

X. Emery

## Properties and limitations of sequential indicator simulation

**Abstract** The sequential indicator algorithm is a widespread geostatistical simulation technique that relies on indicator (co)kriging and is applicable to a wide range of datasets. However, such algorithm comes up against several limitations that are often misunderstood. This work aims at highlighting these limitations, by examining what are the conditions for the realizations to reproduce the input parameters (indicator means and correlograms) and what happens with the other parameters (other two-point or multiple-point statistics). Several types of random functions are contemplated, namely: the mosaic model, random sets, models defined by multiple indicators and isofactorial models. In each case, the conditions for the sequential algorithm to honor the model parameters are sought after. Concurrently, the properties of the multivariate distributions are identified and some conceptual impediments are emphasized. In particular, the *prior* multiple-point statistics are shown to depend on external factors such as the total number of simulated nodes and the number and locations of the samples. As a consequence, common applications such as a flow simulation or a change of support on the realizations may lead to hazardous interpretations.

**Keywords** Nonparametric geostatistics · Indicator kriging · Mosaic model · random sets · isofactorial models · multiple-point statistics

### 1 Introduction

The sequential indicator algorithm

Stochastic simulations are increasingly used for uncertainty modeling in geosciences. Among the vast toolbox

of available methods, the sequential algorithm is undoubtedly one of the most widespread in geostatistical applications, for its conceptual simplicity and straightforwardness. In such algorithm, the values are simulated in turn, conditionally to the original data and the previously simulated values. In case of a Gaussian random field, the conditional distributions are Gaussian-shaped and their first moments are derived from a simple kriging system. However, when a nongaussian random function has to be simulated (for instance, a categorical variable representing a lithological facies, or a continuous variable with spatially correlated extreme values, so that a Gaussian model is ill-suited), the conditional distributions usually cannot be expressed analytically, which raises a problem in the sequential algorithm. A shortcut consists in estimating these conditional distributions by a nonparametric method, e.g. indicator kriging (Alabert, 1987). In the following, let  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  be the locations to simulate. Here is the principle of the “sequential indicator algorithm”:

- 1 set  $i = 1$
- 2 estimate the conditional cumulative distribution function (in short, *ccdf*) at location  $\mathbf{x}_i$  by indicator kriging from the previously simulated values  $\{Y(\mathbf{x}_1), \dots, Y(\mathbf{x}_{i-1})\}$  and the original data
- 3 correct for order-relation deviations and draw a random variable from the corrected *ccdf*;
- 4 set  $i = i + 1$  and go to 2) until all the locations are simulated.

The advantages of this algorithm are the possibility to handle directly conditional simulations, to account for structural patterns such as a spatial correlation of the extreme values, and to incorporate soft information (Alabert, 1987; Deutsch and Journel, 1992, p 151). It can be applied to simulate continuous or categorical variables, e.g. geologic facies types. In contrast, by using an indicator kriging instead of a conditional expectation (which defines the “true” conditional distributions), the algorithm loses accuracy. However, so far, little is

---

X. Emery  
 Department of Mining Engineering,  
 University of Chile, Avenida Tupper 2069, Santiago, Chile  
 E-mail: xemery@cec.uchile.cl

known about the consequences of this issue on the properties of the realizations.

The following sections consider a stationary framework and resort to simple (co)kriging. This is a very favorable situation in practice: the means of the indicator variables are known, hence so is the univariate distribution of the model. The moving neighborhood restriction required for practical implementation of the sequential paradigm will not be examined, but it must be kept in mind since it deteriorates the reproduction of the model parameters (Emery, 2005). The post-processing step (interpolation between the thresholds used for indicator kriging and correction of order-relation violations) is not contemplated either. Instead, we will focus on the following concern: what are the properties of the images obtained when running the sequential indicator algorithm? Such question is crucial to define the use that can be made of the realizations. It can be split into two problems:

- 1) what are the conditions to honor the input parameters?
- 2) what happens with the parameters that are not specified in input (in particular, the so-called *multiple-point statistics*)?

These points will be examined in detail through the study of several classes of stochastic models, namely: mosaic model, random sets, multiple-indicator models and isofactorial models.

## 2 Mosaic model

The mosaic model is the most favorable case to perform indicator kriging, since all the simple and cross indicator covariances are proportional to a single correlogram  $\rho(\mathbf{h})$ . In such model, two values  $Y(\mathbf{x})$  and  $Y(\mathbf{x} + \mathbf{h})$  are equal or independent, with probabilities  $\rho(\mathbf{h})$  and  $1 - \rho(\mathbf{h})$  respectively. The space is therefore divided into random cells: the values are constant in each cell and independent from one cell to another one (Rivoirard, 1994, p 23; Chilès and Delfiner, 1999, p 384).

In the following, the univariate distribution is assumed to be continuous, so that two equal values necessarily belong to the same cell. Let  $F(\cdot)$  be the prior cumulative distribution function (in short, *cdf*) and  $f(\cdot)$  the prior probability density function (*pdf*).

### 2.1 Conditions to honor the univariate and bivariate distributions

The univariate distribution is characterized by the expected values of the indicator variables, whereas the bivariate distributions only depend on the correlogram.

#### Proposition 1

The simulated values honor the univariate and bivariate distributions if and only if the indicator kriging weights

are always positive and their sum is less than or equal to one.

*Proof:* This result can be established by induction on the number of simulated values. Consider a set of data  $\{Y_1, \dots, Y_n\}$  at locations  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  such that, for any integers  $i, j$  in  $\{1, \dots, n\}$  and any real numbers  $y, y'$ :

$$\begin{aligned} E\{I(Y_i; y)\} &= \text{Prob}(Y_i < y) = F(y) \\ \text{cov}\{I(Y_i; y), I(Y_j; y')\} &= \alpha(y, y') \rho(\mathbf{x}_i - \mathbf{x}_j) \end{aligned} \quad (1)$$

In this formula,  $I(Y_i; y)$  is a binary function equal to one if  $Y_i < y$ , zero otherwise. The covariance between  $I(Y_i; y)$  and  $I(Y_j; y')$  is proportional to the correlogram  $\rho(\mathbf{h})$  (mosaic property); the proportionality factor is  $\alpha(y, y') = F(y \wedge y') - F(y)F(y')$ . Let us estimate the *ccdf* at a location  $\mathbf{x}_{n+1}$  by simple indicator kriging:

$$\begin{aligned} [I(Y_{n+1}; y)]^{\text{SK}} &= \sum_{i=1}^n \lambda_i I(Y_i; y) + \lambda_m F(y) \\ \text{with } \lambda_m &= 1 - \sum_{i=1}^n \lambda_i \text{ (weight of the mean).} \end{aligned} \quad (2)$$

The kriging weights do not depend on the threshold  $y$ , since all the indicators have the same correlogram. At each data value  $Y_i$ , the *ccdf* presents a jump with amplitude equal to the kriging weight  $\lambda_i$  assigned to the datum, while the weight of the mean  $\lambda_m$  is given to the prior *cdf*. To ensure the monotonicity of the *ccdf*, all these weights must be nonnegative, hence the condition in proposition 1 is necessary. Under such condition, the simulated value  $Y_{n+1}$  is equal to  $Y_i$  with probability  $\lambda_i$  or independent of  $\{Y_1, \dots, Y_n\}$  with probability  $\lambda_m$ . Accordingly, for any integer  $j$  in  $\{1, \dots, n\}$  and any real numbers  $y, y'$ , one has:

$$\begin{aligned} \text{Prob}(Y_{n+1} < y) &= \sum_{i=1}^n \lambda_i \text{Prob}(Y_i < y) + \lambda_m F(y) \\ &= F(y) \\ \text{cov}\{I(Y_{n+1}; y), I(Y_j; y')\} &= \sum_{i=1}^n \lambda_i \text{cov}\{I(Y_i; y), I(Y_j; y')\} \\ &= \alpha(y, y') \sum_{i=1}^n \lambda_i \rho(\mathbf{x}_i - \mathbf{x}_j) \\ &= \alpha(y, y') \rho(\mathbf{x}_{n+1} - \mathbf{x}_j) \end{aligned} \quad (3)$$

The last identity stems from the simple kriging equations. All the simple and cross indicator covariances remain proportional to  $\rho(\mathbf{h})$ , hence the mosaic model is reproduced.

