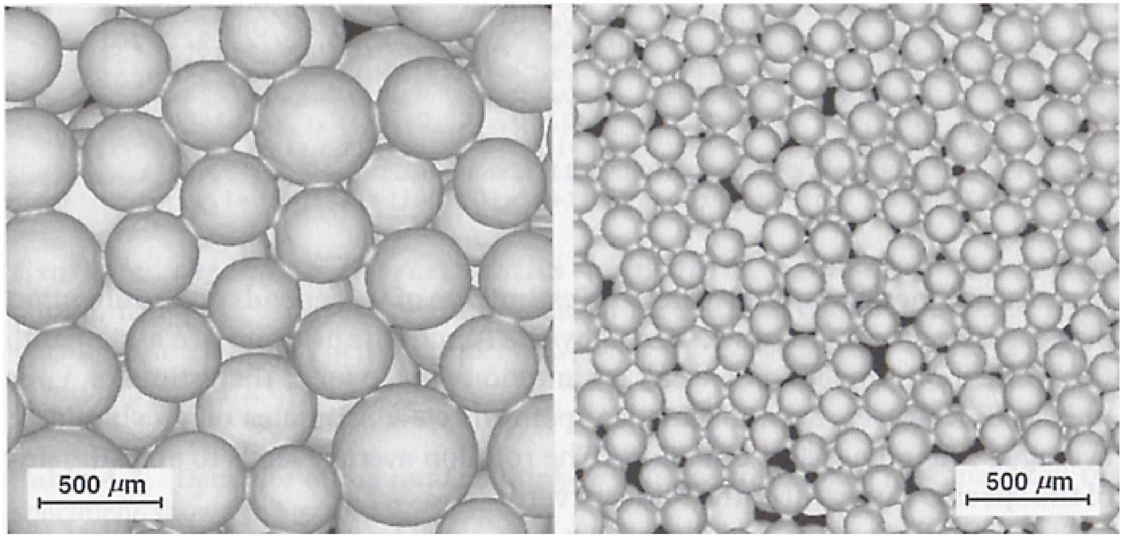
Bedingte Simulation in der stochastischen Geometrie

1 – Das Boolesche Modell
2 – Das "random token"-Modell
3 – Die Boolesche Zufallsfunktion
4 – Literatur

