IV. Déconvolution d'images, modèles markoviens, algorithmes stochastiques

Modélisation markovienne en imagerie

Résumé des épisodes précédents

- Règle de Bayes $\quad f(\boldsymbol{x} \,|\, \boldsymbol{z}) \propto f(\boldsymbol{z} \,|\, \boldsymbol{x}) f(\boldsymbol{x}), \quad f(\boldsymbol{x}) \propto \exp\{-\Omega(\boldsymbol{x})\}$

- Loi *a priori* gaussienne = régularisation quadratique

$$\Omega(\boldsymbol{x}) = (\boldsymbol{x} - \boldsymbol{m}_{\boldsymbol{x}})^{\mathrm{t}} \mathbf{R}_{\boldsymbol{x}}^{-1} (\boldsymbol{x} - \boldsymbol{m}_{\boldsymbol{x}}), \qquad \textit{e.g.}, \;\; \Omega_{\mathrm{PTT}(\boldsymbol{x})} = \|\boldsymbol{D}^{(q)} \boldsymbol{x}\|^2$$

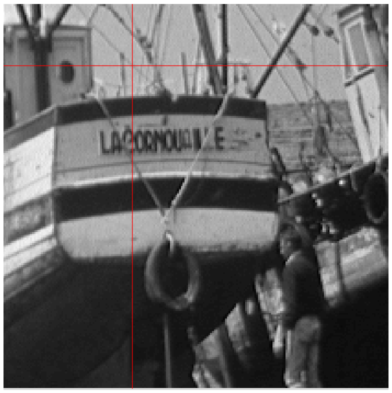
- Lois *a priori* blanches non gaussiennes

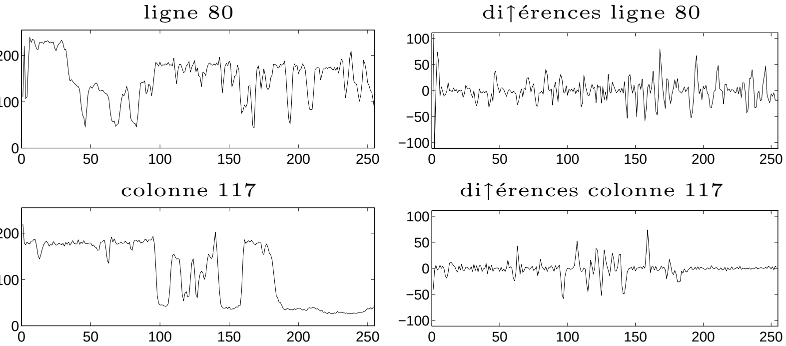
– Critères convexes

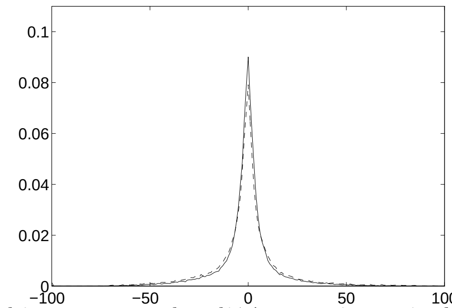
$$\Omega(\boldsymbol{x}) = \sum_i G(x_i), \qquad G_p(x) = |x|^p, \; G_{21}(x) = \sqrt{s^2 + x^2}$$

– Variables cachées : modèle Bernoulli-Gaussien $\boldsymbol{x} = (\boldsymbol{r}, \boldsymbol{q})$

Modèles pour les images ?







histogramme des di↑érences interpixels

Champs de Gibbs-Markov

$$f(\boldsymbol{x}) = \frac{1}{Z} \exp -\Omega(\boldsymbol{x}),$$

$$f_T(\boldsymbol{x}) = \frac{1}{Z(T)} \exp\{-\Omega(\boldsymbol{x})/T\}$$

$$\left(\begin{array}{lcc} \bullet & \Omega \text{ séparable} \quad \Longleftrightarrow & X \text{ « bruit blanc »} \\ & \Omega(\boldsymbol{x}) = \sum_{s \in \mathcal{S}} G_s(x_s) & f(\boldsymbol{x}) = \prod_{s \in \mathcal{S}} f_s(x_s) \end{array} \right)$$

