High-Order Block Support Spatial Simulation Method and Its Application at a Gold Deposit

Joao Pedro de Carvalho¹ · Roussos Dimitrakopoulos¹ ·
Ilnur Minniakhmetov¹

Received: 5 October 2018 / Accepted: 14 January 2019 / Published online: 20 February 2019
© The Author(s) 2019

Abstract High-order sequential simulation methods have been developed as an alternative to existing frameworks to facilitate the modeling of the spatial complexity of non-Gaussian spatially distributed variables of interest. These high-order simulation approaches address the modeling of the curvilinear features and spatial connectivity of extreme values that are common in mineral deposits, petroleum reservoirs, water aquifers, and other geological phenomena. This paper presents a new high-order simulation method that generates realizations directly at the block support scale conditioned to the available data at point support scale. In the context of sequential high-order simulation, the method estimates, at each block location, the cross-support joint probability density function using Legendre-like splines as the set of basis functions needed. The proposed method adds previously simulated blocks to the set of conditioning data, which initially contains the available data at point support scale. A spatial template is defined by the configuration of the block to be simulated and related conditioning values at both support scales, and is used to infer additional high-order statistics from a training image. Testing of the proposed method with an exhaustive dataset shows that simulated realizations reproduce major structures and high-order relations of data. The practical intricacies of the proposed method are demonstrated in an application at a gold deposit.

✉ Joao Pedro de Carvalho
joao.decarvalho@mail.mcgill.ca

Roussos Dimitrakopoulos
roussos.dimitrakopoulos@mcgill.ca

Ilnur Minniakhmetov
ilnur.minniakhmetov@mail.mcgill.ca

¹ COSMO – Stochastic Mine Planning Laboratory, Department of Mining and Materials Engineering, McGill University, FDA Building, 3450 University Street, Montreal, QC H3A 0E8, Canada

Keywords Sequential high-order simulation · Block support · Cross-support joint probability density function · High-order spatial statistics

1 Introduction

Stochastic simulation methods are used to quantify the spatial uncertainty and variability of pertinent attributes of natural phenomena in geosciences and geoengineering. Initial simulation methods were based on Gaussian assumptions and second-order statistics of corresponding random field models (Journel and Huijbregts 1978; David 1988; Goovaerts 1997). To address the limits of such Gaussian approaches, multiple point statistics (MPS)-based simulation methods were introduced (Guardiano and Srivastava 1993; Strebelle 2002; Zhang et al. 2006; Arpat and Caers 2007; Remy et al. 2009; Mariethoz et al. 2010; Mariethoz and Caers 2014; Mustapha et al. 2014; Chatterjee et al. 2016; Li et al. 2016; Zhang et al. 2017) to remove distributional assumptions, as well as to enable the reproduction of complex curvilinear and other geologic features by replacing the random field model with a framework built upon extraction of multiple point patterns from a training image (TI) or geological analogue. The main limitations of MPS methods are that they do not explicitly account for high-order statistics, nor do they provide consistent mathematical models as they generate TI-driven realizations. Previous studies have shown resulting realizations that comply with the TI used but do not necessarily reproduce the spatial statistics inferred from the data (Osterholt and Dimitrakopoulos 2007; Goodfellow et al. 2012). As an alternative, to address the above limitations, a high-order simulation (HOSIM) framework has been proposed as a natural generalization of the second-order-based random field paradigm (Dimitrakopoulos et al. 2010; Mustapha and Dimitrakopoulos 2010a, b, 2011; Minniakhmetov and Dimitrakopoulos 2017a, b; Minniakhmetov et al. 2018; Yao et al. 2018). The HOSIM framework does not make any assumptions about the data distribution, and the resulting realizations reproduce the high-order spatial statistics of the data. Similar to the MPS and most Gaussian simulation approaches, HOSIM methods generate realizations at the point support scale, whereas in most major areas of application, simulated realizations must be at the block support scale. Typically, the change of support scale needed is addressed by generating simulated realizations on a very dense grid of nodes that is then postprocessed to generate realizations at the block support size needed. This is a computationally demanding process, as related configurations may require extremely dense grids with on the order of many millions to billions of nodes. Thus, there is a need for computationally efficient methods that simulate directly at the block support scale.

In the context of conventional second-order geostatistics, direct block support simulation has been proposed. An approach termed “direct block simulation” was presented by Godoy (2003), which discretizes each block into several internal nodes, but only stores a single block value in memory for the next group simulation. This mechanism drastically reduces the amount of data stored in memory and saves considerable computational effort. The sequential direct block simulation method was expanded by Boucher and Dimitrakopoulos (2009) to incorporate multiple correlated variables by applying min/max autocorrelation factors. An explicit change of the support model

and direct simulation at block support scale were used by Emery (2009). Although efficient, these methods carry all the limitations of a Gaussian simulation framework, and the related spatial connectivity is limited to two-point spatial statistics, thus they remain unable to characterize non-Gaussian variables, complex nonlinear geological geometries, and the critically important connectivity of extreme values (Journel 2018). Alternatives are, therefore, needed.

