Echantillonnage exact de distributions de Gibbs d'énergies sous-modulaires

Exact sampling of Gibbs distributions with submodular energies *

Marc Sigelle[1] and Jérôme Darbon [2]

[1] Institut TELECOM TELECOM ParisTech CNRS UMR 5141, 37-39 rue Dareau 75014 Paris Cedex 13 France

[2] CMLA, ENS Cachan, CNRS, UniverSud, 61 Av. Président Wilson, F-94230 Cachan, France

[2] Department of Mathematics, University of California, Los Angeles, CA, USA.

* CNRS PEPS INS2II "Approches combinatoires et stochastiques pour le traitement des images sur architectures parallèles embarquées"
Financement de l'Office of Naval Research ONR-N00014-11-1-0749

Plan

- **motivation**
- **recalls**

  Markov Random Field (MRF)

  Markov Chain Monte Carlo $\leftarrow$ used for sampling MRF
- **exact sampling: Coupling From the Past**

  principle

  monotony
- **our contribution**

  $\rightarrow$ this works as well for submodular energies !
- **conclusion and perspectives**

Motivation

- **sampling a MRF prior distribution**

  → testing a MRF energy model

- **sampling a MRF posterior distribution: denoising, deblurring . . .**

  → statistical estimate based on samples of the distribution

  Louchet and Moisan (2008, 2012)

- **hyperparameter estimation for prior/posterior models**

  → (iterative) hyperparameter update based on current samples

Recalls: Markov Random Field (MRF)

- **definition**

$$x \in \Omega$$ <u>finite</u> (ex: $2^{(256\times 256)}$)

$$\pi(X = x) = \frac{\exp - U(x)}{Z}$$ Gibbs distribution

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

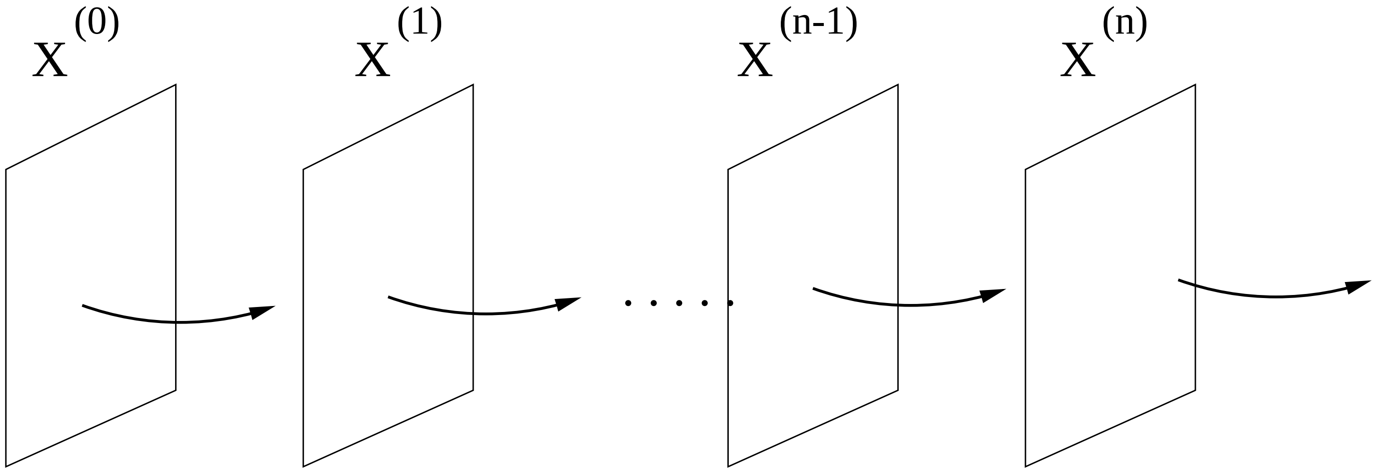
$$U_{(s,t)}(x) = V(x_s, x_t)$$ clique energy

$$Z = \sum_{x \in \Omega} \exp - U(x)$$ partition function

- **Z unavailable $\rightarrow$ impossible to draw (sample) directly in Ω**

Recalls: Markov Chain Monte Carlo (MCMC)

- Gibbs / Metropolis sampler



- ergodicity → sampler ”convergence”
- burn-up period → for how long ?
- when to assess this ”convergence” ?

