



## Introduction aux Champs de Markov

Florence Tupin  
Télécom ParisTech

## Plan du cours - première partie

- Introduction
- Voisinage et cliques
- Définition d'un champ de Markov
- Théorème d'Hammersley-Clifford et probabilités conditionnelles locales
- Echantillonneurs (Gibbs, Metropolis)
- Optimisation

# Introduction

- **Historique**

- Physique statistique (organisation des cristaux)
- Article de Geman et Geman (84)

- **Idée fondamentale des champs de Markov**

introduire des relations contextuelles en traitement d'images

un voisinage local suffit pour des images naturelles

## A priori dans les images naturelles : le contexte spatial



# Introduction

- **utilisation en imagerie bas niveau**
  - Restauration
  - Segmentation
  - Détection de contours
  - Compression
  - Détection du mouvement
- **utilisation en imagerie haut niveau**
  - Reconnaissance d'objets
  - Mise en correspondance de graphes

# Champs de Markov

- **modèle probabiliste de l'image**

$S = \{s\} \subset \mathbf{Z}^d$  ensemble de sites (fini)

$x_s \in E$  espace des niveaux de gris

( $E = \{0..255\} \{0..q - 1\}$  (système de labels)  $\mathbf{R}$ )

$X_s$  variable aléatoire associée à  $s$

$X = \{X_s\}_{s \in S}$  champ aléatoire

$x = \{x_s\}_{s \in S} = \{x_s\} \cup x^s$  configuration (image)

$\Omega = E^{|S|}$  espace des configurations

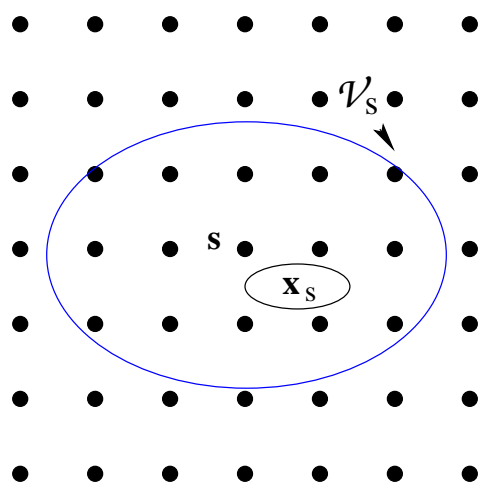
- **probabilités :**

$P(X_s = x_s)$  probabilité locale

$P(X = x) = P(X_1 = x_1, X_2 = x_2 \dots X_s = x_s \dots)$  loi globale (jointe)

$\Pr(X_s = x_s / X_t = x_t, t \neq s)$  probabilité conditionnelle (locale)

# Contexte spatial dans les images naturelles



$s$  : site

$\mathcal{V}_s$  : voisinage (spatial) de  $s$

- régions homogènes en niveaux de gris

$x_s \leftrightarrow$  Intensités aux sites voisins

Moyenne

- régions texturées

$x_s \leftrightarrow$  Intensités aux sites voisins (!!)

Contraste, Entropie

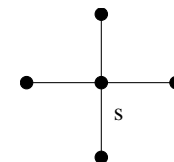
réalisation globale de l'image	$\Leftrightarrow$	voisinage local
probabilité globale	$\Leftrightarrow$	probabilités locales [ conditionnelles ]

# Topologie pour les champs de Markov

## ○ système de voisinage - définition

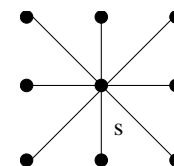
voisinage du site  $s$  :  $\mathcal{V}_s$

propriétés :  $s \notin \mathcal{V}_s$   $s \in \mathcal{V}_r \Leftrightarrow r \in \mathcal{V}_s$



$\mathcal{V} = \{\mathcal{V}_s\}_{s \in S}$  système de voisinage

$x \rightarrow V_s = \{x_r\}_{r \in \mathcal{V}_s}$  configuration locale



## ○ cliques

$c \subset S$  est une clique /  $\mathcal{V}$  ssi :

– card  $(c) = 1$  (singleton)

– card  $(c) \geq 2$  and  $\forall r \neq s \in c \Rightarrow r, s$  voisins

○ notations  $c = (r, s, t, \dots)$ ;  $\mathcal{C} = \{c\}$



# Topologie pour les champs de Markov (suite)

- 4-connexité

4-connexité



ordre 1



ordre 2

- 8-connexité

8-connexité



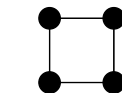
ordre 1



ordre 2

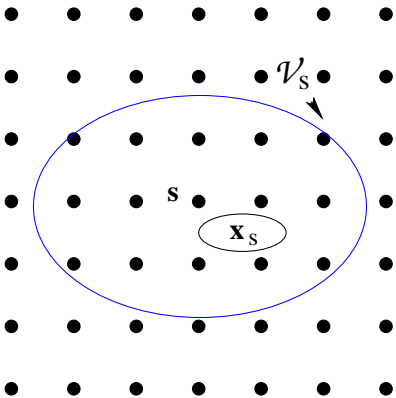


ordre 3



ordre 4

# Champ de Markov : définition



$s$  : site

$\mathcal{V}_s$  : voisinage (spatial) de  $s$

$$\begin{aligned} \Pr(X_s = x_s / \{X_r = x_r\}, r \neq s) &= \Pr(X_s = x_s / \{X_r = x_r\}, r \in \mathcal{V}_s) \\ &= \Pr(X_s = x_s / V_s) \end{aligned}$$

