

# Random Sets Simulation

Christian Lantuéjoul & Thomas Romary

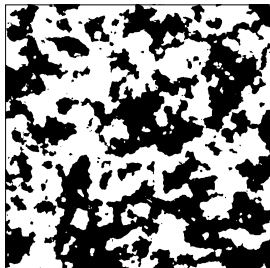
July 5, 2014



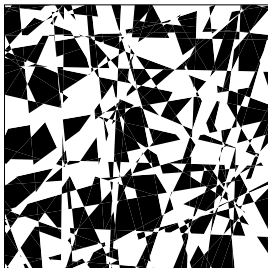
# Outline

- 1 Statistical characterization of random sets
- 2 Gaussian excursion sets
  - Definition
  - Conditional simulation
  - Plurigaussian random functions
- 3 Boolean model
  - Introduction
  - Definitions and properties
  - Conditional simulation algorithm

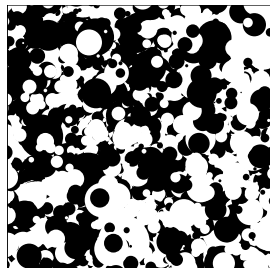
## On the sensitivity of the spatial distribution



Gaussian excursion set



Poisson polygons



Dead leaves model

These three stochastic models have exactly the same trivariate distributions

## Random set



### Remark

The spatial distribution of the indicator function of the random set  $X$ ,  $Y(x) = 1_{x \in X}$  ( $x \in \mathbb{R}^d$ ), does not always answer the question

### Counterexample

The spatial distribution of the random set made of a uniform point in a bounded domain  $D$  says nothing at all

$$P\{Y(x_1) = y_1, \dots, Y(x_n) = y_n\} = \begin{cases} 1 & \text{if } y_1 = \dots = y_n = 0 \\ 0 & \text{otherwise} \end{cases}$$

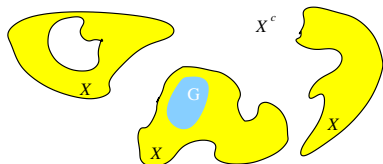
# Random set

## Idea

Instead of probing  $X$  with finite sets of points, probe  $X$  with bigger subsets. **Open subsets** can be considered.

$$G \subset X$$

$$G \cap X \neq \emptyset$$

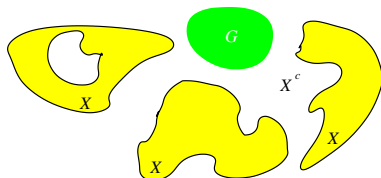


$$G \subset X \iff G \subset \text{Int}(X)$$

Random **open** set theory

$$G \cap X^c \neq \emptyset$$

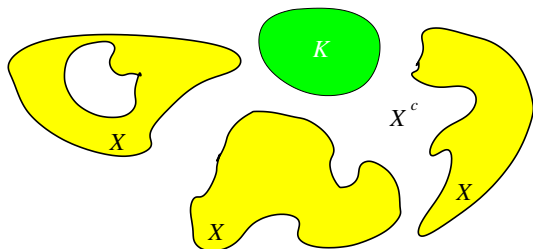
$$G \subset X^c$$



$$G \subset X^c \iff G \cap \text{Cl}(X) = \emptyset$$

Random **closed** set theory

## Random closed set



The statistical properties of a random closed set  $X$  are specified by its **hitting functional**  $T$ , or equivalently, by its **avoiding functional**  $Q$ , i.e. the set of values taken by

$$T(K) = P\{X \cap K \neq \emptyset\}$$

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

for each **compact** (bounded and closed) subset  $K$  of  $\mathbb{R}^d$

# Outline

## 1 Statistical characterization of random sets

## 2 Gaussian excursion sets

- Definition
- Conditional simulation
- Plurigaussian random functions

## 3 Boolean model

- Introduction
- Definitions and properties
- Conditional simulation algorithm

# Gaussian excursion set

## Basic ingredients

- a standardized gaussian random field  $Y$  with covariance  $C$
- a numerical value  $\lambda$

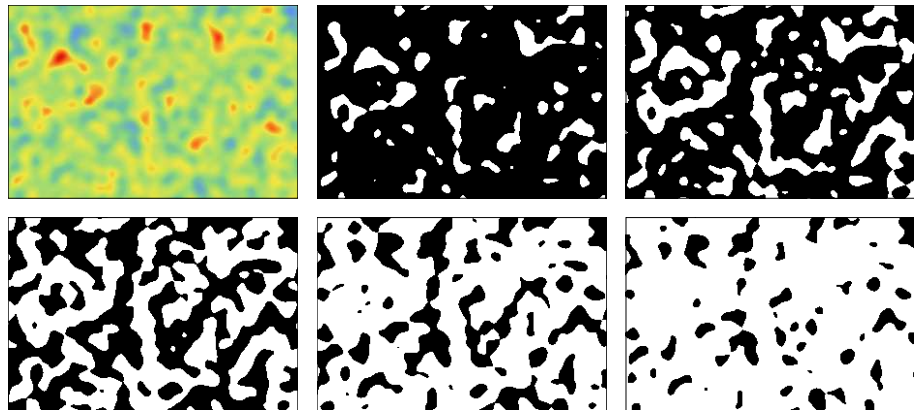
## Definition

The excursion set of  $Y$  at level  $\lambda$  is the set of points where  $Y$  takes values greater or equal to  $\lambda$

