Discrete approximation of the linear Boolean model of heterogeneous materials

John C. Handley*
Xerox Corporation, 800 Phillips Road, MS 128-25E, Webster, New York 14580-9701

(Received 21 May 1999)

The Boolean model is a random set process in which random shapes are positioned according to outcomes of a Poisson process. Two- and three-dimensional versions of the model characterize structures of certain heterogeneous materials. Linear transects of the Boolean model produce a one-dimensional Boolean model that summarizes some material properties. Two functions from linear transects, clump-length and lineal path distributions, provide information on material phase connectivity. Computation of these distributions is notoriously difficult. We provide a discrete approximation to the one-dimensional convex-grain Boolean model that yields stable, linear-time, recursive algorithms to approximate these functions. Computer simulations demonstrate accuracy and speed. [S1063-651X(99)03111-6]

PACS number(s): 02.70.−c, 47.55.Mh, 07.05.Kf, 05.20.−y

The Boolean model consists of a closed random set (shape) process (called a grain process) with an independent point process (called a germ process). Outcomes of the closed random set, consisting of random shapes, are then translated to outcomes of the point process. The result is a random set of possibly overlapping shapes. This model has applications in materials science, microscopy, geostatistics, communications, and the theory of queues [1–4]. The Boolean model is the most tractable, nontrivial random set model because it inherits many properties from its underlying Poisson process. In particular, the intersection of a line with a Boolean model in $\mathbf{R}^d$, $d \geqslant 2$, is a linear Boolean model, where the point process is a Poisson process on the line and one-dimensional (1D) shapes are line segments created by intersecting a line with higher-dimensional sets. If one restricts higher-dimensional shapes to be convex, then the induced linear model has convex shapes as well (grains are single line segments). This linear model is much studied because of its interpretation as the $M/G/\infty$ queueing system: the point process governs (memoryless) arrivals, lengths of line segments are (general) random service times, and each arrival finds one of an infinite set of servers. (See [2], Chap. 2, for a thorough discussion of the linear Boolean model.) In materials science, quantities of the induced 1D model help to characterize heterogeneous materials described by two- or three-dimensional Boolean models.

When a line intersects a Boolean model, it produces gaps and clumps where it intersects voids and shapes. Convex random shapes of the higher-dimensional model intersect a line to produce random-length line segments. These are the random shapes of the linear Boolean model and every clump is a union of these line segments. The probability distribution of segment lengths and the intensity of the Poisson process completely characterize the linear Boolean model, and from these we wish to compute G, the probability distribution of random clump lengths [2]. The probability that a segment of length z is completely covered by shapes of a Boolean model is called the lineal path function, $L(z)$. Distributions G and L are difficult to compute numerically [2,5]. The contribution of this work is to demonstrate that a discrete approximation of the linear Boolean model provides simple recursive formulas to compute these functions.

Consider a Boolean model composed of a Poisson point process with intensity ρ and disks with random radii (Fig. 1). Denote the radius probability density function by ϕ and its ith noncentral moment by

$$M_i = \int_0^\infty r^i \phi(r)\,dr. \tag{1}$$

The induced linear Boolean model has germ process intensity $\lambda = 2\rho M_1$ and the segment length distribution [2,5]

$$C(x) = \frac{1}{M_1}\int_{x/2}^\infty \sqrt{r^2 - x^2/4}\,\phi(r)\,dr. \tag{2}$$

Using the Steiner formula, the mean segment length is [4]

$$\alpha = \frac{\pi M_2}{2M_1}. \tag{3}$$

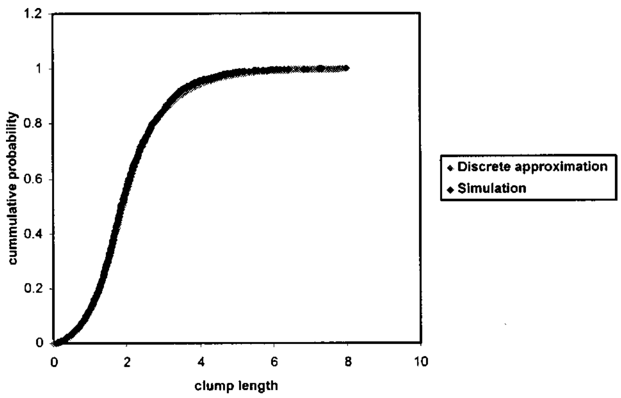
The vacancy (proportion of space not covered) of the linear (and thus spatial) Boolean model is $v = \exp\{-\eta\}$, where $\eta = \alpha\lambda$ is called the reduced density of the model.