Coupling From the Past (CFTP)
J. D. Propp and D.B. Wilson (1996)

James D. Propp and D. B. Wilson. “Exact sampling with coupled Markov chains and applications to statistical mechanics.” *Random Structures and Algorithms*, 9(1,2):223–252, 1996.

Annotated Bibliography of Perfectly Random Sampling with Markov Chains
http://dimacs.rutgers.edu/~dbwilson/exact/

Coupling From the Past (CFTP) (followed)

- **looking backward what can happen at fixed time** 0

  $\xi \in \Omega$ may be obtained from many "paths"

- **many "simultaneous" Markov chains** $(|\Omega|)$

  launched at some time $(t < 0)$

- **random map** $f_t : \Omega \mapsto \Omega \quad Markov(t)$

  from time $t \to t+1$

- **random map** $F_t = f_{-1} \; o \; f_{-2} \ldots \; o \; f_{t+1} \; o \; f_t$

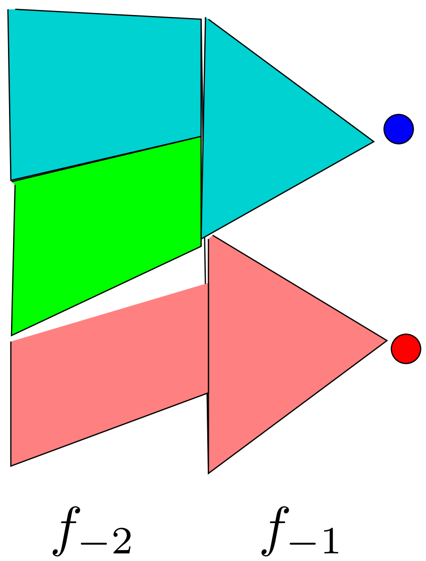
  from time $t \to 0$

- **recursive relation** $F_t = F_{t+1} \; o \; f_t$

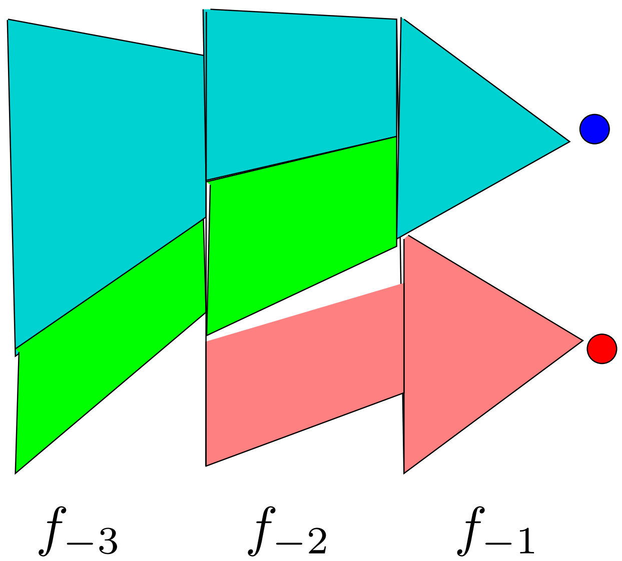
- $F_t =$ **constant** $\Rightarrow$ **coalescence**

Coupling From the Past (CFTP) (followed)

