

EXACT SIMULATION OF A BOOLEAN MODEL

CHRISTIAN LANTUÉJOUL✉

MinesParisTech, 35 rue Saint-Honoré, 77305 Fontainebleau, France
e-mail: christian.lantuejoul@mines-paristech.fr

(Received February 27, 2013; revised April 4, 2013; accepted May 5, 2013)

ABSTRACT

A Boolean model is a union of independent objects (compact random subsets) located at Poisson points. Two algorithms are proposed for simulating a Boolean model with non uniformly bounded objects in a bounded domain. The first one applies only to stationary models. It generates the objects prior to their Poisson locations. Two examples illustrate its applicability. The second algorithm applies to stationary and non-stationary models. It generates the Poisson points prior to the objects. Its practical difficulties of implementation are discussed. Both algorithms are based on importance sampling techniques, and the generated objects are weighted.

Keywords: Boolean model, importance sampling, Minkowsky functionals, Steiner formula.

INTRODUCTION

The Boolean model is certainly one of the most currently used random set models in mathematical morphology, stochastic geometry and spatial statistics. It is defined as the union of a family of independent random compact subsets (denoted for short as “objects”) located at the points of a locally finite Poisson process. It is stationary if the objects are identically distributed (up to their location) and the Poisson process is homogeneous, and non-stationary otherwise.

Despite its widespread use, it seems that little attention has been paid to the following problem: How to perform *exact simulations* of a Boolean model in a bounded domain?

The solution to that problem is not straightforward, unless the objects are uniformly bounded. The difficulty lies in that the intersection of a Boolean model and a bounded domain is also a Boolean model, but its parameters are different. The more remote the object to the domain, the less chance they have to hit it. On the other hand, the larger the objects, the more chance they have to hit the domain.

After a brief reminder on the Boolean model, this problem is investigated. Although most emphasis is placed on the stationary case because of its possible connections with stereology, the general case is also treated in a second part of the paper. Two examples serve to illustrate the algorithms and their implementations.

Here is the set of notation that will be used throughout the paper. Unless specified, the workspace is the d -dimensional Euclidean space \mathbb{R}^d with

Lebesgue measure ν_d . The origin of \mathbb{R}^d is denoted by o . More generally, points are denoted by lower case letters, subsets by capital letters and family of subsets by calligraphic letters. If $x \in \mathbb{R}^d$ and $A \subset \mathbb{R}^d$, $\tau_x A$ stands for the translation of A w.r.t. vector $\vec{o}x$. The dilation of A by another subset B is defined as

$$\delta_B A = \{x \in \mathbb{R}^d : \tau_x B \cap A \neq \emptyset\}.$$

DEFINITION AND MAIN PROPERTIES

The basic ingredients of a Boolean model are

- (i) a *Poisson point process* \mathcal{P} with an intensity function $\theta = (\theta_x, x \in \mathbb{R}^d)$ that is assumed to be locally integrable.
- (ii) a family of nonempty and *mutually independent compact random subsets* $(A_x, x \in \mathbb{R}^d)$. A_x is called the *object* located at x . If it takes simple shapes, then its statistical properties can be specified by elementary descriptors (e.g., distribution of its radius for a disk). For more intricate shapes, the hitting functional of A_x can be considered (Matheron, 1975). It assigns each compact subset K of \mathbb{R}^d (in short $K \in \mathcal{K}$) the probability that it intersects with A_x ,

$$T_x(K) = P\{A_x \cap K \neq \emptyset\}, \quad K \in \mathcal{K}. \quad (1)$$

Definition 1 A Boolean model is a union of objects located at Poisson points,

$$X = \bigcup_{x \in \mathcal{P}} A_x.$$

Our objective is to simulate the Boolean model in a compact domain, say Z . Insofar as this mainly consists of simulating the objects of X that intersect with Z , it is crucial to assume their number to be almost surely finite. Under this assumption, it can be shown (Lantuéjoul, 2002) that the number of objects hitting any compact subset K of Z is Poisson distributed with mean

$$\vartheta(K) = \int_{\mathbb{R}^d} \theta_x T_x(K) dx, \quad K \in \mathcal{K}(Z). \quad (2)$$