Probabilité

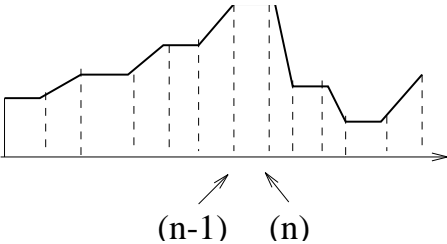
Globale

$\leftrightarrow$

Probabilité

Locale

- extension de la notion de chaîne de Markov



## Théorème d'Hammersley-Clifford :

$P(X = x) > 0 \quad \forall x \in \Omega$  est un champ de Markov ssi

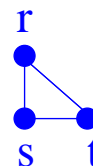
$$P(X = x) = \frac{\exp -U(x)}{Z} \quad \text{distribution de Gibbs}$$

$$U(x) = \sum_{c \in \mathcal{C}} U_c(x) \quad \text{énergie globale}$$

$$U_c(x) = U_c(x_s, s \in c) \quad \text{potentiel de cliques}$$

$$Z = \sum_{y \in \Omega} \exp -U(y) \quad \text{fonction de partition}$$

○ **exemple : cliques**



$$U(x) = A \sum_{s \in S} x_s + B \sum_{(r,s)} x_r x_s + C \sum_{(r,s,t)} x_r x_s x_t$$

non-stationnarité possible :  $A \rightarrow A_s$  ,  $B \rightarrow B_{rs}$  ...

○ fait important : énergie  $U(x)$  basse  $\Leftrightarrow$  probabilité  $P(X = x)$  forte

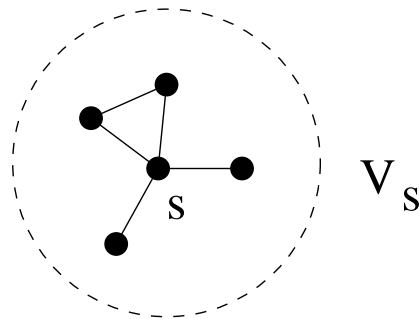
# Forme des probabilités conditionnelles locales

$$\Pr(X_s = x_s / V_s) = \frac{1}{Z^s} \exp -U(x_s / V_s)$$

avec :

$$U(x_s / V_s) = \sum_{c \subset C, s \in c} U_c(x) \quad \text{énergie conditionnelle locale}$$

$$Z^s = \sum_{\xi \in E} \exp -U(\xi / V_s) \quad \text{fonction de partition locale}$$



○ ⇒ forme locale de la loi de Gibbs

## Probabilités conditionnelles locales (suite)

### ○ démonstration

$$\begin{aligned} \Pr(X_s = x_s / X^s = x^s) &= \frac{\Pr(X_s = x_s, X^s = x^s)}{\Pr(X^s = x^s)} \\ &= \frac{\Pr(X_s = x_s, X^s = x^s)}{\sum_{\xi \in E} \Pr(X_s = \xi, X^s = x^s)} = \frac{\Pr(X = x)}{\sum_{\xi \in E} \Pr(X = x')} \end{aligned}$$

○ → maintenant posons  $U(x) = U(x_s / V_s) + \sum_{c \in C, s \notin c} U_c(x)$

↙  $\mathcal{W}$

$$\Pr(X_s = x_s / X^s = x^s) = \Pr(X_s = x_s / V_s) = \frac{\exp - U(x_s / V_s)}{\sum_{\xi \in E} \exp - U(\xi / V_s)} \quad \swarrow Z^s$$

# Probabilités conditionnelles locales : exemple

champ binaire ( $E = \{0, 1\}$ )

- **Voisinage** 4-connexité

- **Potentiels**

cliques d'ordre 2

$$V_{c=(s,t)}(x_s, x_t) = -\beta 1_{x_s=x_t}$$

- **Probabilités conditionnelles locales**

$$U(x_s = 0, V_s = (0, 0, 0, 1)) = -3\beta$$

$$U(x_s = 1, V_s = ((0, 0, 0, 1)) = -\beta$$

$$P(X_s = 0|V_s) = \frac{\exp(3\beta)}{\exp(\beta) + \exp(3\beta)}$$

$$P(X_s = 1|V_s) = \frac{\exp(\beta)}{\exp(\beta) + \exp(3\beta)}$$

# Echantillonnage de champs de Markov

- **problème**

$X$  ( $P(X=x)$ ) étant définis (système de voisinage, potentiels des cliques)

comment tirer une réalisation selon la loi  $P(X)$  ?

