *Equation of state calculations by fast computing machines*, J. Chem. Phys. 21 (1953), pp. 1087-1092.

46. D. MUMFORD, *The statistical description of visual signals*, in Proceedings of the Third International Congress on Industrial and Applied Mathematics, 1995, K. Kirchgassner, O. Mahrenholtz, and R. Mennicken, ed., Akademie Verlag, Hamburg, Germany, 1996.

47. E. PARDO-IGờZQUIZA AND P. A. DOWD, *VARIOG2D: a computer program for estimating the semi-variogram and its uncertainty*, Computers & Geosciences, 27 (2001), pp. 549-561.

48. W. H. PRESS, S. A. TEUKOLSKY, W. T. VETTERING AND B. P. FLANNERY, *Numerical Recipes in Fortran, Vol. 1*, 2nd edition, Cambridge University Press, New York, 1992.

49. G. RICKAYZEN, *GreenÕs Functions and Condensed Matter*, Academic Press, Toronto, Canada, 1980, p. 322.

50. R. L. SMITH, *Spatial Statistics in Environmental Science*, in Nonlinear and Nonstationary Signal Processing, W. J. Fitzgerald, R. L. Smith, A. T. Walden and P. C. Young, eds., Chapter 5, Cambridge University Press (2000), pp. 152-183.

51. R. L. SMITH, *COMS Course in Environmental Statistics*, Chapter 2, http://www.stat.unc.edu/postscript/rs/envstat/env.html, (2000), pp.61-68.

52. N. WATSON, *A Treatise on the Theory of Bessel Functions*, Cambridge University Press, 2nd edition, New York, 1980.

53. G. WINKLER, *Image Analysis, Random Fields and Dynamic Monte Carlo Methods: A Mathematical Introduction*, Springer-Verlag, New York, 1995.

54. M. YAGLOM, *Correlation Theory of Stationary and Related Random Functions I*, Springer-Verlag, New York, 1987.

TABLES

Covariance Model	Φ_s^*	N_{iter}	T_{cpu} (sec)	η_0^*	η_1^*	ξ^*
RF with Gaussian covariance	6.6×10^{-14}	59	8.5	0.82	–1.96	1.56
RF with Exponential covariance	8.3×10^{-18}	89	7.3	26.12	8.94	0.77
RF with Hole covariance	1.35×10^{-6}	115	14.9	0.38	–1.99	1.73

Table 1: Values of the minimum distance metric, Φ_s^* the number of iterations N_{iter}, the CPU time, T_{cpu}, and of the optimal Gibbs model parameters $(\eta_0^*, \eta_1^*, \xi^*)$ for L=100.

Covariance Model	$\overline{S_0}$	$\overline{S_1}$	$\overline{S_2}$
Sample with Gaussian covariance	0.93	0.37	0.40
Optimal Gibbs RF	0.85	0.35	0.36
Sample with Exponential covariance	0.96	0.73	7.67
Optimal Gibbs RF	0.80	0.61	5.01
Sample with Hole covariance	0.92	0.30	0.19
Optimal Gibbs RF	0.68	0.23	0.16

Table 2: Energy moments of the control samples and the Gibbs probability density states based on the model parameters shown in Table 1.

Covariance Model	Φ_s^*	N_{iter}	T_{cpu} (sec)	η_0^*	η_1^*	ξ^*
RF with Gaussian covariance	1.1×10^{-13}	65	10.6	0.88	-1.96	1.60
RF with Exponential covariance	3.0×10^{-17}	91	7.3	29.20	10.47	0.85
RF with Hole covariance	4.6×10^{-14}	63	11.6	0.66	-1.98	1.74

Table 3: Values of the minimum distance metric Φ_s^* the number of iterations N_{iter}, the CPU time, T_{cpu}, and of the optimal Gibbs model parameters $(\eta_0^*, \eta_1^*, \xi^*)$ for L=200.

