

High-Order Spatial Simulation Using Legendre-Like Orthogonal Splines

Ilnur Minniakhmetov¹  ·
Roussos Dimitrakopoulos¹ · Marcelo Godoy²

Received: 28 June 2017 / Accepted: 31 March 2018 / Published online: 17 May 2018
© The Author(s) 2018

Abstract High-order sequential simulation techniques for complex non-Gaussian spatially distributed variables have been developed over the last few years. The high-order simulation approach does not require any transformation of initial data and makes no assumptions about any probability distribution function, while it introduces complex spatial relations to the simulated realizations via high-order spatial statistics. This paper presents a new extension where a conditional probability density function (cpdf) is approximated using Legendre-like orthogonal splines. The coefficients of spline approximation are estimated using high-order spatial statistics inferred from the available sample data, additionally complemented by a training image. The advantages of using orthogonal splines with respect to the previously used Legendre polynomials include their ability to better approximate a multidimensional probability density function, reproduce the high-order spatial statistics, and provide a generalization of high-order simulations using Legendre polynomials. The performance of the new method is first tested with a completely known image and compared to both the high-order simulation approach using Legendre polynomials and the conventional sequential Gaussian simulation method. Then, an application in a gold deposit demonstrates the advantages of the proposed method in terms of the reproduction of histograms, variograms, and high-order spatial statistics, including connectivity measures. The C++ course code of the high-order simulation implementation presented

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s11004-018-9741-2>) contains supplementary material, which is available to authorized users.

✉ Ilnur Minniakhmetov
ilnur.minniakhmetov@mail.mcgill.ca

¹ COSMO—Stochastic Mine Planning Laboratory, McGill University, Montreal, QC H3A 0E8, Canada

² Newmont Mining Corporation, Denver, CO, USA

herein, along with an example demonstrating its utilization, are provided online as supplementary material.

Keywords Stochastic simulation · Orthogonal splines · High-order spatial statistics · Non-Gaussian distribution · Spatial complexity

1 Introduction

Geostatistical simulations are used to quantify the uncertainty of spatially distributed attributes of interest describing mineral deposits, petroleum reservoirs, hydrogeological horizons, environmental contaminants, and other spatially variant natural phenomena. Since the 1990s, multiple-point spatial simulation (MPS) methods and variations (Guardiano and Srivastava 1993; Strebelle 2002; Journel 2005, 2018; Zhang et al. 2006; Arpat and Caers 2007; Chugunova and Hu 2008; de Vries et al. 2009; Mariethoz and Renard 2010; Mariethoz et al. 2010; Straubhaar et al. 2011; De Iaco and Maggio 2011; Honarkhah 2011; Strebelle and Cavelius 2014; Chatterjee et al. 2012; Lochbühler et al. 2014; Mustapha et al. 2014; Rezaee et al. 2013; Toftaker and Tjelmeland 2013; Zhang et al. 2017; others) have been developed to advance the simulation methods beyond the past generation of second-order spatial statistics, which were typically based on Gaussian processes (e.g., Journel and Huijbregts 1978; David 1988; Goovaerts 1998; Chilès and Delfiner 1999). A limitation of MPS approaches is that they are largely algorithmic and do not consistently account for the high-order spatial relations in the available sample data. Patterns and complex spatial relations are derived from the so-termed training images (TIs), or geological analogues, rather than from sample data; this is a critical topic for relatively data-rich type applications, where data statistics have been shown to not be reproduced by simulated realizations based on MPS methods (Osterholt and Dimitrakopoulos 2018; Goodfellow et al. 2012). To address some of these limits, high-order simulation techniques for complex and non-Gaussian spatially distributed variables have also been developed (Mustapha and Dimitrakopoulos 2010, 2011; Mustapha et al. 2011; Tamayo-Mas et al. 2016; Minnakhmetov and Dimitrakopoulos 2017a), based on generating conditional distributions through Legendre polynomials (Lebedev 1965) and high-order spatial cumulants. Yao et al. (2018) developed a new computational model to significantly reduce the computation cost of the method. The high-order simulation approach does not require any transformation of initial data and makes no assumptions about the related probability distribution function. The approach reproduces high-order spatial statistics of sample data and a training image. The high-order spatial statistics are shown to capture directional multiple-point periodicity, connectivity of extreme values, and complex spatial architecture (Dimitrakopoulos et al. 2010). However, polynomial approximations do not always converge to analytic functions (Runge 1901; Boyd and Ong 2009; Fornberg and Zuev 2007). In addition, the high-order polynomials are very sensitive to rounding errors for values near the endpoints of an approximation domain; therefore, even if the interpolants converge in theory, they will diverge rapidly when computed (Platte et al. 2011). This is critical for the simulation of extreme values, as they are located at the endpoints of an approximation domain. In an effort to improve upon the limitations

of polynomial approximation, a spline approximation of complex multidimensional functions is considered here (Piegl 1989; Hughes et al. 2005; Ruiu et al. 2016).

Splines are piecewise-defined polynomial functions in which pieces are connected by some condition of smoothness. The places where these pieces meet are called knots, and two adjacent knots form a knot interval, hereafter referenced simply as an interval. Knot locations have a significant impact on the quality and flexibility of approximation, particularly in the approximation of functions with discontinuities (López de Silanes et al. 2001; Sinha and Schunck 1992) and functions with locally high gradients (Malagù et al. 2014). Furthermore, through the proper choosing of the knot sequence, splines can accurately approximate very complex functions, such as the shapes of three-dimensional objects in a computer-aided geometric design (Hoschek and Lasser 1993; Park and Lee 2007). Therefore, splines are chosen herein to approximate complex multidimensional joint distributions. The most commonly used mathematical formulation in different applications of splines are B-splines (from basis spline). The construction of B-splines is straightforward and simple to implement; however, the high-order simulation framework proposed by Mustapha and Dimitrakopoulos (2010) assumes the orthogonality of basis functions, and, therefore, splines in the form of B-splines are not suitable for a high-order spatial simulation approach.

In this paper, Legendre-like splines (Wei et al. 2013) are used, which are shown to be orthogonal and can be easily integrated in the high-order simulation framework. There are two user-defined parameters used for the constructing of Legendre-like splines: the order of splines and the maximum number of knots. In practice, cubic splines (order 3) are commonly used (Hughes et al. 2005; Piegl 1989), as they provide efficient smooth approximation. For the cubic splines, the first four Legendre-like splines are defined at two endpoint knots of the approximation domain and Legendre polynomials up to order 3. Next, Legendre-like splines are constructed by adding an additional knot per Legendre-like spline until the user-defined maximum number of knots is reached. Increasing the number of knots improves the approximation and describes more complex relations in the available data in the same way that the high-order polynomials capture the complex behavior of the function to be approximated. Thus, the maximum number of knots reflects the maximum order of high-order spatial relations that can be calculated from the available data. This spline approach aims to improve the estimation of the conditional probability density function (cpdf) and overcome the limitations of polynomial approximations. In addition, the proposed approach provides a general framework for high-order simulation techniques. For example, by using only one interval for spline construction, the technique becomes the one proposed by Mustapha and Dimitrakopoulos (2010, 2011).

The paper is organized as follows. First, the high-order simulation framework is outlined. Then, two systems of basis functions are outlined: Legendre polynomials (Lebedev 1965) and Legendre-like orthogonal splines. In the following section, the capabilities of both systems are compared using a fully known dataset to demonstrate the advantages of orthogonal splines in simulating connected high values. Next, the proposed approach is applied to a gold deposit and compared with the sequential Gaussian simulation approach in terms of the reproduction of histograms,

variograms, high-order spatial statistics, and the connectivity of high values. Discussion and conclusions follow. Supplementary material available online provides the C++ course code of the high-order sequential simulation implementation detailed in Sect. 2.

2 Sequential High-Order Simulation

Let $Z(\mathbf{u}_i)$ be a stationary ergodic random field indexed in R^n , where $\mathbf{u}_i \in D \subseteq R^n$ ($n = 1, 2, 3$), $i = 1 \dots N$ and where N is the number of points in a discrete grid $D \subseteq R^n$. Random variables indexed on the grid $D \subseteq R^n$ are denoted by $Z_i \equiv Z(\mathbf{u}_i)$, whereas their outcomes are denoted by $z_i = z(\mathbf{u}_i)$. The focus of high-order simulation techniques is to simulate the realization of the random field $Z(\mathbf{u}_i)$ for all nodes of a grid D with a given set of conditioning data $\mathbf{d}_n = \{z(\mathbf{u}_\alpha), \alpha = 1 \dots n\}$.

The joint probability density function $f(\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_N; z_0, z_1, \dots, z_N | \mathbf{d}_n)$ of the random field $Z(\mathbf{u}_i)$ can be decomposed into the product of conditional univariate distributions using the basic concept of sequential simulation (Journel and Alabert 1989; Journel 1994; Dimitrakopoulos and Luo 2004)

$$\begin{aligned} & f(\mathbf{u}_1, \dots, \mathbf{u}_N; z_1, \dots, z_N | \mathbf{d}_n) \\ &= f(\mathbf{u}_2, \dots, \mathbf{u}_N; z_2, \dots, z_N | z_1, \mathbf{d}_n) f(\mathbf{u}_1; z_1 | \mathbf{d}_n) \\ &= f(\mathbf{u}_3, \dots, \mathbf{u}_N; z_3, \dots, z_N | z_1, z_2, \mathbf{d}_n) f(\mathbf{u}_2; z_2 | z_1, \mathbf{d}_n) f(\mathbf{u}_1; z_1 | \mathbf{d}_n) \\ &= \prod_{i=2}^N f(\mathbf{u}_i; z_i | z_1, \dots, z_{i-1}, \mathbf{d}_n) f(\mathbf{u}_1; z_1 | \mathbf{d}_n). \end{aligned} \quad (1)$$

Accordingly, the random path of visiting all grid nodes is defined first. Then, starting from the first node in the random path, the value z_i is simulated based on the estimated cpdf $f(\mathbf{u}_i; z_i | z_1, \dots, z_{i-1}, \mathbf{d}_n)$. Finally, the simulated value is added to the set of conditional data, and the process is repeated until all grid nodes in the random path are visited. Eventually, any resulting simulation represents a realization of the complex joint distribution $f(\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_N; z_0, z_1, \dots, z_N | \mathbf{d}_n)$.

Without loss of generality, let \mathbf{u}_0 be the first node in the random path. According to Bayes' rule (Lee 2012)

$$f(\mathbf{u}_0; z_0 | \mathbf{d}_n) = \frac{f(\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_n; z_0, \mathbf{d}_n)}{f(\mathbf{u}_1, \dots, \mathbf{u}_n; \mathbf{d}_n)}, \quad (2)$$

where $f(\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_n; z_0, \mathbf{d}_n)$ is a joint probability density function and $f(\mathbf{u}_1, \dots, \mathbf{u}_n; \mathbf{d}_n)$ can be calculated as

$$f(\mathbf{u}_1, \dots, \mathbf{u}_n; \mathbf{d}_n) = \int f(\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_n; \xi_0, \mathbf{d}_n) d\xi_0. \quad (3)$$

In this paper, the joint probability density function $f(\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_n; z_0, \mathbf{d}_n)$ is approximated using Legendre polynomials and Legendre-like orthogonal splines.