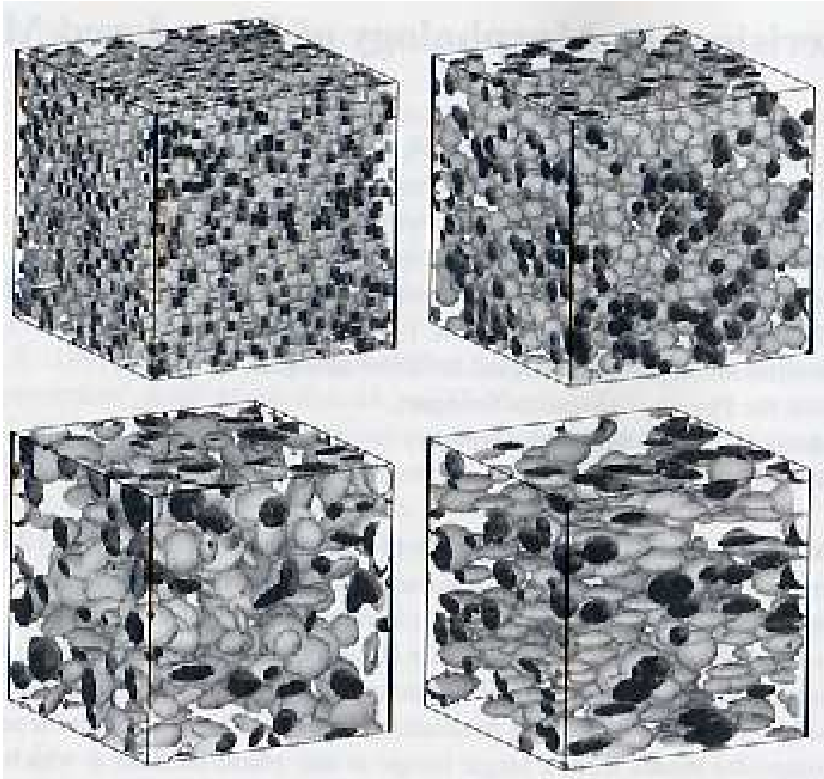
M.Knubel
21.7.2003

1 Das Boolesche Modell

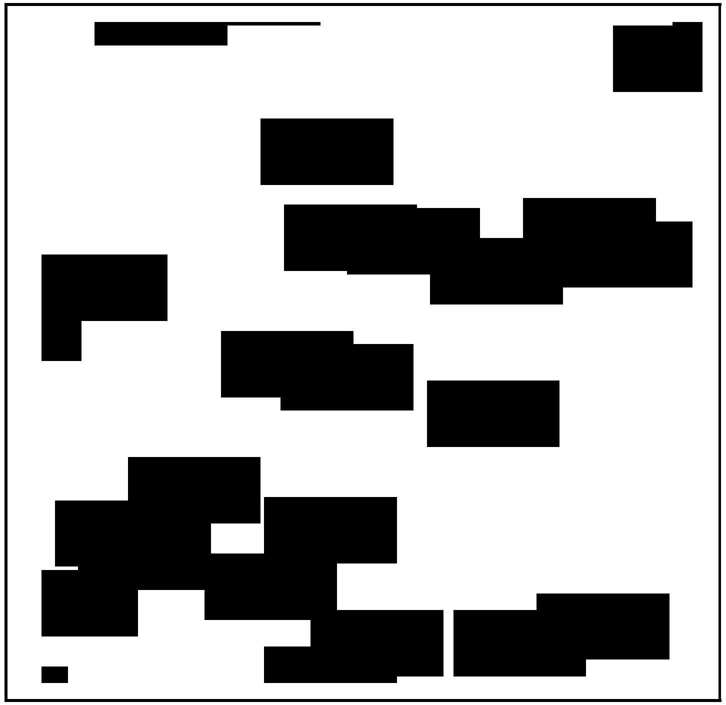
1.1 Motivation



Kupferpulver



Boolesches Modell in 3D



Boolesches Modell in 2D

1.2 Definition und grundlegende Eigenschaften

Zwei grundlegende und unabhängige Bausteine des Booleschen Modells:

1. Eine Menge von Keimen, die durch einen Poisson-Punktprozess $\mathbf{P}$ mit Intensitätsfunktion $\theta = \theta(x), x \in \mathbf{R}^d$ erzeugt werden.

2. Eine Familie $\{A(x), x \in \mathbf{R}^d\}$ von unabhängigen, zufälligen, nichtleeren, kompakten Teilmengen des $\mathbf{R}^d$. Wir bezeichnen die Menge $A(x)$ als Objekt plaziert im Punkt x und definieren das Kapazitätsfunktional

$$T_x := T_x(K) = P\{A(x) \cap K \neq \emptyset\}$$

Definition 1. *Ein Boolesches Modell X ist die Vereinigung aller Objekte platziert in den Poisson-Keimen*

$$X = \bigcup_{x \in \mathbf{P}} A(x)$$

- Problem: Das Boolesche Modell ist im allgemeinen keine abgeschlossene Menge.
- Abgeschlossenheit ist allerdings garantiert, falls die Menge aller Punkte in P endlich ist.

Sei also $N(K)$ die Anzahl der Objekte, die die kompakte Teilmenge K treffen, mit

$$N(K) = \sum_{x \in \mathbf{P}} \mathbf{1}_{\{K \cap A(x) \neq \emptyset\}}$$

Durch einfache Rechnung kann man zeigen, dass gilt

$$v(K) = E(N(K)) = \int_{\mathbf{R}^d} \theta(x) T_x(K) dx \leq \infty$$

Definition 2. *Ein Boolesches Modell ist von endlicher Ordnung, falls gilt*

$$v(K) < \infty, \quad \forall K \in \mathbf{K}$$

wobei $\mathbf{K}$ *die Menge aller kompakten Teilmengen des* $\mathbf{R}^d$ *ist.*