- Ω énergie de Gibbs $\Longleftrightarrow$ X markovien

$$\Omega(\boldsymbol{x}) = \sum_{c \in \mathcal{C}} G_c(\boldsymbol{x}) \qquad f(x_s \,|\, x_r,\, r \in \mathcal{S} \setminus s) = f(x_s \,|\, x_r,\, r \in \mathcal{V}_s)$$

$$\Omega(\boldsymbol{x}) < \infty \; \forall \boldsymbol{x} \qquad \mathcal{V}_s = \{r \in c \ni s,\, r \neq s\}$$

$c \in \mathcal{C} \subset \mathcal{P}(\mathcal{S})$ clique, voisinage du site s

G_c potentiel associé

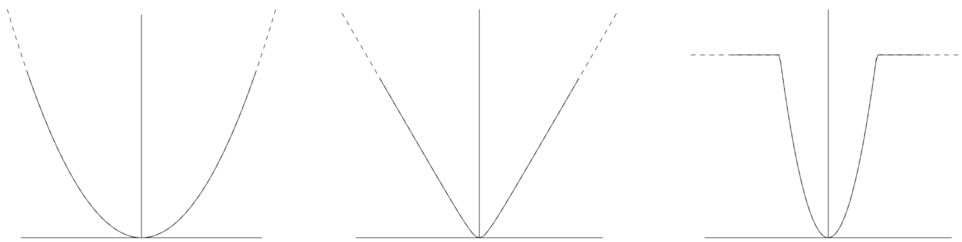
(Théorème de Hammersley-Clifford)

Exemple : champ de Markov d'ordre 1, *i.e.*, *aux plus proches voisins*

$$\Omega(\boldsymbol{x}) = \sum_{\{r,s\}\in\mathcal{C}} G(x_r - x_s)$$

$$\mathcal{C} = \{\{r,s\}, |r-s| = 1\} = \left\{ \bigcirc\;\bigcirc,\ \begin{array}{c}\bigcirc\\ \bigcirc\end{array} \right\} \iff \mathcal{N}_s = \left\{ \begin{array}{ccc} & \bigcirc & \\ \bigcirc & \bullet & \bigcirc \\ & \bigcirc & \end{array} \right\}$$

$G_{22}(x) = x^2$ $\quad$ $G_{21}(x) = \sqrt{t^2 + x^2} - t$ $\quad$ $G_{20}(x) = \min\{x^2, t^2\}$



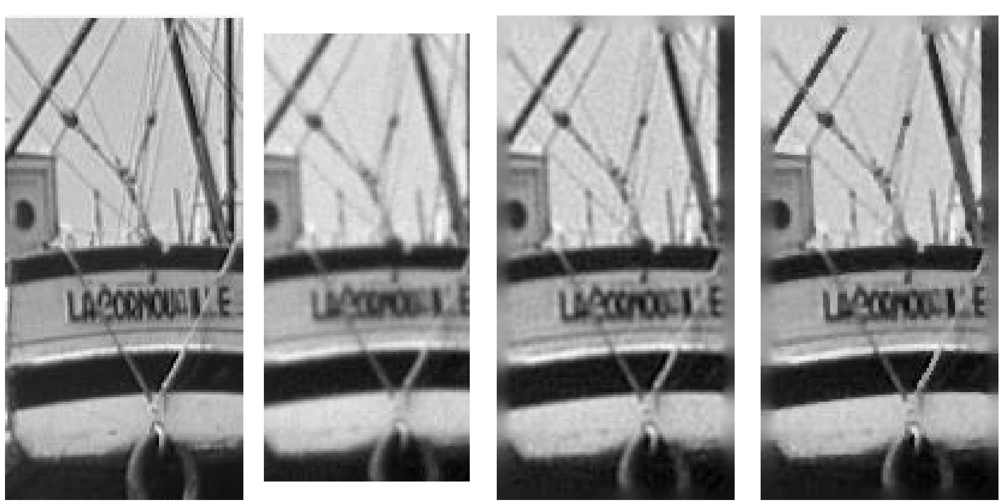
Critère et optimisation convexes

$$\text{Exemple (ordre 1) : } G_{21}(\boldsymbol{x}) = \sum_{\{r,s\},\|r-s\|=1} \sqrt{s^2 - (x_r - x_s)^2}$$