High-order sequential simulation methods use high-order spatial cumulants to describe complex geologic configurations and high-order connectivity. At the same time, simulated realizations remain consistent with respect to the statistics of the available data, while capitalizing on the additional information that TIs can provide. These high-order spatial cumulants are described by Dimitrakopoulos et al. (2010) as combinations of moment statistical parameters. A high-order simulation algorithm was proposed by Mustapha and Dimitrakopoulos (2010a), where the conditional probability density functions (cpdf) are approximated by Legendre polynomials and high-order spatial cumulants. A template is defined based on the central node to be simulated and the nearest conditioning data. The replicates of this configuration are obtained from both the data and TI, and are used as input for the calculation of the Legendre coefficients in the cpdf approximation. Advantages of this method lie in the absence of assumptions on the distribution of the data and in being a data-driven approach. The Legendre polynomial was replaced by Legendre-like splines as the basis function for the estimation of conditional probabilities by Minniakhmetov et al. (2018). Results show a more stable approximation of the related cpdf. Improving upon the computational performance, Yao et al. (2018) proposed a new approach, where the calculation of the cpdf is simplified and no explicit calculation of cumulants is required. Although effective, the methods described above are performed at point support scale.

This paper presents a new method that generates high-order stochastic simulations directly at the block support scale. The technique considers overlapping grids representing a study area at two support scales, viz. point and block, where the simulation process is implemented at the latter. In the sequential simulation process followed, only the initial point support data and previously simulated blocks are added to the set of conditioning values, thus drastically reducing the number of elements stored in memory. The block to be simulated and the nearest conditioning data, at the point or block support scale, define the spatial configuration of the template used. Similarly, the TI is represented at both support scales to provide replicates of related spatial template configurations. The conditional cross-support joint density function estimated at each block is approximated by Legendre-like splines.

The remainder of the paper is organized as follows: First, the proposed model for high-order block support simulation is presented. Subsequently, a case study in a controlled environment assesses the performance of the current approach. Next, the method is applied to an actual gold deposit to demonstrate its practical aspects. Conclusions follow.

2 High-Order Block Support Simulation

2.1 Sequential Simulation

In the following description, the index V relates to elements at the block support, while P represents point support. Consider a stationary and ergodic non-Gaussian random field (RF) $Z_P(u_j)$ in R^n , where u_j defines the location of nodes j in the domain $D \subseteq R^n$. Now, consider a transformation function that takes the above point support RF to the block support RF. Any upscaling function can be applied, but assume Eq. (1) for simplicity

$$Z_V(v) = \frac{1}{|V|} \int_{u_j \in v} Z_P(u_j) du_j. \quad (1)$$

$Z_V(v_i)$ is also a RF, indexed as $v_i \in D \subseteq R^n$, $i = 1, \dots, N_V$, where N_V represents the total number of blocks to be simulated within the domain $D \subseteq R^n$. $Z_V(v_i)$ is the upscaled RF from $Z_P(u_j)$ considering all nodes u_j that are discretized within the block centered in v_i , where V is the volume.

The outcomes from the above RFs are denoted as $z_j^P = Z_P(u_j)$ and $z_i^V = Z_V(v_i)$, respectively, for the point and block support RF $Z_P(u_j) = Z_j^P$ and $Z_V(v_i) = Z_i^V$. Herein, the objective is to simulate a realization of the RF Z_i^V given the set of initial conditioning values at point support scale denoted as $d_P = \{z_1^P, \dots, z_{N_P}^P\}$, N_P being the total of the conditioning point support values. According to the sequential simulation theory in the geostatistical field, the joint probability density function (jpdf) $f_{Z_1^V, \dots, Z_{N_V}^V}$ can be decomposed into the products of the respective univariate distributions (Johnson 1987; Journel and Alabert 1989; Journel 1994; Goovaerts 1997; Dimitrakopoulos and Luo 2004)

$$\begin{aligned} f_{Z_1^V, \dots, Z_{N_V}^V}(z_1^V, z_2^V, \dots, z_{N_V}^V | d_P) &= f_{Z_1^V}(z_1^V | d_P) f_{Z_2^V, \dots, Z_{N_V}^V}(z_2^V, \dots, z_{N_V}^V | d_P, z_1^V) \\ &= f_{Z_1^V}(z_1^V | d_P) f_{Z_2^V}(z_2^V | d_P, z_1^V) \dots f_{Z_{N_V}^V}(z_{N_V}^V | d_P, z_1^V, z_2^V, \dots, z_{N_V-1}^V). \end{aligned} \quad (2)$$