Proposition 1. $N(K)$ *folgt einer Poisson-Verteilung mit Erwartungswert* $\upsilon(K)$

Der Beweis basiert auf der erzeugenden Funktion und man kann zeigen:

$$E\{s^{N(K)}\} = exp\{(s-1)\int_{\mathbf{R}^d} \theta(x)T_x(K)dx\} = exp\{(s-1)\upsilon(K)\}$$

Aufgrund von

$$P\{X \cap K = \emptyset\} = P\{N(K) = 0\}$$

gilt:

Folgerung 1. *Das Kapazitätfunktional des Booleschen Modells ist*

$$P\{X \cap K \neq \emptyset\} = 1 - e^{-\upsilon(K)}, \qquad K \in \mathbf{K}$$

Algebraische und stereologische Stabiltätseigenschaften:

- die Vereinigung von zwei unabhängigen Booleschen Modellen ist ein Boolesches Modell
- die Dilatation eines Booleschen Modells mit einer nichtleeren, kompakten Teilemenge des $\mathbf{R}^d$ ist ein Boolesches Modell
- der Schnitt eines Booleschen Modells mit einer kompakten Teilmenge des $\mathbf{R}^d$ ist ein Boolesches Modell
- der Schnitt eines Booleschen Modells mit einer i-dimensionalen Ebene ist ein Boolesches Modell

1.3 Die unbedingte Simulation

- D ist eine nichtleere, kompakte Teilmenge des $\mathbf{R}^d$
- $X \cap D$ ist ein Boolesches Modell mit Intensitätsfunktion $\theta(.)T_.(D)$
- die Anzahl der Objekte in $X \cap D$ ist Poisson verteilt mit Erwartungswert $v(D)$ (vgl. Proposition 1)

Definition 3. *Ein typisches Objekt von $X \cap D$ ist ein Objekt, dass "gleichmäsig" aus allen Objekten von $X \cap D$ ausgewählt wurde. Sein Kapazitätsfunktional ist:*

$$T(K) = \frac{1}{v(D)} \int_{\mathbf{R}^d} \theta(x) T_x(D \cap K) dx = \frac{v(D \cap K)}{v(D)}$$

Proposition 2. *$X \cap D$ hat dieselbe Verteilung wie die Vereinigung von N unabhängigen typischen Objekten, wobei die Zahl N Poisson verteilt ist mit Erwartungswert $\upsilon(D)$.*

Beweis: Sei Y so eine Vereinigung. Für das Kapazitästfunktional von Y gilt:

$$\begin{aligned} P\{Y \cap K \neq \emptyset\} &= 1 - \sum_{n=0}^{\infty} e^{-\upsilon(D)} \frac{\upsilon(D)^n}{n!} [1 - T(K)]^n \\ &= 1 - exp\{-\upsilon(D) T(K)\} \\ &= 1 - exp\{- \int_{\mathbf{R}^d} \theta(x) T_x(D \cap K) dx\} \end{aligned}$$

und dies ist genau das Kapazitätsfunktional von $X \cap D$.

Algorithmus zur Simulation eines Booleschen Modells:

1. Sei $X = \emptyset$
2. Generiere $N \sim Poisson(v(D))$
3. Wenn $N = 0$ gebe X aus
4. Generiere N unabhängige, typische Objekte $A(x_i) \sim T,\ i = 1, ..., N$
5. Setze $X = \bigcup_{i=1}^{N} A(x_i)$

Algorithmus zur Simulation eines typischen Objekts:

1. Generiere $x_i,\ P\{x_i \in C\} = \int_C \theta(y) T_y(D) dy$
2. Generiere $A \sim T_{x_i}(D \cap .)/T_{x_i}(D)$
3. Liefere A

Eine andere Alternative:

1. Generiere $x_i,\ P\{x_i \in C\} = \int_C \theta(y) T_y(D) dy$
2. Generiere $A \sim T_{x_i}(.)$
3. Wenn $A \cap D = \emptyset$ gehe zu 2
4. Liefere $A \cap D$

Ein iterativer Algorithmus zur Simulation eines Booleschen Modells mit Hilfe des Metropolis-Algorithmus:

1. Sei $\Phi = \emptyset$
2. Generiere eine Zufallsvariable U mit Werten $1, -1, 0$ mit den zugehörigen Wahrscheinlichkeiten

$$p_1 = \frac{v(D)}{2(v(D) + \#\Phi + 1)}, \quad p_{-1} = \frac{\#\Phi}{2(v(D) + \#\Phi)}, \quad p_0 = 1 - p_1 - p_{-1}$$

3. Wenn $U = 1$, dann generiere $A \sim T$ und setze $\Phi = \Phi \cup A$
4. Wenn $U = -1$, dann generiere $A \sim U(\Phi)$ und setze $\Phi = \Phi \backslash A$
5. gehe zu 2

Erzeugung einer Poisson-Verteilung mit Hilfe des Metropolis Algorithmus.

Sei Q eine Übergangsmatrix eines Geburts- und Todesprozess mit der Geburts- und Todeswahrscheinlichkeit $q = \frac{1}{2}$. Es gilt also Q ist irreduzibel, aperiodisch und symetrisch. Sei nun p unsere zu simulierende Poisson-Verteilung, dann ist mit

$$P(x,y) = Q(x,y)\frac{p(y)}{p(x)+p(y)} \quad , x \neq y$$

eine Übergangsmatrix $\mathbf{P}$ definiert. x, y sind Zustände der Markov-Kette.

Es gilt nun, da $p > 0$ für alle Zustände der Markov-Kette, dass P ebenfalls irreduzibel, aperiodisch ist, und ausserdem gilt noch:

$$\sum_x p(x)P(x,y) = \sum_x p(y)P(y,x) = p(y)\sum_x P(y,x) = p(y)$$

das heisst also P ist die Übergangsmatrix einer Markov-Kette mit Grenzverteilung p.
Sei nun p Poisson verteilt mit Erwarungswert $v(D)$ dann ist leicht zu sehen, dass

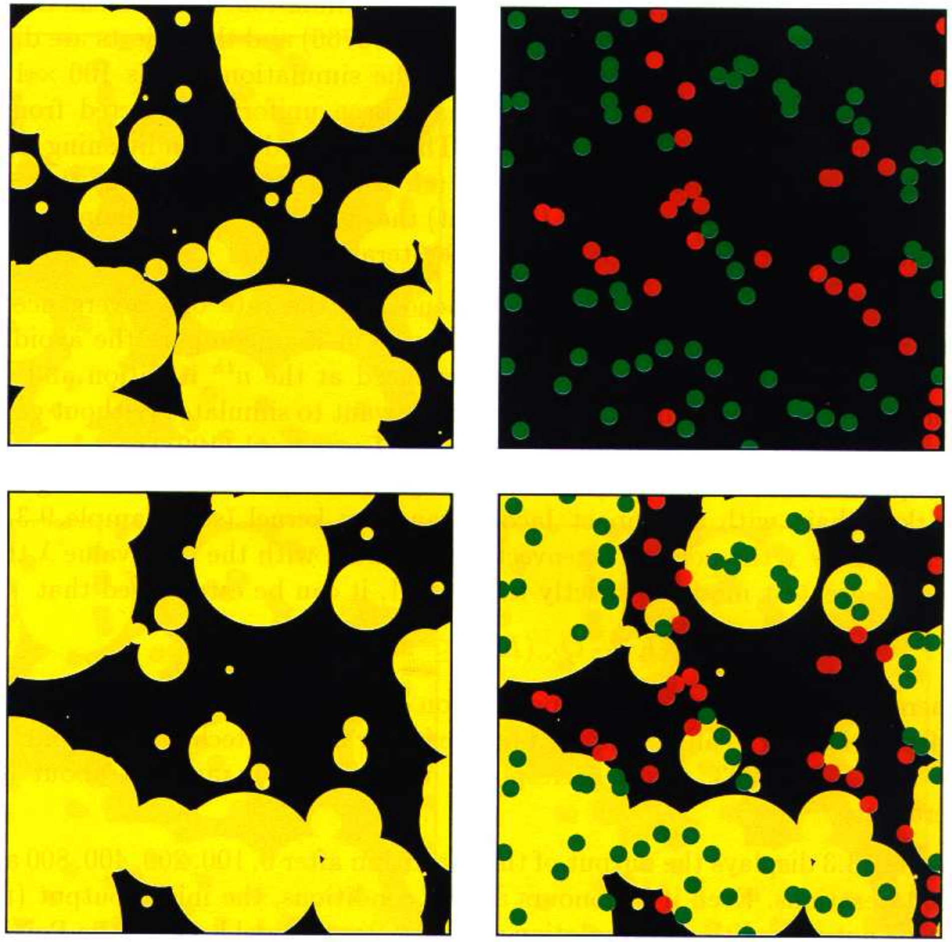
$$P(x,x+1) = p_1 \ und \ P(x,x-1) = p_{-1}$$