$$X_{\lambda}(x) = \begin{cases} 1 & \text{if } Y(x) \geq \lambda \\ 0 & \text{if } Y(x) < \lambda \end{cases}$$

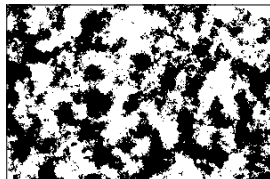


## Examples at various levels

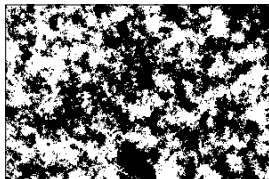


Gaussian random field (gaussian covariance) and its excursion sets at levels 1, 0.5, 0, -0.5 and -1

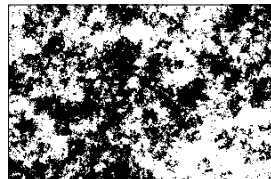
# Examples for different covariance functions



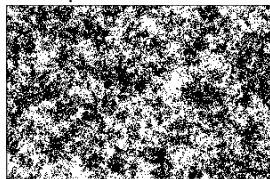
Spherical



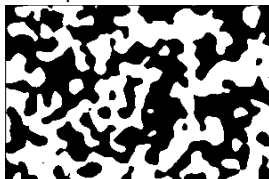
Exponential



Hyperbolic



Stable



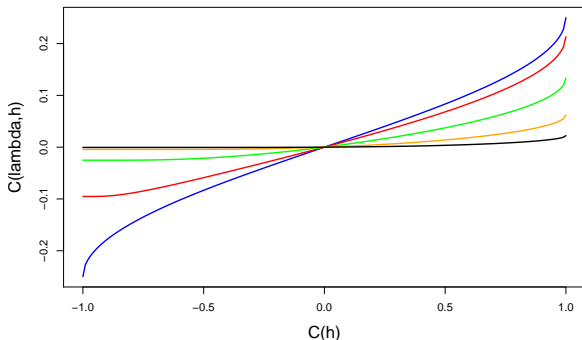
Gaussian



Cardinal sine

## Relationship between covariance functions

$$C_{\lambda}(h) = \frac{1}{2\pi} \int_0^{C(h)} \exp\left(\frac{-\lambda^2}{1+x}\right) \frac{dx}{\sqrt{1-x^2}} \quad h \in \mathbb{R}^d$$



Covariance of excursion sets at levels 0,  $\pm 0.5$ ,  $\pm 1$ ,  $\pm 1.5$  and  $\pm 2$  versus covariance of the underlying Gaussian random function

# Behaviour at the vicinity of the origin

Value at 0

$$C(0) = 1 \qquad C_\lambda(0) = G(\lambda)[1 - G(\lambda)]$$

Behaviour for small  $|h|$

$$C_\lambda(0) - C_\lambda(h) \approx \frac{1}{\pi\sqrt{2}} \sqrt{C(0) - C(h)} \exp\left(-\frac{\lambda^2}{2}\right)$$

Physical meaning of  $C'_\lambda(0)$  (isotropic case)

1D  $-2C'_\lambda(0)$  mean number of boundary points per unit length  
( $N_L$ )

2D  $-\pi C'_\lambda(0)$  mean boundary length per unit area ( $L_A$ )

3D  $-4C'_\lambda(0)$  mean boundary surface area per unit volume ( $S_V$ )

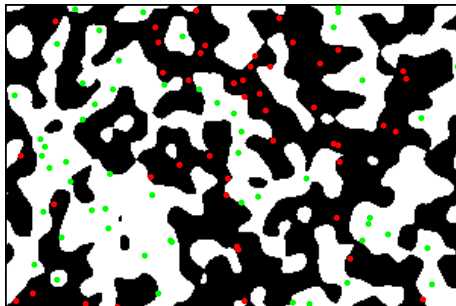
$$\frac{1}{2}N_L = \frac{1}{\pi}L_A = \frac{1}{4}S_V$$