2.1 Approximation of a Joint Probability Density Using Orthogonal Functions

Let $f(z)$ be a probability density function of a random variable Z defined on $[a, b]$ and let $\varphi_1(z), \varphi_2(z), \dots$ be a complete system of orthogonal functions in $[a, b]$, then $f(z)$ can be approximated by the finite number ω of functions $\varphi_1(z), \varphi_2(z), \dots \varphi_\omega(z)$

$$f(z) \approx \sum_{m=0}^{\omega} L_m \varphi_m(z), \quad (4)$$

where L_m are coefficients of approximation. The system of functions $\varphi_1(z), \varphi_2(z), \dots \varphi_\omega(z)$ is orthogonal

$$\int_a^b \varphi_k \varphi_m(z) dz = \delta_{km}, \quad (5)$$

where $\delta_{mk} = \begin{cases} 1, & m = k \\ 0, & m \neq k \end{cases}$ is the Kronecker delta, and, therefore, $\forall k = 0 \dots \omega$

$$\int_a^b \varphi_k(z) f(z) dz \approx \int_a^b \varphi_k \sum_{m=0}^{\omega} L_m \varphi_m(z) dz = \sum_{m=0}^{\omega} L_m \int_a^b \varphi_k \varphi_m(z) dz = \sum_{m=0}^{\omega} L_m \delta_{mk} = L_k. \quad (6)$$

By definition

$$E[\varphi_k(z)] = \int_a^b \varphi_k(z) f(z) dz, \quad (7)$$

where E stands for mathematical expectation. The coefficients L_m can be estimated from the available data. Similarly, a joint probability density function $f(z_0, z_1, \dots, z_n)$ of a set of random variables Z_0, Z_1, \dots, Z_n defined on $[a, b] \times [a, b] \times \dots \times [a, b]$ can be approximated as

$$f(z_0, z_1, \dots, z_n) \approx \sum_{m_0=0}^{\omega_0} \sum_{m_1=0}^{\omega_1} \dots \sum_{m_n=0}^{\omega_n} L_{m_0, m_1, \dots, m_n} \varphi_{m_0}(z_0) \varphi_{m_1}(z_1) \dots \varphi_{m_n}(z_n). \quad (8)$$

Coefficients L_{m_0, m_1, \dots, m_n} are obtained from the orthogonality property

$$\begin{aligned} & \int_a^b \int_a^b \cdots \int_a^b \varphi_{k_0}(z_0) \varphi_{k_1}(z_1) \cdots \varphi_{k_n}(z_n) f(z_0, z_1, \dots, z_n) |dz| \approx \\ & \sum_{m_0=0}^{\omega_0} \sum_{m_1=0}^{\omega_1} \cdots \sum_{m_n=0}^{\omega_n} L_{m_0, m_1, \dots, m_n} \int_a^b \int_a^b \cdots \int_a^b \varphi_{k_0}(z_0) \varphi_{m_0}(z_0) \cdots \varphi_{k_n}(z_n) \varphi_{m_n}(z_n) |dz| \\ & = \sum_{m_0=0}^{\omega_0} \sum_{m_1=0}^{\omega_1} \cdots \sum_{m_n=0}^{\omega_n} L_{m_0, m_1, \dots, m_n} \delta_{m_0 k_0} \delta_{m_1 k_1} \cdots \delta_{m_n k_n} = L_{k_0, k_1, \dots, k_n}, \forall k_0, k_1, \dots, k_n = 0 \dots \omega, \end{aligned} \quad (9)$$

where $|dz| = dz_0 dz_1 \cdots dz_n$. By definition

$$\begin{aligned} & E[\varphi_{k_0}(z_0) \varphi_{k_1}(z_1) \cdots \varphi_{k_n}(z_n)] \\ & = \int_a^b \int_a^b \cdots \int_a^b \varphi_{k_0}(z_0) \varphi_{k_1}(z_1) \cdots \varphi_{k_n}(z_n) f(z_0, z_1, \dots, z_n) |dz|. \end{aligned} \quad (10)$$

When considering the spatial locations $\mathbf{u} = \{\mathbf{u}_0, \mathbf{u}_1, \dots, \mathbf{u}_n\}$ of random variables Z_0, Z_1, \dots, Z_n , the coefficients L_{k_0, k_1, \dots, k_n} can be estimated from available data using Eqs. (9) and (10) by calculating

$$\begin{aligned} L_{k_0, k_1, \dots, k_n} & \approx E[\varphi_{k_0}(z_0) \varphi_{k_1}(z_1) \cdots \varphi_{k_n}(z_n)] \\ & \approx \frac{1}{N_{h_1, h_2, \dots, h_n}} \sum_{k=1}^{N_{h_1, h_2, \dots, h_n}} \varphi_{k_0}(z_0^k) \varphi_{k_1}(z_1^k) \cdots \varphi_{k_n}(z_n^k), \end{aligned} \quad (11)$$

where values $z_i^k, i = 0 \dots n$ are taken from the available data $z_i^k \in \mathbf{d}_n$ and the given training image, and separated by lags $\mathbf{h}_i = \mathbf{u}_i - \mathbf{u}_0, i = 1 \dots n$.

Finally, high-order sequential simulations are generated using the following algorithm:

Algorithm A.1

1. Define a random path for visiting all unsampled nodes on the simulation grid.
2. For each node \mathbf{u}_0 in the path:
 - a. Find the closest sampled grid nodes $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$.
 - b. Calculate lags $\mathbf{h}_i = \mathbf{u}_i - \mathbf{u}_0, i = 1 \dots n$ for unsampled location \mathbf{u}_0 .
 - c. Scan the initial data and find values $z_k^i, i = 0 \dots n$ separated by lags $\mathbf{h}_i = \mathbf{u}_i - \mathbf{u}_0, i = 1 \dots n$.
 - d. Calculate the coefficients L_{k_0, k_1, \dots, k_n} using Eq. (11).
 - e. Build the cpdf $f(\mathbf{u}_0; z_0 | z_1, \dots, z_n)$ for the random variable Z_0 at the unsampled location \mathbf{u}_0 given the conditioning data z_1, \dots, z_n at the corresponding neighbors $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n$ using Eqs. (2) and (8).
 - f. Draw a uniform random value in $[0, 1]$ to generate a simulated value z_0 from the conditional distribution $f(\mathbf{u}_0; z_0 | z_1, \dots, z_n)$.

- g. Add z_0 to the set of sample data and the previously simulated values.
 3. Repeat Steps 2a–g for the next points in the random path defined in Step 1.

2.2 Legendre Polynomials

Mustapha and Dimitrakopoulos (2010) proposed using a Legendre series as a set of basis functions $\varphi_1(z), \varphi_2(z), \dots$. The Legendre polynomial P_k of order k is defined as in Lebedev (1965)

$$P_k = \frac{1}{2^k k!} \left(\frac{d^k}{dz^k} \right) [(z^2 - 1)^k], \quad -1 \leq z \leq 1. \quad (12)$$

The set of Legendre polynomials $\{P_k(z)\}_k$ forms a complete basis set on the interval $[-1, 1]$, and, accordingly, the function $f(\mathbf{u}_0; z_0|z_1, \dots, z_n)$ can be approximated using Eqs. (2) and (8). By their construction, the order of Legendre polynomials corresponds to the order of high-order spatial statistics of the probability function $f(\mathbf{u}_0; z_0|z_1, \dots, z_n)$. However, there are practical limitations when using Legendre polynomials for the approximation of functions in multidimensional space. This is discussed in Sect. 3.

2.3 Legendre-Like Orthogonal Splines Approximation

In the present work, Legendre-like splines (Wei et al. 2013) are used as a set of basis functions. These splines are constructed using Legendre polynomials and the linear combination of B-splines. B-splines of order r in a variable $t \in [a, b]$ are piecewise polynomials defined over the domain

$$T = \underbrace{\{a, a, \dots, t_0 = a\}}_{r+1} < t_1 \leq t_2 \leq \dots \leq t_{m_{\max}} < \underbrace{t_{m_{\max}+1} = b, b, \dots, b}_{r+1}. \quad (13)$$

The points $t_i, i = 0 \dots m_{\max}$ are called knots. Each piece, a B-spline of order r , can be derived using de Boor's formula (de Boor 1978)

$$B_{i,0} = \begin{cases} 1, & t_i \leq t \leq t_{i+1} \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

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

B-splines do not form an orthogonal basis; however, Wei et al. (2013) introduced orthogonal splines based on the combination of B-splines and a set of knot sequences.

The first $r + 1$ splines are defined as the Legendre polynomials up to order r

$$S_k(t) = P_k(t), \quad k = 0 \dots r. \quad (16)$$

The subsequent splines are constructed on subsets $T_m = \{t_{i,m}\}_{i=-r}^{r+m+1}$, $m = 1 \dots m_{\max} - 1$ of the knot sequence T , where the $t_{i,m}$ are defined as follows

$$t_{i,m} = \begin{cases} a, & -r \leq i \leq 0 \\ t_i, & 1 \leq i \leq m \\ b, & m + 1 \leq i \leq m + r + 1. \end{cases} \quad (17)$$

For example, the first and second subsets are $T_1 = \{\underbrace{a, a, \dots, a}_{r+1} < t_1 < \underbrace{b, b, \dots, b}_{r+1}\}$ and $T_2 = \{\underbrace{a, a, \dots, a}_{r+1} < t_1 \leq t_2 < \underbrace{b, b, \dots, b}_{r+1}\}$, respectively. Let $B_{i,r,m}(t)$ be a B-spline of order r on the knot sequence T_m

$$B_{i,0,m} = \begin{cases} 1, & t_{i,m} \leq t \leq t_{i+1,m} \\ 0, & \text{otherwise} \end{cases} \quad (18)$$

$$B_{i,r,m}(t) = \frac{t - t_{i,m}}{t_{i+r-1,m} - t_{i,m}} B_{i,r-1,m}(t) + \frac{t_{i+r,m} - t}{t_{i+r,m} - t_{i+1,m}} B_{i+1,r-1,m}(t), \quad (19)$$

then, the remaining Legendre-like splines $S_k(t)$, $k = r+2 \dots r+m_{\max}$ are determined by

$$S_{r+m}(t) = \frac{d^{r+1}}{dt^{r+1}} f_m(t), \quad m = 1 \dots m_{\max}, \quad (20)$$

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

$$f_m(t) = \det \begin{pmatrix} B_{-r,2r+1,m}(t) & B_{-r+1,2r+1,m}(t) & \cdots & B_{-r+m-1,2r+1,m}(t) \\ B_{-r,2r+1,m}(t_1) & B_{-r+1,2r+1,m}(t_1) & \vdots & B_{-r+m-1,2r+1,m}(t_1) \\ \vdots & \vdots & \ddots & \vdots \\ B_{-r,2r+1,m}(t_{m-1}) & B_{-r+1,2r+1,m}(t_{m-1}) & \cdots & B_{-r+m-1,2r+1,m}(t_{m-1}) \end{pmatrix}. \quad (21)$$

The examples of orthogonal splines of order $r = 3$ and the knot sequence $T = [-1, -1, -1, -1, -0.6, -0.2, 0.2, 0.6, 1, 1, 1, 1]$ are presented in Figs. 1 and 2. The first $r+1$ splines are defined on the knot sequence with only one interval $[-1, 1]$ and are, thus, simply Legendre polynomials up to the order r (Fig. 1). For each subsequent spline, the knot sequence is updated by adding a knot from the initial knot sequence T , e.g., the fifth spline (Fig. 2a) is defined by Eq. (20) on two intervals $[-1, -0.6]$ and $[-0.6, -1]$, or the knot sequence $T_1 = [-1, -1, -1, -1, -0.6, 1, 1, 1, 1]$. It should be noted that Eq. (20) is obtained from the condition of orthogonality in respect to all previous splines. The following three splines (Fig. 2b–d) are defined on the knot sequence $T_2 = [-1, -1, -1, -1, -0.6, -0.2, 0.2, 0.6, 1, 1, 1, 1]$, $T_3 = [-1, -1, -1, -1, -0.6, -0.2, 0.2, 0.6, 1, 1, 1, 1]$, and $T_4 = T$.