- **solutions**

deux algorithmes possibles :

- échantillonneur de Gibbs
- algorithme de Metropolis

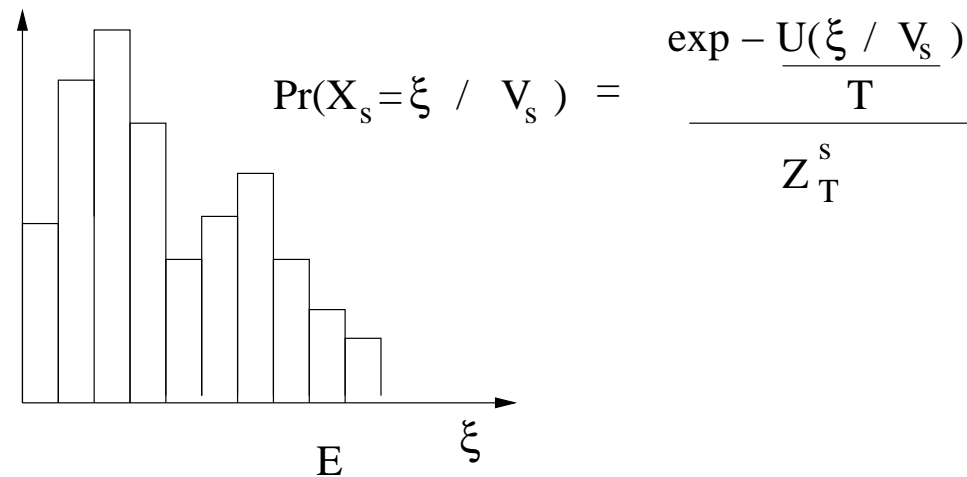
## Exemple : échantillonneur de Gibbs

- **principe** construction d'une séquence de configurations  $x(n)$  en visitant chaque site (tiré aléatoirement)

- **changement local** :  $x = \{x_s, x^s\} \rightarrow x' = \{\xi, x^s\}$

$$\frac{\Pr(X = x')}{\Pr(X = x)} = \frac{\Pr(X_s = \xi / X^s = x^s) \Pr(X^s = x^s)}{\Pr(X_s = x_s / X^s = x^s) \Pr(X^s = x^s)} = \frac{\Pr(X_s = \xi / V_s)}{\Pr(X_s = x_s / V_s)}$$

- $\Rightarrow$  **tirage selon la loi conditionnelle locale**





## Échantillonneur de Metropolis

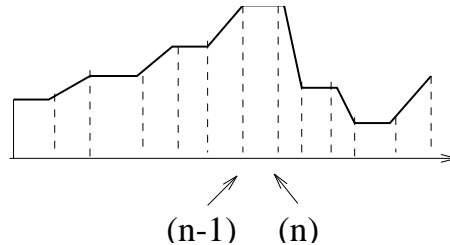
- **principe** tirage d'un site  $s$  et d'un nouvel état du site  $x'_s$

calcul de la variation énergétique entre la configuration  $x = (x_s, x_t, t \neq s)$  et  $x' = (x'_s, x_t, t \neq s)$

- si  $\Delta U = U(x') - U(x) < 0$  le changement de label du site est accepté
- sinon, le changement est accepté ou rejeté en tirant selon la probabilité  $p = \exp(-\Delta U), 1 - p$