According to Eq. (2), each block v^k is simulated based on the estimation of the conditional cross-support probability density function $f_{Z_k^V}(z_k^V | d_P, z_1^V, z_2^V, \dots, z_{k-1}^V)$, which according to Bayes' rule (Stuart and Ord 1987) is

$$f_{Z_k^V}(z_k^V | d_P, z_1^V, z_2^V, \dots, z_{k-1}^V) = \frac{f_Z(d_P, z_1^V, z_2^V, \dots, z_{k-1}^V, z_k^V)}{\int_D f_Z(d_P, z_1^V, z_2^V, \dots, z_{k-1}^V, z_k^V) dv_k}, \quad (3)$$

where $Z = Z_1^P, \dots, Z_{N_P}^P, Z_1^V, \dots, Z_{N_V}^V$. It is sufficient to approximate only the cross-support joint probability density function $f_Z(d_P, z_1^V, z_2^V, \dots, z_{k-1}^V, z_k^V)$. In this paper, this cross-support joint probability density function is approximated using Legendre-like orthogonal splines (Wei et al. 2013; Minniakhmetov et al. 2018).

2.2 Joint Probability Density Function Approximation

For simplicity, let $f(z)$ be the pdf of a random variable Z defined in $\Omega = [a, b]$ and let $\varphi_1(z), \varphi_2(z), \dots$ be a set of orthogonal functions defined in the same space Ω . Then, a fixed number ω of those orthogonal functions can approximate $f(z)$ (Lebedev 1965; Mustapha and Dimitrakopoulos 2010a; Minniakhmetov et al. 2018; Yao et al. 2018), when multiplied by the coefficients L_i

$$f(z) \approx \sum_{i=0}^{\omega} L_i \varphi_i(z). \quad (4)$$

Since the sets of functions are orthogonal

$$\int_a^b \varphi_i(z) \varphi_j(z) dz = \delta_{ij}, \quad (5)$$

where δ_{ij} is the Kronecker delta indexed by i and j , such that it take a unitary value if $i = j$ and 0 otherwise, using the definition of the expected value of one for a basis function

$$E[\varphi_i(z)] = \int_a^b \varphi_i(z) f(z) dz. \quad (6)$$

Replacing $f(z)$ as in Eq. (4) yields

$$\begin{aligned} E[\varphi_i(z)] &\approx \int_a^b \varphi_i(z) \sum_{j=0}^{\omega} L_j \varphi_j(z) dz = \sum_{j=0}^{\omega} L_j \int_a^b \varphi_j(z) \varphi_i(z) dz \\ &= \sum_{j=0}^{\omega} L_j \delta_{ij} = L_i. \end{aligned} \quad (7)$$

The coefficient L_i can be obtained experimentally from an available sample, thus $f(z)$ is approximated by Eq. (4).

Moving to the multivariate cross-support case, at every block location v^k the cross-support jpfd $f_Z(d_P, z_1^V, z_2^V, \dots, z_{k-1}^V, z_k^V)$ can be defined in a similar sense. Considering in practice that not all the samples are included as conditioning, n_V and n_P are the maximum number of elements at block support and point support scale, respectively, in the calculation. Hereinafter, the above cross-support jpfd is referred to as $f(z_0^V, \dots, z_{n_V}^V, z_1^P, \dots, z_{n_P}^P)$ to simplify the notation and ensure better understanding of the variables in both the block and point support layers. Also note that, without loss of generality, z_0^V is the value to be simulated at location v_0 . The cross-support jpfd is defined in the domain $[a, b]^{n_V+1} \times [a, b]^{n_P}$. Note that the interval for the block support is not necessarily the same as for the point support. This also applies

to the basis functions $\varphi_j(z)$, which could be discretized differently for both supports. Similarly to the univariate case, the cross-support jpjdf can be approximated as

$$f\left(z_0^V, \dots, z_{n_V}^V, z_1^P, \dots, z_{n_P}^P\right) \approx \sum_{k_0^V}^{\omega} \dots \sum_{k_{n_V}^V}^{\omega} \sum_{k_1^P}^{\omega} \dots \sum_{k_{n_P}^P}^{\omega} \left[L_{k_0^V \dots k_{n_V}^V k_1^P \dots k_{n_P}^P} \varphi_{k_0^V}\left(z_0^V\right) \dots \varphi_{k_{n_V}^V}\left(z_{n_V}^V\right) \varphi_{k_1^P}\left(z_1^P\right) \dots \varphi_{k_{n_P}^P}\left(z_{n_P}^P\right) \right]. \quad (8)$$

The coefficients $L_{i \dots j k \dots l}$ can be calculated experimentally, since they can be obtained from the orthogonality property of the basis functions. Following the definition of the expected value of a basis function, this is expressed as

$$\begin{aligned} E \left[\varphi_i\left(z_0^V\right) \dots \varphi_j\left(z_{n_V}^V\right) \varphi_k\left(z_1^P\right) \dots \varphi_l\left(z_{n_P}^P\right) \right] \\ = \int_a^b \dots \int_a^b \int_a^b \dots \int_a^b \varphi_i\left(z_0^V\right) \dots \varphi_j\left(z_{n_V}^V\right) \varphi_k\left(z_1^P\right) \\ \dots \varphi_l\left(z_{n_P}^P\right) f\left(z_0^V, \dots, z_{n_V}^V, z_1^P, \dots, z_{n_P}^P\right) dz_0^V \dots dz_{n_V}^V dz_1^P \dots dz_{n_P}^P. \end{aligned} \quad (9)$$