# Outline

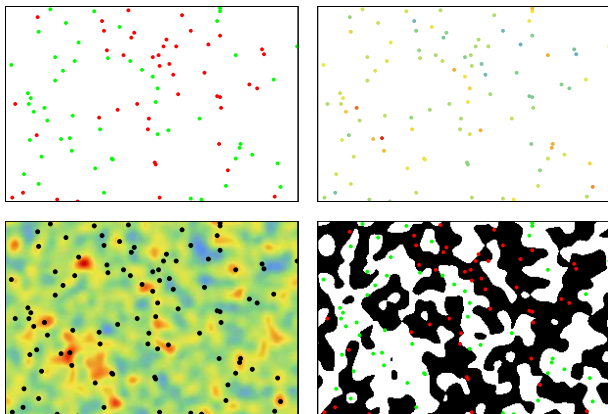
- 1 Statistical characterization of random sets
- 2 Gaussian excursion sets
  - Definition
  - **Conditional simulation**
  - Plurigaussian random functions
- 3 Boolean model
  - Introduction
  - Definitions and properties
  - Conditional simulation algorithm

## Conditional simulation – Presentation of the problem

How to produce realizations of a gaussian excursion set with covariance  $C$  and level  $\lambda$ , in such a way that each component contains a finite number of prespecified points?



# Conditional simulation algorithm



Conditioning data set (TL). Conditional simulation of  $Y$  at the data points only ([Gibbs sampler](#)) (TR). Conditional simulation of  $Y$  (BL). Threshold at level  $\lambda$  (BR).

# Gibbs sampler

## Problem

Simulate a standardized gaussian random vector  $(Y(x_\alpha), \alpha \in A)$  subject to the condition

$$Y(x_\alpha) \in I_\alpha \quad \alpha \in A$$

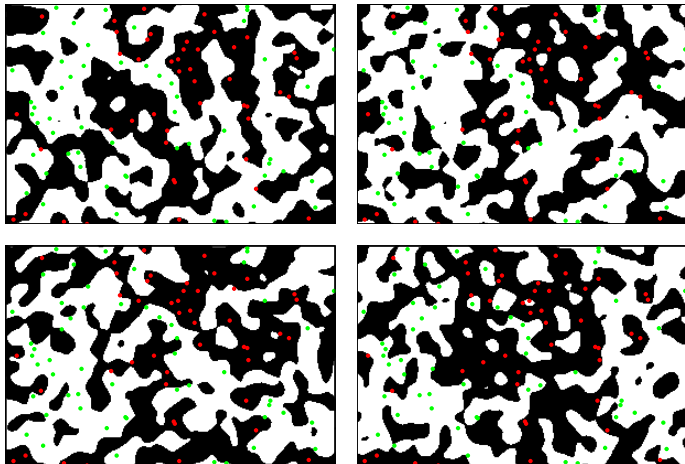
## Algorithm

- (i) *for each*  $\alpha \in A$ , *generate*  $y(x_\alpha) \sim \mathcal{N}(\cdot | I_\alpha)$
- (ii) *select*  $\alpha \sim \mathcal{U}(A)$
- (iii) *generate*  $y(x_\alpha) \sim \mathcal{N}(y_{A^\alpha}(x_\alpha), \sigma_{A^\alpha}^2(x_\alpha))$
- (iv) *if*  $y(x_\alpha) \notin I_\alpha$ , *goto* (iii)
- (v) *goto* (ii)

$y_{A^\alpha}(x_\alpha)$  and  $\sigma_{A^\alpha}^2(x_\alpha)$  respectively stand for the kriging estimate and the kriging variance of point  $x_\alpha$  using  $Y(x_\beta), \beta \neq \alpha$



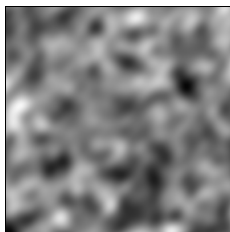
## Four conditional simulations



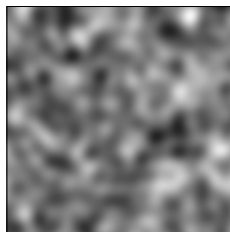
# Plurigaussian random functions

## Basic ingredients

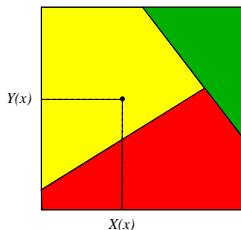
- two (independent) gaussian random functions  $X$  and  $Y$  on  $\mathbb{R}^d$
- a colored partition of  $\mathbb{R}^2$



$X$



$Y$

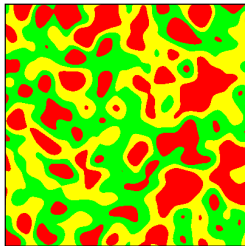
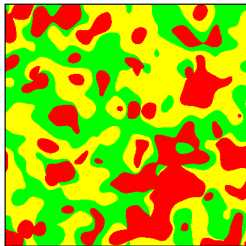
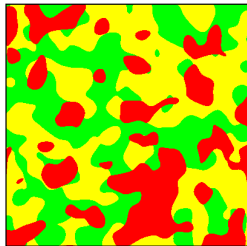
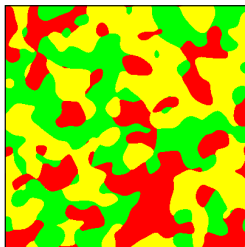
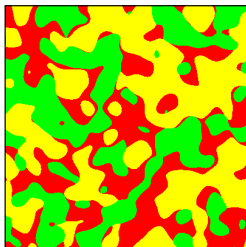
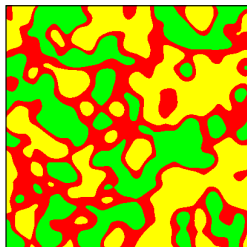