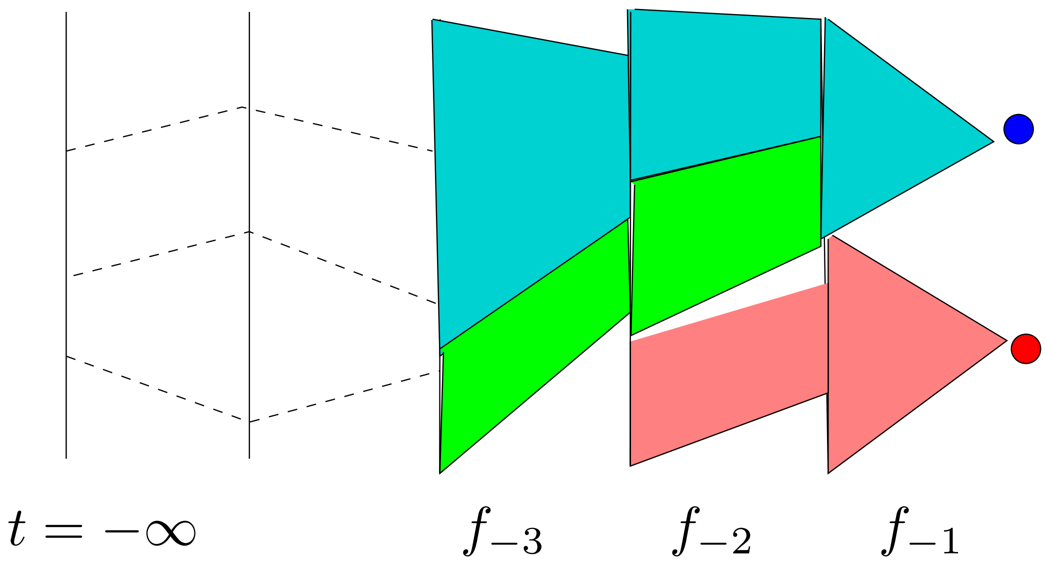
Coupling From the Past (CFTP) (followed)



Coupling From the Past (CFTP) (followed)



Coupling From the Past (CFTP) (followed)



CFTP : main results (Propp and Wilson (1996))