application : restauration d'images, en minimisant

$$J(\boldsymbol{x}) = \|\boldsymbol{z} - \mathbf{H}\boldsymbol{x}\|^2 + \alpha G_{21}(\boldsymbol{x}), \quad \boldsymbol{x} \in [0, x_{max}]^{\mathcal{S}}$$

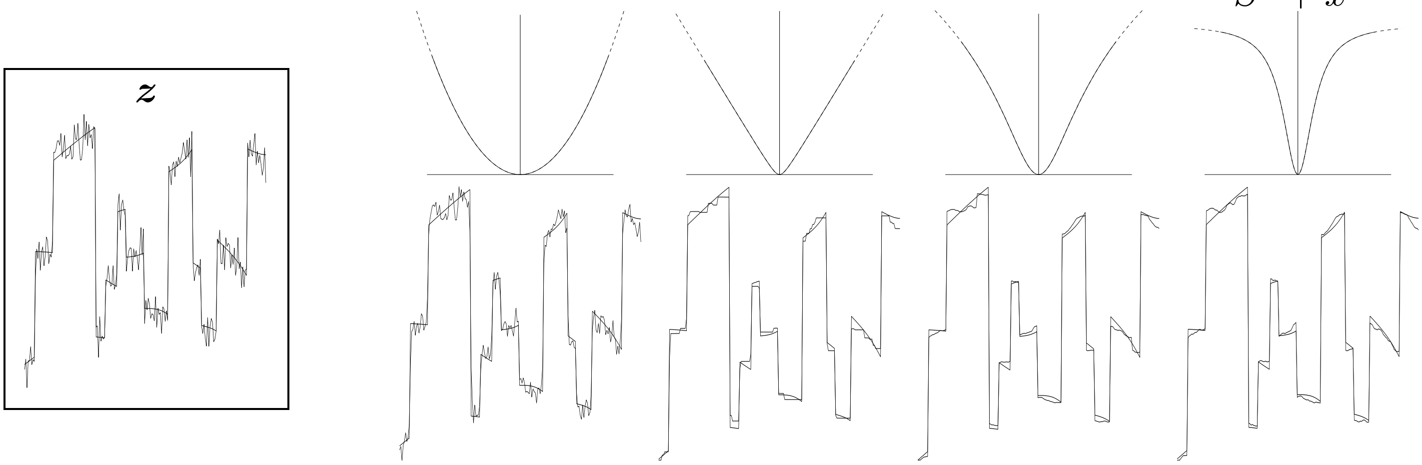
image originale — image observée — résultat gaussien G_{22} — résultat G_{21}



Test comparatif (débruitage : $\mathbf{H} = \mathbf{I}$)

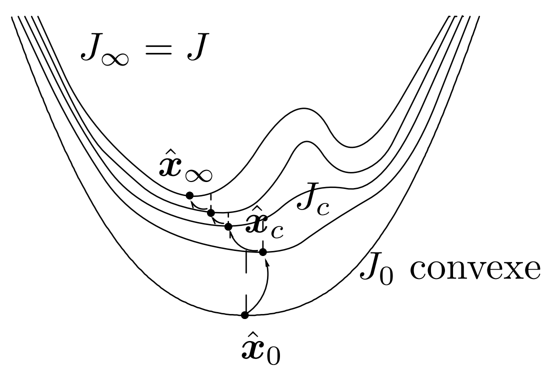
$$\widehat{\boldsymbol{x}} = \arg\min_{\boldsymbol{x}} \|\boldsymbol{z} - \boldsymbol{x}\|^2 + \alpha \sum_m G(x_m - x_{m-1})$$