#### Remark

If an ordinary kriging (with unknown mean) is used instead of a simple kriging, the simulated values no

longer honor a mosaic distribution: the algorithm cannot generate a value different from the original data because the weight of the mean is zero. Furthermore, the simulated indicator correlogram will not match the theoretical model.

### Proposition 2

Except for the pure nugget effect, no stationary correlogram defined in  $\mathbf{R}^d$  with  $d > 1$  fulfills the conditions in proposition 1. This result has been proved by Matheron (1986, p 30) and shows that, even in the mosaic model, the conditions to honor the input parameters (univariate and bivariate distributions) are very restrictive.

In one dimension, the only known examples of stationary correlograms that satisfy the conditions of proposition 1 are the exponential model and the completely monotone functions, i.e. functions such that the sign of the successive derivatives alternates. Bernstein's theorem (Widder, 1941, p 161) states that a completely monotone function is proportional to the Laplace transform of a probability density, hence it is a mixture of exponential correlograms.

## 2.2 Multivariate distribution of the simulated values

### Proposition 3

The joint distribution of the values  $\{Y_1, \dots, Y_n\}$  simulated at locations  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  depends on the visiting sequence. By an iterated application of Bayes' theorem, this distribution can be expressed as follows:

$$f_{\sigma,n}(y_1, \dots, y_n) = \prod_{k=1}^n \left[ \sum_{\ell < k} \lambda_{\ell,k}^{(\sigma)} \delta_{y_k=y_\ell} + (1 - \sum_{\ell < k} \lambda_{\ell,k}^{(\sigma)}) f(y_k) \right] \quad (4)$$

where

$\sigma = \{i_1, \dots, i_n\}$  is a permutation of  $\{1, \dots, n\}$  that represents the visiting sequence;

$\{\lambda_{\ell,k}^{(\sigma)}, \ell < k\}$  are the weights assigned to the locations  $\{\mathbf{x}_\ell, \ell < k\}$  when kriging  $\mathbf{x}_k$ .

Equation 4 expresses that the  $k^{\text{th}}$  simulated value is equal to datum  $Y_{i_\ell}$  with probability  $\lambda_{\ell,k}^{(\sigma)}$ , or independent of all the previous values with the complementary probability.

To examine the properties of such multivariate distribution, we will present a simple example. Consider a mosaic with an exponential correlogram on four aligned and regularly spaced nodes  $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$  and assume that the visiting sequence is  $\{\mathbf{x}_2, \mathbf{x}_1, \mathbf{x}_4, \mathbf{x}_3\}$ . Using the screening effect property of the exponential correlogram, the probabilities for three consecutive values to be equal are easily expressed as a function of the correlation coefficient  $\rho$  between two consecutive nodes:

$$\text{Prob}[Y_1 = Y_2 = Y_3] = \rho \left\{ \frac{\rho^2 \times 2\rho}{1 + \rho^2} + \frac{(1 - \rho^2) \times \rho}{1 + \rho^2} \right\} = \rho^2 \quad (5)$$

whereas

$$\text{Prob}[Y_2 = Y_3 = Y_4] = \rho^2 \frac{2\rho}{1 + \rho^2} \quad (6)$$

The trivariate probability is not shift-invariant. This observation leads to the following proposition.

### Proposition 4

In general, the multiple-point statistics are not stationary, i.e. they change when the multiple-point configurations are shifted. The reason is that the multivariate distribution given in Eq. 4 is closely related to the kriging configurations, which are not shift-invariant. As an exception, the bivariate distributions are stationary and mosaic-type (proposition 1).

To avoid artifacts when a moving neighborhood is used, it is often advised to randomize the visiting sequence, which does not alter the bivariate distributions (independent of the ordering of the locations to simulate). If  $S_n$  stands for the set of permutations of  $\{1, \dots, n\}$  (set of equiprobable visiting sequences), the multivariate distribution (Eq. 4) becomes:

$$f_n(y_1, \dots, y_n) = \frac{1}{n!} \sum_{\sigma \in S_n} f_{\sigma,n}(y_1, \dots, y_n) \quad (7)$$

Even in this case, the multiple-point statistics are not necessarily stationary for the same reason as in proposition 4 (the kriging configurations are not shift-invariant, in particular when comparing the edge and the center of the simulated domain). This result means that the probability of occurrence of an event at a given multiple-point configuration depends on whether this configuration is located on the border of the simulated image or at its center. To avoid such edge effect, the simulation should be performed on a domain much greater than the one that really interests.

### Proposition 5

The joint distribution of a set of values  $\{Y_1, \dots, Y_n\}$  at locations  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  depends on whether other locations have been simulated and, if so, how many. In other words, the probability of occurrence of an event also depends on the resolution of the realization (grid mesh) and on the size of the simulated domain, both parameters being arbitrarily defined by the user.

*Proof:* Let us come back to the example of the four aligned nodes and calculate the probability to have  $Y_1 = Y_3 \neq Y_2$ , i.e. the probability that  $\mathbf{x}_2$  splits a cell containing  $\mathbf{x}_1$  and  $\mathbf{x}_3$ . This kind of statistics has a growing interest in geostatistical studies, since it is related to facies connectivity. For instance in a vein-type deposit,

**Table 1** Distribution of  $\{Y_1, Y_2, Y_3\}$  when  $Y_4$  is *not* simulated

Visiting sequence	Prob $[Y_1 = Y_3 \neq Y_2]$	Average probability
$x_1 - x_2 - x_3$	0	$\frac{\rho^2 (1-\rho)^2}{3(1+\rho^2)}$
$x_2 - x_1 - x_3$		
$x_2 - x_3 - x_1$		
$x_3 - x_2 - x_1$		
$x_1 - x_3 - x_2$	$\frac{\rho^2 (1-\rho)^2}{1+\rho^2}$	
$x_3 - x_1 - x_2$		

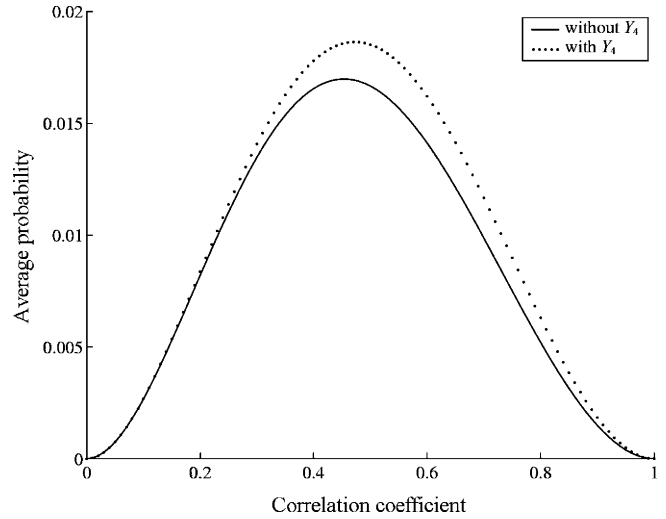
**Table 2** Distribution of  $\{Y_1, Y_2, Y_3\}$  when  $Y_4$  is simulated