In this work, the initial values are linearly transformed into the $[-1, 1]$ values range and divided into m_{\max} intervals. There are two parameters that have a significant impact on the quality of approximation: the maximum number of intervals m_{\max}

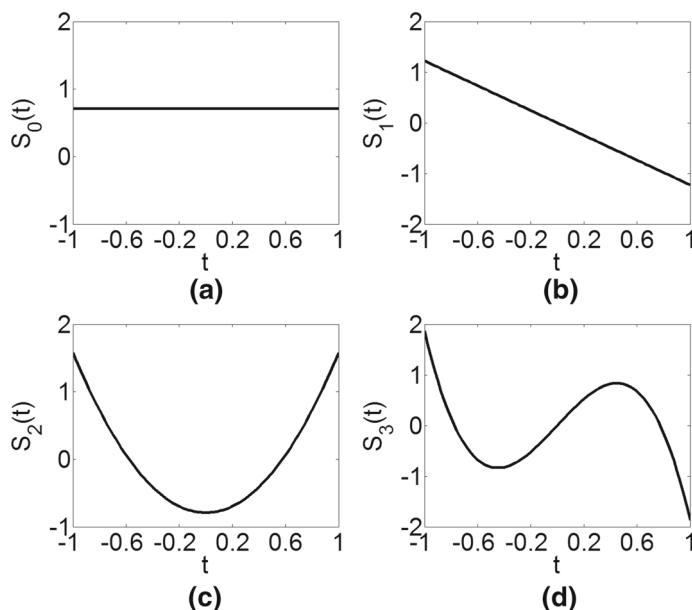


Fig. 1 The first four Legendre-like splines over the knot sequence T

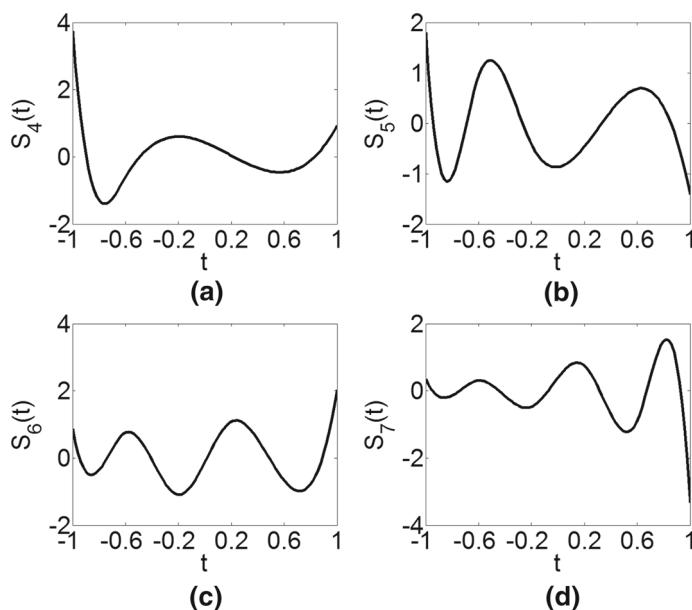


Fig. 2 The last four Legendre-like splines over the knot sequence T

and the order of splines r . These parameters reflect the maximum order of high-order spatial statistics that can be captured from the available data. The first parameter is the order of splines r . High values of the order r lead to a non-stable approximation (Runge

1901; Platte et al. 2011) and high computational costs, whereas low values affect the continuity and smoothness of approximation. For example, zero-order splines are good for the approximation of the stepwise function because each spline is a constant function. Splines with $r = 1$ are used for the approximation of continuous, but not smooth, functions, i.e., a polygonal line. In practice, cubic splines, i.e., $r = 3$, are commonly used (Hughes et al. 2005; Piegl 1989). The second parameter is the maximum number of intervals m_{\max} . Low values of m_{\max} lead to an approximation that is close to a polynomial case, for which limitations are discussed in Sect. 4. For example, an approximation with $m_{\max} = 1$ corresponds to the Legendre polynomial approximation presented by Mustapha and Dimitrakopoulos (2010). High values of m_{\max} result in overfitting or poor predictive performance, as it overreacts to minor fluctuations in the data. In addition, an approximation with a high value of m_{\max} affects the variability of the simulations because it directly samples values from the initial available data and pastes them into simulations. To choose values of r and m_{\max} , different measures are tested. The widely known Kolmogorov–Smirnov statistics test (Stephens 1974) that indicates whether two data samples come from the same distribution is not utilized herein because the related quantile–quantile plot is reproduced well for a very wide range of r and m_{\max} values; thus, such a statistical test does not provide guidance on selecting suitable parameters. Other approaches, including comparing high-order spatial statistics maps or connectivity properties, are hard to quantify by a single number and complex to implement. In this work, a simple and fast measure of the quality of approximation is used. This quality of approximation is expressed in terms of the number of grid nodes where splines fail to approximate the conditional distribution. At these nodes, the high-order simulation method produces numerical artefacts, such as outliers or noise values. Outliers can be easily detected by comparing the value at nodes with their local neighborhood average value. The average number of outliers is calculated for different values of r and m_{\max} (Fig. 3). According to Fig. 3, as the number of intervals m_{\max} is increased, the quality of approximation improves. At the same time, increasing the order of splines r decreases the quality of approximation. For cubic splines (order $r = 3$), the reasonable number of intervals is 30, as it provides the same quality of approximation as 50, demonstrates better predictive performance, and is computationally less expensive. The corresponding order of high-order spatial statistics is $m_{\max} + r = 33$.

3 Testing the Simulation Approach with a Fully Known Dataset

The high-order simulation method presented in the previous section is tested with a fully known dataset obtained from an image of a fracture network downloaded from a texture synthesis website (<http://br.depositphotos.com/5211338/stock-photo-dry-terrain.html>) and displayed in Fig. 4. Gray-scale values of the image are transformed to the $[0, 1]$ domain. The reference image (Fig. 5a) and the TI (Fig. 5b) are taken from different parts of the image and have sizes 150×150 and 200×400 grid nodes, respectively. The dataset (Fig. 5c) is generated from the reference image with a uniform random sampling of 225 values (1% of the image points).

Fig. 3 The number of outliers depending on the order of splines r and the maximum number of intervals m_{\max} . Low number of outliers corresponds to good quality of approximation

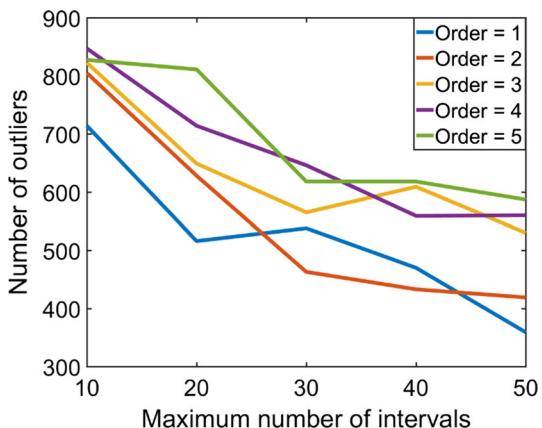
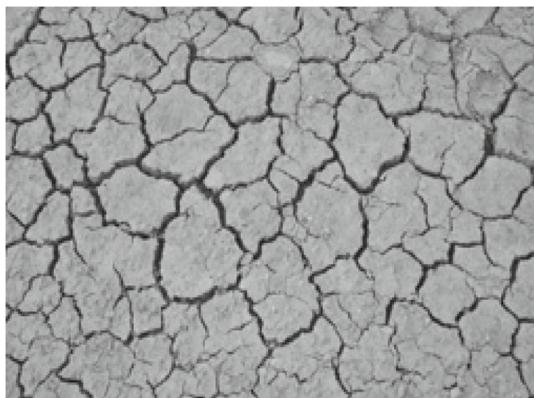


Fig. 4 Image of a fracture network (public dataset)



Three different systems of functions for the high-order simulation approach (*hosim*) and the sequential Gaussian simulation (*sgsim*) method are compared next: (a) Legendre-like splines ($r = 3, m_{\max} = 30$, the corresponding order of high-order spatial statistics is $m_{\max} + r = 33$), (b) *sgsim* method, (c) Legendre polynomials of order 10 (the corresponding order of high-order spatial statistics is 10), and (d) Legendre polynomials of order 20 (the corresponding order of high-order spatial statistics is 20). Hereafter, the system of functions based on splines and Legendre polynomials are correspondingly called *hosim-splines* and *hosim-polynomials*. The simulation using *hosim-splines* (Fig. 6a) shows a stable reproduction of spatially connected structures. Simulations using *hosim-polynomials* (Fig. 6c, d) have less connected features than the simulation using splines. *sgsim* (Fig. 6b) fails to reproduce the spatial continuity of high values. Table 1 shows the average value, median, and variance for the sample data, reference image, TI, and simulated realizations; note that only *hosim-polynomials* of order 20 are included in the comparisons that follow. All methods reproduce well the low-order statistics of the sample data and the TI.

Figure 7 shows the quantile–quantile (QQ) plots of ten simulated realizations of each simulation approach with the sample data. The QQ plots for the simulations using

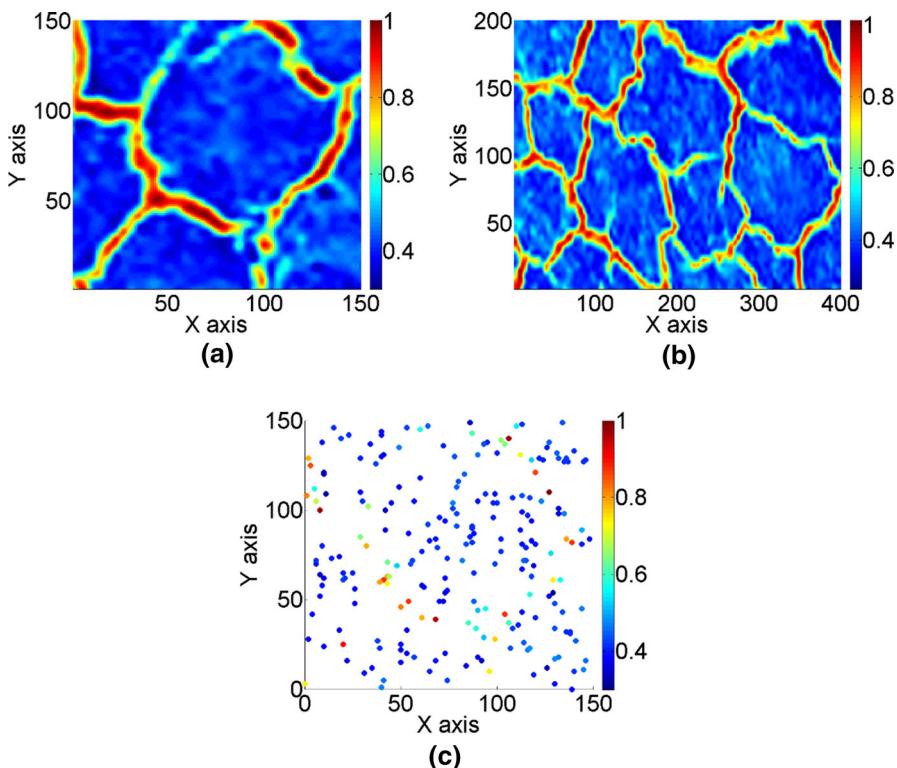


Fig. 5 Fracture network: **a** reference image; **b** training image; **c** dataset from reference image

hosim-splines are represented by red lines. The 45° black line represent QQ plots of the sample data with the sample data. The blue line represents the QQ plot of the reference image with the sample data. The green line represents the QQ plot of the training image with the sample data. The QQ plots for the simulations using *sgsim* are depicted by gray dashed lines and the QQ plots for the simulations using *hosim-polynomials* are shown by gray solid lines. Overall, the QQ plots of simulations using *hosim-splines* are consistent with the QQ plot of the sample data and the reference image, whereas QQ plots of simulations using *hosim-polynomials* and *sgsim* slightly deviate from the QQ plot of the sample data.