Covariance Model	$\overline{S_0}$	$\overline{S_1}$	$\overline{S_2}$
Sample with Gaussian covariance	0.98	0.38	0.39
Optimal Gibbs RF	0.96	0.37	0.36
Sample with Exponential covariance	0.99	0.70	7.21
Optimal Gibbs RF	0.86	0.60	4.70
Sample with Hole covariance	1.00	0.33	0.26
Optimal Gibbs RF	0.90	0.30	0.23

Table 4: Energy moments of the control samples and the Gibbs probability density states based on the model parameters shown in Table 3.

FIGURE CAPTIONS

Figure 1: Plot of the Spartan Gibbs spectral density for $\eta_0 = 1$, $\xi = 1$, $\eta_1 > 0$.

Figure 2: Plot of the Spartan Gibbs spectral density for $\eta_0 = 1$, $\xi = 1$, $\eta_1 < 0$.

Figure 3: Covariance function for positive shape coefficient ($\eta_1 = 1$).

Figure 4: Covariance function for negative values of the shape coefficient.

Figure 5: Plots of the continuum and lattice spectral densities for $\eta_0 = 1$, $\xi = 1$ and $\eta_1 = 0.2$

Figure 6: Plots of the continuum and lattice spectral densities for $\eta_0 = 1$, $\xi = 1$ and $\eta_1 = -1.2$.

Figure 7: Plot of the CPU time (in seconds) versus the number of iterations of the Nelder-Mead search algorithm on a square lattice with *L*=100 nodes per side.

Figure 8: Plot of the CPU time (in seconds) versus the sample size (in nodes per side).

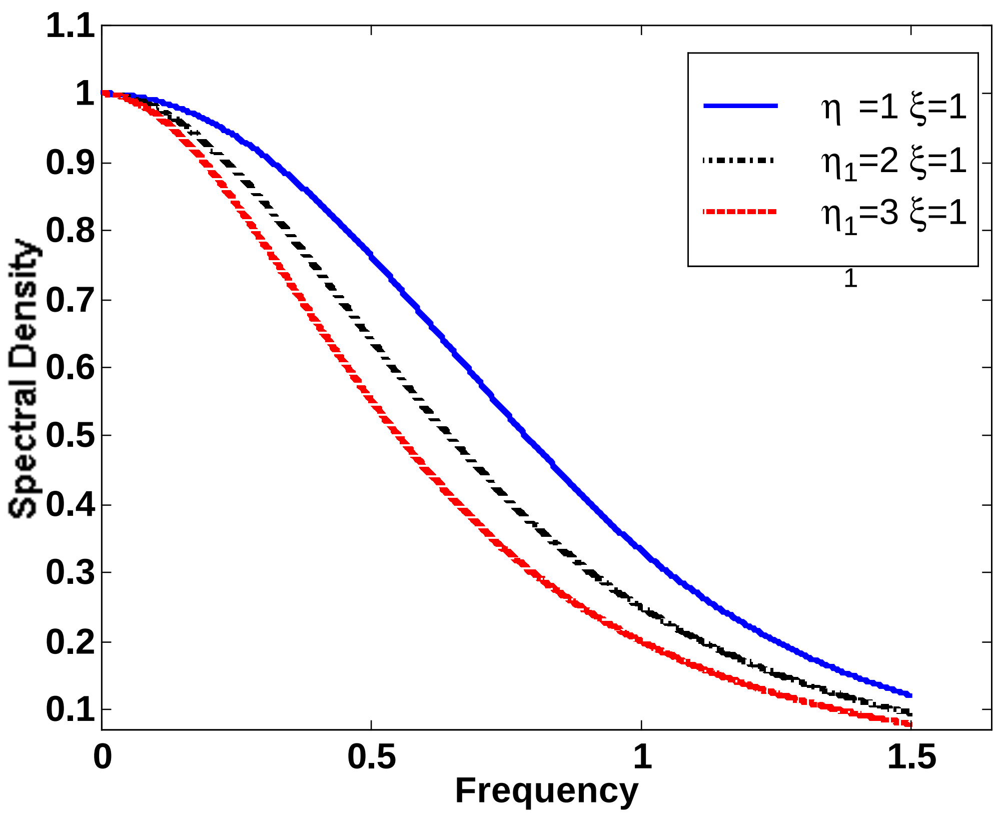
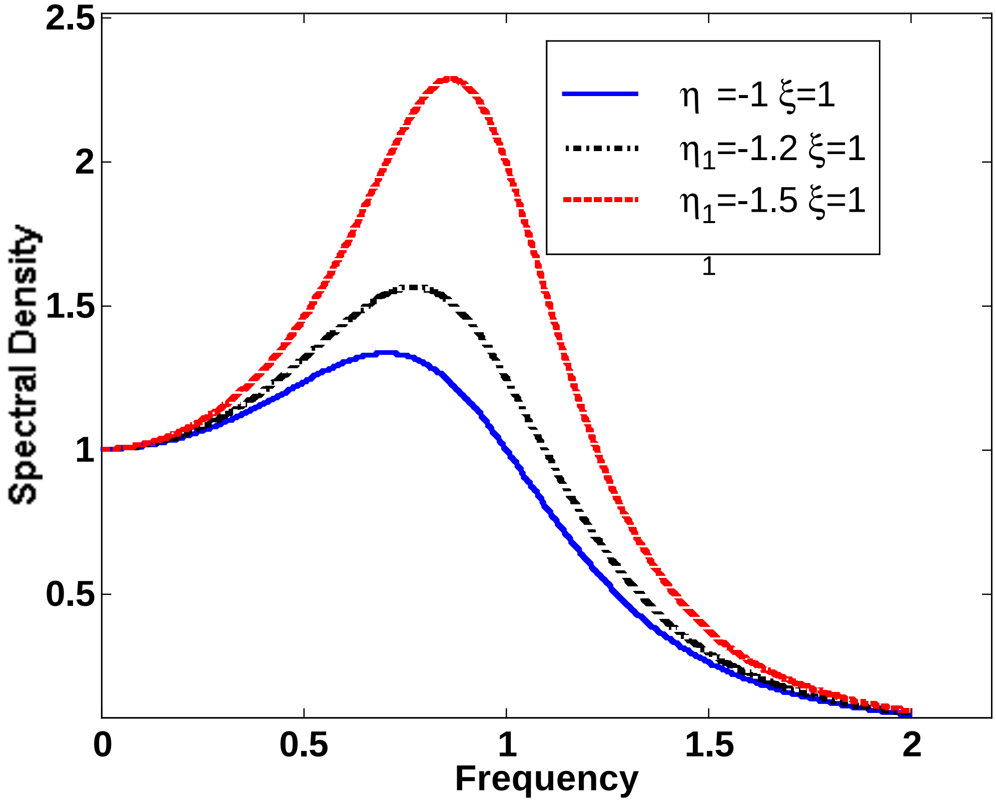