Replacing $f\left(z_0^V, \dots, z_{n_V}^V, z_1^P, \dots, z_{n_P}^P\right)$ as in Eq. (8) yields

$$\begin{aligned} E \left[\varphi_i\left(z_0^V\right) \dots \varphi_j\left(z_{n_V}^V\right) \varphi_k\left(z_1^P\right) \dots \varphi_l\left(z_{n_P}^P\right) \right] \\ \approx \int_a^b \dots \int_a^b \int_a^b \dots \int_a^b \varphi_i\left(z_0^V\right) \dots \varphi_j\left(z_{n_V}^V\right) \varphi_k\left(z_1^P\right) \dots \varphi_l\left(z_{n_P}^P\right) \\ \times \sum_{k_0^V}^{\omega} \dots \sum_{k_{n_V}^V}^{\omega} \sum_{k_1^P}^{\omega} \dots \sum_{k_{n_P}^P}^{\omega} \left[L_{k_0^V \dots k_{n_V}^V k_1^P \dots k_{n_P}^P} \varphi_{k_0^V}\left(z_0^V\right) \dots \varphi_{k_{n_V}^V}\left(z_{n_V}^V\right) \varphi_{k_1^P}\left(z_1^P\right) \dots \varphi_{k_{n_P}^P}\left(z_{n_P}^P\right) \right] dz_0^V \dots dz_{n_V}^V dz_1^P \dots dz_{n_P}^P \\ = \sum_{k_0^V}^{\omega} \dots \sum_{k_{n_V}^V}^{\omega} \sum_{k_1^P}^{\omega} \dots \sum_{k_{n_P}^P}^{\omega} \left[L_{k_0^V \dots k_{n_V}^V k_1^P \dots k_{n_P}^P} \int_a^b \dots \int_a^b \int_a^b \dots \int_a^b \varphi_i\left(z_0^V\right) \varphi_{k_0^V}\left(z_0^V\right) \dots \varphi_j\left(z_{n_V}^V\right) \varphi_{k_{n_V}^V}\left(z_{n_V}^V\right) \varphi_k\left(z_1^P\right) \varphi_{k_1^P}\left(z_1^P\right) \dots \varphi_l\left(z_{n_P}^P\right) \varphi_{k_{n_P}^P}\left(z_{n_P}^P\right) \right] \\ dz_0^V \dots dz_{n_V}^V dz_1^P \dots dz_{n_P}^P \\ = \sum_{k_0^V}^{\omega} \dots \sum_{k_{n_V}^V}^{\omega} \sum_{k_1^P}^{\omega} \dots \sum_{k_{n_P}^P}^{\omega} \left[L_{k_0^V \dots k_{n_V}^V k_1^P \dots k_{n_P}^P} \delta_{i k_0^V} \dots \delta_{j k_{n_V}^V} \delta_{k k_1^P} \delta_{l k_{n_P}^P} \right] = L_{i \dots j k \dots l}. \end{aligned} \quad (10)$$

Fig. 1 Example template τ with conditioning data capturing values at both point and block support scales

Now, to determine $L_{i...jk...l}$, the expected value from Eq. (10) is calculated from replicates of the training image according to a template defined from the simulation grid and sampling data.

Let $\tau = [v_0, \dots, v_{n_V}, u_1, \dots, u_{n_P}]$ be a template as in Fig. 1, where v_0 and v_1 represent locations of block support, and u_1, u_2 , and u_3 represent point support locations. v_0 is the location of the block to be simulated, and n_P and n_V are, respectively, the total number of points and blocks used as conditioning. In the figure, the grids at point and block support scale appear separated, but this is for visualization purposes only. In reality, they overlap with each other, and the distance between the layers is set to 0. τ is defined considering limited conditioning values, which are chosen in order of Euclidean proximity from the central block to be simulated. Having the specified template τ , the TI is scanned, and the replicates of such a template are retrieved. Note that τ has elements that belong to the point and block support scales. Similarly, the TI must be available at both scales. Therefore, assuming a TI input at the point support scale, it is rescaled to block support scale, and both are retrieved during the simulation process, each in its respective layer.

The algorithm for the block support high-order simulation method is as follows:

1. Upscale the TI from point to block support scale.
2. Define a random path to visit all the unsampled block locations on the simulation grid.
3. At each v^0 block location:
  - (a) Find the nearest conditioning point and block support values.
  - (b) Obtain the template τ according to the configuration of the central block and related conditioning values at both support scales.
  - (c) Scan the training images, searching for replicates of the template τ and corresponding values.
  - (d) Calculate all the spatial cross-support coefficients $L_{i...jk...l}$ using Eq. (10).
  - (e) Derive the conditional cross-support jpdf $f_{Z_0^V}(z_0^V | d_P, z_1^V, z_2^V, \dots, z_k^V)$ according to Eqs. (8) and (3).
  - (f) Draw a uniform value from $[0, 1]$ to sample z_0^V from the conditional cumulative distribution derived from the above.