In particular, the avoiding functional of $X \cap Z$, defined for each $K \in \mathcal{K}(Z)$ as

$$Q_{X \cap Z}(K) = P\{(X \cap Z) \cap K = \emptyset\} = P\{X \cap K = \emptyset\}$$

is equal to

$$Q_{X \cap Z}(K) = \exp(-\vartheta(K)) \quad K \in \mathcal{K}(Z). \quad (3)$$

STATIONARY CASE

ALGORITHM

In this case, the intensity function is constant, say θ , and all objects have the same hitting functional, up to a translation

$$T_x(K) = T_o(\tau_{-x}K). \quad (4)$$

In what follows, it is convenient to write

$$T_o(K) = \int_{\mathcal{K}} dF(A) 1_{A \cap K \neq \emptyset},$$

where F (symbolically) denotes the distribution of the parameters of A_o . Then we have

$$T_x(K) = \int_{\mathcal{K}} dF(A) 1_{\tau_x A \cap K \neq \emptyset}.$$

Applying this formula to Eq. 2 and permuting integrals, the mean number of objects hitting Z becomes

$$\vartheta(Z) = \theta \int_{\mathcal{K}} dF(A) v_d(\delta_A Z) = \theta E\{v_d(\delta_A Z)\}, \quad (5)$$

which gives the following expression for the avoiding functional of $X \cap Z$.

$$Q_{X \cap Z}(K) = \exp\left(-\theta \int_{\mathcal{K}} dF(A) v_d(\delta_A K)\right).$$

Let us write it slightly differently

$$Q_{X \cap Z}(K) = \exp\left(-\vartheta(Z) \int_{\mathcal{K}} dF^Z(A) \frac{v_d(\delta_A K)}{v_d(\delta_A Z)}\right), \quad (6)$$

which involves a weighted version F^Z of F

$$dF^Z(A) = \frac{dF(A) v_d(\delta_A Z)}{E\{v_d(\delta_A Z)\}}. \quad (7)$$

An interpretation of Eq. (6) is as follows: $X \cap Z$ is the union of a Poisson number (mean $\vartheta(Z)$) of independent objects of the form $\tau_x A$ where A is distributed like F^Z and x is a uniform point over $\delta_A Z$. Hence the simulation algorithm:

Algorithm 1

- (i) set $X = \emptyset$;
- (ii) generate $n \sim \mathcal{P}(\vartheta(Z))$;
- (iii) if $n = 0$, return $X \cap Z$ and stop;
- (iv) generate $A \sim dF^Z$;
- (v) generate $x \sim \mathcal{U}(\delta_A Z)$;
- (vi) put $X = X \cup \tau_x A$, $n = n - 1$ and goto (iii).

The main difficulty with this algorithm is step (iv): How to simulate the weighted distribution dF^Z ? The next section shows that interesting simplifications arise when the objects are convex.

CONVEX OBJECTS

Such algorithmic simplifications actually arise only when the simulations are performed within a ball-shaped domain. Accordingly, it is advantageous to firstly extend the simulation field to a ball, then perform the simulations in the extended domain, and finally restrict the simulations produced to the actual simulation field. The choice of the ball is unimportant, as long as it encloses the simulation field. One possibility is the ball circumscribed to the simulation field.

In what follows, the simulation field Z is assumed to be a ball, say $B(o, \rho)$. Then Steiner formula applies and shows that the mean number of objects hitting Z depends only on the expected Minkowski functionals of the objects:

$$\vartheta(Z) = \theta \sum_{k=0}^d \binom{d}{k} E\{W_k(A)\} \rho^k. \quad (8)$$

Moreover, dF^Z can be simulated as a mixture of distributions of objects weighted by their Minkowski functionals:

$$\begin{aligned} dF^Z(A) &= dF(A) \frac{\sum_{k=0}^d \binom{d}{k} W_k(A) \rho^k}{\sum_{k=0}^d \binom{d}{k} E\{W_k(A)\} \rho^k} \\ &= \sum_{k=0}^d p_k dF_k(A), \end{aligned}$$

with

$$dF_k(A) = \frac{dF(A)W_k(A)}{E\{W_k(A)\}} \quad k = 0, \dots, d, \quad (9)$$

and

$$p_k = \frac{\binom{d}{k} E\{W_k(A)\} \rho^k}{\sum_{l=0}^d \binom{d}{l} E\{W_l(A)\} \rho^l} \quad k = 0, \dots, d. \quad (10)$$

Using this mixture of weighted distributions, the algorithm for simulating dF^Z becomes

Algorithm 2

- (i) generate $k \sim p$;
- (ii) generate $A \sim dF_k(A)$;
- (iii) return A and stop.

The following two examples show how this algorithm can be implemented.

Example 1

The objects are balls and their radii follow independent exponential distributions with mean $1/a$

$$P\{R > r\} = \exp(-ar) .$$

Starting from Eq. 5, the mean number of objects hitting Z is

$$\vartheta(Z) = \theta E\{\omega_d(\rho + R)^d\} = \frac{\theta d! \omega_d}{a^d} \sum_{k=0}^d \frac{(a\rho)^k}{k!} ,$$

where $\omega_d = \pi^{d/2} / \Gamma(d/2 + 1)$ is the d -volume of the unit ball.