Visiting sequence	Prob $[Y_1 = Y_3 \neq Y_2]$	Average probability
$x_1 - x_2 - x_3 - x_4$	0	$\frac{\rho^2 (1-\rho)^2 (2+\rho^2+2\rho^3)}{6(1+\rho^2)(1+\rho^3)}$
$x_1 - x_2 - x_4 - x_3$		
$x_2 - x_1 - x_3 - x_4$		
$x_2 - x_1 - x_4 - x_3$		
$x_2 - x_3 - x_1 - x_4$		
$x_2 - x_3 - x_4 - x_1$		
$x_2 - x_4 - x_1 - x_3$		
$x_2 - x_4 - x_3 - x_1$		
$x_3 - x_2 - x_1 - x_4$		
$x_3 - x_2 - x_4 - x_1$		
$x_3 - x_4 - x_2 - x_1$		
$x_4 - x_2 - x_1 - x_3$		
$x_4 - x_2 - x_3 - x_1$		
$x_4 - x_3 - x_2 - x_1$		
$x_1 - x_3 - x_2 - x_4$	$\frac{\rho^2 (1-\rho)^2}{1+\rho^2}$	
$x_1 - x_3 - x_4 - x_2$		
$x_3 - x_1 - x_2 - x_4$		
$x_3 - x_1 - x_4 - x_2$		
$x_3 - x_4 - x_1 - x_2$	$\frac{\rho^4 (1-\rho)^2 (1+\rho)}{(1+\rho^2)(1+\rho^3)}$	
$x_4 - x_3 - x_1 - x_2$		
$x_1 - x_4 - x_2 - x_3$		
$x_4 - x_1 - x_2 - x_3$		
$x_1 - x_4 - x_3 - x_2$	$\frac{\rho^2 (1-\rho)^2}{1+\rho^3}$	
$x_4 - x_1 - x_3 - x_2$		

given two high-graded samples, the miner looks for the probability to find high grades at intermediate locations. Now, the previous probability depends on whether the fourth value  $Y_4$  is simulated or not (Tables 1 and 2). The discrepancy between both cases is not negligible, especially if  $\rho$  lies between 0.4 and 0.8 (Fig. 1). A similar conclusion prevails if intermediate locations are considered, e.g. the midpoints of  $[x_1, x_2]$  and  $[x_2, x_3]$ : the three-point statistics of  $\{Y_1, Y_2, Y_3\}$  depend on the resolution of the realization.

*Consequence*

Although the univariate and bivariate distributions of the model are honored, the higher-order statistics calculated over the realizations are meaningless. For instance, the properties of the values simulated in a given block depend on whether other blocks are simulated or not (!). . . Therefore, in rigor, any use of the realizations



**Fig. 1** Trivariate probability in function of the correlation coefficient between two consecutive nodes

that involves the multiple-point statistics (change of support, flow simulation, etc.) should be avoided. This statement considerably restricts the use of the sequential algorithm.

2.3 Conditioning to a set of original data

To compel the simulation to honor a set of hard data  $\{z_1, \dots, z_p\}$  at locations  $\{u_1, \dots, u_p\}$ , a common approach is to consider these data as if they were previously simulated values and use them in each subsequent indicator kriging. Such procedure means that the visiting sequence begins with the data locations.

*Proposition 6*

The prior distributions of the conditional simulations (i.e. the expected distributions of  $\{Y_1, \dots, Y_n\}$  when the data locations are simulated first, without conditioning to the data values) are no longer representative of those of the non-conditional simulations. In particular, they depend on the number of conditioning data and on their locations. Therefore, the set of models that can be reproduced changes when a sample is added, removed or shifted.

*Proof:* Assuming that the data locations constitute the beginning of the visiting sequence, the prior multivariate distribution of the whole set  $\{Z_1, \dots, Z_p; Y_1, \dots, Y_n\}$  is

$$f_{p,n}(z_1, \dots, z_p; y_1, \dots, y_n) = \frac{1}{p!n!} \sum_{\sigma \in S_p \times S_n} f_{\sigma,p,n}(z_1, \dots, z_p; y_1, \dots, y_n) \tag{8}$$

where  $f_{\sigma,p,n}$  is the distribution obtained following the sequence  $\sigma \in S_p \times S_n$  (Eq. 4). Hence, the prior distribution of the values  $\{Y_1, \dots, Y_n\}$  is

$$\frac{1}{p! n!} \sum_{\sigma \in S_p \times S_n} \int f_{\sigma,p,n}(z_1, \dots, z_p; y_1, \dots, y_n) dz_1 \dots dz_p \quad (9)$$

There is no reason why this expression should coincide with Eq. 7. In general, it will depend on the number of data and on their locations, via the sequence  $\sigma$  in  $S_p \times S_n$  (see the example below).

*Remark: prior and posterior distributions*

In stochastic simulation, the *posterior* distributions are altered when a sample is added, because they reflect the knowledge one has about the spatial phenomenon: collecting one or several data at specific locations reduces the uncertainty on the unsampled neighboring values, which are correlated with these data. However, the *prior* distributions should remain unchanged since they describe the global properties of the phenomenon and are not conditioned to the values of the nearby samples. Now, the sequential algorithm transgresses this basic rule, leading to the absurd conclusion that the phenomenon under study changes when it is sampled! In other words, the probability of occurrence of an event is likely to change if one decides to make an extra sample, *even before knowing the value that will be measured in that sample*.

An example should clarify the situation. Let us consider the previous configuration with four aligned nodes  $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}$ . If no sample is taken, the probability that  $\mathbf{x}_2$  splits a cell containing  $\mathbf{x}_1$  and  $\mathbf{x}_3$  is equal to (Table 2):

$$Prob[Y_1 = Y_3 \neq Y_2] = \frac{\rho^2 (1 - \rho)^2 (2 + \rho^2 + 2\rho^3)}{6(1 + \rho^2)(1 + \rho^3)} \quad (10)$$

With one sample, the algorithm only considers the sequences starting with the associated location. Hence, the former probability becomes:

- if  $\mathbf{x}_1$  is sampled:

$$Prob[Y_1 = Y_3 \neq Y_2 | Y_1 = y_1] = \frac{\rho^2 (1 - \rho)^2 (3 + 2\rho^2 + 3\rho^3)}{6(1 + \rho^2)(1 + \rho^3)}$$