- (g) Add z_0^V to the simulation grid at block support scale so that it can be a conditioning value for the next block.
4. Repeat steps 2 and 3 for additional realizations.

2.3 Approximation of a Joint Probability Density Using Legendre-Like Orthogonal Splines

The current paper uses Legendre-like splines (Wei et al. 2013; Minniakhmetov et al. 2018) as a means to obtain the basis function mentioned above. In short, those splines are a combination of Legendre polynomials (Lebedev 1965) up to order r and linear combinations of B-splines (de Boor 1978). B-splines are a particular class of piecewise polynomials (splines) connected by some condition of continuity, and by themselves do not form an orthogonal basis. Thus, as introduced in Wei et al. (2013), the first $r+1$ splines are the Legendre polynomials, which can be defined as (Lebedev 1965)

$$\varphi_r = \frac{1}{2^r r!} \left(\frac{d^r}{dz^r} \right) \left[(z^2 - 1)^r \right], \quad -1 \leq z \leq 1. \quad (11)$$

The additional functions are constructed given the domain

$$T = \underbrace{\{a, a, \dots, t_0 = a < t_1 \leq t_2 \leq \dots \leq t_{m_{\max}} < t_{m_{\max}+1} = b, b, \dots, b\}}_{r+1}, \quad (12)$$

where the elements t_i are referred to as knots and $m_{\max}$ represents the maximum number of knots; note that Minniakhmetov et al. (2018) present a study on how to choose $m_{\max}$ to obtain computationally stable polynomial approximations. The final Legendre-like splines are defined as

$$\varphi_{r+m}(t) = \frac{d^{r+1}}{dt^{r+1}} f_m(t), \quad m = 1 \dots m_{\max}. \quad (13)$$

$f_m(t)$ is the determinant of the following matrix:

$$f_m(t) = \det \begin{pmatrix} B_{-r, 2r+1, m}(t) & B_{-r+1, 2r+1, m}(t) & \cdots & B_{-r+m-1, 2r+1, m}(t) \\ B_{-r, 2r+1, m}(t_1) & B_{-r+1, 2r+1, m}(t_1) & \vdots & B_{-r+m-1, 2r+1, m}(t_1) \\ \vdots & \vdots & \ddots & \vdots \\ B_{-r, 2r+1, m}(t_{m-1}) & B_{-r+1, 2r+1, m}(t_{m-1}) & \cdots & B_{-r+m-1, 2r+1, m}(t_{m-1}) \end{pmatrix}, \quad (14)$$

which is constructed from the auxiliary splines $B_{i, r, m}(t)$ of order r , obtained according to the recursive rule

$$B_{i, 0, m} = \begin{cases} 1, & t_{i, m} \leq t \leq t_{i+1, m} \\ 0, & \text{otherwise} \end{cases},$$

$$B_{i, r, m}(t) = \frac{t - t_{i, m}}{t_{i+r-1, m} - t_{i, m}} B_{i, r-1, m}(t) + \frac{t_{i+r, m} - t}{t_{i+r, m} - t_{i+1, m}} B_{i+1, r-1, m}(t). \quad (15)$$

Fig. 2 Exhaustive image V **a** at point support scale, **b** at block support scale, and **c** 234 samples from the image in **a**

These auxiliary functions are defined on the knot sequence $T_m = \{t_{i,m}\}_{i=-r}^{r+m+1}$, $m = 1 \dots m_{\max} - 1$, and the term $t_{i,m}$ is defined as

$$t_{i,m} = \begin{cases} a, & -r \leq i \leq 0 \\ t_i, & 1 \leq i \leq m \\ b, & m+1 \leq i \leq m+r+1 \end{cases}. \quad (16)$$

3 Testing with an Exhaustive Dataset

The method outlined above is tested using the two-dimensional image of the Walker Lake dataset (Isaaks and Srivastava 1989). This exhaustive dataset comprises two correlated variables U and V with sizes of 260×300 pixels. Random stratified sampling is used to retrieve 234 values from or 0.3 % of the exhaustive image V to be used as the dataset in the direct block simulation of V , to test the proposed method. The full dataset V is converted from the point to a block support representation by averaging over 5×5 pixels. This block support version is referred to here as the fully known reference image and is used for comparisons. Figure 2 shows V at the point and block support scale, as well as the dataset to be used. The image U is chosen as the training image in the simulation process. Figure 3 presents the TI at both point and block support (5×5 unit size) scales. To help the method find more meaningful spatial patterns of

Fig. 3 Training image U at **a** point support scale and **b** block support scale

Fig. 4 Histogram of data, reference, and training image at point support scale

Table 1 Basic statistics of dataset, training image, and fully known image at point support scale

Basic statistic	Dataset	Reference image	Training image
Average	277.0	278	278.5
Median	193.5	221.3	193.6
Variance	68,926	62,423	70,385

the potential conditioning templates, the histogram of the TI is matched to that of the dataset. Histograms of the exhaustively known image, TI, and dataset are displayed in Fig. 4, and basic statistics are presented in Table 1.