1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
Spectral Density
0
0.5
1
1.5
Frequency
η =1 ξ=1
η1=2 ξ=1
η1=3 ξ=1
1

Figure 1



2.5
2
1.5
1
0.5
Spectral Density
0
0.5
1
1.5
2
Frequency
η =-1 ξ=1
η1=-1.2 ξ=1
η1=-1.5 ξ=1
1

Figure 2

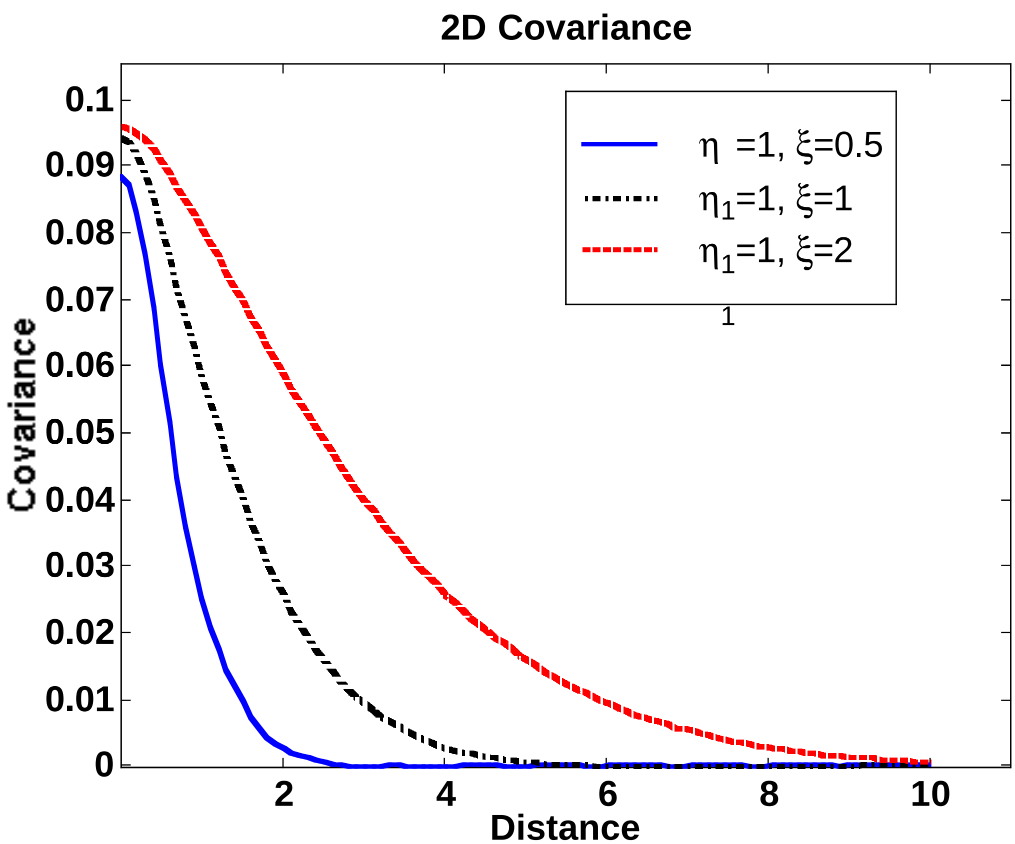

2D Covariance
η =1, ξ=0.5
η1=1, ξ=1
η1=1, ξ=2
1
0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
Covariance
2
4
6
8
10
Distance