Figure 8 shows variograms along the north–east and north–west directions; that is, directions of the main continuity of high values, calculated for the simulations using *hosim-splines* (the red lines), the sample data (dots), the reference image (the blue line), the TI (the green line), the simulations using *hosim-polynomials* (the gray solid lines), and the simulations using *sgsim* (the gray dashed lines). All techniques demonstrate reasonable reproduction of the second-order statistics of the sample data.

For the calculation of the third-order and fourth-order spatial statistics, the estimations of the high-order moment are used (Dimitrakopoulos et al. 2010)

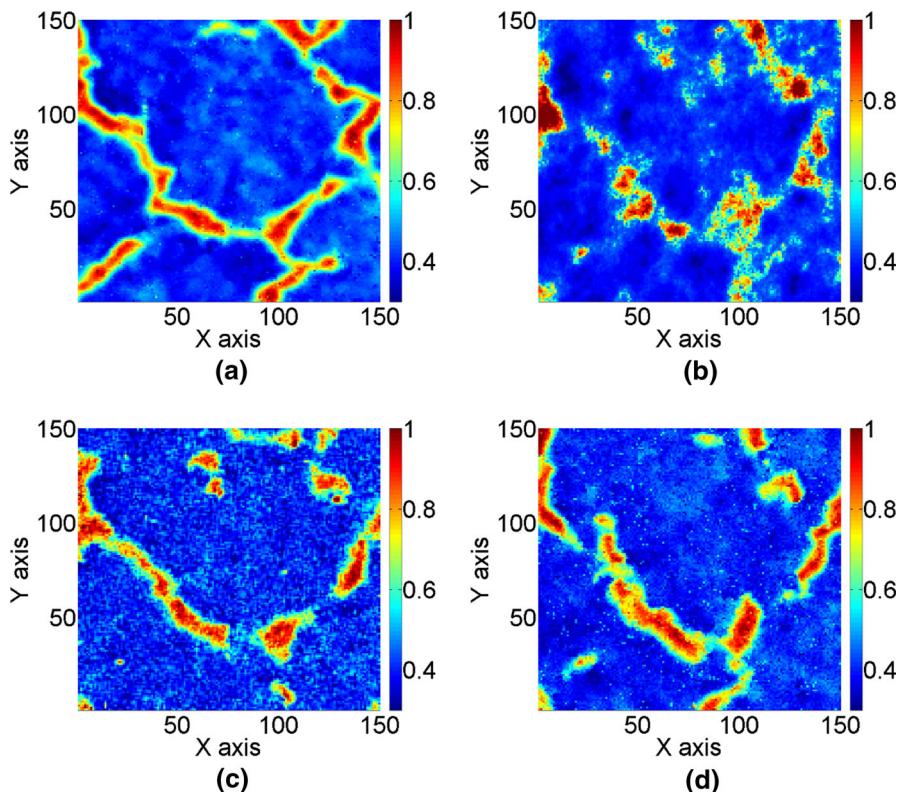


Fig. 6 The simulation results are as follows: **a** the simulation using *hosim-splines*, **b** the simulation using *sgsim*, and **c, d** the simulations using *hosim-polynomials* of orders 10 and 20, respectively

Table 1 The basic statistics of the sample data, reference image, training image, and simulations

Data and simulations	Average	Median	Variance
The sample data	0.47	0.41	0.021
The reference image	0.48	0.42	0.022
The training image	0.48	0.42	0.023
<i>hosim-splines</i>	0.48	0.42	0.020
<i>hosim-polynomials</i>	0.46	0.41	0.018
<i>sgsim</i>	0.46	0.41	0.019

$$m_3(\mathbf{h}_1, \mathbf{h}_2) = \frac{1}{N_{h_1 h_2}} \sum_{i=0}^{N_{h_1 h_2}} Z(\mathbf{u})Z(\mathbf{u} + \mathbf{h}_1)Z(\mathbf{u} + \mathbf{h}_2), \quad (22)$$

$$m_4(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) = \frac{1}{N_{h_1 h_2 h_3}} \sum_{i=0}^{N_{h_1 h_2 h_3}} Z(\mathbf{u})Z(\mathbf{u} + \mathbf{h}_1)Z(\mathbf{u} + \mathbf{h}_2)Z(\mathbf{u} + \mathbf{h}_3), \quad (23)$$

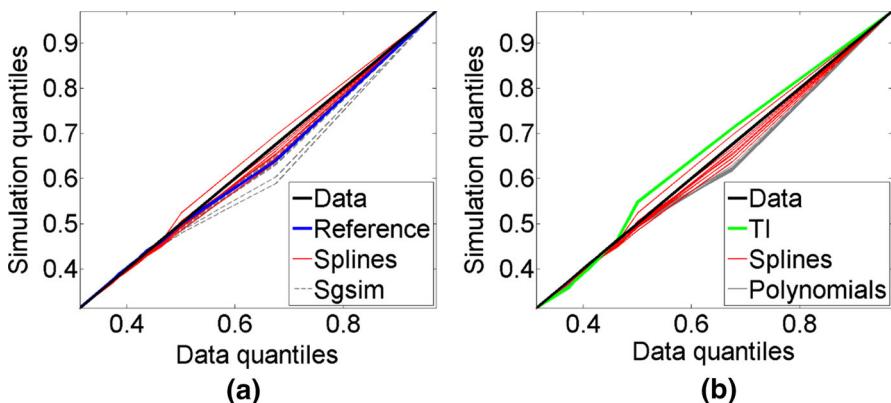


Fig. 7 Quantile–quantile (QQ) plots of simulations with the sample data. The blue line is the reference image; the green line is the training image; the red lines are the simulations using splines; the gray solid lines are the simulations using Legendre polynomials; and the gray dashed lines are the simulations using *sgsim*

where $N_{h_1 h_2}$ is the number of elements of replicates found on lags \mathbf{h}_1 and \mathbf{h}_2 , and $N_{h_1 h_2 h_3}$ is the number of elements of replicates found on distances \mathbf{h}_1 , \mathbf{h}_2 , and \mathbf{h}_3 . To highlight the connectivity property along the north–east (NE) and north–west (NW) directions, the third-order moments are calculated for binary images with a cut-off value of 0.82 (95th percentile). An example of a binary image is shown in Fig. 9a. The third-order spatial statistics are estimated based on a template with directional vectors along the NE and NW directions (Fig. 9b), i.e., $\mathbf{h}_1 = (idx, idy)$ and $\mathbf{h}_2 = (-jdx, jdy)$, respectively, where $i, j = 1 \dots 30$ and the lag discretization along x and y is $dx = dy = 1$ pixel. The physical meaning of the third-order moment of the binary image is straightforward—it is the probability of having high values at the three points separated by lags \mathbf{h}_1 and \mathbf{h}_2 (Minniakhmetov and Dimitrakopoulos 2017b). The red–orange values represent the average sizes of connected high values along the NE and NW directions. In the third-order indicator moment map of the reference image, the average sizes of the interconnected high values are 10 and 20 pixels along the NE and NW directions, respectively.

The third-order moments are calculated for simulations and averaged to account for differences between the realizations. The high-order simulation technique using *hosim-splines* and *hosim-polynomials* (Fig. 10b, e) reproduce the third-order moment map of the sample data (Fig. 10a), the reference image (Fig. 10c), and the TI (Fig. 10d), as can be seen from the similar size of the red–orange value areas in the corresponding figures. The moment map of the simulation using *sgsim* (Fig. 10f) does not reproduce connectivity along the NE and NW directions; the size of the red-shaded area is 8×10 pixels, compared to a 10×20 -pixel area in the reference image's moment map (Fig. 10c).

The fourth-order spatial statistics are estimated for binary images with a cut-off value of 0.82 (95th percentile) based on a template with directional vectors NE $\mathbf{h}_1 = (idx, idy)$, NW $\mathbf{h}_2 = (-jdx, jdy)$, and south–west (SW) $\mathbf{h}_3 = (-kdx, -kdy)$, where $i, j, k = 1 \dots 30$ and lag discretization along x and y is $dx = dy = 1$ pixels. Similarly to

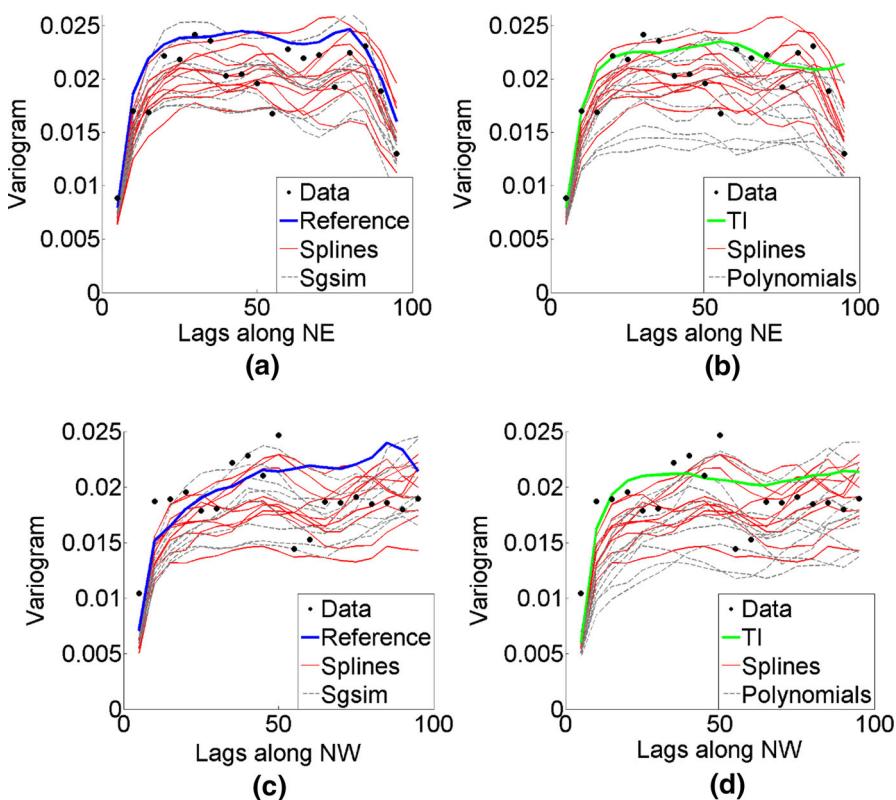


Fig. 8 Variograms along the north–east direction (top subfigures **a, b**) and north–west direction (bottom subfigures **c, d**) of the sample data (dots), reference image (blue line), training image (green line), simulations using *hosim-splines* (red lines), simulations using *hosim-polynomials* (gray lines), and simulations using *sgsim* (gray dashed lines)