# Échantillonnage d'un champ de Markov

- chaîne de Markov

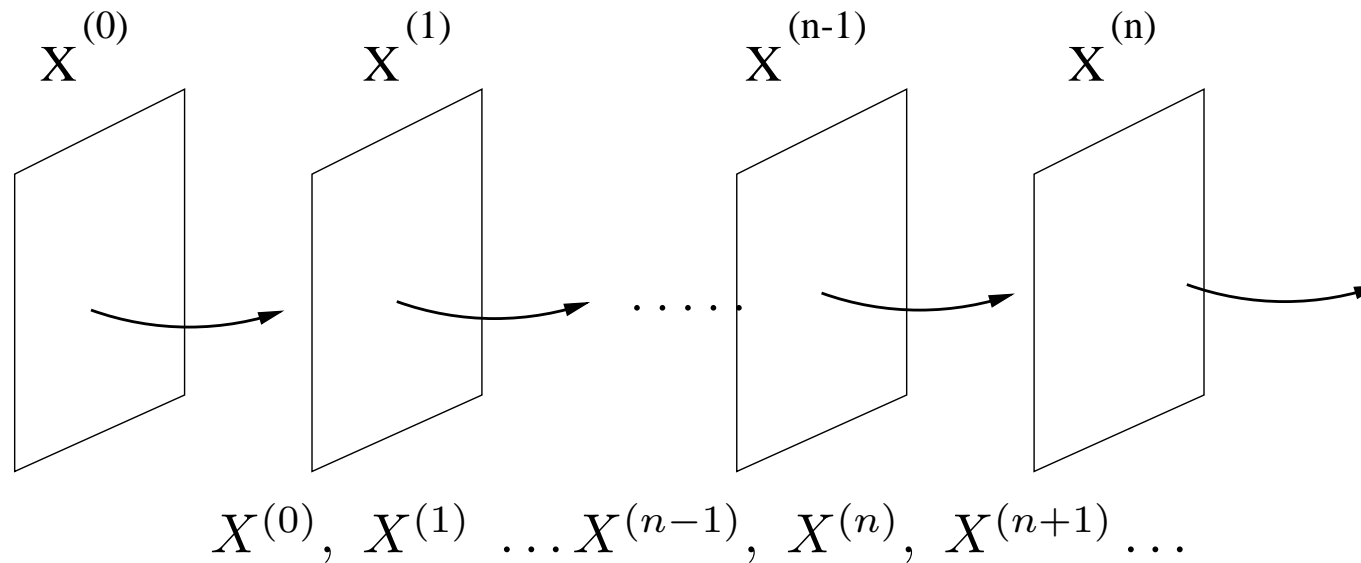


$$\begin{aligned} \Pr(X^{(n)} = x^{(n)} / X^{(0)} = x^{(0)}, X^{(1)} = x^{(1)} \dots X^{(n-1)} = x^{(n-1)}) \\ = \Pr(X^{(n)} = x^{(n)} / X^{(n-1)} = x^{(n-1)}) \end{aligned}$$

- noyau de transition

$$Q_n(x, y) = \Pr(X^{(n)} = y / X^{(n-1)} = x) \quad x \rightarrow y$$

- chaîne de Markov homogène :  $Q_n(x, y)$  indépendant de  $n$
- Echantillonnage : chaîne de Markov de variables d'images !



- échantillonnage homogène : trouver  $Q(x, y)$  tel que

$$\lim_{n \rightarrow +\infty} \Pr(X^{(n)} = x) = P(X = x)$$

# Champs de Markov prototypes

- **modèle d'Ising**

$$U(x) = -\beta \sum_{c=(s,t)} x_s x_t - B \sum_{s \in S} x_s \quad E = \{-1, +1\}$$

- **modèle de Potts**

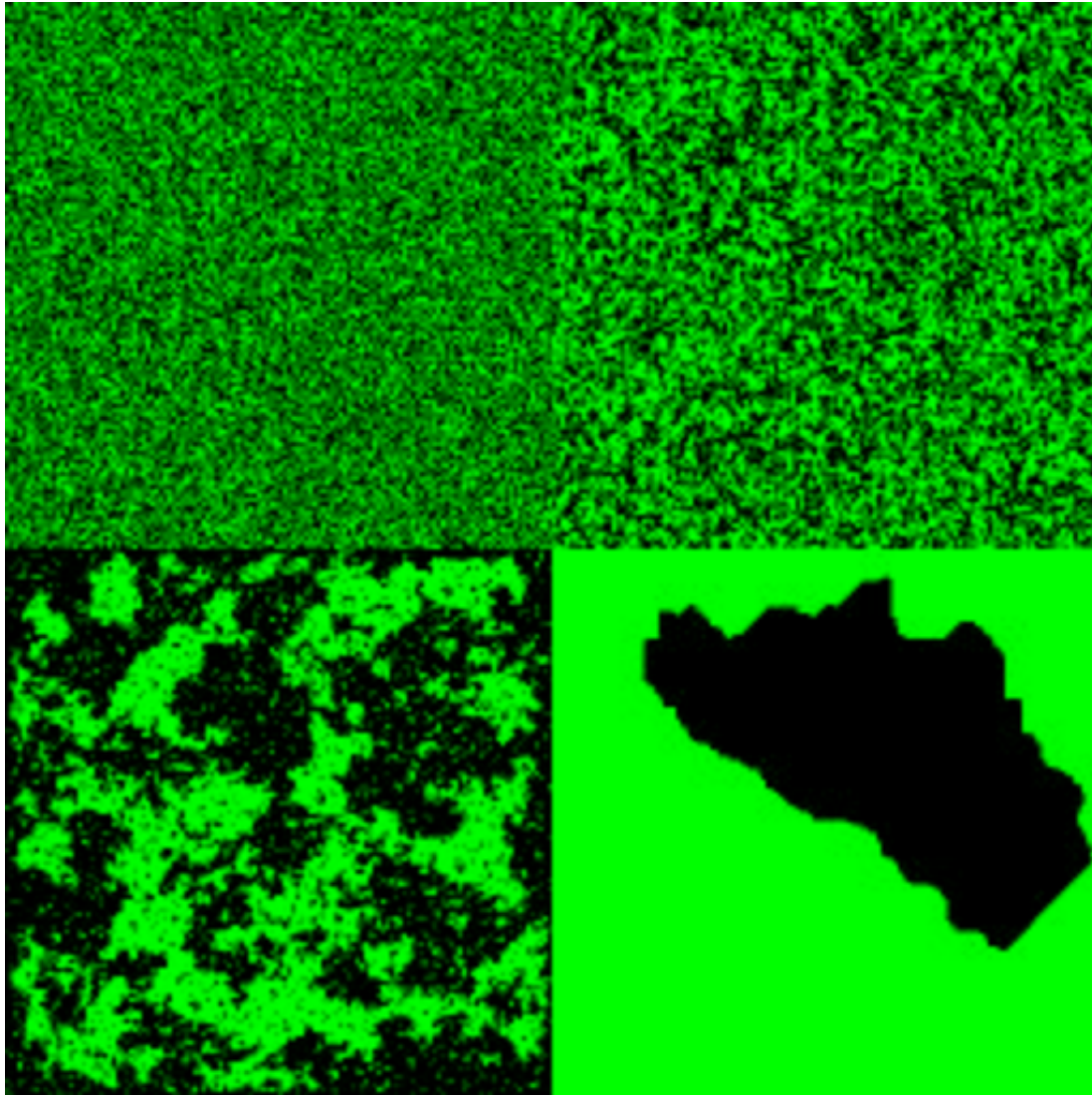
$$\text{définissons } \Delta(a, b) = \begin{cases} +1 & \text{si } a = b \\ -1 & \text{si } a \neq b \end{cases}$$