The next step is the simulation of $dF_k(A)$. Actually, it can also be written $dF_k(r)$ because the only random element of A is its radius. If $A = B(o, r)$, then $W_k(A) = \omega_d r^{d-k}$ and $E\{W_k(A)\} = \omega_d (d-k)! / a^{d-k}$. Plugging these values into Eq. 9, we obtain

$$dF_k(r) = \frac{a \exp(-ar) (ar)^{d-k}}{(d-k)!} .$$

This is a gamma distribution with parameter $d - k + 1$ and scale factor a . A simple way to simulate it is to consider $-\ln(u_1 \cdots u_{d-k+1})/a$ where u_1, \dots, u_{d-k+1} are independent uniform values on $]0, 1[$.

As an illustration, Fig. 1 shows a simulation of a Boolean model of discs with exponential radii. The displayed simulation field is a 7×5 rectangle. The Poisson intensity is $\theta = 10$ and the parameter a of the

exponential distribution has been chosen equal to 7.22 so as to provide a background proportion of 30%.

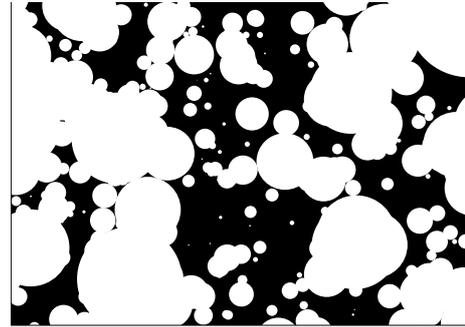


Fig. 1. Simulation of a Boolean model of discs with exponential radii.

Example 2

Here is a somewhat more elaborate example, even if the Boolean model considered is only two-dimensional. The objects are typical Poisson polygons derived from a network of Poisson lines with intensity λ , see Fig. 2.

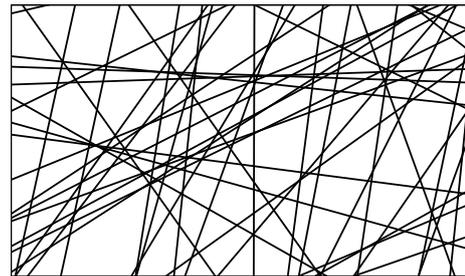


Fig. 2. Realization of a Poisson line process. The polygons delimited by the lines are Poisson polygons.

These polygons have been extensively studied in Miles (1969) and Matheron (1975). In particular their expected Minkowski functionals are

$$E\{W_0(A)\} = \frac{1}{\pi\lambda^2} \quad E\{W_1(A)\} = \frac{1}{\lambda} \quad E\{W_2(A)\} = \pi .$$

Using these expected values, Eq. (8) gives

$$\vartheta(Z) = \frac{\theta}{\pi\lambda^2} (1 + \pi\lambda\rho)^2 .$$

Not so simple is the simulation of $dF^Z(A) = p_0 F_0(A) + p_1 F_1(A) + p_2 F_2(A)$. The explicit values for the weights are

$$p_0 = \frac{1}{(1 + \pi\lambda\rho)^2} , \quad p_1 = \frac{2\pi\lambda\rho}{(1 + \pi\lambda\rho)^2} ,$$

$$p_2 = \frac{\pi^2\lambda^2\rho^2}{(1 + \pi\lambda\rho)^2} .$$

With probability p_0 , a polygon must be simulated from F_0 . The standard procedure to do this consists of generating Poisson lines sequentially by increasing

distance from the origin. The procedure is continued until the generation of additional lines no longer affects the polygon containing the origin.

With probability p_1 a polygon must be simulated from F_1 . To do this, a polygon is first generated from F_0 , and then split by a uniformly oriented line through the origin. It remains to select at random one of the two polygons thus delimited (see Fig. 3).

With probability p_2 a polygon must be simulated from F_2 . The algorithm proposed by Miles (1974) consists of taking the intersection between a polygon generated from F_0 and a cone delimited by two random rays emanating from the origin (see Fig. 3). Both rays are uniformly oriented and separated by an angle with p.d.f. $f(\varphi) = \varphi \sin \varphi / \pi$.

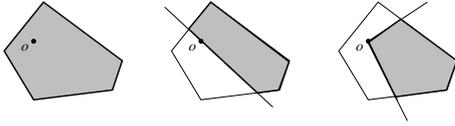


Fig. 3. *Weighted polygons generated from distributions F_0 (left), F_1 (middle) and F_2 (right).*

To illustrate the construction, a simulation of a Boolean model of Poisson polygons is depicted in Fig. 4. With a Poisson intensity θ and a Poisson line intensity λ respectively set to 10 and 1.625, the background proportion is close to 30%.