$G(x) =$	x^2	$\sqrt{S^2 + x^2}$	$\ln(S^2 + x^2)$	$\dfrac{x^2}{S^2 + x^2}$



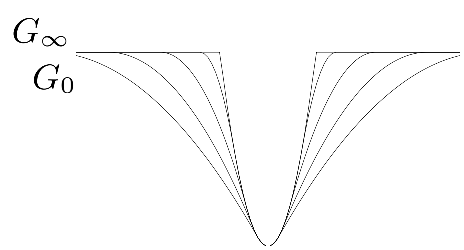
$\|\widehat{\boldsymbol{x}} - \boldsymbol{x}^*\|_1 \,/\, \|\boldsymbol{x}^*\|_1 =$ (minimal en α, S)	12,5%	6,2%	5,67%	5,6%

Non convexité graduelle (GNC)



■ **G_{20} : le cas du débruitage** [Blake et Zisserman **1987**]

$$J_c(\boldsymbol{x}) = \beta \|\boldsymbol{z} - \boldsymbol{x}\|^2 + \sum_{\{r,s\} \in \mathcal{C}} G_c(x_r - x_s)$$

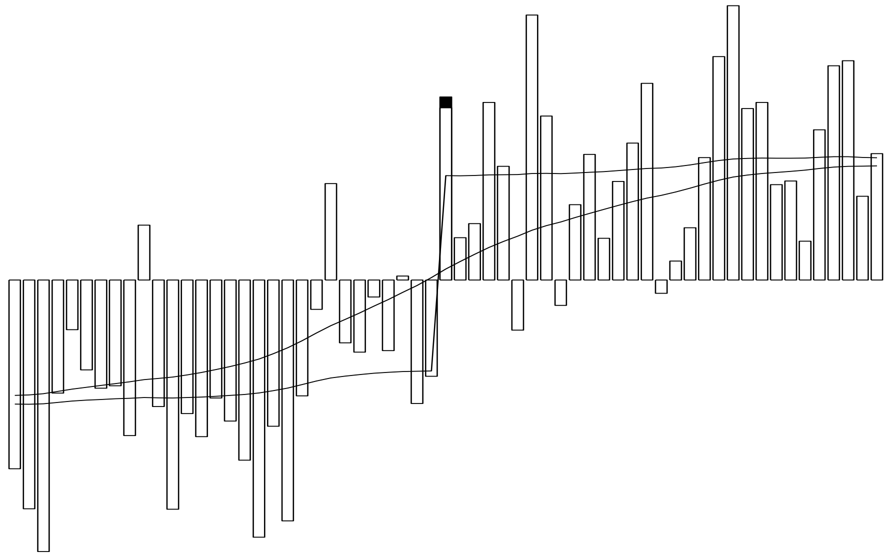


Stabilité de $\widehat{\boldsymbol{x}}(\boldsymbol{z})$

J convexe $\Longrightarrow \widehat{\boldsymbol{x}}(\boldsymbol{z})$ continue [Bouman et Sauer 1993]

$\Longrightarrow$ stabilité au sens de Hadamard

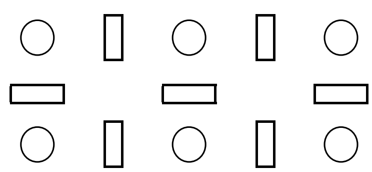
Contre-exemple pour J non convexe :
(d'après [Li et coll. 1995] ; voir aussi [Künsch 1994])



Principe de détection, variables cachées

- [Blake et Zisserman **1987**]

$$\min_{\boldsymbol{x}} \beta \left\|\boldsymbol{z} - A(\boldsymbol{x})\right\|^2 + \sum_{\{r,s\}\in\mathcal{C}} \min\left\{(x_r - x_s)^2,\, S^2\right\}$$
$$= \min_{\boldsymbol{x},\boldsymbol{l}} \beta \left\|\boldsymbol{z} - A(\boldsymbol{x})\right\|^2 + \sum_{\{r,s\}\in\mathcal{C}} (1 - l_{rs})(x_r - x_s)^2 + l_{rs} S^2$$