$$U(x) = -\beta \sum_{c=(s,t)} \Delta(x_s, x_t) - B \sum_{s \in S} \Delta(x_s, 0) \quad E = \{0..q - 1\}$$

- **modèles markoviens gaussiens**

$$U(x) = \beta \sum_{c=(s,t)} (x_s - x_t)^2 + \alpha \sum_{s \in S} (x_s - \mu_s)^2 \quad E = \mathbf{R}$$

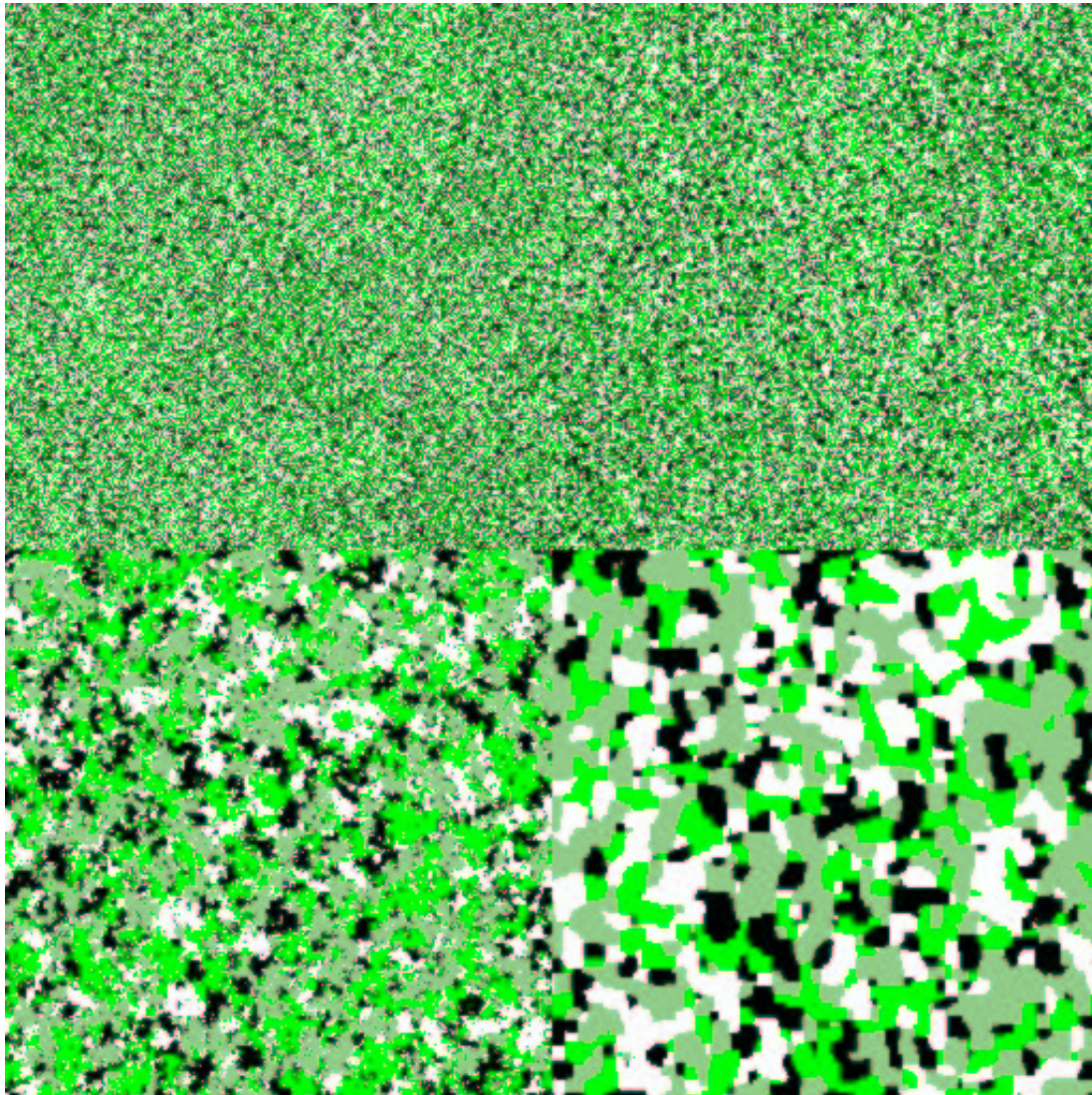
# Modèle d'Ising en 4-connexité et $B = 0$