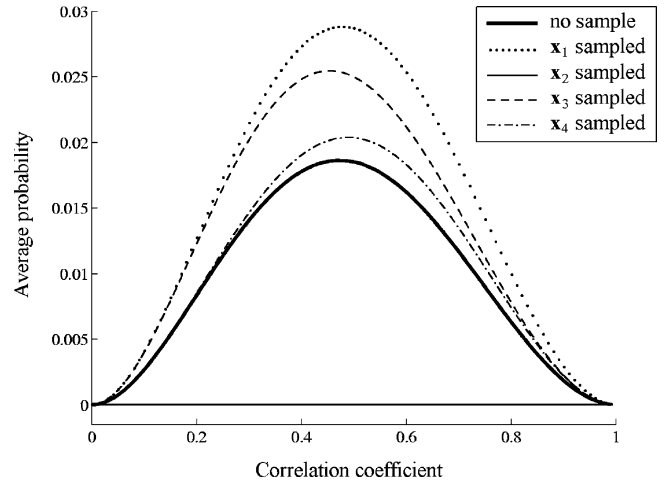
- if  $\mathbf{x}_2$  is sampled:  $Prob[Y_1 = Y_3 \neq Y_2 | Y_2 = y_2] = 0$
- if  $\mathbf{x}_3$  is sampled:

$$Prob[Y_1 = Y_3 \neq Y_2 | Y_3 = y_3] = \frac{\rho^2 (1 - \rho)^2}{2(1 + \rho^2)}$$

- if  $\mathbf{x}_4$  is sampled:

$$Prob[Y_1 = Y_3 \neq Y_2 | Y_4 = y_4] = \frac{\rho^2 (1 - \rho)^2 (1 + \rho^2 + \rho^3)}{3(1 + \rho^2)(1 + \rho^3)}$$

Since all these posterior probabilities do not depend on the sampled values, they are also prior probabilities (i.e. prior to the knowledge of the sampled values). Consequently, the answer to the question “what is the



**Fig. 2** Effect of the sampling on the *prior* trivariate probability

probability that  $\mathbf{x}_2$  splits a cell containing  $\mathbf{x}_1$  and  $\mathbf{x}_4$ ?” depends on whether a sample is made or not and, if so, where (Fig. 2), whatever the sampled value. . .

### 3 Random sets

A random set is a model whose realizations are subsets of the space. It can be represented by an indicator variable:

$$Y(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \text{ belongs to the random set} \\ 0 & \text{otherwise} \end{cases} \quad (11)$$

#### 3.1 Conditions to honor the mean and correlogram

*Proposition 7*

The simulated values honor the univariate and bivariate distributions of the random set (mean and correlogram of the indicator variable) if and only if the indicator kriging estimate always lies between 0 and 1. This particular case of proposition 1 has been proved by Chilès and Delfiner (1999, p 522).

*Conditions on the kriging weights*

Let  $m$  be the mean value of the indicator and  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  a set of locations. The simple kriging at  $\mathbf{x}_i$  from the other locations can be written as follows:

$$[Y(\mathbf{x}_i)]^{SK} = \sum_{\substack{j=1 \\ j \neq i}}^n \lambda_j Y(\mathbf{x}_j) + (1 - \sum_{\substack{j=1 \\ j \neq i}}^n \lambda_j) m \quad (12)$$

Let  $\Sigma^+$  and  $\Sigma^-$  be the sum of positive and negative weights respectively. The bounds of the estimate  $[Y(\mathbf{x}_i)]^{SK}$  are  $[\Sigma^+ + (1 - \Sigma^+ - \Sigma^-)m]$  for the maximum and  $[\Sigma^- + (1 - \Sigma^+ - \Sigma^-)m]$  for the minimum. These bounds lie between 0 and 1 if and only if

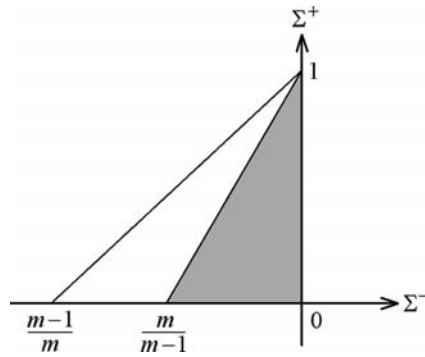


Fig. 3 Allowable area for the indicator kriging weights

$$\begin{cases} 0 \leq \Sigma^+ \leq 1 + \frac{m}{1-m} \Sigma^- \\ (\Sigma^+ - 1) \frac{m}{1-m} \leq \Sigma^- \leq 0 \end{cases} \quad (13)$$

In the plane parametrized by  $\Sigma^+$  and  $\Sigma^-$ , the allowable area is a triangle (Fig. 3) located between the coordinate axes and the line with equation

$$\Sigma^+ = 1 + \max\left(\frac{m}{1-m}, \frac{1-m}{m}\right) \Sigma^- \quad (14)$$

The loosest condition occurs when  $m = 1/2$ , in which case it becomes  $\Sigma^+ - \Sigma^- \leq 1$ . Such inequality means that the sum of the absolute values of the kriging weights must be less than or equal to 1:

$$\sum_{\substack{j=1 \\ j \neq i}}^n |\lambda_j| \leq 1 \quad (15)$$

This relation can be expressed in a matrix form. Let  $\mathbf{C} = [C(\mathbf{x}_i - \mathbf{x}_j)]_{i,j=1..n}$  be the covariance matrix of  $\{Y_1, \dots, Y_n\}$  and  $\mathbf{B} = [b_{ij}]_{i,j=1..n}$  the inverse of  $\mathbf{C}$ . The weights assigned to  $\{\mathbf{x}_j, j \neq i\}$  when kriging  $\mathbf{x}_i$  are  $\{-b_{ij}/b_{ii}, j \neq i\}$ , whereas  $1/b_{ii}$  is the corresponding kriging variance (Dubrule, 1983). Thus, Eq. 15 amounts to writing that  $\mathbf{B}$  is a *diagonal dominant* matrix:

$$\forall i \in \{1, \dots, n\}, b_{ii} \geq \sum_{\substack{j=1 \\ j \neq i}}^n |b_{ij}| \quad (16)$$

According to Gerschgorin's theorem (Bell, 1965), the eigenvalues of  $\mathbf{B}$  lie in the segment  $[0, 2 \max_{i \in \{1..n\}} (b_{ii})]$ , thus the eigenvalues of  $\mathbf{C}$  are greater than

$$\frac{1}{2} \min_{i \in \{1..n\}} \left(\frac{1}{b_{ii}}\right) \quad (17)$$

which is half the minimum kriging variance of one location from all the other ones. These properties have to be checked for any choice of the locations  $\{\mathbf{x}_i, i = 1 \dots n\}$ .

### Conjecture

Although no complete characterization of the correlograms that ensure Eq. 15 has been found, it is conjectured that they are only the pure nugget effect and the completely monotone functions on  $\mathbf{R}_+$ . If true, this conjecture

means that the apparent looser condition with respect to the continuous mosaic (proposition 2) is not effective.

To corroborate the conjecture, numerical tests have been made with other usual correlogram models (spherical, triangular, quadratic, Gaussian...): for each of these models, it is easy to exhibit a configuration that contradicts Eq. 15, even in one dimension and in presence of a nugget effect. For instance, a correlogram composed by a nugget effect with sill 0.95 plus a spherical model with sill 0.05 and range 1, leads to a sum of absolute weights equal to 1.03 when kriging the origin from one thousand data regularly spaced along the segment  $[-1,1]$ .

## 3.2 Autoduality

### Proposition 8

When  $m = 1/2$ , the random set generated by the sequential algorithm is *autodual*, i.e. the 0's and 1's of its indicator can be swapped without modifying the multivariate distribution.