Variables cachées $l_{rs} \in \{0,\, 1\}$ $\rightsquigarrow$ problème de *détection-estimation*

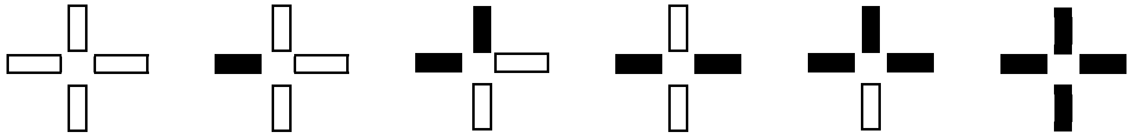
$\boldsymbol{X}$, $\boldsymbol{L}$, champ de Markov composite.

Régularité des contours [Geman et Geman 1984, Jeng et Woods 1991]

$$\min_{\boldsymbol{x},\boldsymbol{l}} \beta \left\|\boldsymbol{z} - A(\boldsymbol{x})\right\|^2 + \sum_{\{r,s\}\in\mathcal{C}} (1 - l_{rs})(x_r - x_s)^2 + \sum_{c\in\mathcal{C}_L} G_c(\boldsymbol{l})$$

Exemple :

$c \in \mathcal{C}_L$



$G_c(\boldsymbol{l})$	0	2,7	1,8	0,9	1,8	2,7

Échantillonnage stochastique

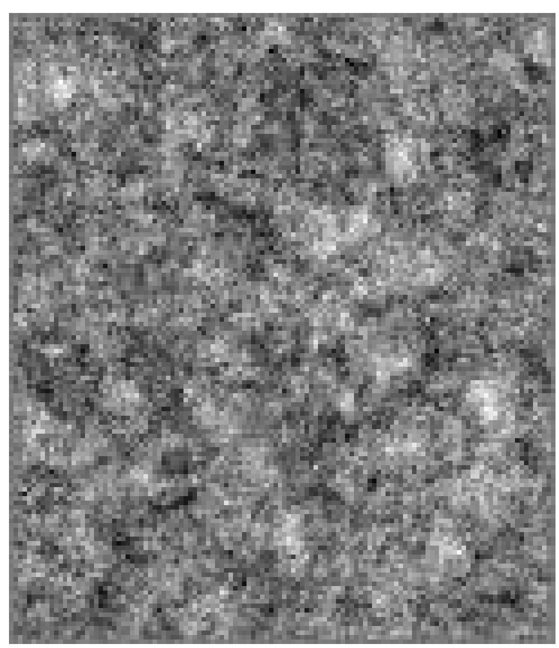
Application en synthèse d'image

- Exemple : modèle de Gibbs aux plus proches voisins

$$\mathcal{C} = \{\{r,\, s\},\, \|r - s\| = 1\}$$

Exemple d'échantillon pour

$$\Phi(\boldsymbol{x}) = \sum_{\{r,s\},\|r-s\|=1} \sqrt{1 + (x_s - x_r)^2}$$



Application à l'inférence

Contexte :

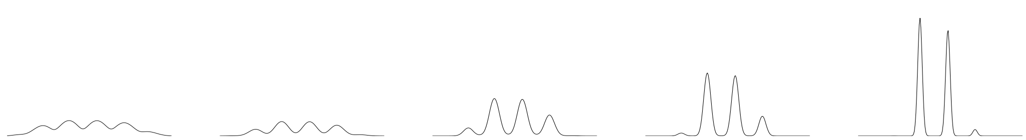
- $\boldsymbol{z} = [z_1,\, \dots,\, z_N]^{\mathrm{t}}$ vecteur observé
- $\boldsymbol{x} = [x_1,\, \dots,\, x_M]^{\mathrm{t}}$ vecteur inconnu
- $\boldsymbol{z}$ est relié à $\boldsymbol{x}$ sous forme probabiliste : on connaît la loi de $(\boldsymbol{X},\, \boldsymbol{Z})$, typiquement, sous la forme $f_{\boldsymbol{X},\boldsymbol{Z}}(\boldsymbol{x},\, \boldsymbol{z}) = f_{\boldsymbol{Z}|\boldsymbol{X}}(\boldsymbol{z} \,|\, \boldsymbol{x}) f_{\boldsymbol{X}}(\boldsymbol{x})$