also erzeugt obiger Algorithmus genau die gewünschte Verteilung!

1.4 Die bedingte Simulation

Es gilt:

- C_0 ist eine endliche Teilmenge von X^c
- C_1 ist eine endliche Teilmenge von X
- Ω_c ist die Menge aller zulässigen Realisierungen der zufälligen Menge Φ, die die Restriktionen C_0 und C_1 erfüllen.



Algorithmus zur Simulation eines bedingten Booleschen Modells

1. Generiere eine zulässige Startmenge $\Phi \in \Omega_C$
2. Generiere eine Zufallsvariable U mit Werten $1, -1, 0$ mit den zugehörigen Wahrscheinlichkeiten

$$p_1 = \frac{v(D)}{(2v(D) + \#\Phi + 1)}, \quad p_{-1} = \frac{\#\Phi}{2(v(D) + \#\Phi)}, \quad p_0 = 1 - p_1 - p_{-1}$$

3. Wenn $U = 1$, dann generiere $A \sim T$. Wenn $A \cap C_0 = \emptyset$, dann setze $\Phi = \Phi \cup A$
4. Wenn $U = -1$, dann generiere $A \sim U(\Phi)$. Wenn $C_1 \subset \Phi \backslash A$ ist, dann setze $\Phi = \Phi \backslash A$
5. gehe zu 2

Warum erzeugt die Akzeptanz- und Verwerfungsmethode im obigen Algorithmus genau die Verteilung des bedingten Booleschen Modells?

Sei nun:

$$p_c(A) = \frac{p(A)}{p(\Omega_c)}, \quad A \in \mathbf{A}_c \ (die\ zu\ \Omega_c\ geh\ddot{o}rende\ \sigma - Algebra)$$

$$P_c(x, A) = P(x, A) + \mathbf{1}_{\{x \in A\}} P(x, \Omega \backslash \Omega_c), \quad x \in \Omega_c,\ A \in \mathbf{A}_c$$

Man kann nun zeigen, dass in unserem Fall P_c irreduzibel, aperiodisch und reversibel ist, und p_c invariant bezüglich P_c ist.

Algorithmus zur Generierung einer zulässigen Startkonfiguration $\Phi \in \Omega_c$

1. Setze $\Phi = \emptyset$ und $C = C_1$
2. Generiere $A \sim T$
3. Wenn $A \cap C_0 \neq \emptyset$ oder $A \cap C = \emptyset$, dann gehe zu 2
4. Setze $\Phi = \Phi \cup A$ und $C = C \backslash A$
5. Wenn $C \neq \emptyset$, dann gehe zu 2
6. liefere Φ

1.5 Stationäres Boolesches Modell

- Sei θ konstante Intensitätsfunktion: $\theta(x) = \theta$.
- Alle $A(x) \stackrel{d}{=} A + x$, wobei A das typische Korn platziert im Ursprung ist, sind unabhängig, identisch verteilt (modulo x).