*Proof:* If the visiting sequence is described by a permutation  $\sigma = \{i_1, \dots, i_n\}$  of  $\{1, \dots, n\}$ , then the multivariate distribution of  $\{Y(\mathbf{x}_1), \dots, Y(\mathbf{x}_n)\}$  is (Eq. 4):

$$f_{\sigma,n}(y_1, \dots, y_n) = \prod_{k=1}^n \left[ \sum_{\ell < k} \lambda_{\ell,k}^{(\sigma)} \delta_{y_k=y_\ell} + (1 - \sum_{\ell < k} \lambda_{\ell,k}^{(\sigma)})/2 \right] \quad (18)$$

where  $\{y_1, \dots, y_n\}$  are values in  $\{0,1\}$ . The autoduality stems from the fact that this distribution is unchanged when substituting  $\{1-y_1, \dots, 1-y_n\}$  for  $\{y_1, \dots, y_n\}$ .

### Consequence

When  $m = 1/2$ , the sequential indicator algorithm fails to simulate non-autodual random sets, for instance the Boolean models in which the "pores" do not play the same role as the "grains" (Lantuéjoul, 2002, p 153). This means that the algorithm is not general and only reproduces a restrictive class of random sets.

## 4 Multiple indicator models

The sequential indicator algorithm is often used to simulate models in which each indicator has its own covariance (Deutsch and Journel, 1992, p 146-152). Hereafter, we will distinguish the cases of continuous and categorical variables.

### 4.1 Continuous variables

The study of the mosaic model has shown several coherence problems with the multivariate distributions

of the simulated values. Such problems also arise in this more general model and will not be stated again. Instead, the following results focus on the bivariate distributions: these are, this time, only *partially* defined by the indicator covariances.

### Proposition 9

The simulation honors the indicator means and covariances if and only if all the *ccdf*'s estimated by indicator kriging are non-decreasing.

The proof of this assertion is similar to the one of proposition 1. Note that the reproduction of the indicator means and covariances is no longer true when order-relation violations occur and have to be corrected.

### Consequence

Under the conditions of proposition 9, the sequential simulation also honors the madogram of the model, since it is the sum of all the indicator variograms (Matheron, 1989, p 30).

Let us now examine the other parameters of the bivariate distributions, specifically the indicator cross-covariances and the covariance of the continuous variable itself.

### Proposition 10

In general, the simulated indicator cross-covariances depend on the visiting sequence and are not stationary (an exception is the above-mentioned mosaic model).

*Proof:* Let us consider a set of initial values  $\{Y_1, \dots, Y_n\}$  at locations  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  and assume that

- a)  $\forall y \in \mathbf{R}, \forall i \in \{1, \dots, n\}$ ,  
 $E\{I(Y_i; y)\} = \text{Prob}(Y_i < y) = F(y)$
- b) for any pair of values  $\{y, y'\}$ , there exists a function  $C_{yy'}(\mathbf{h})$  such that:

$$\forall i, j \in \{1, \dots, n\}, \text{cov}\{I(Y_i; y), I(Y_j; y')\} = C_{yy'}(\mathbf{x}_i - \mathbf{x}_j) \quad (19)$$

So far, the indicator cross-covariances are stationary, as they only depend on the vector separating the pairs of indicator values. The question is: does this situation hold when an extra value is simulated at  $\mathbf{x}_{n+1}$ ? The *ccdf* at that location is estimated by simple indicator kriging:

$$E[I(Y_{n+1}; y) | \text{data}] = \sum_{i=1}^n \lambda_i(y) I(Y_i; y) + \lambda_m(y) F(y)$$

with  $\lambda_m(y) = 1 - \sum_{i=1}^n \lambda_i(y)$ . (20)

With respect to the mosaic model (Eq. 2), the indicator covariances and the kriging weights depend on the thresholds. The previous formulae entail:

$$\begin{aligned} E[I(Y_{n+1}; y) I(Y_j; y')] &= E\{I(Y_j; y') E[I(Y_{n+1}; y) | \text{data}]\} \\ &= E\{I(Y_j; y') [\sum_{i=1}^n \lambda_i(y) I(Y_i; y) + \lambda_m(y) F(y)]\} \\ &= \sum_{i=1}^n \lambda_i(y) C_{yy'}(\mathbf{x}_i - \mathbf{x}_j) + F(y) F(y') \end{aligned} \quad (21)$$

so that

$$\text{cov}\{I(Y_{n+1}; y), I(Y_j; y')\} = \sum_{i=1}^n \lambda_i(y) C_{yy'}(\mathbf{x}_i - \mathbf{x}_j) \quad (22)$$

This expression depends on the kriging configurations (hence, on the visiting sequence) and has no reason to be equal to the value of function  $C_{yy'}$  at lag  $\mathbf{h} = \mathbf{x}_{n+1} - \mathbf{x}_j$ :

$$\text{cov}\{I(Y_{n+1}; y), I(Y_j; y')\} \neq C_{yy'}(\mathbf{x}_{n+1} - \mathbf{x}_j) \quad (23)$$

To corroborate the above inequality, the reader can make the following exercise: consider five aligned and regularly spaced locations ( $\mathbf{x}_1$  to  $\mathbf{x}_5$ ) and perform a simulation with exponential indicator covariance models, so as to produce a screening effect and simplify the calculations. The covariance between  $I(Y_2; y)$  and  $I(Y_4; y')$  is found to differ from the one between  $I(Y_1; y)$  and  $I(Y_3; y')$ , unless the simulation sequence is regular or all the indicator covariances are equal (mosaic model).

### First consequence (non-stationarity)

As was mentioned in the mosaic model for the multiple-point statistics, edge effects are likely to occur on the borders of the simulated image, but this time, they also concern the two-point statistics. In other words, the behavior of a pair of values depends on whether this pair is located on the border of the simulated image or at its center.

### Second consequence (dependence on the sequence)

Except the indicator variograms and the madogram, the other two-point statistics depend on the visiting sequence. In particular, this occurs with the covariance of the variable itself, which is a mixture of all the simple and cross indicator covariances (Chilès and Delfiner, 1999, p 103). Now, the visiting sequence is renewed for each realization (chosen at random among all the possible sequences). Therefore, the regional variogram is likely to differ from one realization to another beyond the range of the expectable “ergodic fluctuations”,

which may explain the poor ergodic properties often observed with sequential indicator simulation (Deutsch and Journel, 1992, p 127–129).

## 4.2 Categorical variables

Categorical variables encode mutually exclusive *lithofacies*. Let us assume that the whole space is partitioned into  $N$  facies and define the variable:

$$Y(\mathbf{x}) = y \text{ if } \mathbf{x} \text{ belongs to the } y^{\text{th}} \text{ facies} \quad (24)$$

Since the codification does not correspond to any real ordering of the facies, it is convenient to work on *standardized* indicators  $i(Y(\mathbf{x}); y)$  (function equal to 1 if  $Y(\mathbf{x}) = y$ , 0 otherwise) rather than *cumulative* indicators (Deutsch and Journel, 1992, p 147). However, the sequential algorithm leads to a coherence problem, as detailed in the next proposition.

### Proposition 11

Unless all the indicators have the same correlogram (mosaic model), the estimated indicators do not necessarily add to 1 and, therefore, cannot be interpreted as a probability system.