The test conducted consists of generating 15 simulated realizations of the V dataset at block support scale, using the data and the training image mentioned above. Note that the maximum number of knots used (Eq. 12) is 50, which provides computationally efficient and stable polynomial approximations, as appropriate. Figure 5 shows three of the simulated realizations generated, and Table 2 presents the statistics related to the average of the 15 simulations, training image, and reference image at block scale. Comparison of Figs. 2b and 5 suggests that the simulations reproduce the main structures of low and high values of the fully known reference image V . The histograms and variograms presented in Figs. 6 and 7 reasonably follow the behavior exhibited by the variogram model from the data and training image, respectively. Note that

Fig. 5 Three example simulated realizations of the Walker Lake reference image V

Table 2 Basic statistics of the average of the simulations, training image, and reference image at block support scale

Basic statistic	Simulations	Training image	Reference image
Average	280.6	278.5	278.0
Median	234.6	237.0	235.4
Variance	51,764	46,518	52,304

Fig. 6 Histograms of the simulations at block support scale, and comparison with reference and training image also at block support scale

the variograms of the data are computed at point scale and rescaled to represent the corresponding volume–variance relation (Journal and Huijbregts 1978).

Spatial cumulants (Dimitrakopoulos et al. 2010) can quantify the spatial relationships between three and more points and are used herein to assess high-order spatial patterns. The third-order cumulant maps are presented along with the template used

Fig. 7 Variograms of simulated realizations, exhaustive image, TI, and variogram from data rescaled to block support variance: **a** WE direction, and **b** NS direction

Fig. 8 Third-order cumulant maps for **a** point support data used, **b** fully known block support image V , **c** training image, and **d** the average map of the 15 simulated realizations

for its calculation in Fig. 8. Figure 9 shows the fourth-order cumulant map, where three slices of the complete cumulant map and the related template are displayed. In both figures, the color ranges from blue to red, representing lower to higher spatial intercorrelation between values. Note that the reference and training image high-order maps are calculated on block support scale, while the cumulant maps related to each simulation are averaged to a single map using the 15 stochastic simulated realizations at block support scale. During the calculation of the high-order spatial statistics from the data, only a few replicates are obtained and Fig. 8a presents a smooth interpolation using B-splines. Regarding the third-order maps, the average of the simulations match the spatial features observed in the data and fully known dataset. It also shares

Fig. 9 Slices of the fourth-order cumulant maps for **a** fully known image *V*, **b** training image, and **c** average map of the 15 simulations, all at block support scale

similarities with the third-order cumulants map from TI; this is somewhat expected as the process captures high-order relations from the TI at block support scale as well. These spatial relations present in the TI end up being present in the realizations as well.

The fourth-order cumulant map reproduces the characteristics that are closer to the TI than the fully known image, as expected. Note that, by explicitly calculating the spatial high-order cumulants, the information received from the training image to infer local cross-support distributions is conditioned to the data.

4 Application at a Gold Deposit

This section applies the proposed method at a gold deposit. The dataset comprises 2300 drillholes spaced approximately in a $35 \times 35 \text{ m}^2$ configuration, covering an area of 4.5 km^2 . The training image is defined on $405 \times 445 \times 43$ grid blocks of size $5 \times 5 \times 10 \text{ m}^3$ and is based on blasthole samples. Both inputs are composited in a 10 m bench and are considered to be at point support scale. Figure 10 presents the drillholes available and the training image at block scale. The deposit to be simulated is represented by 510,800 blocks, each measuring $10 \times 10 \times 10 \text{ m}^3$.

Fifteen simulated realizations are generated; cross-sections from two of them are presented in Fig. 11 to show similarities with the data and TI in the corresponding cross-section in Fig. 10. Notable is the reproduction of a sharp transition from high

Fig. 10 **a** Cross-section of the available drillhole locations, and **b** training image at block support scale

Fig. 11 Cross-section of two simulated realizations

Fig. 12 Histograms of simulated realizations and training image

to low grades. Figure 12 shows the histograms of the simulations and TI at block support scale. Table 3 presents the related statistics. Variograms at block support scale are displayed in Fig. 13, where the data variogram is regularized to reflect the corresponding volume–variance relation. The second-order spatial statistics of the simulations match reasonably with the pattern followed by the data and are close to

Table 3 Basic statistics of the average of the simulations and training image at block support scale and dataset at point scale

Basic statistic	Simulations	TI block support	Data point support
Average	0.63	0.62	0.63
Median	0.40	0.39	0.39
Variance	0.74	0.76	1.81

Fig. 13 Variograms of simulated realizations and training image and data variograms rescaled to represent block variance: WE direction (left) and NS direction (right)

Fig. 14 Third-order cumulant maps, obtained with the template on the left, for the **a** dataset, **b** training image at block support, and **c** average map of the 15 simulations

Fig. 15 Three slices of the fourth-order cumulant maps, obtained with the template at the bottom, for the **a** dataset, **b** training image at block support, and **c** average map of the 15 simulations

those of the TI. Results for third- and fourth-order cumulants and related maps for the data, TI, and simulated realizations are shown in Figs. 14 and 15, respectively. Note that the high-order statistics of the simulated realizations match those of the data and TI.