Es gilt dann also:

$$P\{A(x) \cap K \neq \emptyset\} = P\{A_x \cap K \neq \emptyset\} = P\{A \cap K_{-x} \neq \emptyset\} = T(K_{-x})$$

und das Kapazitätsfunktional von X ergibt sich also zu:

$$P\{X \cap K \neq \emptyset\} = 1 - exp\{-\theta \int_{\mathbf{R}^d} T(K_{-x})dx\}$$

Proposition 3. *Das Kapazitätsfunktional eines stationären Booleschen Modells mit Intensität θ und Objekt A ist:*

$$P\{X \cap K \neq \emptyset\} = 1 - e^{-\theta E\{|A \oplus \check{K}|\}} \qquad K \in \mathbf{K}$$

Beweis

$$
\begin{aligned}
\int_{\mathbf{R}^d} T(K_{-x})dx &= E\{\int_{\mathbf{R}^d} \mathbf{1}_{\{A\cap K_{-x}\neq\emptyset\}}dx\} \\
&= E\{\int_{\mathbf{R}^d} \mathbf{1}_{\{-x\in A\oplus\check{K}\}}dx\} \\
&= E\{|A\oplus\check{K}|\}
\end{aligned}
$$

2 Das "random token"-Modell

2.1 Definition und grundlegende Eigenschaften

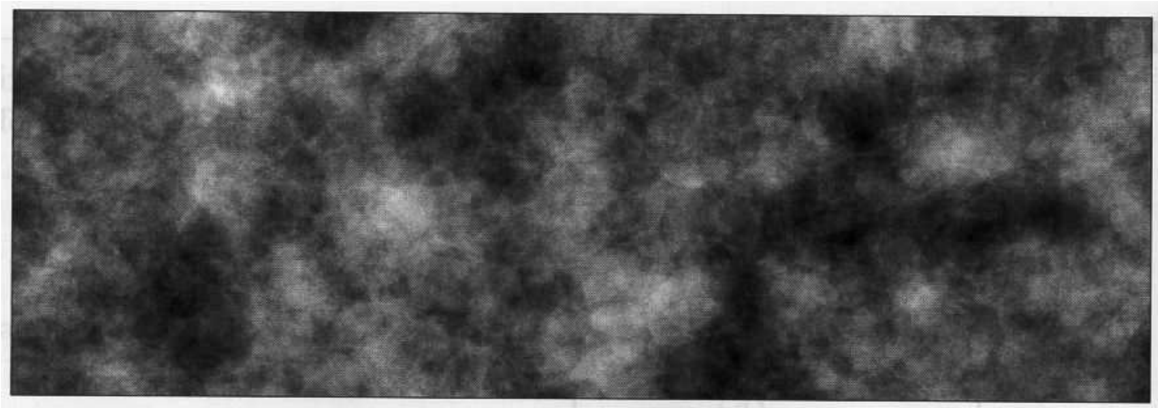
Bausteine des "random token"-Modells

1. Ein stationäres Boolesches-Modell X, wobei $E|A| < \infty$, erzeugt durch einen stationären Poisson-Punktprozess $\mathbf{P}$.

2. Eine Familie von paarweise unabhängigen Gewichten $(\epsilon(x), x \in \mathbf{R}^d)$ für jedes Objekt in X. Diese Gewichte nehmen nur zwei Werte nämlich 1 und -1 mit der gleichen Wahrscheinlichkeit $p = \frac{1}{2}$ an.

Definition 4. *Ein "random token"-Modell ist die gewichtete Summe der Indikatorfunktionen von Objekten platziert an Poisson-Keimen.*

$$Z(y) = \sum_{x \in P} \epsilon(x) \mathbf{1}_{\{y \in A(x)\}}, \qquad y \in \mathbf{R}^d$$



Satz 1. *Sei J eine endliche Menge von Indizes. Dann gilt für die charakteristische Funktion von $(Z(y_j), j \in J)$:*

$$E\{exp\{i\sum_{j\in J} u_j Z(y_j)\}\} = exp\{\theta \sum_{K\subseteq J} [\cos u_K - 1]p_k\}$$

wobei

$$u_K = \sum_{k\in K} u_k, \qquad p_k = \sum_{K\subseteq L\subseteq J} (-1)^{\#L\setminus K} |A_0 \ominus \check{Y}_L|$$

und

$$Y_L = \{y_l, l \in L\}$$

daraus folgt $Z(y)$ nimmt Werte aus $\mathbf{Z}$ an , mit positiver Wahrscheinlichkeit

2.2 Simulation

Die bedingte Simulation von Z in D mit gegebenen Daten $\{Z(c) = z(c), c \in C\}$.