On veut estimer $\boldsymbol{x}$ sous la forme $\widehat{\boldsymbol{x}} = \widehat{\boldsymbol{X}}(\boldsymbol{z})$

① Estimation à coût bayésien séparable :
Si $\{\boldsymbol{X}^{(k)}\}$ est stationnaire ergodique et de loi instantanée $f(\boldsymbol{x} \,|\, \boldsymbol{z})$, alors

$$\widehat{\boldsymbol{x}}^{\mathrm{EAP}}(\boldsymbol{z}) = \mathrm{E}[\boldsymbol{X} \,|\, \boldsymbol{z}] = \lim_{K\to\infty} \frac{1}{K} \sum_{k=1}^{K} \boldsymbol{x}^{(k)}$$

② Estimation MAP (principe du recuit simulé) :
Si la loi instantanée de $\{\boldsymbol{X}^{(k)}\}$ est $f_{T(k)}(\boldsymbol{x} \,|\, \boldsymbol{z})$, avec $T(k) \searrow 0$ et
$f_T(\boldsymbol{x} \,|\, \boldsymbol{z}) \propto (f(\boldsymbol{x} \,|\, \boldsymbol{z}))^{1/T}$, alors : $\quad \lim_{T\to 0} f_T(\boldsymbol{x} \,|\, \boldsymbol{z}) = \delta_{\widehat{\boldsymbol{X}}^{\mathrm{MAP}}}(\boldsymbol{x})$



$f_4(x)$ $\quad f_2(x)$ $\quad f(x)$ $\quad f_{1/2}(x)$ $\quad f_{1/5}(x)$

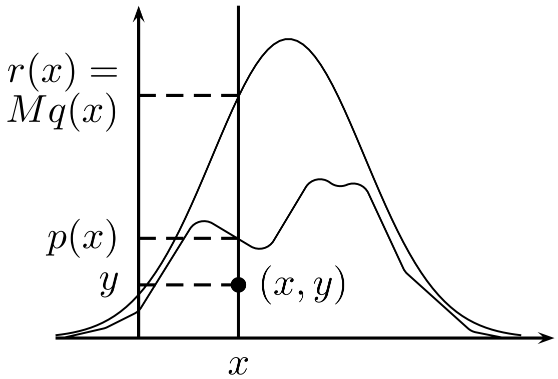
③ Estimation non supervisée (principe de l'augmentation de données) :

$$\{\boldsymbol{X}^{(k)},\, \Theta^{(k)}\} \sim f(\boldsymbol{x},\, \theta \,|\, \boldsymbol{z}) \Longrightarrow \begin{cases} \Theta^{(k)} \sim f(\theta \,|\, \boldsymbol{z}) = \int f(\boldsymbol{x},\, \theta \,|\, \boldsymbol{z})\, \mathrm{d}\boldsymbol{x} \\ \boldsymbol{X}^{(k)} \sim f(\boldsymbol{x} \,|\, \boldsymbol{z}) = \int f(\boldsymbol{x},\, \theta \,|\, \boldsymbol{z})\, \mathrm{d}\theta. \end{cases}$$

Méthodes d'échantillonnage i.i.d. (X faible dimension)

- **directes** :
  - lois uniformes,
  - par transformation : discrètes, normales, lois avec F^{-1} explicite, etc.