*Proof:* Let us estimate one location  $\mathbf{x}$  from a single datum located at  $\mathbf{x} + \mathbf{h}$ . For any  $y$  in  $\{1, \dots, N\}$ , one has:

$$[i(Y(\mathbf{x}); y)]^{SK} = \rho_y(\mathbf{h}) i(Y(\mathbf{x} + \mathbf{h}); y) + [1 - \rho_y(\mathbf{h})] f(y) \quad (25)$$

where  $f(y)$  and  $\rho_y(\mathbf{h})$  stand for the mean and correlogram of the indicator  $i(Y(\mathbf{x}); y)$ . Let  $y_0$  be the index of the facies that prevails at  $\mathbf{x} + \mathbf{h}$ . A summation over all the indices gives:

$$\begin{aligned} \sum_{y=1}^N [i(Y(\mathbf{x}); y)]^{SK} &= \rho_{y_0}(\mathbf{h}) + \sum_{y=1}^N [1 - \rho_y(\mathbf{h})] f(y) \\ &= \rho_{y_0}(\mathbf{h}) + 1 - \sum_{y=1}^N \rho_y(\mathbf{h}) f(y) \end{aligned} \quad (26)$$

This sum is always equal to one if and only if it does not depend on  $y_0$ , i.e. if all the indicators have the same correlogram, Q.E.D.

Assume that the estimated indicators do not add to one. A first solution for correcting this situation is to rescale the kriging weights (e.g. divide them by their sum), but such procedure introduces biases and deteriorates the reproduction of the indicator correlograms. A second possibility is to use a full indicator cokriging, or a disjunctive kriging, instead of a kriging of each indicator. This variant of the sequential indicator algorithm (Emery, 2002) can be implemented to simulate isofactorial models.

## 5 Isofactorial models

### 5.1 What is an isofactorial model?

A stationary random function  $\{Y(\mathbf{x}), \mathbf{x} \in \mathbf{R}^d\}$  with a *cdf*  $F(\cdot)$  and a *pdf*  $f(\cdot)$  have bivariate isofactorial distributions if there exists a family of functions called “factors”  $\{\chi_p, p \in \mathbf{N}\}$  such that (Chilès and Delfiner, 1999, p 394):

1. any function  $\varphi[Y(\mathbf{x})]$  with finite variance can be expanded into the factors:

$$\varphi[Y(\mathbf{x})] = \sum_{p \geq 0} \varphi_p \chi_p[Y(\mathbf{x})]$$

$$\text{with } \forall p \in \mathbf{N}, \varphi_p = \int \varphi(y) \chi_p(y) f(y) dy \quad (27)$$

2. the first factor is constant and equal to one; the other factors have a zero mean and a unit variance:

$$\forall p \in \mathbf{N}^*, \forall \mathbf{x} \in \mathbf{R}^d, \begin{cases} E \{\chi_p[Y(\mathbf{x})]\} = 0 \\ \text{var} \{\chi_p[Y(\mathbf{x})]\} = 1 \end{cases} \quad (28)$$

3. two different factors have no spatial cross-correlation:  $\forall p, q \in \mathbf{N}^*, \forall \mathbf{x}, \mathbf{x} + \mathbf{h} \in \mathbf{R}^d$ ,

$$\text{cov} \{\chi_p[Y(\mathbf{x} + \mathbf{h})], \chi_q[Y(\mathbf{x})]\} = \begin{cases} 0 & \text{if } p \neq q \\ T_p(\mathbf{h}) & \text{otherwise} \end{cases} \quad (29)$$

These properties characterize the distribution of all the pairs  $\{Y(\mathbf{x} + \mathbf{h}), Y(\mathbf{x})\}$ . The bivariate density for a lag separation  $\mathbf{h}$  can be expressed as:

$$f_{\mathbf{h}}(y, y') = f(y)f(y') \left[ 1 + \sum_{p \geq 1} T_p(\mathbf{h}) \chi_p(y) \chi_p(y') \right] \quad (30)$$

A fundamental property is the *exchange relation* between factors (Chilès and Delfiner, 1999, p 396), which is an alternative definition of the isofactorial property:

$$\forall p \in \mathbf{N}^*, E \{\chi_p[Y(\mathbf{x} + \mathbf{h})] | Y(\mathbf{x})\} = T_p(\mathbf{h}) \chi_p[Y(\mathbf{x})] \quad (31)$$

### 5.2 Why isofactorial models?

The isofactorial models are interesting for two reasons.

- 1) They constitute a wide class of models that can adapt to the description of diverse structural patterns, such as an asymmetry in the spatial continuity between high and low values, a destructuring or, on the contrary, a significant spatial correlation of the extreme values (Hu, 1988; Chilès and Delfiner, 1999, p 398-405; Emery, 2002). However, isofactorial models may not be able to describe complex spatial patterns, insofar as they remain bivariate distribution models and do not allow to control the multiple-point distributions (Guardiano and Srivastava, 1993).
- 2) They enable the estimation of the posterior distributions by disjunctive kriging, which relies on the full bivariate distributions of the random function and



not only on some of its indicator correlograms (Matheron, 1976; Chilès and Delfiner, 1999, p 389). Disjunctive kriging is nothing but a simple indicator cokriging: by using the isofactorial property, it reduces the cokriging system to a simple kriging of each factor separately. In brief, isofactorial models conciliate two conflicting aspects: the internal consistency of the model (the full bivariate distributions are taken into account in a coherent way) and the ability to incorporate the structural patterns of the underlying phenomenon. A well-known example is the *bigaussian* model, in which the univariate distribution is normal and the factors are Hermite polynomials (Rivoirard, 1994, p 37). A more general example is the *Hermitian* model, which can be seen as a mixture of bigaussian distributions with a random correlation coefficient (Matheron, 1976, p 229; Chilès and Delfiner, 1999, p 404). Another case is the *mosaic* model seen in the second section of this work.

### 5.3 Conditions to honor the univariate and bivariate distributions

#### *Proposition 12*

The simulated values honor the univariate and bivariate distributions of the model if and only if all the *ccdf*'s estimated by disjunctive kriging are non-decreasing.

*Proof:* As in proposition 1, the proof is established by induction. Consider a set of values  $\{Y_1, \dots, Y_n\}$  located at  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  whose bivariate distributions are isofactorial (Eq. 31):  $\forall p, \in \mathbf{N}^*, \forall i, j \in \{1, \dots, n\}$ ,

$$E[\chi_p(Y_i)|Y_j] = T_p(\mathbf{x}_i - \mathbf{x}_j)\chi_p(Y_j) \quad (32)$$

Let us estimate the *ccdf* at a location  $\mathbf{x}_{n+1}$  by disjunctive kriging (in short, *DK*) and assume this is a non-decreasing function. Then, a value  $Y_{n+1}$  can be drawn according to this *ccdf*, so that:

$$F_{\mathbf{x}_{n+1}}(y|data) = \text{Prob}(Y_{n+1} < y|data) = [I(Y_{n+1};y)]^{DK} \quad (33)$$