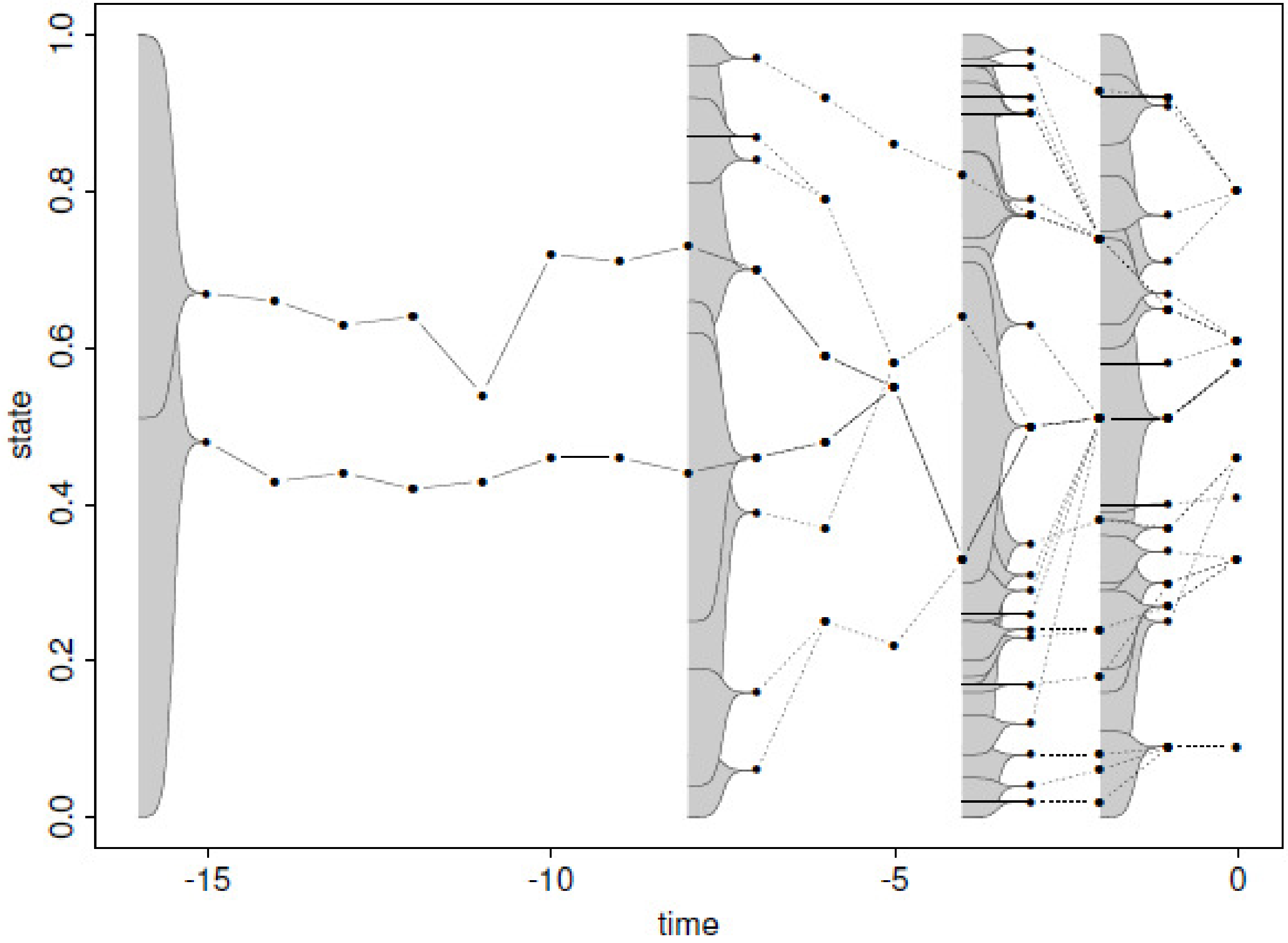
- **with probability 1:**

$\exists$ *finite* time $t = M < 0$ *s.t.* coalescence occurs when starting at time M

- **the state ξ so obtained is a perfect sample of distribution π**

CFTP : a result from Murdoch and Green (1998)

Figure 2: A run of the raw rejection coupler on the toy example. The solid line texture corresponds to paths from $-\infty$ coalescing by time 0.



CFTP : monotony (Propp and Wilson (1996))

- **launch $|\Omega|$ chains ?**
- **partial ordering on Ω**

  $x \preceq y \Leftrightarrow x_s \leq y_s \quad \forall s \in S$

- **2 extremal elements:**

  $\hat{0} \preceq x \preceq \hat{L} \quad \forall x \in \Omega$

- **if $f_t = Markov(t)$ preserves monotony $\forall t$:**

  $\forall x,\ y \in \Omega \quad x \preceq y \Rightarrow \quad f_t(x) \preceq f_t(y)$