Partition

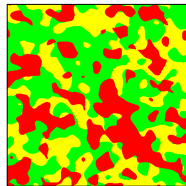
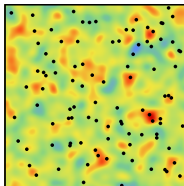
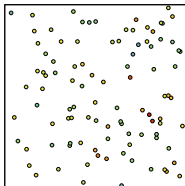
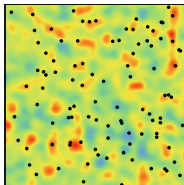
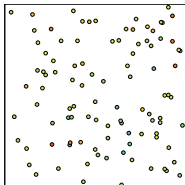
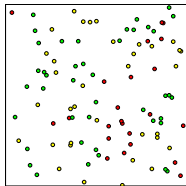
## Definition

The color assigned at each point  $x \in \mathbb{R}^d$  is the color of the cell that contains  $(X(x), Y(x))$

## Examples of plurigaussian random functions



# Conditional simulation algorithm

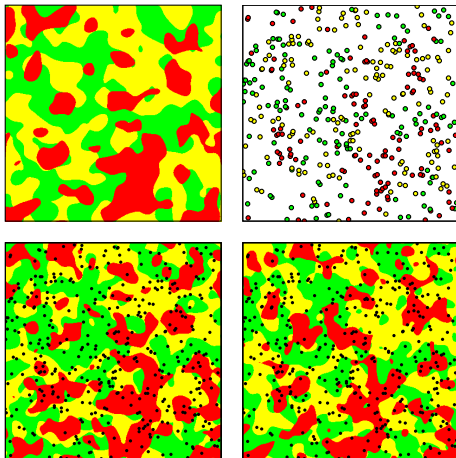


Gibbs sampler

Conditional simulation of  
gaussian random functions

Composition using the flag

## Examples of conditional simulation

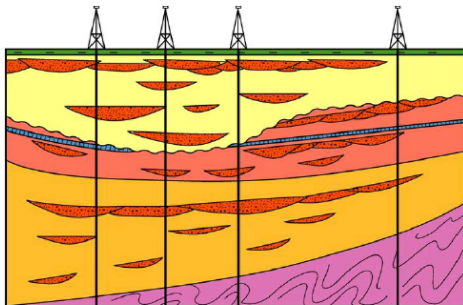


Non conditional simulation (TL). 400 conditional data points (TR). Two conditional simulations (BL and BR)

# Outline

- 1 Statistical characterization of random sets
- 2 Gaussian excursion sets
  - Definition
  - Conditional simulation
  - Plurigaussian random functions
- 3 Boolean model
  - Introduction
  - Definitions and properties
  - Conditional simulation algorithm

# Motivation in Reservoir engineering



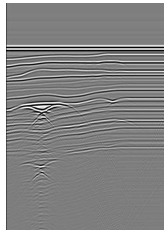
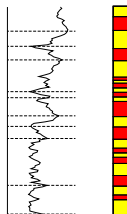
## Case of a fluvial reservoir

- Lenses or channels of **sandstone** are porous and may contain oil
- The **clay** acts as a barrier for oil flow

The reservoir geometry is necessary to perform fluid flow simulations. How to predict it from the available data ?

# Data integration

- Diagraphic interpretation :  
gives facies information along  
each well
- Seismic interpretation :  
gives facies proportion inside the  
reservoir
- Well test :  
give connexity properties

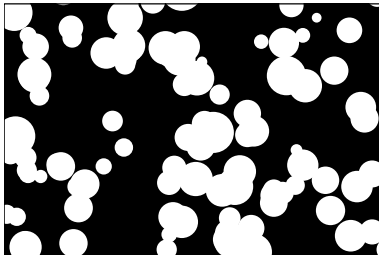
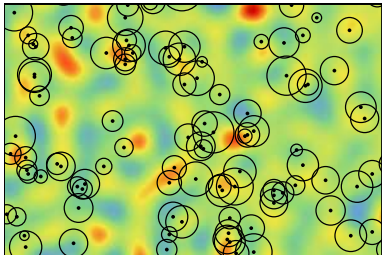
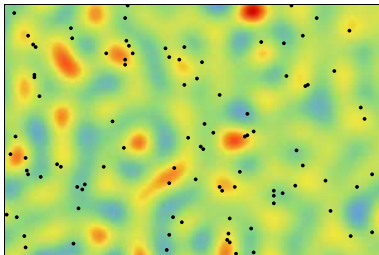
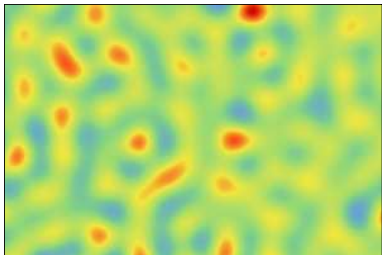