Because disjunctive kriging is unbiased, the univariate distribution is honored:

$$\text{Prob}(Y_{n+1} < y) = E\{[I(Y_{n+1};y)]^{DK}\} = F(y) \quad (34)$$

Concerning the bivariate distributions, we will need to expand the disjunctive kriging of the indicator as follows (Eq. 27):

$$F_{\mathbf{x}_{n+1}}(y|data) = \sum_{q \geq 0} \left[ \int_{-\infty}^y \chi_q(u) f(u) du \right] [\chi_q(Y_{n+1})]^{SK} \quad (35)$$

where the superscript *SK* refers to a simple kriging. Because of the orthogonality of the factors (Eq. 29), this entails, for any  $p > 0$ :

$$\begin{aligned} E[\chi_p(Y_{n+1})|data] &= \int \chi_p(y) F_{\mathbf{x}_{n+1}}(dy|data) \\ &= \sum_{q \geq 0} \int \chi_q(y) \chi_p(y) f(y) dy [\chi_q(Y_{n+1})]^{SK} \\ &= [\chi_p(Y_{n+1})]^{SK} = \sum_{i=1}^n \lambda_i^p \chi_p(Y_i) \end{aligned} \quad (36)$$

where  $\{\lambda_i^p, i = 1 \dots n\}$  are the weights assigned to  $\{\chi_p(Y_i), i = 1 \dots n\}$  when kriging  $\chi_p(Y_{n+1})$ . Now, for any integers  $p \in \mathbf{N}^*$  and  $j \in \{1, \dots, n\}$ , one has:

$$E[\chi_p(Y_{n+1})|Y_j] = E[\chi_p(Y_{n+1})|Y_1, \dots, Y_n|Y_j] \quad (37)$$

To use this formula, it is necessary to calculate  $E[\chi_p(Y_{n+1})|Y_1, \dots, Y_n]$ , which has been done in Eq. 36. Therefore, we obtain:

$$\begin{aligned} E[\chi_p(Y_{n+1})|Y_j] &= \sum_{i=1}^n \lambda_i^p E[\chi_p(Y_i)|Y_j] \\ &= \sum_{i=1}^n \lambda_i^p T_p(\mathbf{x}_i - \mathbf{x}_j) \chi_p(Y_j) \\ &= T_p(\mathbf{x}_{n+1} - \mathbf{x}_j) \chi_p(Y_j) \end{aligned} \quad (38)$$

The second equality corresponds to the induction hypothesis (Eq. 32), whereas the last equality stems from the simple kriging system. Equation (38) shows that, if the estimated *ccdf* at  $\mathbf{x}_{n+1}$  is non-decreasing, all the bivariate distributions of the set  $\{Y_1, \dots, Y_n, Y_{n+1}\}$  are still isofactorial, with the factors  $\{\chi_p, p \in \mathbf{N}\}$  and covariances  $\{T_p(\mathbf{h}), p \in \mathbf{N}^*\}$ . Reciprocally, if the estimated *ccdf* is non-monotone and has to be corrected, the previous properties (Eq. 33 to 38) are no longer fulfilled, hence the univariate and bivariate distributions may not be reproduced. Now, the condition of monotonicity of the estimated *ccdf*'s is quite restrictive, as illustrated hereafter.

#### *Proposition 13 (bigaussian distribution)*

In the isotropic case, the bigaussian distribution is never reproduced exactly except for

- 1) a pure nugget effect model;
- 2) a one-dimensional markovian model, when the visiting sequence is regular.

The proof is given in Appendix A. Such proposition can be extended to other diffusive isofactorial models, such as the bigamma, negative binomial and Poisson distributions (Hu, 1988; Chilès and Delfiner, 1999, p 401–404): only the markovian models can be reproduced exactly. A diffusive model is defined by the fact that the correlogram of the  $p^{\text{th}}$  factor is equal to the correlogram of the first factor raised to power  $p$  (Chilès and Delfiner, 1999, p 409).

## Consequence

Aside from the markovian case, the sequential isofactorial algorithm cannot simulate a multigaussian random field and, more generally, a diffusion-type process. *A fortiori*, neither can the sequential indicator algorithm, since an indicator kriging is always less accurate than a disjunctive kriging.

## 6 Conclusions

In addition to the difficulties linked to the sequential paradigm, in particular concerning the moving neighborhood restriction required for its practical implementation, several conceptual problems arise from the indicator formalism.

- 1) The conditions to honor exactly the input parameters are extremely restrictive: they are often limited to the one-dimensional space and to some specific indicator correlograms. In all the other cases, the reproduction of the model parameters is only approximate.
- 2) The probability of occurrence of an event at a given multiple-point configuration depends on a) whether this configuration is located on the edge of the simulated image or at its center, b) the total number of simulated nodes (which includes the grid mesh of the simulation and the size of the simulated domain), and c) the number and configuration of the available samples, *prior to the knowledge of their values*. Therefore, the multiple-point statistics are baseless and should be considered as “undefined”. This rule also applies in the general case of multiple-indicator models for all the two-point statistics that are not fully characterized by the indicator correlograms (e.g. the variogram of the variable itself).

For many authors, one of the purposes of stochastic simulation is to go beyond the model input parameters (in this case, the indicator correlograms) and use the realizations in operations that involve multiple-point statistics, e.g. upscaling or flow simulation. Now, such approach is debatable: because of the previous remark, the multiple-point statistics are not only model-dependent, but also implementation-dependent. Accordingly, one may wish to limit the use of the realizations to operations whose results are fully controlled by the model parameters. This point of view considerably restricts the scope of the algorithm, for instance it does not allow to perform a change of support on the realizations. One way to solve the change-of-support problem consists in simulating directly the block values without resorting to a point-support simulation, but this solution requires defining a joint distribution between point and block values (Emery, 2002, p 96).

All these impediments appear as the counterpart of the flexibility of the method, which intends to be all-purpose. Actually, the sequential indicator algorithm should be qualified as “stochastic imaging” rather than

“stochastic simulation”. The realizations (images) do not refer to a fully specified random function model and their properties depend on external factors such as the total number of simulated nodes.

**Acknowledgements** The author would like to acknowledge the reviewers for their helpful comments and the sponsoring by Codelco-Chile for supporting this research.

## Appendix A

In this appendix, we give the proof of Proposition 13. Let us consider two data  $\{Y_1, Y_2\}$  at locations  $\{\mathbf{x}_1, \mathbf{x}_2\}$  and simulate the value  $Y_3$  at location  $\mathbf{x}_3$ . The estimated *ccdf* is (Rivoirard, 1994, p 57):

$$[I(Y_3; y)]^{DK} = F(y) + \sum_{p \geq 1} \frac{1}{\sqrt{p}} \chi_{p-1}(y) f(y) [\chi_p(Y_3)]^{SK} \quad (\text{A1})$$

$$\text{with } [\chi_p(Y_3)]^{SK} = \lambda_1^p \chi_p(Y_1) + \lambda_2^p \chi_p(Y_2).$$

This function is required to be non-decreasing, hence its derivative must be non-negative whatever the values of  $Y_1$  and  $Y_2$ :

$$f(y) \{1 + \sum_{p \geq 1} \chi_p(y) [\lambda_1^p \chi_p(Y_1) + \lambda_2^p \chi_p(Y_2)]\} \geq 0 \quad (\text{A2})$$

A necessary condition is found by assuming that both data are equal to a same scalar  $y'$ :

$$f(y, y') = f(y) f(y') [1 + \sum_{p \geq 1} (\lambda_1^p + \lambda_2^p) \chi_p(y) \chi_p(y')] \geq 0 \quad (\text{A3})$$

The integral of  $f(y, y')$  over  $\mathbf{R}^2$  is equal to 1 since all the terms of order  $p \geq 1$  have a zero integral (Eq. 28). Thus, what is sought after is a condition for  $f(y, y')$  to be a probability density (positive function with a unit integral) with an isofactorial expression (Eq. 30). According to Matheron (1976, p 233), this occurs if and only if the coefficients  $\{a_p, p \in \mathbf{N}\}$  defined by

$$a_0 = 1 \text{ and } \forall p \geq 1, a_p = \lambda_1^p + \lambda_2^p \quad (\text{A4})$$

constitute the successive moments of a random variable in  $[-1, 1]$ . Now, in the bigaussian model, the covariance of the  $p^{\text{th}}$  factor is equal to the correlogram of the Gaussian random function raised to power  $p$ :