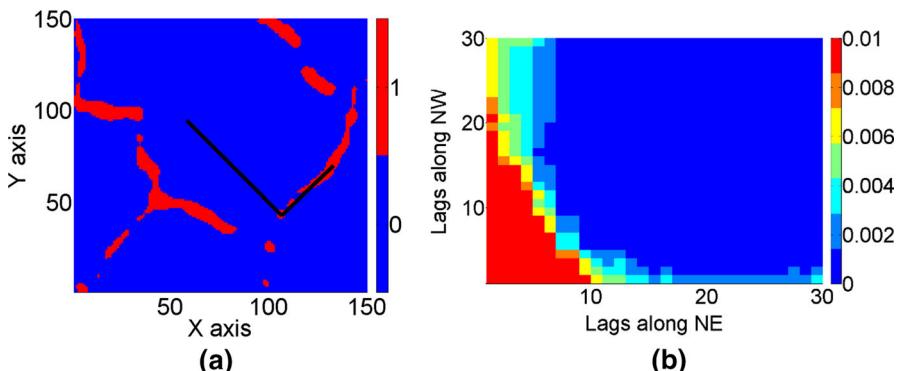


Fig. 9 The third-order indicator moment calculation for the reference image: **a** the binary values of the reference image for a cut-off value of 0.82 (95th percentile), the black lines represent L-type template for calculation of the third-order spatial statistics; **b** the third-order spatial moments of the binary image

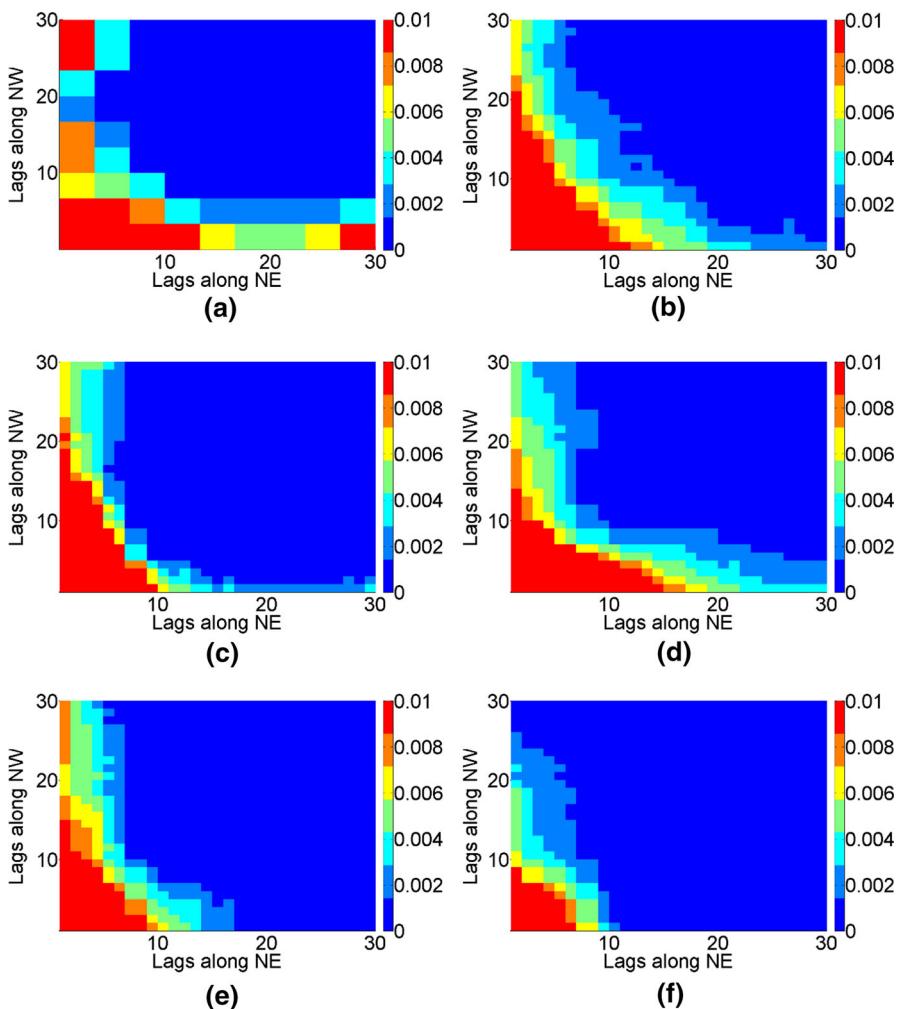


Fig. 10 The third-order indicator moments: **a** sample data samples, **b** the simulation using *hosim-splines*, **c** the reference image, **d** the training image, **e** a simulation using *hosim-polynomials*, and **f** a simulation using *sgsim*

the third order, the fourth-order moments are calculated for simulations and averaged to account for differences in the various realizations. The red–orange areas along the axes of the fourth-order spatial statistics (Fig. 11) represent the high values along the NE, NW, and SW directions. According to Fig. 11, the fourth-order moment map for the simulation using *hosim-splines* (Fig. 11b) reproduce the sizes of fractures along the NE, NW, and SW directions in the fourth-order moment of the sample data (Fig. 11a), the reference image (Fig. 11c), and the TI (Fig. 11d). The fourth-order moment map of the simulation using *hosim-polynomials* (Fig. 11e) shows a smaller connectivity of fractures along the NE and SW directions. The spatial statistics map of the simulation

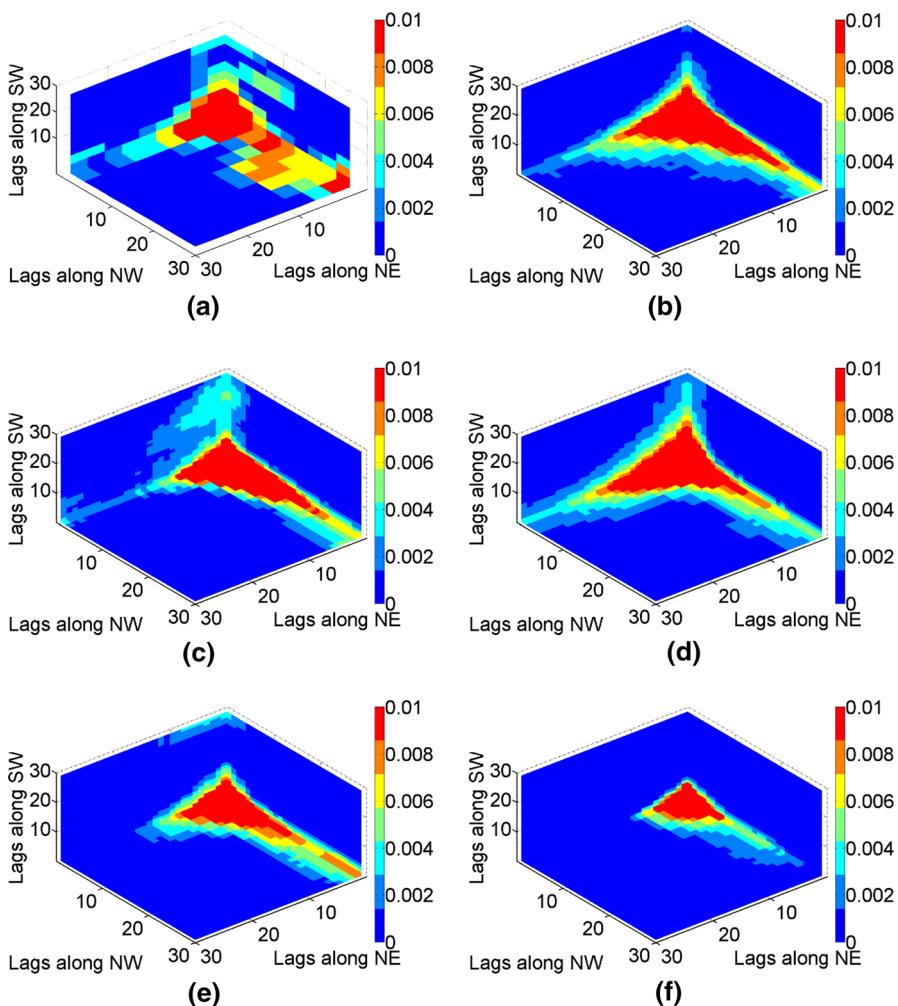


Fig. 11 The fourth-order indicator moments: **a** sample data samples, **b** the simulation using *hosim-splines*, **c** the reference image, **d** the training image, **e** a simulation using *hosim-polynomials*, and **f** a simulation using *sgsim*

using *sgsim* (Fig. 11f) does not reproduce the connectivity of fractures along the NE and SW directions.

The connectivity of high values is measured using the function presented by Journel and Alabert (1989). As in the above examples, the cut-off value is equal to 0.82 (95th percentile). Figure 12 shows P50 statistics of the connectivity measure along the NE (top subfigures) and NW directions (bottom subfigures). The P50 statistics of connectivity are calculated for the simulations using the proposed techniques (red solid line), *hosim-polynomials* (gray solid line), and *sgsim* method (gray dashed line). The connectivity measures of the reference image (blue line) and the TI (green line) falls within the P10 and P90 statistics of the connectivity measure in the simulations

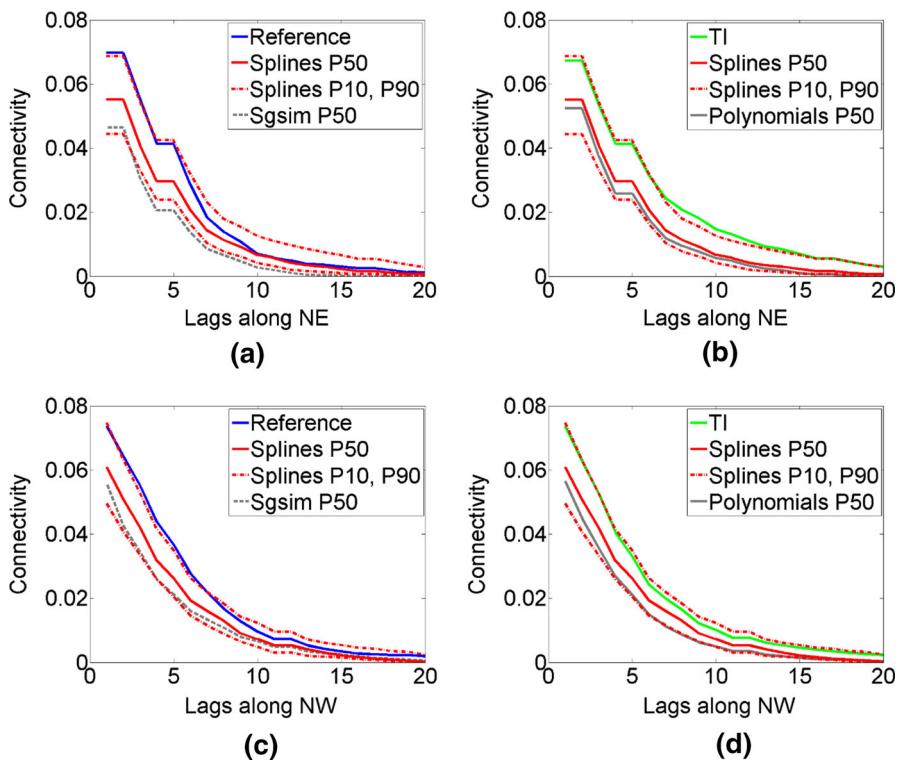


Fig. 12 The connectivity functions along the north–east (top subfigures) and north–west directions (bottom subfigures) 95th percentile: reference image (blue line), training image (green line), P50 statistics for simulations using *hosim-splines* (red solid line), P10 and P90 statistics for simulations using *hosim-splines* (red dash-dot line), P50 statistics for simulations using *hosim-polynomials* (gray solid line), and P50 statistics for simulations using *sgsim* (gray dashed line)

using *hosim-splines* (red dash-dot lines), whereas the connectivity of the simulations using *hosim-polynomials* and *sgsim* is lower, on average, than the connectivity of the TI and the reference image.