# Outline

- 1 Statistical characterization of random sets
- 2 Gaussian excursion sets
  - Definition
  - Conditional simulation
  - Plurigaussian random functions
- 3 Boolean model
  - Introduction
  - **Definitions and properties**
  - Conditional simulation algorithm

## Construction of a boolean model



# Definition of a boolean model

## Basic ingredients

- 1 A **Poisson point process**  $\mathcal{P}$  with intensity function  $\theta = \left(\theta(x), x \in \mathbb{R}^d\right)$ .
- 2 A family  $\left(A(x), x \in \mathbb{R}^d\right)$  of **non empty, compact, independent random sets** (the **objects**). The statistical properties of each object  $A(x)$  can be specified by its **hitting functional**

$$T_x(K) = P\{A(x) \cap K \neq \emptyset\} \quad K \in \mathcal{K}(\mathbb{R}^d)$$

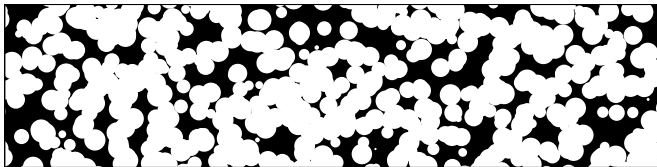
## Definition

A boolean model is the union of the objects situated at the Poisson points

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

$$X = \text{foreground} \quad X^c = \text{background}$$

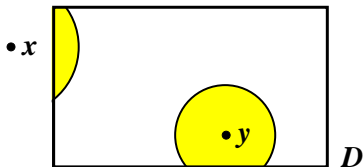
# Examples



# Intersection of a boolean model and a domain

Let  $X$  be a boolean model with parameters  $(\theta, T)$ , and let  $D$  a compact. Then  $X \cap D$  is also a boolean model. Its parameters  $(\theta^{(D)}, T^{(D)})$  are

- **Poisson intensity**  $\theta^{(D)}(x) = \theta(x)T_x(D) \quad x \in \mathbb{R}^d$
- **Hitting functional**  $T_x^{(D)}(K) = \frac{T_x(K)}{T_x(D)} \quad K \in \mathcal{K}(D)$



# Typical objects

## Hypothesis

The Poisson intensity  $\theta^{(D)}$  has a **finite** integral, denoted by  $\vartheta(D)$

## Consequences

- $X \cap D$  is made of a finite number (Poissonian) of objects
- the function  $x \longrightarrow \theta^{(D)}(x)/\vartheta(D)$  is a pdf

## Definition

Let  $\dot{x} \sim \theta^{(D)}/\vartheta(D)$ . The compact  $A(\dot{x})$  is called a **typical object**. Its pdf is

$$T^{(D)}(K) = \int_{\mathbb{R}^d} \frac{\theta^{(D)}(x)}{\vartheta(D)} \frac{T_x(K)}{T_x(D)} dx = \frac{\int_{\mathbb{R}^d} \theta(x) T_x(K) dx}{\int_{\mathbb{R}^d} \theta(x) T_x(D) dx}$$

## Property

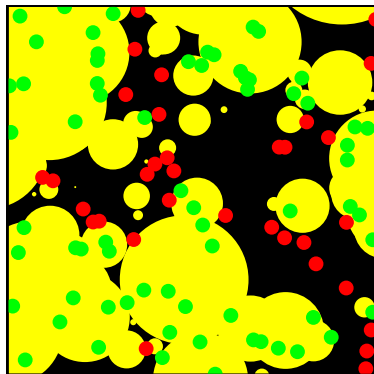
$X \cap D$  is the union of a **Poisson** number (of mean  $\vartheta(D)$ ) of **independent** typical objects

# Outline

- 1 Statistical characterization of random sets
- 2 Gaussian excursion sets
  - Definition
  - Conditional simulation
  - Plurigaussian random functions
- 3 Boolean model
  - Introduction
  - Definitions and properties
  - Conditional simulation algorithm

## Problem

How to produce realizations of  $X$  with the simulation domain  $D$  such that two finite sets of points must belong to  $X$  and  $X^c$  ?