Fig. 4. *Simulation of a Boolean model of Poisson polygons.*

NON STATIONARY CASE

ALGORITHM

Let us start again with Eq. 3 that provides the avoiding functional of $X \cap Z$, and let us write it like

$$Q_{X \cap Z}(K) = \exp\left(-\vartheta(Z) \frac{\vartheta(K)}{\vartheta(Z)}\right),$$

provided that $\vartheta(Z) > 0$ (if $\vartheta(Z) = 0$, then $X \cap Z = \emptyset$ a.s.). On the other hand, Eq. 2 can be used to derive the following expansion

$$\frac{\vartheta(K)}{\vartheta(Z)} = \int_{\mathbb{R}^d} \frac{\theta_x T_x(Z)}{\vartheta(Z)} \frac{T_x(K)}{T_x(Z)} dx,$$

which shows that $\vartheta(K)/\vartheta(Z)$ is the hitting functional of some object of Z , say A^Z . More precisely, A^Z is located according to the pdf

$$f(x) = \frac{\theta_x T_x(Z)}{\vartheta(Z)}, \quad x \in \mathbb{R}^d. \quad (11)$$

Moreover, the conditional hitting functional of A^Z given its location x is equal to

$$\frac{T_x(K)}{T_x(Z)} = P\{(A_x \cap Z) \cap K \neq \emptyset \mid A_x \cap Z \neq \emptyset\}. \quad (12)$$

A^Z is called a *typical object hitting Z* .

The interest of typical objects lies in the following property (Lantuéjoul, 2002):

Proposition 1 *$X \cap Z$ is the union of a Poisson number (mean $\vartheta(Z)$) of typical objects hitting Z .*

Indeed, we can readily check that the avoiding functional of such a union of typical objects coincides with that of $X \cap Z$:

$$\sum_{n=0}^{\infty} \exp(-\vartheta(Z)) \frac{\vartheta^n(Z)}{n!} \left(1 - \frac{\vartheta(K)}{\vartheta(Z)}\right)^n = \exp(-\vartheta(K)).$$

From this proposition, the following algorithm is derived for simulating a Boolean model, even non stationary, in the domain Z .

Algorithm 3

- (i) set $X = \emptyset$;
- (ii) generate $n \sim \mathcal{P}(\vartheta(Z))$;
- (iii) if $n = 0$, return X and stop;
- (iv) generate a typical object A^Z hitting Z ;
- (v) put $X = X \cup A^Z$, $n = n - 1$ and goto (iii).

PRACTICAL IMPLEMENTATION

Algorithm 3 calls for several remarks. Step (ii) assumes that $\vartheta(Z)$ is explicitly known, but this is not always the case because the integral of $\theta_x T_x(Z)$ may not be mathematically tractable. Moreover, step (iv) requires simulating the pdf (Eq. 11) that specifies the location of the typical objects hitting Z . This distribution may have a complicated expression. Step (iv) also involves generating conditionally typical objects given their locations. It is possible to generate objects A_x sequentially till one is produced that hits Z . This algorithm can be easily implemented but lacks efficiency.

In order to alleviate these difficulties, rejection and coupling methods can be resorted to. A typical approach is to consider another Boolean model X' that *dominates* X , in the sense that its ingredients satisfy the following three properties:

- (i) its Poisson intensity function θ' satisfies $\theta'_x \geq \theta_x$ (at each point x)
- (ii) its population of objects (A'_x) satisfies $A'_x \supset A_x$
- (iii) there exists an algorithm to conditionally simulate A_x given A'_x .

Then the idea is firstly to generate the objects of X' hitting Z . Then each generated object A'_x is saved with probability θ_x/θ'_x and discarded with the complementary probability $1 - \theta_x/\theta'_x$. Finally, each

remaining object A'_x is replaced by a new object A_x generated using the conditional algorithm mentioned in (iii). A simulation of X in Z is given by the union of these new objects in Z .

Hence, a Boolean model can effectively be generated whenever it is dominated by another Boolean model that is numerically tractable and can be simulated. There is no general rule for building such a dominating model. This must be done on a case by case basis.

ACKNOWLEDGEMENTS

The topic of this paper was presented at the S4G Conference, June 25-28, 2012 in Prague, Czech Republic. The author thanks P. Calka for fruitful discussions about the simulation of weighted Poisson polygons.

REFERENCES

- Lantuéjoul C (2002). Geostatistical simulations: Models and algorithms. Berlin: Springer.
- Matheron G (1975). Random sets and integral geometry. New York: Wiley.
- Miles RE (1969). Poisson flats in Euclidean spaces. Adv Appl Prob 1:211-37.
- Miles RE (1974). A synopsis of Poisson flats in Euclidean spaces. In: Harding EF, Kendall DG, eds. Stochastic geometry. New York: Wiley. pp. 202-27.