A	B
C	D

- A : *Image aléatoire* :  $\beta = 0$    - B : *régularisation faible* :  $\beta = 0.2$
- C : *régularisation "critique"* :  $\beta \approx 0.44$    - D : *régularisation forte* :  $\beta = 4.0$

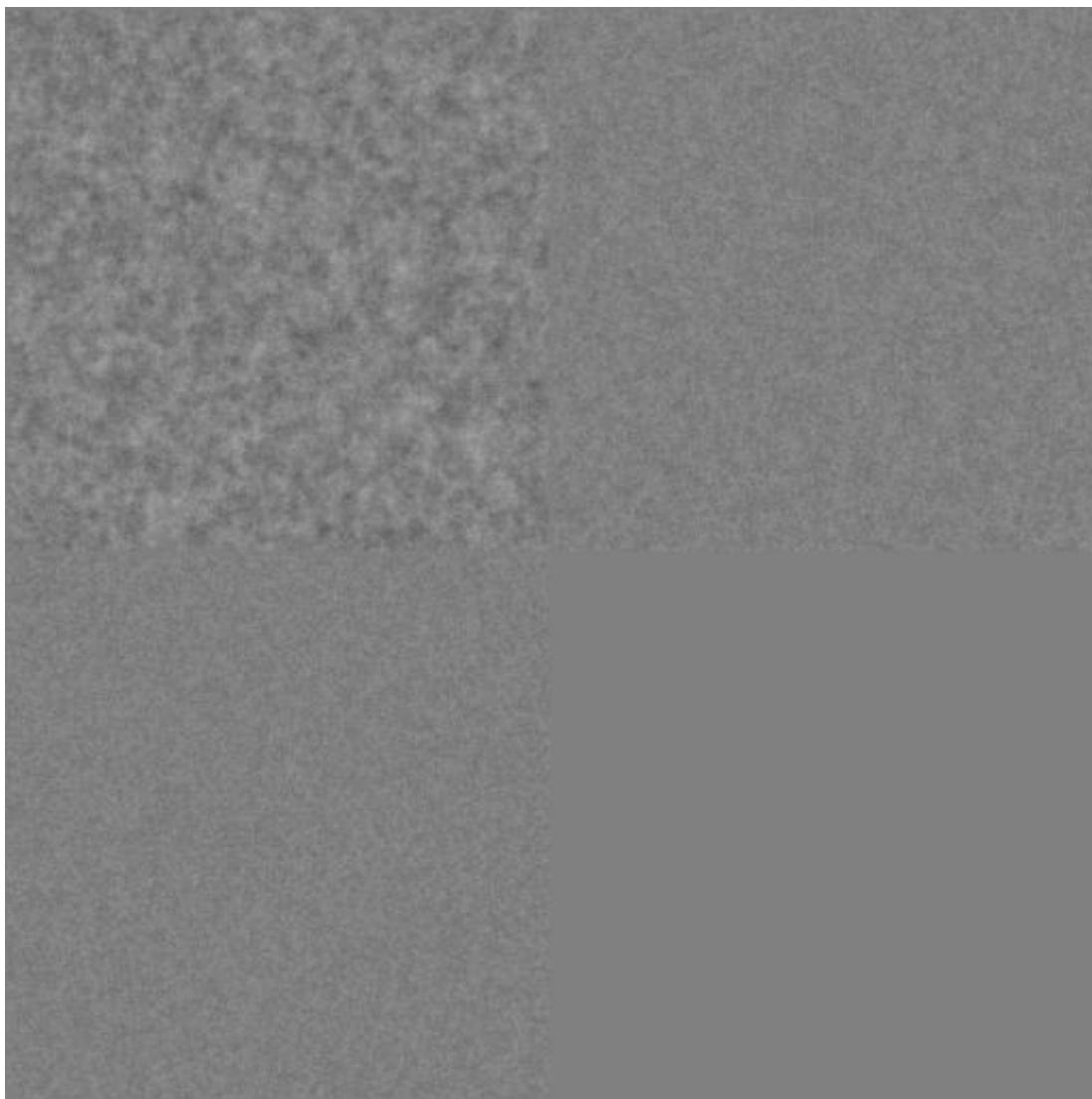
# Modèle de Potts en 4-connexité et $B = 0$ ( $q = 4$ )



A	B
C	D

- A : *Image aléatoire* :  $\beta = 0$    - B : *régularisation faible* :  $\beta = 0.2$
- C : *régularisation "critique"* :  $\beta \approx 1,099$    - D : *régularisation forte* :  $\beta = 4.0$

# Modèle markovien gaussien en 4-connexité



$$U(x) = \beta \sum_{c=(s,t)} (x_s - x_t)^2 + \alpha \sum_{s \in S} (x_s - \mu_s)^2$$