- $\in X$
- $\in X^c$

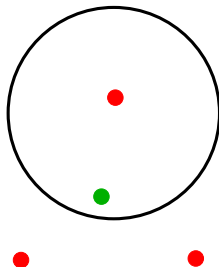


## Remarks on compatibility

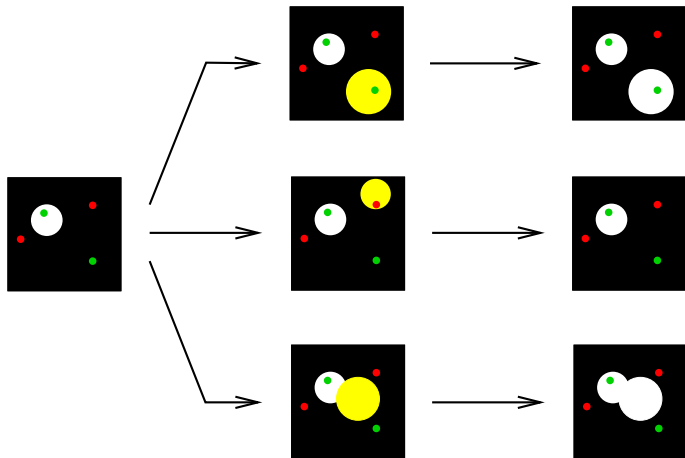
Contrarily to the Gaussian random functions, a boolean model **can be not** compatible with the data

### Example

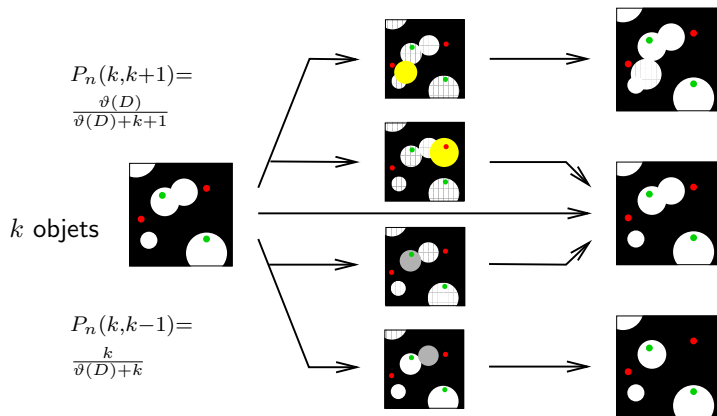
Consider a boolean model made of circular objects with fixed radius and suppose that a data point is surrounded by background points:



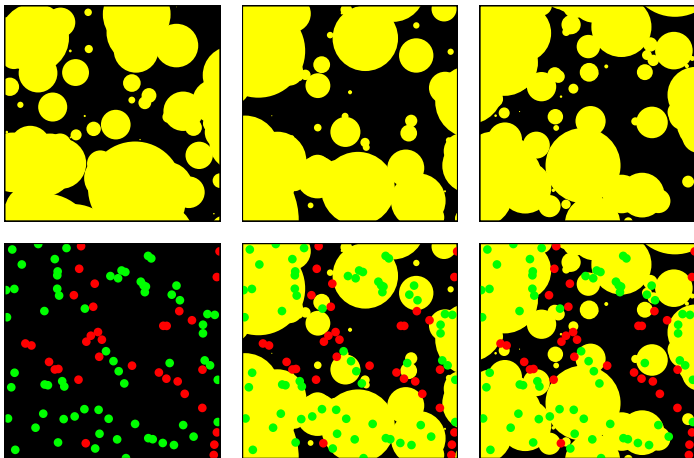
# Initialization



# One iteration of conditional simulation

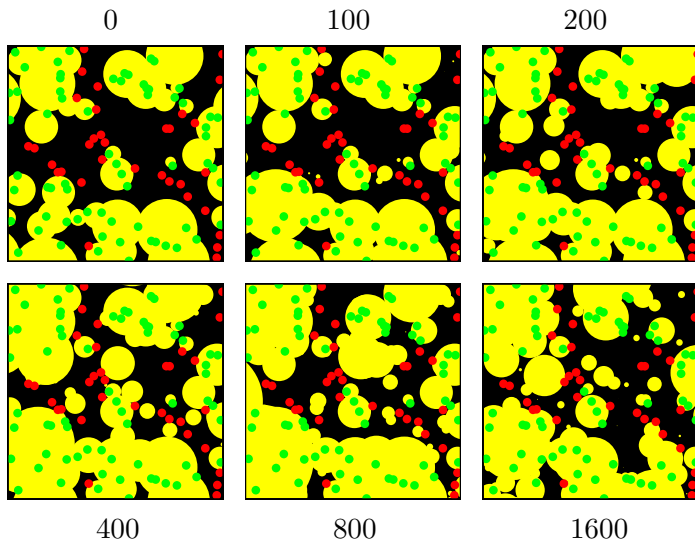


## Example of conditional simulation



Stationary collean model of disks (Poisson intensity 0.00766; the radii follows an exponential distribution with mean 5). Simulation domain is  $100 \times 100$ .