4 Application at a Gold Deposit

Data from a gold deposit are used as a case study to demonstrate the intricacies and advantages of the high-order spatial simulation method described above. In addition, the method is compared with the *sgsim* approach for the reproduction of histograms, variograms, high-order spatial statistics, and the connectivity of high and extreme values.

The deposit is 2 km by 2 km wide and extends to a depth of 500 m. Sample data are available from 288 exploration drillholes. Blast-hole data are also available for the deposit and used in the construction of a training image. The three-dimensional TI is defined on $405 \times 445 \times 86$ grid blocks of size $5 \times 5 \times 5 \text{ m}^3$. The simulation grid

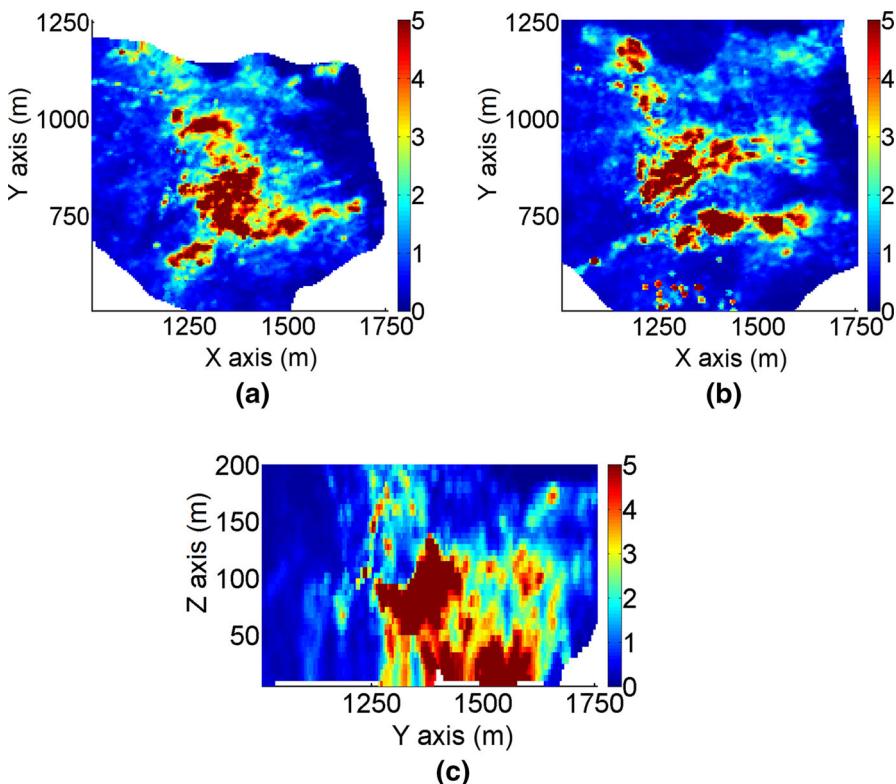


Fig. 13 Sections of the training image from part of the deposit: **a** horizontal section at $Z = 50$ m, **b** horizontal section at $Z = 100$ m, and **c** the vertical profile. The colors indicate grades in g/t

coincides with the grid of the training image. The simulation of grades is performed using the proposed method with cubic splines $r = 3$ and a maximum number of intervals $m_{\max} = 30$. Examples of horizontal sections and a vertical profile for the orebody area are shown in Figs. 13 and 14. High grades are located in the south-eastern sector of the deposit, predominantly in the bottom part.

The two-dimensional sections in Fig. 14 show that the simulation using the proposed method reproduces the spatial distribution of grades and the continuity of high grades. The areas with high values in Fig. 14 are in good agreement with the drillhole data (Fig. 15) and the TI (Fig. 13). The simulation using *sgsim* (Fig. 16) exhibits a greater number of disconnected structures with high values and sparsely distributed outliers.

These observations are confirmed by a quantitative analysis in a subsequent validation by (1) mean and variance comparison, (2) QQ plots between drillhole data and simulated values, (3) variogram validation, (4) high-order spatial cumulant validation, and (5) connectivity measure. Table 2 shows the average value, median, and variance for the drillhole data, the TI, and the simulations. Both methods reproduce well the low-order statistics of the drillhole data and the TI.

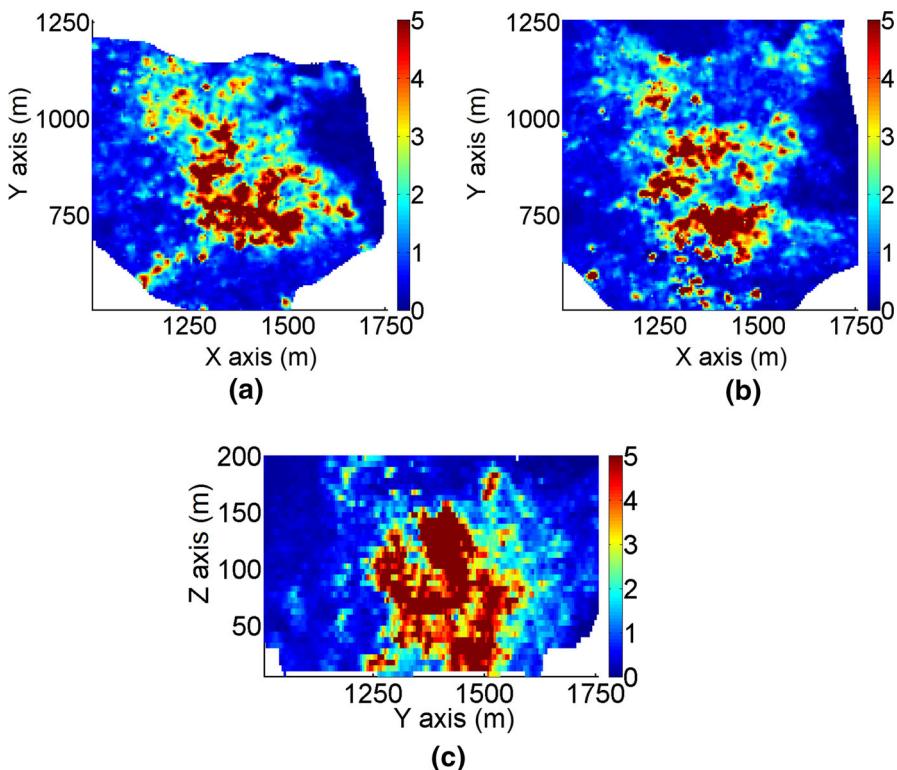


Fig. 14 Sections of the simulation using *hosim-splines* from part of the deposit: **a** horizontal section at $Z = 50$ m, **b** horizontal section at $Z = 100$ m, and **c** the vertical profile. The colors indicate grades in g/t

Figure 17 shows the QQ plots of the simulated realizations and the drillhole data. Quantiles of simulations using *hosim-splines* are shown by red lines. Quantiles of simulations using the *sgsim* method are shown by gray lines. In addition, the QQ plots of the training image and the drillhole data are depicted by the blue line. The closer these curves are to the 45° black line in the graph, the better they reproduce the distribution of the drillhole data. Both methods provide simulations consistent with the drillhole data in terms of distributions. Figure 18 presents variograms for the north–south and east–west directions. Simulations using *hosim-splines* (red lines) share the second-order statistics of drillhole data and the TI (blue lines). The simulations using the *sgsim* method (gray lines) preserve the second-order statistics of the drillhole data (black dots).

Applying a zero-mean transformation, the third-order cumulants can be calculated using Eq. (22) with lags $\mathbf{h}_1 = (idx, 0)$, $\mathbf{h}_2 = (0, jdy)$ indexed by $i = 1 \dots 7$, $j = 1 \dots 7$, where dx and dy are distances between drillholes, that is, 100 m \times 100 m. Figure 19 shows the comparison of cumulant maps for sample data, the TI, and the simulations. The values along axes reflect variograms along their corresponding directions because the third-order moment $E(Z^2(\mathbf{x})Z(\mathbf{x} + \mathbf{h}))$ has similar spatial relations as the second-order moment $E(Z(\mathbf{x})Z(\mathbf{x} + \mathbf{h}))$. However, the square term $Z^2(\mathbf{x})$ in

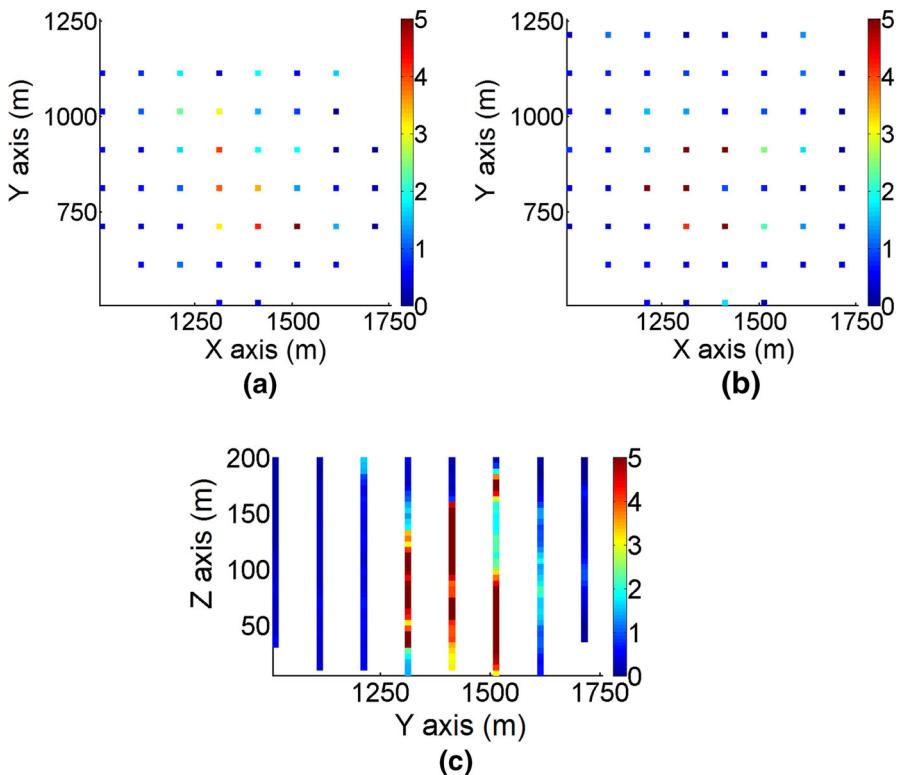


Fig. 15 Sections of the exploration drillhole data from part of the deposit: **a** horizontal section at $Z = 50$ m, **b** horizontal section at $Z = 100$ m, and **c** the vertical profile. The colors indicate grades in g/t

$E(Z^2(\mathbf{x})Z(\mathbf{x} + \mathbf{h}))$ affects the absolute value of the statistics and, moreover, combines both negative and positive correlations of $Z(\mathbf{x})$ due to the square operation. Thus, in addition to analyzing the values along the axes, it is important to compare the area of $[200; 400] \times [200; 400]$ on the third-order cumulant maps. The simulations using the proposed *hosim-splines* method (Fig. 19c) reproduce red areas along the x – y axes and yellow–green areas in the cumulant map of the drillhole data (Fig. 19a) and the TI (Fig. 19b). These areas reflect the size of connected high grades and are equal to approximately 400 m along the x -axis and 300 m along the y -axis. The cumulant map for the simulation using *sgsim* (Fig. 19d) neither reproduce the magnitude of the red area along the x and y axes nor the values at the area of $[200; 400] \times [200; 400]$.