Following [5] we present examples where radii are log-normally distributed,

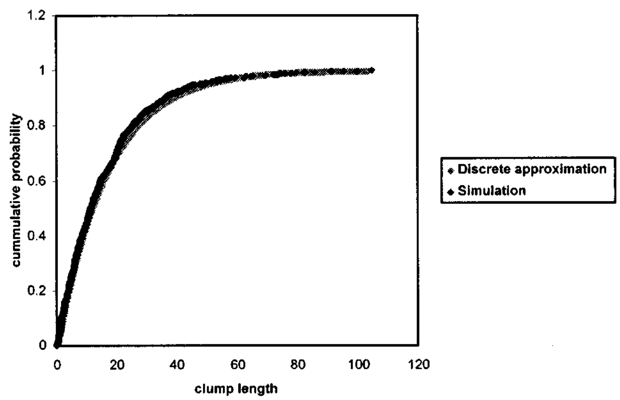
$$\phi(r) = \frac{1}{r\beta\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}\left[\frac{\ln(r/R_0)}{\beta}\right]^2 \right\} \tag{4}$$

FIG. 1. Random disk Boolean model with log-normally distributed (dimensionless) radii, $\beta = 0.25$ and $R_0 = 1$, with spatial intensity $\rho = 0.1$ on the left and $\rho = 1.0$ on the right.

*Electronic address: John.Handley@crt.xerox.com



(a)



(b)

FIG. 2. Plots of approximated ($n=100$) and simulated clump-length distributions for random disk Boolean models with log-normally distributed (dimensionless) radii, $\beta=0.25$ and $R_0=1$. Spatial intensities are $\rho=0.1$ (a) and $\rho=1.0$ (b).

with parameters R_0 and β. When $\beta=0$, we define the disk radius to be constant $r=R_0$. For examples presented here, we assume $R_0=1$. For the log-normal distribution,

$$M_i=R_0^i \exp\{i^2\beta^2/2\}. \tag{5}$$

On the line, the clump-length distribution G has Laplace-Stieltjes transform [2]

$$\gamma(s)=1+s/\lambda-\left(\lambda\int_0^\infty \exp\left[-st-\lambda\int_0^t[1-C(x)]dx\right]dt\right)^{-1}, \tag{6}$$

where C is the segment-length distribution. A closed form for G is known only in the special case where segment lengths are constant.

The discrete version of the model starts with a Bernoulli marking process on the discrete line: a point is marked with probability p and markings are independent. The number of points marked in an interval of n points has a binomial distribution. Segment lengths are governed by a random variable that is independent of the marking process and has discrete distribution C. Define $F(x)=1-p+pC(x)$, $x=0,1,\ldots$, and $C(0)\equiv 0$ so that $F(0)=1-p=q$. Let K be the length of a clump in this model. A recursive formula for the discrete probability density function of K is [6]

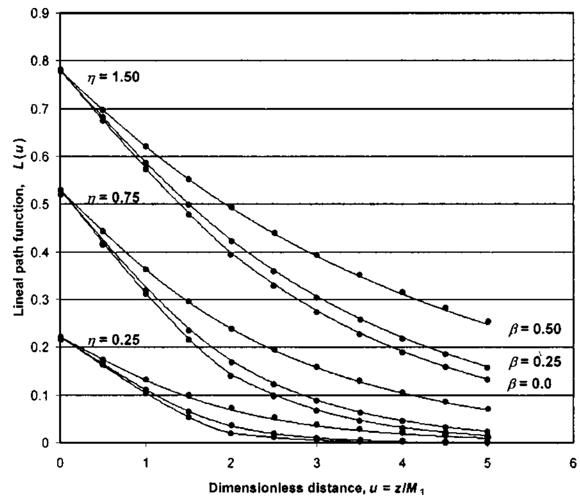


FIG. 3. Graphs showing approximated lineal path functions and simulated values for random disk Boolean models with log-normally distributed radii, $R_0=1.0$.

$$P(K=m)=\sum_{j=1}^{m}[F(m)-F(j-1)]\prod_{i=1}^{j-1}F(i-1)P(K=m-j) \tag{7}$$

for $m=1,2,\ldots$, where $P(K=0)=q/p$ starts the recursion. Approximate a Poisson process by a binomial process in the usual way with $p=\lambda/n$ and approximate a segment of length x by $[nx]$ pieces of length $1/n$. Equation (7) can be rewritten

$$P(K=x)=\sum_{j=1}^{[nx]}\left[F(x)-F\left(\frac{j-1}{n}\right)\right]\prod_{i=1}^{j-1}F\left(\frac{i-1}{n}\right)\times P(K=x-j/n). \tag{8}$$

As n increases, the binomial distribution converges to the Poisson distribution with parameter λ, the discrete segment length distribution converges to the continuous version, and Eq. (8) approximates its continuous counterpart:

$$P(K=x)\approx\frac{G(x+1/2n)-G(x-1/2n)}{n}. \tag{9}$$

Equation (8) is easily implemented in software and, by storing intermediate products and distribution values, it is linear in time and storage as a function of n. Figure 2 shows graphs of approximated and simulated clump length distributions from the linear section of a disk model with log-normally distributed radii. Computations for clump lengths up to 8 units long took 5 sec and up to length 104 units took 53 sec on a 200-MHz computer.