## Different iterations



Conditional simulation at different iterations

# Formal presentation of the problem

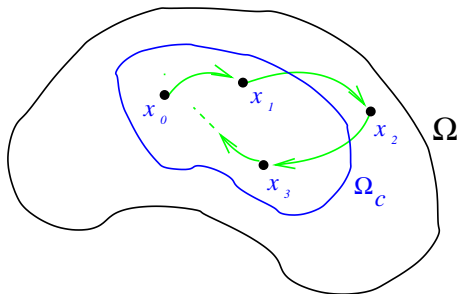
$\Omega$  Set of all possible states

$p$  probability on  $\Omega$

$\Omega_c$  set of **authorized** states

$p_c$  **conditional** probability induced by  $p$  on  $\Omega_c$

$$p_c(x) = \frac{p(x)}{p(\Omega_c)} \quad x \in \Omega_c$$

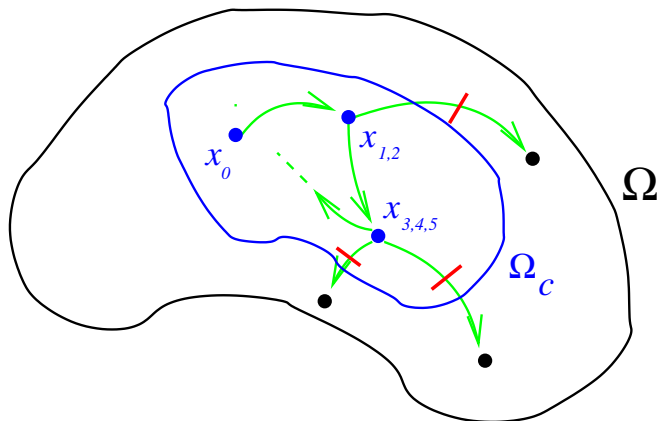


Suppose that we can simulate  $p$  as the limit distribution of a Markov chain with **known** kernel  $P$ :

$$\lim_{n \rightarrow \infty} P^n(x, y) = p(y) \quad x, y \in \Omega$$

How to simulate from  $p_c$ ?

## Idea : Restriction of a Markov kernel

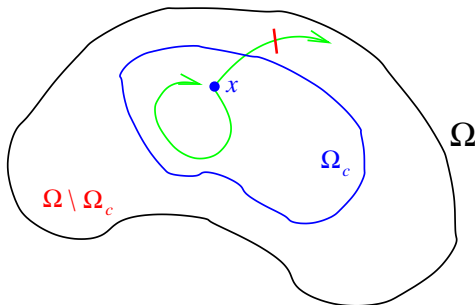


**Principle** Run  $P$  and forbid to exit from the space of authorized states

## Restriction of a Markov kernel

The generated sequence of random states behaves like a Markov chain on  $\Omega_c$  with transition kernel

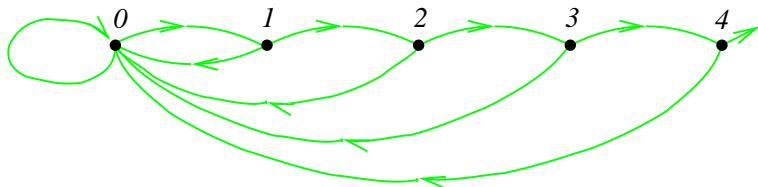
$$P_c(x, y) = \begin{cases} P(x, y) & \text{if } y \neq x \\ P(x, x) + P(x, \Omega \setminus \Omega_c) & \text{if } y = x \end{cases}$$



Do we have  $\lim_{n \rightarrow \infty} P_c^n(x, y) = p_c(y) \quad \forall x, y \in \Omega_c$ ?



## Example

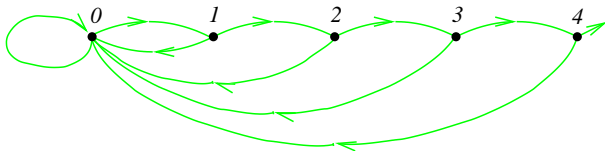


$$P(x, x+1) = P(x, 0) = \frac{1}{2}$$

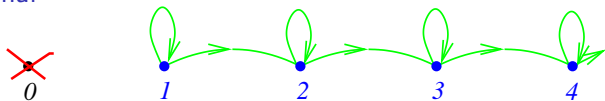
The limit distribution is  $p(x) = \frac{1}{2^{x+1}} \quad x \in \mathbb{N}$

Example: case  $\Omega_c = \{1, 2, 3, \dots\}$

Non conditional



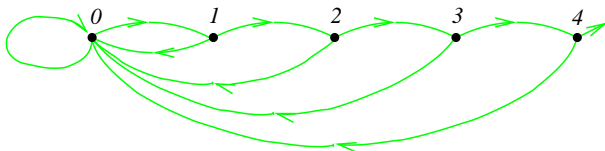
Conditional



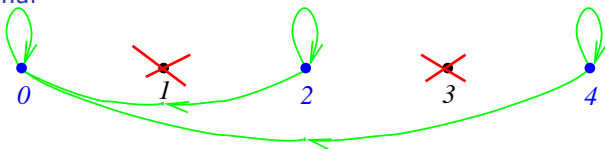
$$\lim_{n \rightarrow \infty} P_c^n(x, y) = 0 \quad x, y \in \Omega_c$$