Figure 3

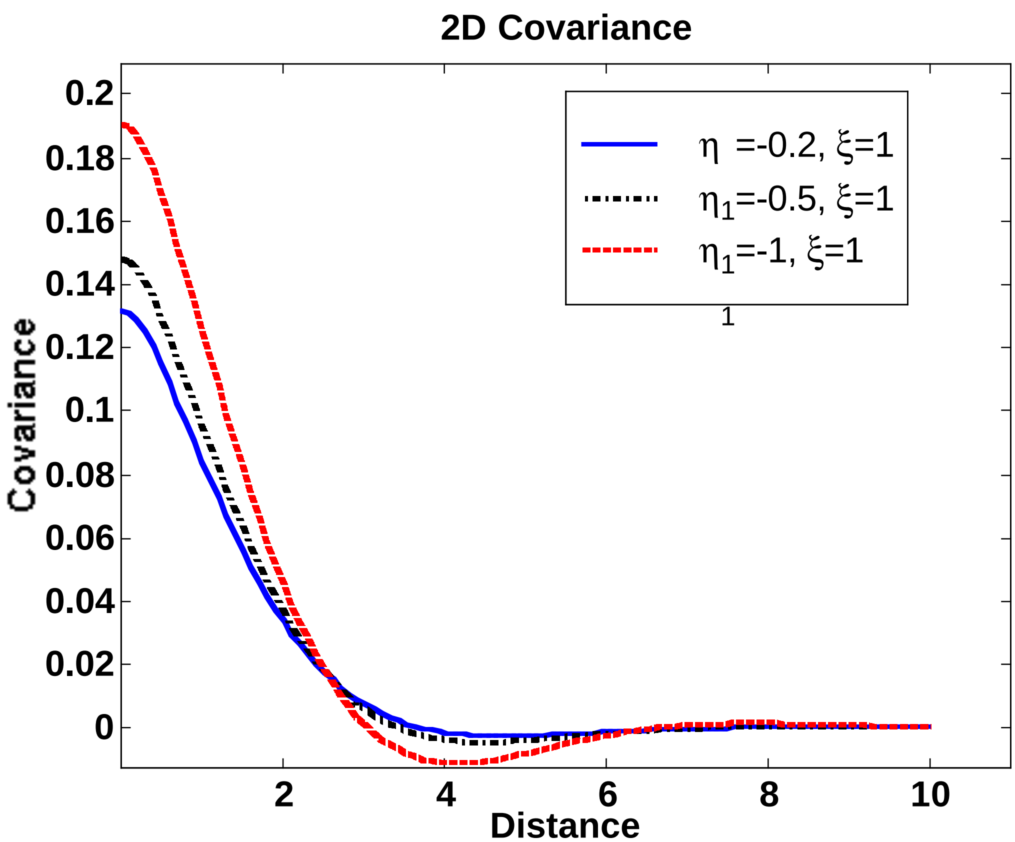