The fourth-order cumulants are calculated using the following equation

$$c_4(\mathbf{h}_1, \mathbf{h}_2, \mathbf{h}_3) = \frac{1}{N_{h_1 h_2 h_3}} \sum_{i=0}^{N_{h_1 h_2 h_3}} Z(\mathbf{u})Z(\mathbf{u} + \mathbf{h}_1)Z(\mathbf{u} + \mathbf{h}_2)Z(\mathbf{u} + \mathbf{h}_3) - m_2(\mathbf{h}_1)m_2(\mathbf{h}_2) - m_2(\mathbf{h}_1)m_2(\mathbf{h}_3) - m_2(\mathbf{h}_2)m_2(\mathbf{h}_3), \quad (24)$$

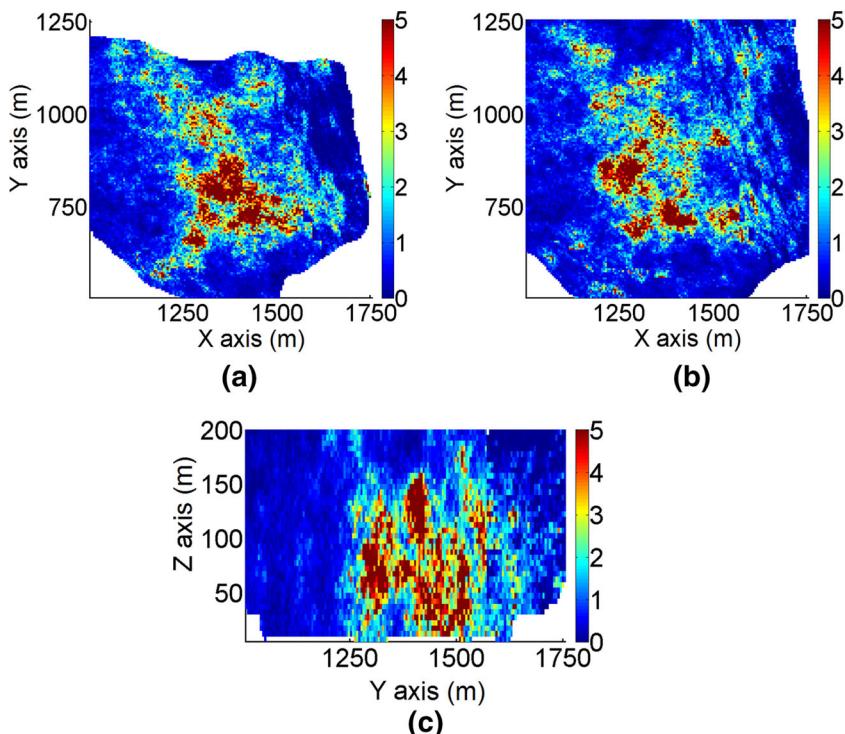
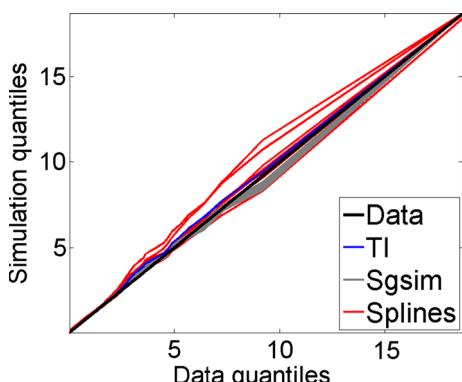


Fig. 16 Sections of the simulation using the *sgsim* method from part of the deposit: **a** horizontal section at $Z=50$ m, **b** horizontal section at $Z=100$ m, and **c** vertical profile. The colors indicate grades in g/t

Table 2 The basic statistics of the drillhole data, training image, and simulations

Data and simulations	Average	Median	Variance
The drillhole data	0.63	0.38	0.69
The training image	0.66	0.42	0.79
<i>hosim</i>	0.65	0.40	0.76
<i>sgsim</i>	0.61	0.39	0.67

Fig. 17 Quantile–quantile (QQ) plots of sample data (black line), training image (TI, blue line), simulations using *hosim-splines* (red lines), and simulations using *sgsim* (gray lines)



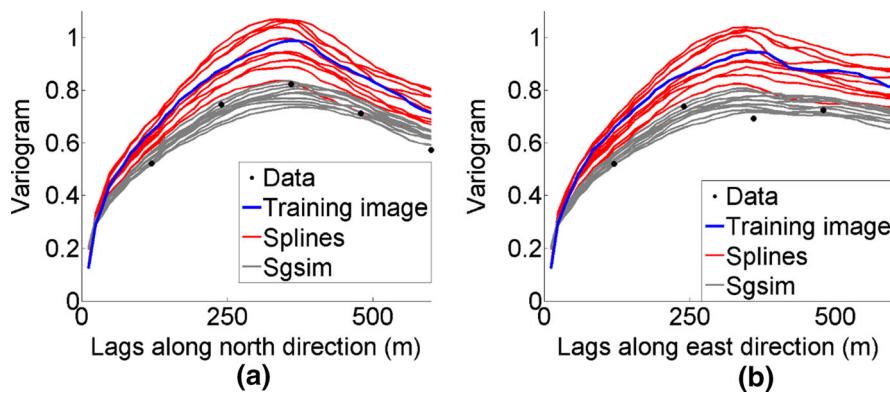


Fig. 18 Variograms of the hard data (dots), training image (blue lines), simulations using *hosim-splines* (red lines), and simulations using *sgsim* (gray lines) along the north-south (a) and east–west (b) directions

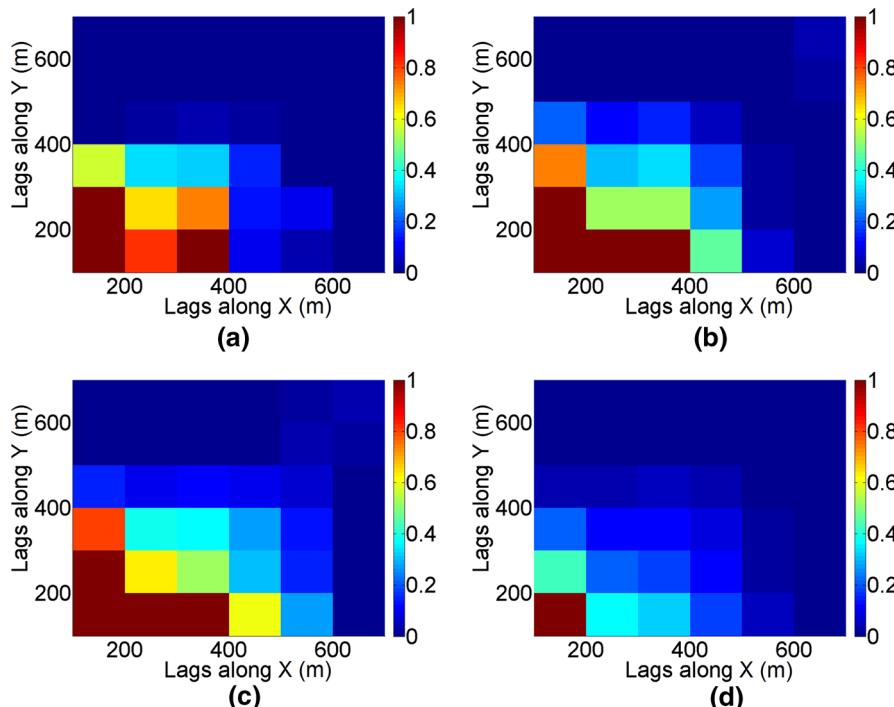


Fig. 19 Third-order spatial cumulant maps of the **a** drillhole data, **b** training image (TI), **c** a simulation using *hosim-splines*, and **d** a simulation using *sgsim*

where $N_{h_1 h_2 h_3}$ is the number of elements of replicates found on distances \mathbf{h}_1 and \mathbf{h}_2 , and $m_2(\mathbf{h})$ is the second-order moment along direction \mathbf{h} , which is equal to the covariance for a zero-mean random field. The lags $\mathbf{h}_1 = (idx, 0)$, $\mathbf{h}_2 = (0, jdy)$, and $\mathbf{h}_3 = (0, kdz)$ are indexed by $i = 1 \dots 7$, $j = 1 \dots 7$, and $k = 1 \dots 7$, where, dx , dy , and dz are distances between data samples, that is, $100 \text{ m} \times 100 \text{ m} \times 5 \text{ m}$. The high-order

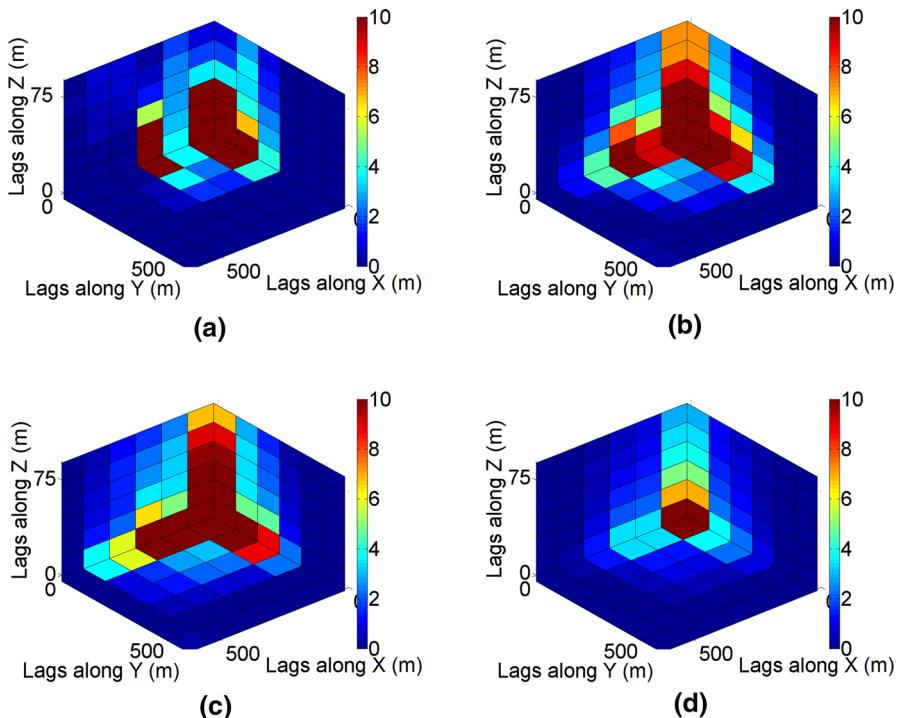


Fig. 20 The fourth-order spatial cumulant maps of the **a** drillhole data, **b** training image (TI), **c** a simulation using *hosim-splines*, and **d** a simulation using *sgsim*

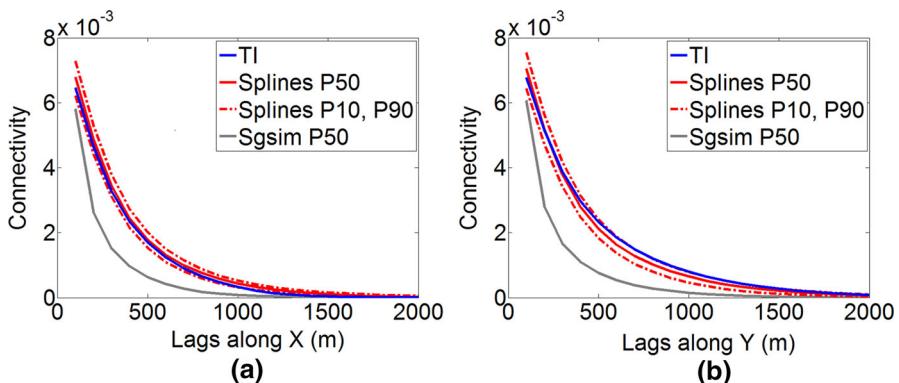


Fig. 21 The connectivity along *x* (left subfigure) and *y* (right subfigure) for the 99th percentile: the training image (TI, blue lines), P50 statistics for *hosim* simulations (red solid lines), P10 and P90 statistics for *hosim* simulations (red dashed lines), and P50 statistics for simulations using *sgsim* (gray lines)

cumulants calculated reflect the complex structures of orebodies (Dimitrakopoulos et al. 2010). According to Figs. 19 and 20, the size of connected structures is reproduced in simulations using the proposed method (Figs. 19c, 20c). This can also be traced in the vertical profiles (Figs. 13c, 14c, 15c). The fourth-order cumulant map of the

simulation using *sgsim* (Fig. 20d) has a rather small red area in comparison with structures in the cumulant maps of the drillhole data (Fig. 20a) and the TI (Fig. 20b).