Initialisierungsalgorithmus:

1. Setze $X_0 = \emptyset$. Setze $\tilde{z}(c) = z(c)$ und $\tilde{z}(c) = z(c) - z_k(c)$
2. Wenn $|\tilde{z}(c)| = 0, \forall c \in C$, dann liefere X_0
3. Wähle $c_0 \sim U\{c \in C | \tilde{z}(c) \neq 0\}$, generiere dann ein Objekt $A(x)$ mit $c_0 \in A(x)$, und setze $\epsilon(x) = sign(\tilde{z}(c_0))$
4. Wenn $|\tilde{z}(c) - \epsilon(x)| \leq |\tilde{z}(c)|$ für alle $c \in C \cap A(x)$, dann füge $(A(x), \epsilon(x))$ zu X_0 und setze ausserdem $\tilde{z}(c) = \tilde{z}(c) - \epsilon(x)$ für alle $c \in C \cap A(x)$

5. Gehe zu 2

Sei B eine kompakte Teilmenge des $\mathbf{R}^d$. Betrachte nun alle Verschiebungen dieser Menge, die die Beobachtungsstruktur D schneiden. **Algorithmus:**

1. Sei $X = X_0$ (erzeugt mit dem vorherigen Algorithmus)
2. Generiere $x \sim U(D \oplus \check{B})$. Setze $Y = \{(A(y), \epsilon(y) \in X | u \in B_x\}$
3. Generiere $N \sim Poisson(\theta|B|)$. Generiere eine Menge Z von N unabhängigen gewichteten Objekten mit Keimen in B_x
4. Wenn $\sum_{(X \backslash Y) \cup Z} \epsilon(x) \mathbf{1}_{\{c \in A(x)\}} = z(c)$ für alle $c \in C$, dann setze $X = (X \backslash Y) \cup Z$
5. Gehe zu 2

3 Die Boolesche Zufallsfunktion

3.1 Defintion und grundlegende Eigenschaften

Bausteine der Booleschen Zufallsfunktionen:

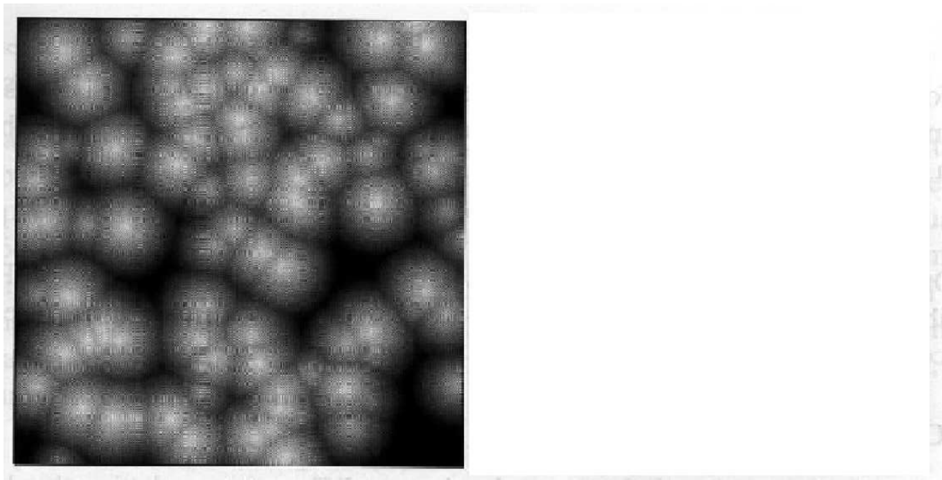
1. Ein stationäres Boolesches Modell X, wobei das erwartete Volumen von A endlich sein soll, erzeugt von einem stationären Poisson-Punktprozess $\mathbf{P}$.