- **par réjection** [Press et coll. 1992]



① Trouver une densité q et $M > 0$ tel que $\forall x,\ r(x) = Mq(x) \geqslant p(x)$.
② Tirer $X = x$ suivant q.
③ Tirer $Y = y$ suivant $\mathcal{U}([0,\ r(x)])$.
④ Accepter x si $y \leqslant p(x)$, sinon retour en ②

Méthodes d'échantillonnage MCMC (*Monte-Carlo Markov chain*)

$\{\boldsymbol{X}^{(k)}\}$ est une chaîne de Markov : $f(\boldsymbol{x}^{(k)} \mid \boldsymbol{x}^{(k-1)}, \ldots, \boldsymbol{x}^{(k)}) = f(\boldsymbol{x}^{(k)} \mid \boldsymbol{x}^{(k-1)})$

- Théorème 1 :

  Soit $\{\boldsymbol{X}^{(k)}\}_{k \in \mathbb{N}}$ une chaîne de Markov homogène « ergodique » (irréductible récurrente positive de période 1) dont le noyau de transition φ vérifie la *condition d'équilibre*

$$\varphi(\boldsymbol{x}' \mid \boldsymbol{x}) f(\boldsymbol{x}) = \varphi(\boldsymbol{x} \mid \boldsymbol{x}') f(\boldsymbol{x}'),$$

  alors la densité de probabilité de $\boldsymbol{X}^{(\infty)}$ est f.

- Théorème 2 (loi des grands nombres) :

$$\text{si } \mathrm{E}_f\big[\psi^2\big] < \infty, \quad \lim_{K \to \infty} \frac{1}{K} \sum_{k=1}^{K} \psi(\boldsymbol{x}^{(k)}) = \mathrm{E}_f\big[\psi\big] \text{ p.s.}$$

MCMC 1 : Algorithme de Métropolis [MÉTROPOLIS et coll. 1953]

- Soit $f(\boldsymbol{x})$ à échantillonner et $g(\boldsymbol{x}' \,|\, \boldsymbol{x})$ un *noyau de proposition* symétrique

  ⓪ Configuration courante : $\boldsymbol{x}$

  ① Proposer $\boldsymbol{x}'$ par échantillonnage de $g(\boldsymbol{x}' \,|\, \boldsymbol{x})$;

  ② Si $f(\boldsymbol{x}') \geqslant f(\boldsymbol{x})$, $\quad \boldsymbol{x}'$ remplace $\boldsymbol{x}$;
  Si $f(\boldsymbol{x}') < f(\boldsymbol{x})$, $\quad P(\boldsymbol{x}'$ remplace $\boldsymbol{x}) = f(\boldsymbol{x}')/f(\boldsymbol{x})$;
  retour en ⓪ pour l'itération suivante

- Application à $\boldsymbol{X}$ champ de Gibbs : $\quad f(\boldsymbol{x}) = \dfrac{1}{Z} \exp - \sum_c G_c(\boldsymbol{x})$

  Noyau de proposition :

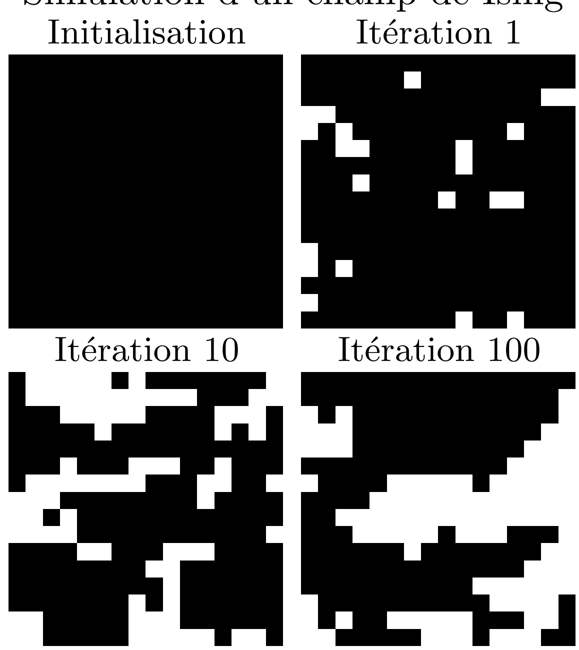
$$\begin{aligned} g(\boldsymbol{x}' \,|\, \boldsymbol{x}) &= h(x'_s \,|\, x_s)/\mathrm{Card}\,\mathcal{S} && \text{si } \forall r \neq s,\ x'_r = x_r \\ &= 0 && \text{sinon} \end{aligned}$$