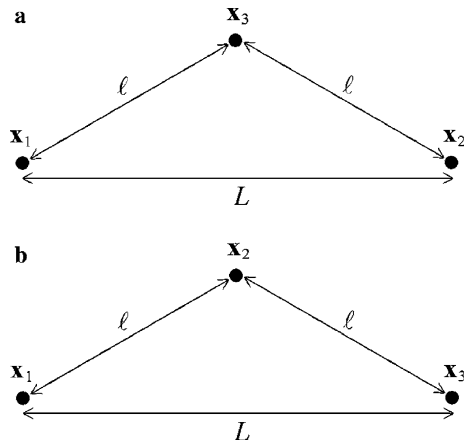
$$\forall p \in \mathbf{N}^*, T_p(\mathbf{h}) = \rho^p(\mathbf{h}) \quad (\text{A5})$$

Hence, the coefficients  $\{a_p, p \in \mathbf{N}\}$  are found to be

$$\forall p \geq 1, a_p = \lambda_1^p + \lambda_2^p = \frac{\rho_{13}^p + \rho_{23}^p}{1 + \rho_{12}^p} \quad (\text{A6})$$

in which  $\rho_{ij}$  stands for  $\rho(\mathbf{x}_i - \mathbf{x}_j)$ ,  $\forall i, j \in \{1, 2, 3\}$ .

A theorem known as “classical moment problem” (Hausdorff, 1921; Widder, 1941, p 100) states that a series of real numbers  $\{a_p, p \in \mathbf{N}\}$  are the moments of a



**Fig. 4** The data and simulated locations are the vertices of an isosceles triangle

random variable in  $\mathbb{R}$  if and only if for every finite set  $\{\alpha_0, \dots, \alpha_p\}$  of scalars, the following inequality is true:

$$S(\alpha_0, \dots, \alpha_p) = \sum_{i,j=0}^p \alpha_i \alpha_j a_{i+j} \geq 0 \quad (\text{A7})$$

In particular, for  $p = 1$ :

$$S(\alpha_0, \alpha_1) = \alpha_0^2 a_0 + 2\alpha_0 \alpha_1 a_1 + \alpha_1^2 a_2 \quad (\text{A8})$$

This is a polynomial of degree 2 with respect to  $\alpha_0$  and  $\alpha_1$ ; it is always positive if its discriminant is negative or zero:

$$\Delta = \frac{2(\rho_{13} - \rho_{23}\rho_{12})(\rho_{23} - \rho_{13}\rho_{12})}{(1 + \rho_{12})^2(1 + \rho_{12}^2)} \leq 0 \quad (\text{A9})$$

Let us examine the case when  $\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}$  are the vertices of an isosceles triangle.

- If  $\mathbf{x}_3$  is equidistant from  $\mathbf{x}_1$  and  $\mathbf{x}_2$  (Fig. 4a), then:

$$\Delta = \frac{2\rho^2(\ell)[1 - \rho(L)]^2}{[1 + \rho(L)]^2[1 + \rho^2(L)]} \quad (\text{A10})$$

The condition  $\Delta \leq 0$  is never met, unless the correlogram is a pure nugget effect ( $\rho(l) = 0$  for all  $l \neq 0$ ).

- If  $\mathbf{x}_2$  is equidistant from  $\mathbf{x}_1$  and  $\mathbf{x}_3$  (Fig. 4b), the condition becomes:

$$\Delta = \frac{2\rho(\ell)[\rho(L) - \rho^2(\ell)][1 - \rho(L)]}{[1 + \rho(\ell)]^2[1 + \rho^2(\ell)]} \leq 0 \quad (\text{A11})$$

Now, a similar reasoning when both data  $\{Y_1, Y_2\}$  are equal to  $y'$  and  $-y'$  respectively leads to another necessary condition:

$$\Delta' = -\frac{2\rho(\ell)[\rho(L) - \rho^2(\ell)][1 - \rho(L)]}{[1 - \rho(\ell)]^2[1 + \rho^2(\ell)]} \leq 0 \quad (\text{A12})$$

The only possible correlograms are the pure nugget effect and the exponential model in a one-dimensional

space ( $L = 2\ell$ ), for which  $\rho(2\ell) = \rho^2(\ell)$  for all  $\ell$ . Both cases are also sufficient conditions:

- 1) The nugget effect simulates mutually independent random values (white noise process).
- 2) The exponential correlogram in one dimension corresponds to a Markov process: each factor also has an exponential correlogram, so that the last simulated value screens all the previous ones in the disjunctive kriging system. Hence, the distribution of the value to simulate given the last simulated value is directly derived from the bivariate distribution (Eq. 30), and the monotonicity of the estimated *ccdf* is guaranteed. However, the simulation must be performed according to a regular sequence.

## References

- Alabert F (1987) Stochastic imaging of spatial distributions using hard and soft information, Master's thesis, Department of Applied Earth Sciences, Stanford University, Stanford, 198 p.
- Bell HE (1965) Gerschgorin's theorem and the zeros of polynomials. *Amer. Math. Monthly* 72: 292–295
- Chilès JP, Delfiner P (1999) Geostatistics: Modeling spatial uncertainty. Wiley, New York, 696 p.
- Deutsch CV, Journel AG (1992) GSLIB: Geostatistical software library and user's guide. Oxford University Press, New York, 340 p.
- Dubrulle O (1983) Cross-validation of kriging in a unique neighborhood. *Math. Geology* 15(7): 687–699
- Emery X (2002) Conditional simulation of nongaussian random functions. *Math. Geology* 34(1): 79–100
- Emery X (2005) Testing the correctness of the sequential algorithm for simulating Gaussian random fields. *Stochastic Environmental Research and Risk Assessment*, in press
- Guardiano FB, Srivastava RM (1993) Multivariate geostatistics: beyond bivariate moments. In: Soares A (ed.) *Geostatistics Tróia'92*. Kluwer Academic, Dordrecht, p. 133–144
- Hausdorff F (1921) Summationsmethoden und Momentfolgen. *Mathematische Zeitschrift* 9: 74–109 & 280–299
- Hu LY (1988) Mise en oeuvre du modèle gamma pour l'estimation des distributions spatiales, Doctoral thesis, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau, 142 p.
- Lantuéjoul C (2002) Geostatistical simulation, models and algorithms. Springer-Verlag, Berlin, 256 p.
- Matheron G (1976) A simple substitute for conditional expectation: the disjunctive kriging. In: David M, Huijbregts CJ (eds.) *Advanced Geostatistics in the Mining Industry*. Reidel, Dordrecht, p. 221–236
- Matheron G (1986) Sur la positivité des poids de krigeage. Internal report N-30/86/G, Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau, 44 p.
- Matheron G (1989) The internal consistency of models in geostatistics. In: Armstrong M (ed.) *Geostatistics*. Kluwer Academic, Dordrecht, p. 21–38
- Rivoirard J (1994) Introduction to disjunctive kriging and nonlinear geostatistics. Clarendon Press, Oxford, 181 p.
- Widder DV (1941) The Laplace transform. Princeton University Press, Princeton, 406 p.