The connectivity along the x and y axes is analyzed using the connectivity measure presented by Journel and Alabert (1989). The cut-off value is equal to 5 ppm (99th percentile). The P10, P50, and P90 statistics of connectivity measures are calculated for simulations using the *hosim-splines* method and depicted by red lines in Fig. 21. Solid lines represent the P50 of connectivity measures, whereas dashed lines show the P10 and P90 statistics. The connectivity of the simulations using the proposed method (red lines) remains close to the connectivity measure of the TI (blue lines). The P50 statistics of connectivity measure calculated using *sgsim* simulations (gray lines) is quite far from the connectivity of the TI. Thus, despite reproducing the histograms and variograms, Gaussian simulation methods fail to reproduce an important property of the connectivity of high values.

5 Conclusions

This paper presents a novel approach for the high-order simulation of continuous variables based on Legendre-like orthogonal splines. Splines are flexible tools for the approximation of complex probability density functions. Using different knot sequences, orders of splines, and smoothness of piecewise polynomials, it is possible to obtain a stable approximation that reproduces the spatial connectivity of the extreme values. The simulations are consistent with the spatial statistics of the sample data and share the high-order spatial statistics of the available data and the training image.

The proposed approach is also compared with the conventional second-order approach sequential Gaussian simulation and the high-order simulation method using Legendre polynomials. The approach using splines exhibits a more stable approximation of the conditional probability density function (cpdf) and a better representation of the spatial connectivity of extreme values. The applied connectivity measure confirms the results obtained by analyzing the high-order statistics and demonstrates the limitations of Gaussian simulation methods in the characterization of a mineral deposit. In addition, the proposed approach provides a general framework for high-order simulation techniques. For example, by using just one interval for spline construction, the technique reproduces the method proposed by Mustapha and Dimitrakopoulos (2010, 2011).

Further research will address the simulation of categorical variables using splines of order zero and the simulation of multiple correlated continuous and discrete variables within a general framework. In addition, the adaptive knot sequence will be investigated for a better approximation of the cpdf.

Acknowledgements This work was funded from NSERC Collaborative Research and Development Grant CRDPJ 411270, entitled “Developing new global stochastic optimization and high-order stochastic models for optimizing mining complexes with uncertainty”, NSERC Discovery Grant 239019, and mining industry partners of the COSMO Stochastic Mine Planning Laboratory (AngloGold Ashanti, Barrick Gold, BHP, De Beers Canada, Kinross Gold, Newmont Mining, and Vale). The authors would also like to thank Newmont Mining for providing and allowing the use of their dataset for this publication.

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 Geosci* 39(2):177–203
- Boyd JP, Ong JR (2009) Exponentially-convergent strategies for defeating the Runge phenomenon for the approximation of non-periodic functions, Part I: single-interval schemes. *Commun Comput Phys* 5:484–497
- Chatterjee S, Dimitrakopoulos R, Mustapha H (2012) Dimensional reduction of pattern-based simulation using wavelet analysis. *Math Geosci* 44(3):343–374
- Chièze J-P, Delfiner P (1999) Geostatistics: modeling spatial uncertainty. Wiley, New York
- Chugunova TL, Hu LY (2008) Multiple-point simulations constrained by continuous auxiliary data. *Math Geosci* 40(2):133–146
- David M (1988) Handbook of applied advanced geostatistical ore reserve estimation. Elsevier, Amsterdam
- de Boor C (1978) A practical guide to splines. Springer, Berlin
- De Iaco S, Maggio S (2011) Validation techniques for geological patterns simulations based on variogram and multiple-point statistics. *Math Geosci* 43(4):483–500
- de Vries LM, Carrera J, Falivene O, Gratacós O, Slooten LJ (2009) Application of multiple point geostatistics to non-stationary images. *Math Geosci* 41(1):29–42
- 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):567–591
- 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
- Fornberg B, Zuev J (2007) The Runge phenomenon and spatially variable shape parameters in RBF interpolation. *Comput Math Appl* 54(3):379–398
- Goodfellow R, Consuegra FA, Dimitrakopoulos R, Lloyd T (2012) Quantifying multi-element and volumetric uncertainty, Coleman McCready deposit, Ontario, Canada. *Comput Geosci* 42:71–78
- Goovaerts P (1998) Geostatistics for natural resources evaluation. Oxford, New York
- Guardiano FB, Srivastava RM (1993) Multivariate geostatistics: beyond bivariate moments. In: Soares A (ed) Geostatistics Tróia '92. Quantitative Geology and Geostatistics, vol 5. Springer, Dordrecht, pp 133–144
- Honarkhah M (2011) Stochastic simulation of patterns using distance-based pattern modeling. Ph.D. dissertation, Stanford University, Stanford
- Hoschek J, Lasser D (1993) Fundamentals of computer aided geometric design. AK Peters, London
- Hughes TJR, Cottrell JA, Bazilevs Y (2005) Isogeometric analysis: CAD, finite elements, NURBS, exact geometry and mesh refinement. *Comput Methods Appl Mech Eng* 194(39–41):4135–4195
- Journel AG (1994) Modelling uncertainty: some conceptual thoughts. In: Dimitrakopoulos R (ed) Geostatistics for the next century. Kluwer, Dordrecht, pp 30–43
- Journel AG (2005) Beyond covariance: the advent of multiple-point geostatistics. In: Leuanthong O, Deutsch CV (eds) Geostatistics Banff 2004. Springer, Dordrecht, pp 225–233
- 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, Cham, pp 47–55. https://doi.org/10.1007/978-3-319-69320-0_5
- Journel AG, Alabert F (1989) Non-Gaussian data expansion in the earth sciences. *Terra Nova* 1:123–134
- Journel AG, Huijbregts CJ (1978) Mining geostatistics. Academic Press, London
- Lebedev NN (1965) Special functions and their applications. Prentice Hall, New Jersey
- Lee PM (2012) Bayesian statistics: an introduction. Wiley, New York
- Lochbühler T, Pirot G, Straubhaar J, Linde N (2014) Conditioning of multiple-point statistics facies simulations to tomographic images. *Math Geosci* 46(5):625–645
- López de Silanes MC, Parra MC, Pasadas M, Torrens JJ (2001) Spline approximation of discontinuous multivariate functions from scattered data. *J Comput Appl Math* 131(1–2):281–298

- Malagù M, Benvenuti E, Duarte CA, Simone A (2014) One-dimensional nonlocal and gradient elasticity: assessment of high order approximation schemes. *Comput Methods Appl Mech Eng* 275(15):138–158
- Mariethoz G, Renard P (2010) Reconstruction of incomplete data sets or images using direct sampling. *Math Geosci* 42(3):245–268
- Mariethoz G, Renard P, Straubhaar J (2010) The direct sampling method to perform multiple-point geostatistical simulations. *Water Resour Res.* <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
- 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, Cassiraga E, Vargas-Guzmán JA (eds) *Geostatistics Valencia 2016*. Springer, Cham, pp 287–301
- Mustapha H, Dimitrakopoulos R (2010) High-order stochastic simulations for complex non-Gaussian and non-linear geological patterns. *Math Geosci* 42(5):457–485
- 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
- Mustapha H, Dimitrakopoulos R, Chatterjee S (2011) Geologic heterogeneity representation using high-order spatial cumulants for subsurface flow and transport simulations. *Water Resour Res.* <https://doi.org/10.1029/2010wr009515>
- 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
- Osterholt V, Dimitrakopoulos R (2018) 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) *Advances in applied strategic mine planning*. Springer, Cham, pp 335–352. https://doi.org/10.1007/978-3-319-69320-0_22
- Park H, Lee JH (2007) B-spline curve fitting based on adaptive curve refinement using dominant points. *Comput Aided Des* 39(6):439–451
- Piegl L (1989) Modifying the shape of rational B-splines. Part 1: curves. *Comput Aided Des* 21(8):509–518
- Platte RB, Trefethen LN, Kuijlaars AB (2011) Impossibility of fast stable approximation of analytic functions from equispaced samples. *SIAM Rev* 53:308–318
- Rezaee H, Mariethoz G, Konoshloo M, Asghari O (2013) Multiple-point geostatistical simulation using the bunch-pasting direct sampling method. *Comput Geosci* 54:293–308
- Ruiu J, Caumon G, Viseur S (2016) Modeling channel forms and related sedimentary objects using a boundary representation based on non-uniform rational B-splines. *Math Geosci* 48(3):259–284
- Runge C (1901) Über empirische Funktionen und die Interpolation zwischen äquidistanten Ordinaten. *Zeitschrift für Mathematik und Physik* 46(224–243):20
- Sinha SS, Schunck BG (1992) A two-stage algorithm for discontinuity-preserving surface reconstruction. *IEEE Trans Pattern Anal Mach Intell* 14(1):36–55
- Stephens MA (1974) EDF statistics for goodness of fit and some comparisons. *J Am Stat Assoc* 69(347):730–737
- Straubhaar J, Renard P, Mariethoz G, Froidevaux R, Besson O (2011) An improved parallel multiple-point algorithm using a list approach. *Math Geosci* 43(3):305–328
- Strebelle S (2002) Conditional simulation of complex geological structures using multiple-point statistics. *Math Geol* 34(1):1–21
- Strebelle S, Cavelius C (2014) Solving speed and memory issues in multiple-point statistics simulation program SNESIM. *Math Geosci* 46(2):171–186
- Tamayo-Mas E, Mustapha H, Dimitrakopoulos R (2016) Testing geological heterogeneity representations for enhanced oil recovery techniques. *J Petrol Sci Eng* 146:222–240
- Toftaker H, Tjelmeland H (2013) Construction of binary multi-grid Markov random field prior models from training images. *Math Geosci* 45(4):383–409
- Wei Y, Wang G, Yang P (2013) Legendre-like orthogonal basis for spline space. *Comput Aided Des* 45(2):85–92
- Yao L, Dimitrakopoulos R, Gamache M (2018) A new computational model of high-order spatial simulation based on spatial Legendre moments. *Math Geosci* (submitted)

- Zhang T, Switzer P, Journel A (2006) Filter-based classification of training image patterns for spatial simulation. *Math Geol* 38(1):63–80
- Zhang T, Gelman A, Laronga R (2017) Structure- and texture-based fullbore image reconstruction. *Math Geosci* 49(2):195–215