$$\Longrightarrow \quad f(\boldsymbol{x}')/f(\boldsymbol{x}) = \exp \sum_{c \ni s} (G_c(\boldsymbol{x}) - G_c(\boldsymbol{x}'))$$

  Extensions : balayage systématique, partiellement parallélisable

- Exemple

Simulation d'un champ de Ising



MCMC 2 : Échantillonneur de Gibbs [Geman et Geman 1984]

0 Configuration courante : $\boldsymbol{x}$

1 Choisir s « au hasard » ou cycliquement ;

2 $\forall r \neq s,\ x'_r = x_r$; échantillonner x'_s suivant $f(x'_s \mid \boldsymbol{x}_{-s}),\ \ \boldsymbol{x}_{-s} = \{x_r, r \neq s\}$.
retour en 0 pour l'itération suivante

Remarque : MCMC 1 et 2 sont des cas particuliers de l'échantillonneur de Métropolis-Hastings [Hastings 1970].

discussion) », *J. R. Statist. Soc. B*, **36**, n°2, 192-236.

[Besag 1986] J. E. Besag (1986), « On the statistical analysis of dirty pictures (With discussion) », *J. R. Statist. Soc. B*, **48**, n°3, 259-302.

[Blake et Zisserman 1987] A. Blake et A. Zisserman (1987), *Visual reconstruction*, The MIT Press, Cambridge, MA, USA.

[Bouman et Sauer 1993] C. A. Bouman et K. D. Sauer (1993), « A generalized Gaussian image model for edge-preserving MAP estimation », *IEEE Trans. Image Processing*, **2**, n°3, 296-310.

[Geman et Geman 1984] S. Geman et D. Geman (1984), « Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images », *IEEE Trans. Pattern Anal. Mach. Intell.*, **PAMI-6**, n°6, 721-741.

[Hastings 1970] W. K. Hastings (1970), « Monte Carlo sampling methods using Markov Chains and their applications », *Biometrika*, **57**, 97.

[Jeng et Woods 1991] F. C. Jeng et J. W. Woods (1991), « Compound Gauss-Markov random fields for image estimation », *IEEE Trans. Signal Processing*, **39**, n°3, 683-697.

[Künsch 1994] H. R. Künsch (1994), « Robust priors for smoothing and image restoration », *Ann. Inst. Stat. Math.*, **46**, n°1, 1-19.

[Li et coll. 1995] S. Z. Li, Y. H. Huang et J. S. Fu (1995), « Convex MRF potential functions », dans *Proc. IEEE ICIP*, volume 2, 296–299, Washington DC, USA.

[Marroquin et coll. 1987] J. L. Marroquin, S. K. Mitter et T. A. Poggio (1987), « Probabilistic solution of ill-posed problems in computational vision », *J. Amer. Stat. Assoc.*, **82**, 76-89.

[Metropolis et coll. 1953] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller et E. Teller (1953), « Equations of state calculations by fast computing machines », *Journal of chemical physics*, **21**, 1087-1092.

[Press et coll. 1992] W. H. Press, S. A. Teukolsky, W. T. Vetterling et B. P. Flannery (1992), *Numerical recipes in C, the art of scientific computing*, Cambridge Univ. Press, New York, 2nd ed.

[Robert 1992] C. Robert (1992), *L'analyse statistique bayésienne*, Economica, Paris.

[Robert 1996] C. Robert (1996), *Méthodes de Monte-Carlo par chaînes de Markov*, Economica, Paris.

[Winkler 1995] G. Winkler (1995), *Image Analysis, Random Fields and Dynamic Monte Carlo Me-*