Further highlighting the advantages of the proposed direct block high-order simulation method, note that, for this case study, the runtime of the related algorithm was approximately 5.5 h, while the point high-order simulation requires approximately 24 h. Both approaches are tested with the same specifications and computing equipment: Intel® Core™ i7-7700 CPU with 3.60 GHz, 16 GB of RAM, running under Windows 7.

5 Conclusions

This paper presents a new high-order simulation method that simulates directly at block support scale by estimating, at every block location, the cross-support joint probability density function. Legendre-like splines are the set of basis functions used to approximate the above density function. The related coefficients are calculated from replicates of a spatial template employed. The latter template is generated from

the configuration of the block to be simulated and associated conditioning values, whose support can be at both point and block scale. The high-order character of the proposed direct block simulation method ensures that the generated realizations reflect the complex, nonlinear spatial characteristics of the variables being simulated and reproduce the connectivity of extreme values.

The proposed algorithm is tested using an exhaustive image, showing that the different realizations generated can reasonably reproduce spatial architectures observed in the exhaustive image. An application at a gold deposit shows the practical aspects of the method. In addition, it documents that the method works well, while simulated realizations are shown to reproduce the spatial statistics of the available data up to the cumulants of fourth order that were calculated. Further work will focus on improving the computational efficiency, generating training images that are consistent with the high-order relations in the available data, and extending the proposed method to jointly simulate multiple variables.

Acknowledgements This work is funded by the National Science and Engineering Research Council of Canada, Natural Science and Engineering Research Council of Canada (NSERC) CRD Grant CRDPJ 500414-16, the COSMO mining industry consortium (AngloGold Ashanti, Barrick Gold, BHP, De Beers, IAMGOLD, Kinross, Newmont Mining and Vale) NSERC Discovery Grant 239019, and the IAMG by the 2017 Mathematical Geosciences Student Award.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Arpat GB, Caers J (2007) Conditional simulation with patterns. *Math Geol* 39(2):177–203. <https://doi.org/10.1007/s11004-006-9075-3>
- Boucher A, Dimitrakopoulos R (2009) Block simulation of multiple correlated variables. *Math Geosci* 41(2):215–237. <https://doi.org/10.1007/s11004-008-9178-0>
- Chatterjee S, Mustapha H, Dimitrakopoulos R (2016) Fast wavelet-based stochastic simulation using training images. *Comput Geosci* 20(3):399–420. <https://doi.org/10.1007/s10596-015-9482-y>
- David M (1988) Handbook of applied advanced geostatistical ore reserve estimation. Elsevier, Amsterdam
- de Boor C (1978) A practical guide to splines. Springer, Berlin
- Dimitrakopoulos R, Luo X (2004) Generalized sequential Gaussian simulation on group size v and screen-effect approximations for large field simulations. *Math Geol* 36(5):567–590. <https://doi.org/10.1023/B:MATG.0000037737.11615.df>
- Dimitrakopoulos R, Mustapha H, Gloaguen E (2010) High-order statistics of spatial random fields: exploring spatial cumulants for modeling complex non-Gaussian and non-linear phenomena. *Math Geosci* 42(1):65–99. <https://doi.org/10.1007/s11004-009-9258-9>
- Emery X (2009) Change-of-support models and computer programs for direct block-support simulation. *Comput Geosci* 35(10):2047–2056. <https://doi.org/10.1016/j.cageo.2008.12.010>
- Godoy M (2003) The effective management of geological risk in long-term production scheduling of open pit mines. Ph.D. Thesis, University of Queensland, Brisbane, QLD, Australia
- Goodfellow R, Albor Consuegra F, Dimitrakopoulos R, Lloyd T (2012) Quantifying multi-element and volumetric uncertainty, Coleman McCreedy deposit, Ontario, Canada. *Comput Geosci* 42:71–78. <https://doi.org/10.1016/j.cageo.2012.02.018>
- Goovaerts P (1997) Geostatistics for natural resources evaluation. Oxford University Press, New York
- Guardiano FB, Srivastava RM (1993) Multivariate geostatistics: beyond bivariate moments. In: Soares A (ed) *Geostatistics Tróia'92*, vol 1. Springer, Dordrecht, pp 133–144