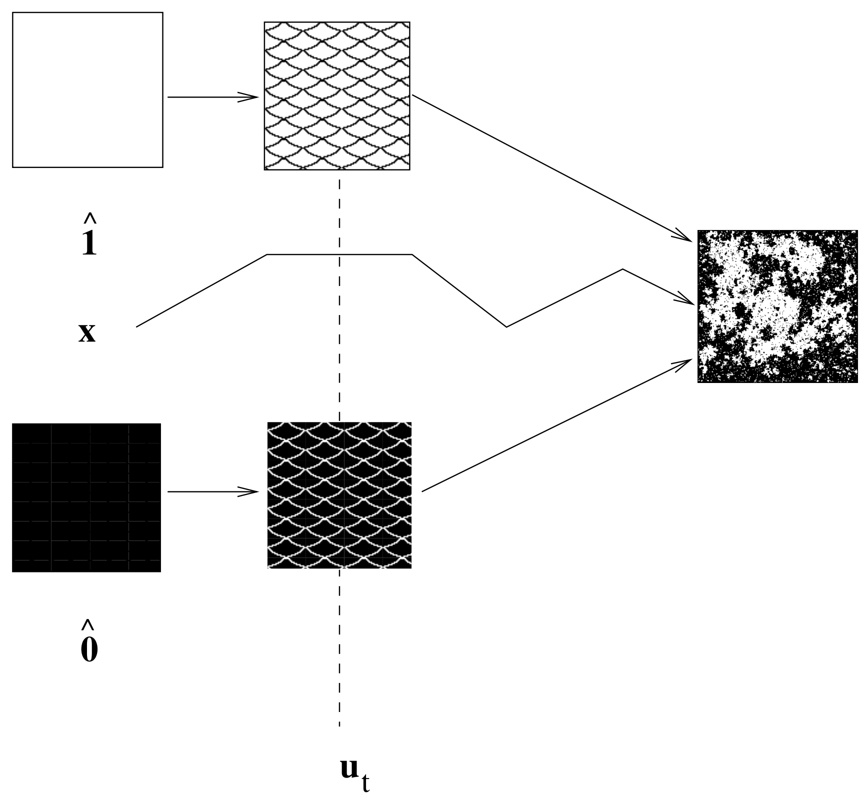
  $\rightarrow$ coupled Gibbs sampler

- **then 2 chains starting from $\hat{0}$ and $\hat{L}$ are enough**

  $\rightarrow$ until they coalesce !

Monotony for the ferromagnetic Ising model

- experiment at Ising critical temperature: $M = 30$ sweeps claimed !



$$\hat{0} \preceq x \preceq \hat{1} \quad \forall x \in \Omega$$

- equivalent monotony condition for binary models

$$\mathcal{N}_s \preceq \mathcal{N}'_s \Rightarrow \pi(X_s = 0 \mid \mathcal{N}_s) \geq \pi(X_s = 0 \mid \mathcal{N}'_s)$$

Our contribution: CFTP extends to submodular energies !

$$U(x_s+1,\ x_t+1)+U(x_s,\ x_t)\leq U(x_s+1,\ x_t)+U(x_s,\ x_t+1)$$

- **condition on interaction energy only**

  $\rightarrow$ independent of attachment to data term !!

- **example : L^2 + (anisotropic) TV**

$$U=\sum_{s\in S}(x_s-y_s)^2+\sum_{(s,t)} w_{st}|\ x_s-x_t\ | \qquad w_{st}\geq 0$$

- **coalescence obtained from $\hat{0}$ and $\hat{L}$**

  $M=100$ sweeps claimed for a 256 greylevel image (Moisan Darbon (2012)) !!

Conclusion and perspectives

- **other models**
- **other estimates**

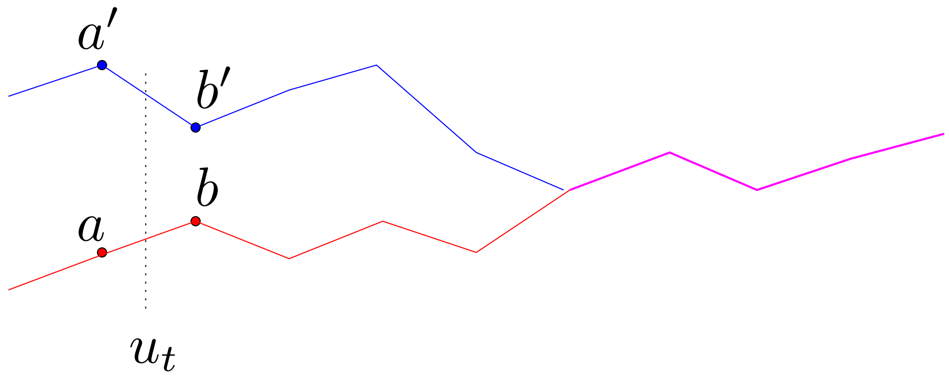
  in progress (Tupin, Darbon, Sigelle (2013))
- **other samplers**

  Metropolis-Hastings $\rightarrow$ done
- **efficient hardware/sofware implementation**

  in progress . . .