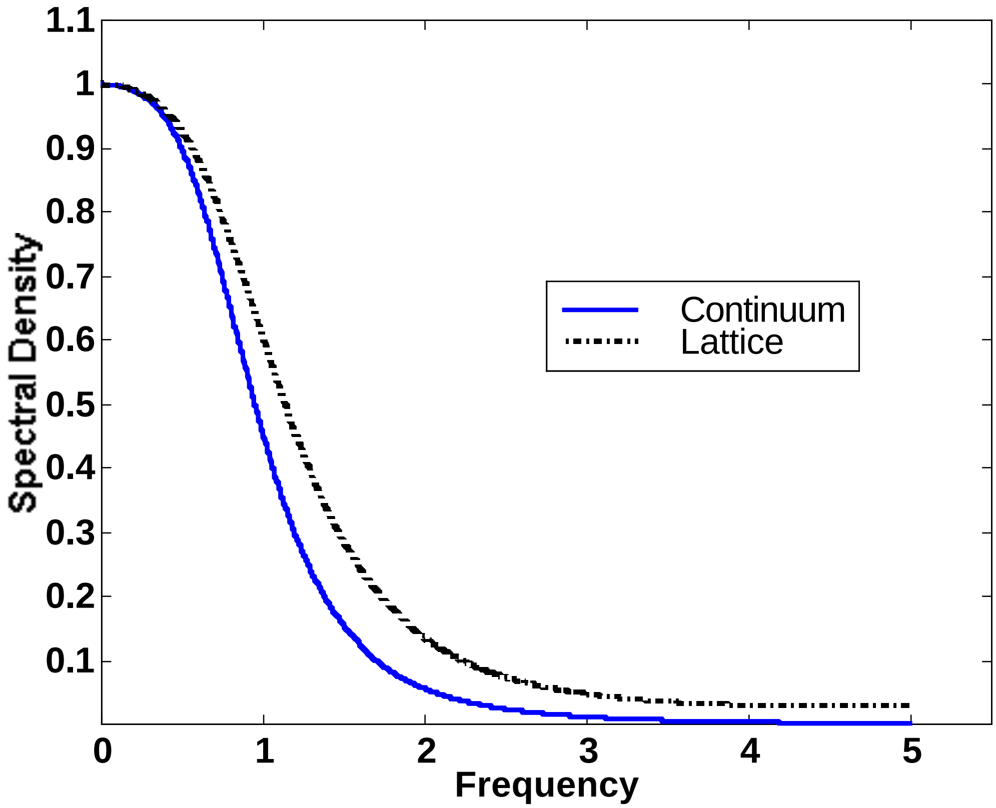
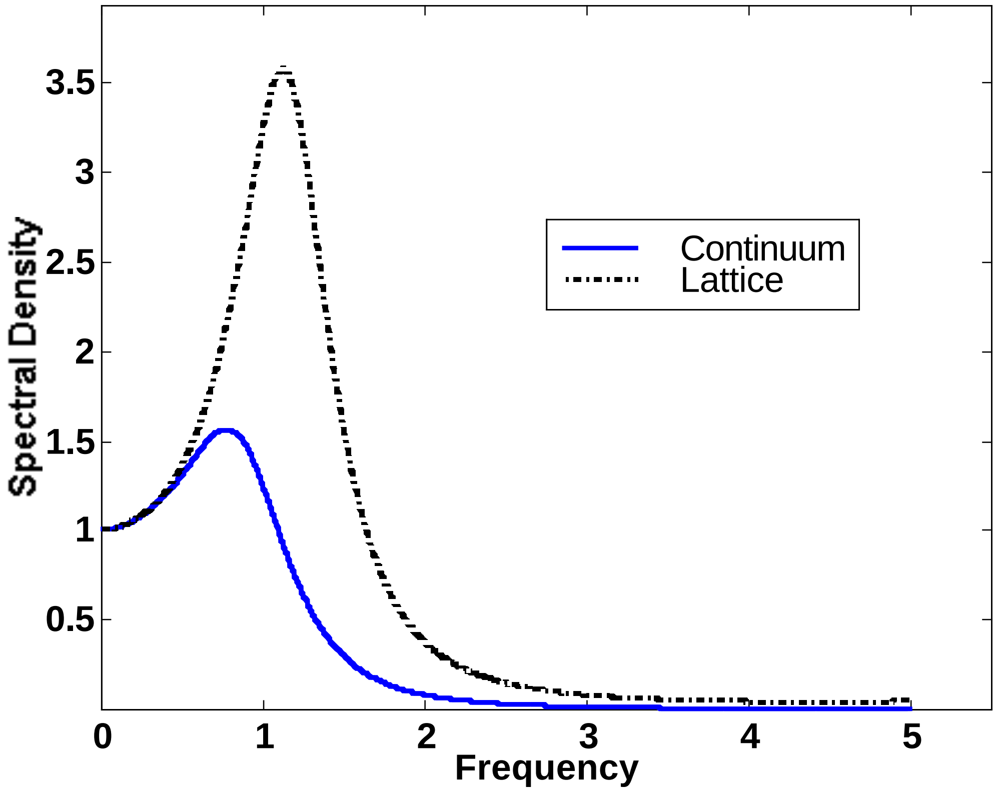