2. Eine Familie von paarweise unabhängigen Gewichten $\{\epsilon(x), x \in \mathbf{R}^d\}$ für jedes Objekt in X. Diese Gewichte folgen der gleichen Verteilung, und nehmen die Werte $z_1 < ... < z_n$ mit den Wahrscheinlichkeiten $p_1, ..., p_n$ an.

Definition 5. *Die Boolesche Zufallsfunktion weist jedem Punkt $y \in \mathbf{R}^d$ das maximale Gewicht über all die Objekte zu, die y enthalten.*

$$Z(y) = \max_{x \in P, y \in A(x)} \epsilon(x), falls \{x \in \mathbf{P} : y \in A(x)\} \neq \emptyset, \quad y \in \mathbf{R}^d$$

$$Z(y) = -\infty, \; sonst$$



Alternativ gilt auch:

$$Z(y) = \max_{\{i|y \in X_i\}} z_i \qquad y \in \mathbf{R}^d$$

wobei X_i ist die Vereinigung aller Objekte mit Gewicht z_i. Dieses X_i ist ein stationäres Boolesches Modell mit Intensität der Keinme $\theta_i = \theta p_i$.

Proposition 4. *Die Verteilung des Maximums einer Booleschen Zufallsfunktion über eine kompakte Teilmenge K des $\mathbf{R}^d$ ist gegeben durch:*

$$P\{\max_{y \in K} Z(y) < z\} = \prod_{\{i|z_i > z\}} e^{-\theta_i E|A_i \oplus \check{K}|}$$

3.2 Simulation

Bedingungen:

- **Bedingungen:**
  $Z(c) = z(c)$ für alle c in einer endlichen Teilmenge C des $\mathbf{R}^d$.

- **Keine Randeffekte:**
  $N(D) \sim Poisson(v(D))$ typische Objekte werden im Beobachtungsfenster D simuliert, $v(D) = \theta E|D \oplus \check{A}|$. Jedes typische Objekt ist $D \cap A(x)$, $A(x)$ ist mit Verteilungsdichte $g(x) = P\{(A + x) \cap D \neq \emptyset\}$ gegeben. Dem Objekt $D \cap A(x)$ ist der Wert $\epsilon(x)$ zugewiesen.

Algorithmus:

1. Generiere $X \in \Omega_c$
2. Generiere eine Zufallsvariable U mit Werten $1, -1, 0$ mit den Wahrscheinlichkeiten:
$$p_1 = \frac{v(D)}{2(v(D) + \#X + 1)}, \quad p_{-1} = \frac{\#X}{2(v(D) + \#X)}, \quad p_0 = 1 - p_1 - p_{-1}$$
3. Wenn gilt $U = 1$, dann generiere ein typisches,gewichtetes Objekt (A, ϵ).Ist $z(c) \geq \epsilon$ für alle $c \in C \cap A$, dann setze $X = X \cup (A, \epsilon)$
4. Wenn gilt $U = -1$, dann generiere $(A, \epsilon) \sim U(X)$. Wenn dann für jedes $c \in A \cap C$ ein $(A\prime, \epsilon\prime) \in X \backslash (A, \epsilon)$ existiert, so dass $z(c) = \epsilon\prime$, dann setze $X = X \backslash (A, \epsilon)$
5. Gehe zu 2

Zusammenfasung der Ergebnisse

- Das Boolesche Modell lässt sich mit Hilfe von Markov-Ketten simulieren, im Paper von Kendall und Thönnes wird sogar ein perfekter Simulationsalgorithmus vorgestellt.
- Anstatt der Akzeptanz- und Verwerfungsmethode gibt es noch die Methode des "Simulated Annealing" zur bedingten Simulation.
- Die Boolesche Zufallsfunktion, lässt sich mit Hilfe von Markov-Ketten simulieren, ob auch ein perfekter Simulationsalgorithmus implementiert werden kann ging aus der benutzten Literatur nicht hervor.

4 Literatur

1. C. Lantujoul, Geostatistical Simulation: Models and Algorithms. Springer, 2002

2. D. Jeulin, Advances in Theory and Applications of Random Sets, World Scientific 1997

3. W. Kendall and E. Thönnes, Perfect Simulations in Stochastic Geometry, Department of Statistics, University of Warwick , Preprint 1998