Recalls: Markov Chain Monte Carlo (MCMC)

- **Doeblin coupling**



$$
\begin{aligned}
\tilde{Q}((a,\ a'),\ (b,\ b')) &= Q(a,\ b)\ .\ Q(a',\ b') && \text{if } a \neq b \\
&= Q(a,\ b) && \text{if } a = a',\ b = b' \\
&= 0 && \text{else}
\end{aligned}
$$

- **marginally each Markov chain follows kernel Q**

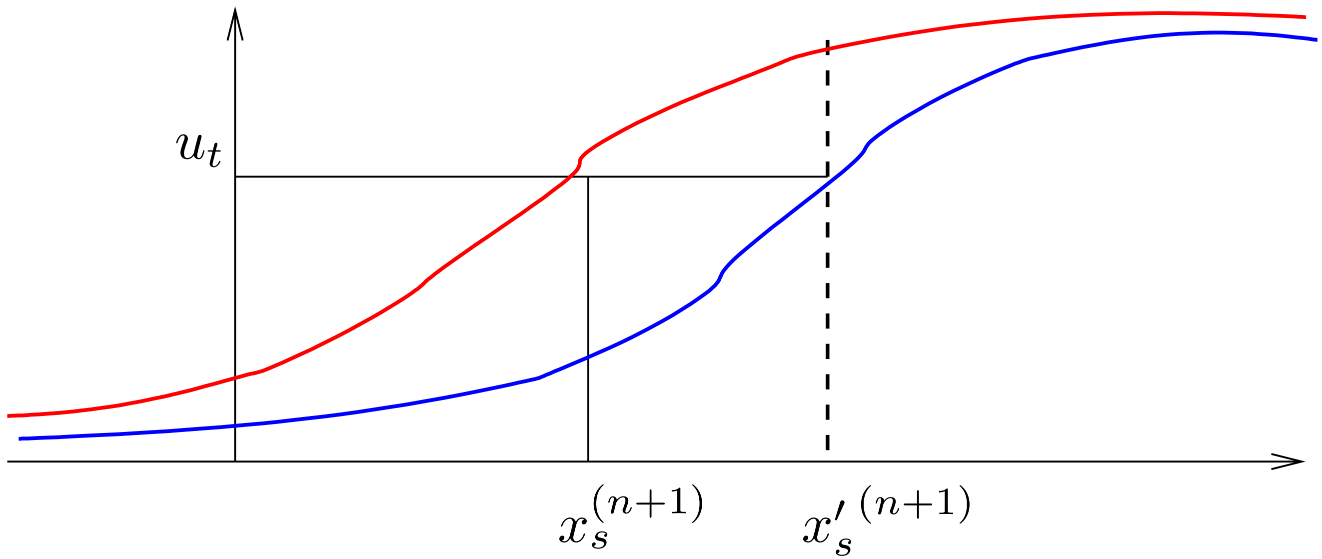
Louchet and Moisan (2008, 2009, 2012)