A	B
C	D

- A :  $\alpha = 5.10^{-4}$    - B :  $\alpha = 5.10^{-3}$
- C :  $\alpha = 2.10^{-3}$    - D :  $\alpha = \infty$  ( $\mu = 127$  pour toutes simulations)

# Modèles markoviens gaussiens

- modèle à pixels indépendants

$$\Pr(X = x) = \left[ \sqrt{\frac{\alpha}{\pi}} \right]^{|S|} \prod_{s \in S} e^{-\alpha (x_s - \mu)^2} \Leftrightarrow \frac{\exp -U(x)}{Z}$$

$$\text{avec } U(x) = \alpha \sum_{s \in S} (x_s - \mu)^2$$

- cas général - modèle auto-normal

$$U(x) = \alpha \sum_{s \in S} (x_s - \mu_s)^2 + \beta \sum_{c=(s,t)} (x_s - x_t)^2$$

↓

moyenne locale

- illumination variable  $\mu_s$

- illumination constante  $\mu_s = \mu = 128$

↓

couplage



## Loi conditionnelle locale en connexité $\gamma$ : gaussienne

$$\Pr(X_s = x_s / V_s) = \frac{1}{z} \exp - \left[ \alpha(x_s - \mu_s)^2 + \beta \sum_{c=(s,t), t \in \mathcal{V}_s} (x_s - x_t)^2 \right]$$

$$= \sqrt{\frac{2(\alpha + \beta \gamma)}{\pi}} \cdot \exp - (\alpha + \beta \gamma) \left[ x_s - \left( \frac{\alpha \mu_s + \beta \sum_{t \in \mathcal{V}_s} x_t}{\alpha + \beta \gamma} \right) \right]^2$$

### ○ espérance conditionnelle

$$\mathbf{E}[X_s / V_s] = \frac{\alpha \mu_s + \beta \sum_{t \in \mathcal{V}_s} x_t}{\alpha + \beta \gamma} = \frac{\alpha \mu + \beta \sum_{t \in \mathcal{V}_s} x_t}{\alpha + \beta \gamma} \rightarrow \text{barycentre } (\mu_s = \mu)$$

### ○ variance conditionnelle

$$\text{var}(X_s / V_s) = \frac{1}{2(\alpha + \beta \gamma)} \rightarrow \text{indépendante de } \mu_s \text{ et de } x_t, t \in \mathcal{V}_s$$

○  $\Rightarrow$  permet un ensemble de statistiques à variable  $X_{\mathbf{V}} = \sum_{t \in \mathcal{V}_s} x_t$  fixée

## Distribution de Gibbs avec température

$$P_T(X = x) = \frac{1}{Z_T} \exp -\frac{U(x)}{T}$$

$$U(x) = \sum_{c \in \mathcal{C}} U_c(x)$$

énergie globale

$$Z_T = \sum_{y \in \Omega} \exp -\frac{U(y)}{T}$$

fonction de partition

# Comportement aux températures limites

## ◦ intuition

$$\frac{P_T(X = y)}{P_T(X = x)} = \exp - \frac{[U(y) - U(x)]}{T} \quad \forall x, y \in \Omega$$

$$T \rightarrow \infty \quad \exp - \frac{[U(y) - U(x)]}{T} \rightarrow 1 \quad \forall x, y \in \Omega \text{ fini}$$

$$T \rightarrow 0 \quad \exp - \frac{[U(y) - U(x)]}{T} \rightarrow 0 \quad \text{si } U(y) > U(x)$$

## ◦ démonstration pour $T \rightarrow \infty$

$$\begin{aligned} P_T(X = x) &= \frac{\exp - \frac{U(x)}{T}}{\sum_{y \in \Omega} \exp - \frac{U(y)}{T}} = \frac{1}{\sum_{y \in \Omega} \exp - \frac{[U(y) - U(x)]}{T}} \\ &\rightarrow \frac{1}{\text{Card } \Omega} \quad \forall x \in \Omega \quad \text{équidistribution sur } \Omega \end{aligned}$$