2D Covariance
0.2
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
Covariance
η =-0.2, ξ=1
η1=-0.5, ξ=1
η1=-1, ξ=1
1
2
4
6
8
10
Distance

Figure 4



1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
Spectral Density
0
1
2
3
4
5
Frequency
Continuum
Lattice

Figure 5



Spectral Density
3.5
3
2.5
2
1.5
1
0.5
0
1
2
3
4
5
Frequency
Continuum
Lattice

Figure 6

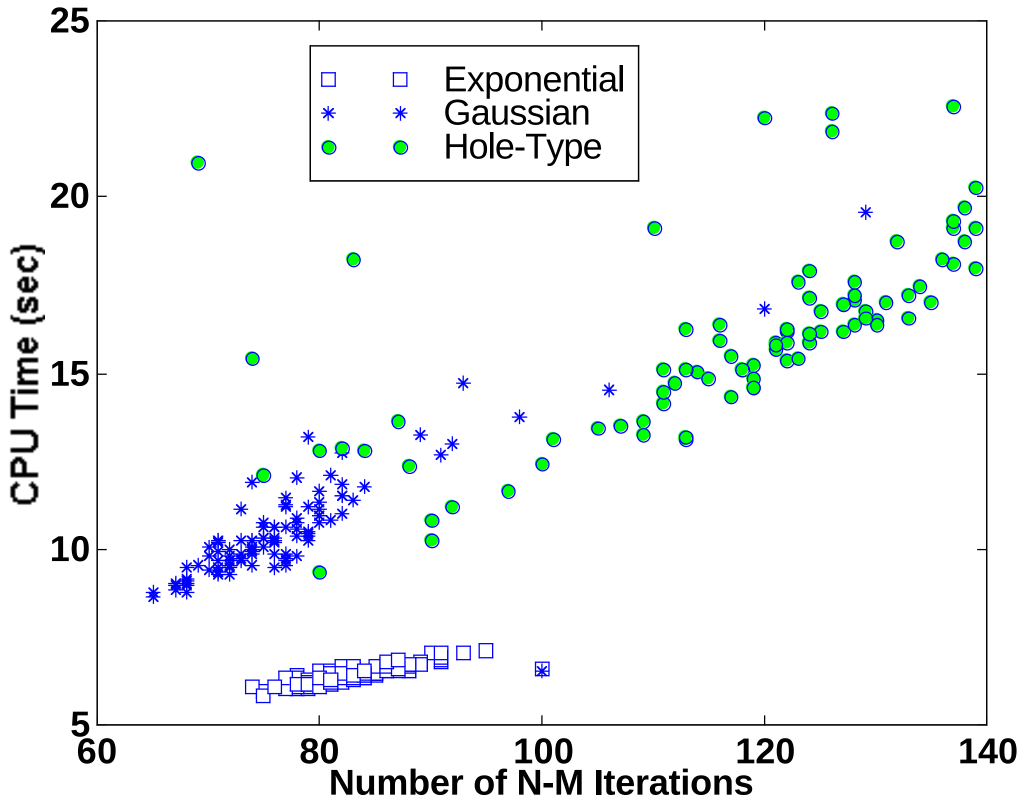
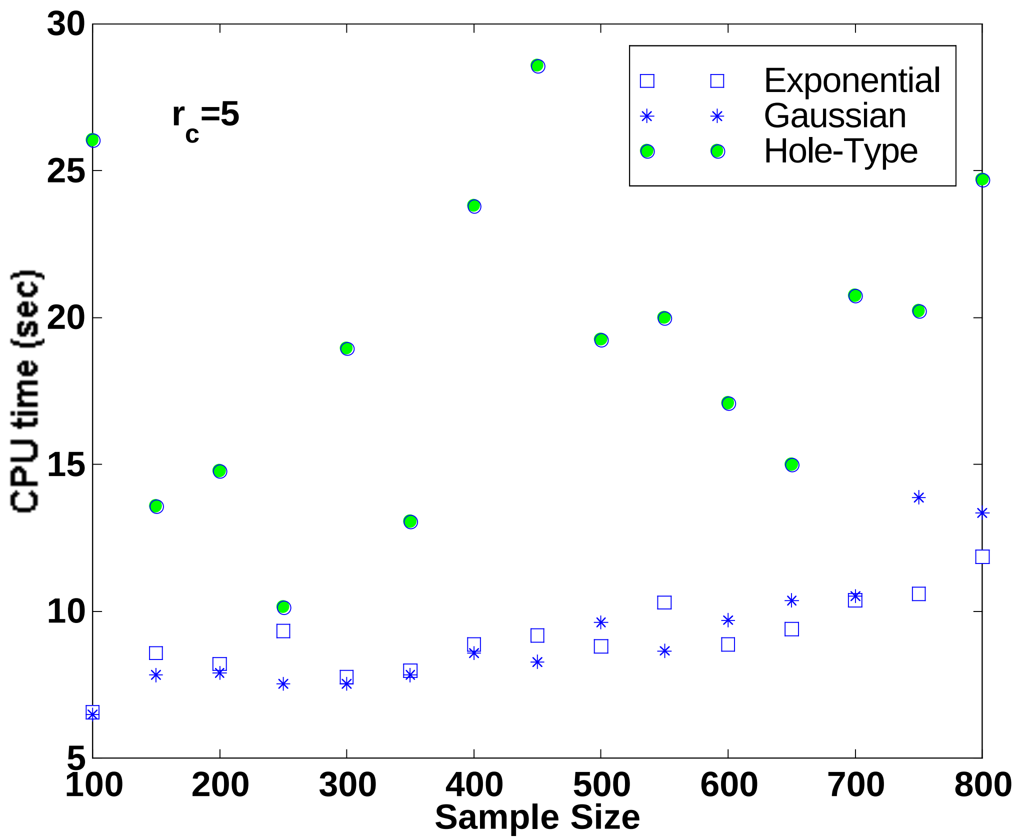

25
20
15
10
5
60
80
100
120
140
CPU Time (sec)
Number of N-M Iterations
Exponential
Gaussian
Hole-Type

Figure 7



30
25
20
15
10
5
100
200
300
400
500
600
700
800
CPU time (sec)
Sample Size
r c=5
Exponential
Gaussian
Hole-Type

Figure 8