Recalls: Gibbs Sampler

- **classical Gibbs sampler**

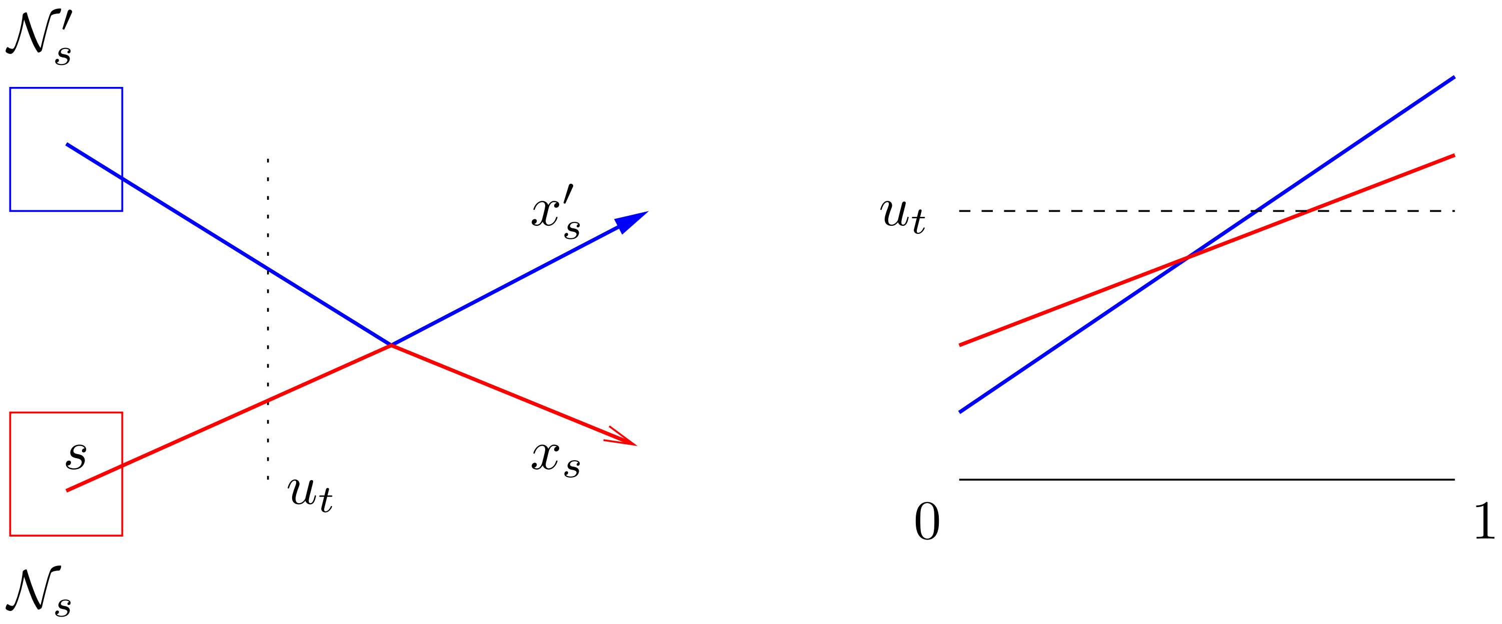
$x^{(n)}$	current configuration	
s	current site	$x_s^{(n+1)} \quad \leftarrow \pi(X_s = . \mid \mathcal{N}_s^{(n)})$
$\mathcal{N}_s^{(n)}$	current neighborhood	

- **coupled Gibbs sampler ($Markov(t)$)**



Recalls: Gibbs Sampler (followed)

- **a remark**



References

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, **57**(1), 97–109.

Lalley, S. P. (2009). *Convergence Rates of Markov Chains*. http://galton.uchicago.edu/~lalley/Courses/313ConvergenceRates.pdf.

Louchet, C. and Moisan, L. (2008). Total Variation denoising using posterior expectation. In *Proceedings of EUropean Signal and Image Processing Conference EUSIPCO*.

Louchet, C. and Moisan, L. (2012). Posterior Expectation of the Total Variation model: Properties and Experiments. http://hal.archives-ouvertes.fr/hal-00764175/.

Murdoch, D. J. and Green, P. J. (1998). Exact sampling from a continuous state space. *Scandinavian Journal of Statistics*, **25**(3), 483–502. http://www.maths.bristol.ac.uk/~mapjg/papers/MurdochGreenSJS.pdf.

Propp, J. G. and Wilson, D. B. (1996). Exact sampling with coupled Markov chains and statistical mechanics. *Random Structures and Algorithms*, **9**(1), 223–252.