Estimating the segment-length distribution from a linear Boolean model realization is given scant attention in the literature, perhaps because it requires the difficult inversion of the clump-length distribution [solving for C in Eq. (6)]. However, the segment-length distribution would aid in characterizing the distribution of particle shapes, none of which may be completely observed due to overlapping. Equation (7) can be inverted to provide an estimated segment-length distribution from the observed clump-length distribution. Gap lengths are geometrically distributed with mean $1/[1-F(0)]$. One could tally a histogram of clump lengths, fit the resulting empirical density function with continuous density $\hat{f}_K$, approximate the discrete density of K,

$$P(K=x)=\int_{x-1/2n}^{x+1/2n}\hat{f}_K(t)dt, \tag{10}$$

and approximate F by

$$\hat{F}(x)=\frac{\hat{P}(K=x)+\Sigma_{j=1}^{[nx]}\hat{F}\left(\frac{j-1}{n}\right)\Pi_{i=1}^{j-1}\hat{F}\left(\frac{i-1}{n}\right)\hat{P}(K=x-j/n)}{\Sigma_{j=1}^{[nx]}\Pi_{i=1}^{j-1}\hat{F}\left(\frac{i-1}{n}\right)\hat{P}(K=x-j/n)}, \tag{11}$$

where $\hat{F}(0)=1-1/\overline{V}$ and $\overline{V}$ is the average gap length.

The lineal function $L(z)$ is the probability that a line segment of length z is completely covered by a Boolean model. For a one-dimensional model, $L(z)$ is the probability of complete coverage. From [2], theorem 2.6, we have the ordinary Laplace transform of the lineal function

$$\pi(s)=s^{-1}-\left(s^2e^{\alpha\lambda}\int_0^\infty \exp\left[-sz-\lambda\int_0^z\{1-G(x)\}dx\right]dz\right)^{-1}, \tag{12}$$

where α is the mean segment length. The difficulty of computing L from Eq. (12) is discussed in [5].

In the discrete setting, denote by H_m the event that a discrete interval of length $m=1,2,...$ is completely covered. Let W denote the event that the Boolean model does not cover a given point. (The probability of this event in the discrete model corresponds to the vacancy in the linear Boolean model.) If C has finite mean [7],

$$P(W)=\prod_{i=0}^{\infty}F(i). \tag{13}$$

Let D_m denote the set of outcomes on a discrete interval of length m such that the segments cover the interval. Event D_m has probability

$$P(D_m)=1-\sum_{j=1}^{m}\prod_{i=1}^{j-1}F(i)P(D_{m-j}), \tag{14}$$

where $P(D_1)\equiv 1-F(0)$ [6,7]. Assuming the probability of an infinite-length covering is zero, the discrete version of the probability of total coverage has a recursive expression [6,7],

$$P(H_m)=P(H_{m-1})-P(W)P(D_m), \tag{15}$$

where the recursion is initiated by $P(H_0)\equiv 1$. Using the discrete approximation to the Boolean model,

$$\hat{L}(z)=1-P(W)\sum_{m=1}^{[nz]}P(D_m), \tag{16}$$

where

$$P(D_m)=1-\sum_{j=1}^{m}\prod_{i=0}^{j-1}F(i/n)P(D_{m-j}). \tag{17}$$

Using Eqs. (16) and (17), Fig. 3 replicates Fig. 3 of [5] in which lineal path functions for several random-radius disk Boolean models are plotted with values from computer simulations. By storing intermediate values, computation time is dominated by numerical integration of the radius probability density function. For each L, where $\beta>0$, calculations from $u=0$ to 20 using $n=100$ took 12 sec on a 200-MHz computer.

This discrete approximation to the convex-grain linear Boolean model provides simple, fast recursive algorithms to compute clump-length and lineal path distributions of spatial Boolean models. These methods can be applied to linear transects of any convex-gain Boolean model, provided the segment-length distribution can be computed.

[1] A. Frey and V. Schmidt, Adv. Performance Anal. **1**, 1 (1998).
[2] P. Hall, *Introduction to the Theory of Coverage Processes* (Wiley, New York, 1988).
[3] I. Molchanov, *Statistics of the Boolean Model for Practitioners and Mathematicians* (Wiley, Chichester, 1997).
[4] D. Stoyan, W. S. Kendall, and J. Mecke, *Stochastic Geometry and Its Applications* (Wiley, New York, 1995).
[5] J. Quintanilla and S. Torquato, Phys. Rev. E **54**, 4027 (1996).
[6] J. C. Handley and E. R. Dougherty, Signal Process. **51**, 147 (1996).
[7] E. R. Dougherty and J. C. Handley, Signal Process. **43**, 1 (1995).