Example : case  $\Omega_c = \{0, 2, 4, \dots\}$

Non conditional



Conditional



$$\lim_{n \rightarrow \infty} P_c^n(x, y) = 1_{y=0} \quad x, y \in \Omega_c$$

## Restriction of a Markov kernel

In general  $\lim_{n \rightarrow \infty} P_c^n(x, y) \neq p_c(y)$

However

- If  $P_c$  is irreducible, then  $P_c$  is automatically aperiodic
- If  $P$  est reversible, then  $P_c$  also is reversible, which implies that  $p_c$  is  $P_c$ -invariant

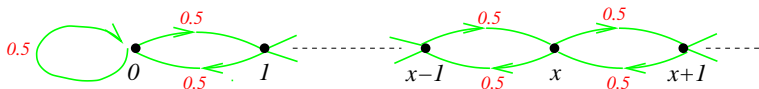
**Conclusion** If  $P$  is **reversible** and if  $P_c$  is **irreducible**, then

$$\forall x, y \in \Omega_c \quad \lim_{n \rightarrow \infty} P_c^n(x, y) = p_c(y)$$

# Simulation of a Poisson distribution

$$p(x) = \exp\{-\vartheta\} \frac{\vartheta^x}{x!} \quad x \in \mathbb{N}$$

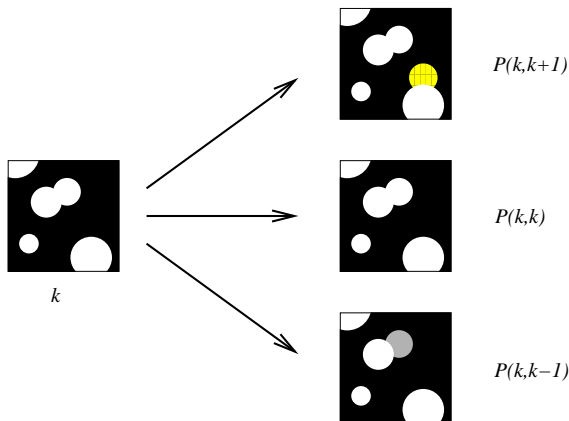
Choice of  $Q$



Transition kernel  $P$

$$P(x, y) = \begin{cases} \frac{x}{\vartheta + x} & \text{if } y = x - 1 \\ \frac{\vartheta}{(\vartheta + x)(\vartheta + x + 1)} & \text{if } y = x \\ \frac{\vartheta}{\vartheta + x + 1} & \text{if } y = x + 1 \end{cases}$$

# An iteration of non conditional simulation

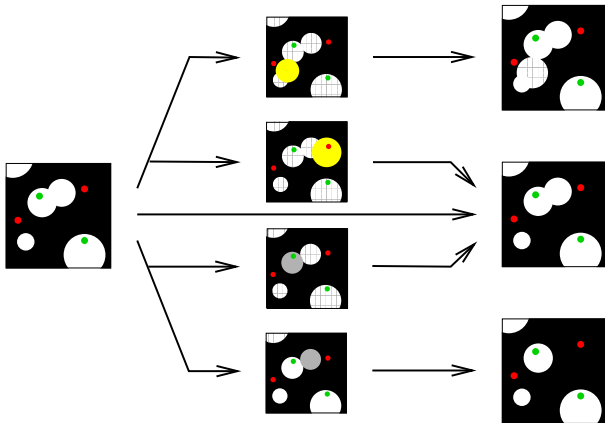


$$P(k, k-1) = \frac{k}{\vartheta + k}$$

$$P(k, k) = \frac{\vartheta}{(\vartheta + k)(\vartheta + k + 1)}$$

$$P(k, k+1) = \frac{\vartheta}{\vartheta + k + 1}$$

## An iteration of conditional simulation



# References

- Armstrong M., Galli A., Le Loch G., Geffroy F. and Eschard R. (2003) - Plurigaussian simulations in geosciences. Springer (Berlin).
- Freulon X. (1992) - Conditionnement du modèle gaussien par des inégalités ou des randomisées. Thèse de doctorat, Ecole des Mines de Paris.
- Galli A. et Gao H. (2001) - "Rate of convergence of the Gibbs sampler in the gaussian case". J. of Math. Geol., vol. 33, pp. 653-677.
- Matheron G. (1975) - Random sets and integral geometry. Wiley, New York.
- Lantuéjoul C. (2001) - Geostatistical simulation: models and algorithms. Springer, Berlin.