- Isaaks EH, Srivastava RM (1989) Applied geostatistics. Oxford University Press, Oxford
- Johnson ME (1987) Multivariate statistical simulation. Wiley, Hoboken
- Journel AG (1994) Modeling uncertainty: some conceptual thoughts. In: Dimitrakopoulos R (ed) Geostatistics for the next century. Springer, Dordrecht, pp 30–43
- Journel AG (2018) Roadblocks to the evaluation of ore reserves - the simulation overpass and putting more geology into numerical models of deposits. In: Dimitrakopoulos R (ed) Advances in applied strategic mine planning. Springer, Heidelberg, pp 47–55
- Journel AG, Alabert F (1989) Non-Gaussian data expansion in the earth sciences. *Terra Nova* 1(2):123–134. <https://doi.org/10.1111/j.1365-3121.1989.tb00344.x>
- Journel AG, Huijbregts CJ (1978) Mining geostatistics. Blackburn, New York
- Lebedev NN (1965) Special functions and their applications. Prentice-Hall, New York
- Li X, Mariethoz G, Lu DT, Linde N (2016) Patch-based iterative conditional geostatistical simulation using graph cuts. *Water Resour Res* 52(8):6297–6320. <https://doi.org/10.1002/2015WR018378>
- Mariethoz G, Caers J (2014) Multiple-point geostatistics: stochastic modeling with training images. Wiley, Hoboken
- Mariethoz G, Renard P, Straubhaar J (2010) The direct sampling method to perform multiple-point geostatistical simulations. *Water Resour Res* 46(11):1–14. <https://doi.org/10.1029/2008WR007621>
- Minniakhmetov I, Dimitrakopoulos R (2017a) Joint high-order simulation of spatially correlated variables using high-order spatial statistics. *Math Geosci* 49(1):39–66. <https://doi.org/10.1007/s11004-016-9662-x>
- Minniakhmetov I, Dimitrakopoulos R (2017b) A high-order, data-driven framework for joint simulation of categorical variables. In: Gómez-Hernández JJ, Rodrigo-Ilarri J, Rodrigo-Clavero ME et al (eds) Geostatistics Valencia 2016. Springer, Cham, pp 287–301
- Minniakhmetov I, Dimitrakopoulos R, Godoy M (2018) High-order spatial simulation using Legendre-like orthogonal splines. *Math Geosci* 50(7):753–780. <https://doi.org/10.1007/s11004-018-9741-2>
- Mustapha H, Dimitrakopoulos R (2010a) High-order stochastic simulation of complex spatially distributed natural phenomena. *Math Geosci* 42(5):457–485. <https://doi.org/10.1007/s11004-010-9291-8>
- Mustapha H, Dimitrakopoulos R (2010b) A new approach for geological pattern recognition using high-order spatial cumulants. *Comput Geosci* 36(3):313–334. <https://doi.org/10.1016/j.cageo.2009.04.015>
- Mustapha H, Dimitrakopoulos R (2011) HOSIM: a high-order stochastic simulation algorithm for generating three-dimensional complex geological patterns. *Comput Geosci* 37(9):1242–1253. <https://doi.org/10.1016/j.cageo.2010.09.007>
- Mustapha H, Chatterjee S, Dimitrakopoulos R (2014) CDFSIM: efficient stochastic simulation through decomposition of cumulative distribution functions of transformed spatial patterns. *Math Geosci* 46(1):95–123. <https://doi.org/10.1007/s11004-013-9490-1>
- Osterholt V, Dimitrakopoulos R (2007) Simulation of orebody geology with multiple-point geostatistics—application at Yandi channel iron ore deposit, WA, and implications for resource uncertainty. In: Dimitrakopoulos R (ed) Orebody modelling and strategic mine planning. AusIMM Spectrum series 14, pp 51–60
- Remy N, Boucher A, Wu J (2009) Applied geostatistics with SGEMS: a user's guide. Cambridge University Press, Cambridge
- Strebelle S (2002) Conditional simulation of complex geological structures using multiple-point statistics. *Math Geol* 34(1):1–21. <https://doi.org/10.1023/A:1014009426274>
- Stuart A, Ord JK (1987) Kendall's advanced theory of statistics, 5th edn. Oxford University Press, New York
- Wei Y, Wang G, Yang P (2013) Legendre-like orthogonal basis for spline space. *CAD Comput Aided Des* 45(2):85–92. <https://doi.org/10.1016/j.cad.2012.07.011>
- Yao L, Dimitrakopoulos R, Gamache M (2018) A new computational model of high-order stochastic simulation based on spatial Legendre moments. *Math Geosci* 50(8):929–960. <https://doi.org/10.1007/s11004-018-9744-z>
- Zhang T, Switzer P, Journel A (2006) Filter-based classification of training image patterns for spatial simulation. *Math Geol* 38(1):63–80. <https://doi.org/10.1007/s11004-005-9004-x>
- Zhang T, Gelman A, Laronga R (2017) Structure- and texture-based fullbore image reconstruction. *Math Geosci* 49(2):195–215. <https://doi.org/10.1007/s11004-016-9649-7>