## Distribution de Gibbs avec température(suite)

### ◦ démonstration pour $T \rightarrow 0$

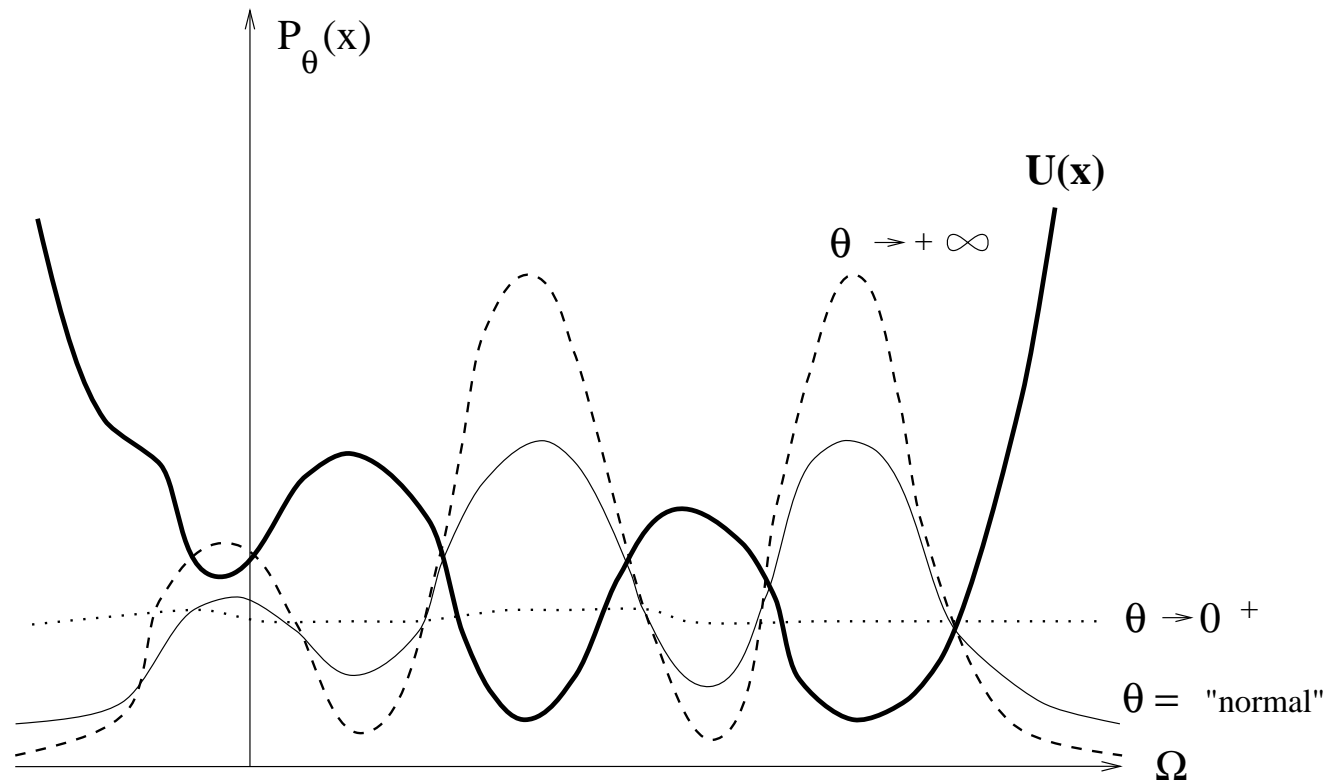
$$U^* = \min_{x \in \Omega} U(x) \quad \Omega^* = \{x \in \Omega \mid U(x) = U^*\}$$

$$P_T(X = x) = \frac{\exp - \frac{[U(x) - U^*]}{T}}{\sum_{y \in \Omega} \exp - \frac{[U(y) - U^*]}{T}} = \frac{\exp - \frac{[U(x) - U^*]}{T}}{\text{Card } \Omega^* + \sum_{y \in \Omega, y \notin \Omega^*} \exp - \frac{[U(y) - U^*]}{T}}$$

$\rightarrow \begin{cases} \frac{1}{\text{Card } \Omega^*} & \text{si } x \in \Omega^* \\ 0 & \text{si } x \notin \Omega^* \end{cases}$       équidistribution sur  $\Omega^*$

(Rappel :  $\exp - \frac{[U(y) - U(x)]}{T} \rightarrow 0$  si  $U(y) > U(x)$  )

## Distribution de Gibbs avec température(suite)



## Le recuit simulé

- **théorème (Geman and Geman 1984)**

- construction d'une séquence d'images par échantillonneurs pour  $P_{T(n)}(X)$  avec  $T(n)$  décroissant lentement et en prenant pour initialisation des échantillonneurs la configuration courante
- La configuration obtenue lorsque la température est proche de 0 est un minimum global de l'énergie
- Conditions : la décroissance de la température doit être très lente (logarithmique) et la température initiale doit être suffisamment élevée pour franchir les barrières énergétiques du paysage

# Le recuit simulé

- **théorème (Geman and Geman 1984)**

$$\text{si } Q_n(x, y) \text{ avec } T_n \searrow 0, T_n \geq \frac{T_0}{\log(1+n)}$$

$$\text{alors } \lim_{n \rightarrow +\infty} \Pr(X^{(n)} = x) = \frac{1}{|\Omega^*|} \delta(x \in \Omega^*) \leftarrow \text{minimum global de l'énergie}$$

construction d'une séquence d'images par échantillonneurs pour  $P_{T(n)}(X)$

avec  $T(n)$  qui suit une décroissance logarithmique

- **condition théorique**

$$T_0 = \Delta U_{max} \text{ Metropolis} \quad \text{---} \quad T_0 = \sum_{s \in S} \delta U(\cdot / V_s)_{max} \text{ Gibbs}$$

- **en pratique :  $T_n = T_0 \alpha^n$  avec :**

$$T_0 \approx \delta U(\cdot / V_s)_{max}, \alpha \approx 0.98$$

# Recuit simulé

