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 ← used for sampling MRF

○ exact sampling: Coupling From the Past

  principle

  monotony

○ our contribution

  → 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
  

- **hyperparameter estimation for prior/posterior models**
  - → (iterative) hyperparameter update based on current samples
Recalls: Markov Random Field (MRF)

- **Definition**

\[ x \in \Omega \text{ finite (ex: } 2^{(256 \times 256)}) \]

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

\[ U(x) = \sum_{c \in C} U_c(x) \] 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 \(\Omega\)**
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)  


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 $f_t : \Omega \mapsto \Omega$ \textit{Markov}(t)
  from time $t \to t + 1$

- random map $F_t = f_{-1} \circ f_{-2} \ldots \circ f_{t+1} \circ f_t$
  from time $t \to 0$

- recursive relation $F_t = F_{t+1} \circ f_t$

- $F_t = \text{constant} \Rightarrow \text{coalescence}$
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 \( \xi \) so obtained is a perfect sample of distribution \( \pi \)
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 $\Omega$
  \[ x \preceq y \iff x_s \preceq y_s \quad \forall s \in S \]

- 2 extremal elements:
  \[ \hat{0} \preceq x \preceq \hat{L} \quad \forall x \in \Omega \]

- if $f_t = \text{Markov}(t)$ preserves monotony $\forall t$:
  \[ \forall x, y \in \Omega \quad x \preceq y \Rightarrow f_t(x) \preceq f_t(y) \]
  \[ \rightarrow \text{coupled Gibbs sampler} \]

- then 2 chains starting from $\hat{0}$ and $\hat{L}$ are enough
  \[ \rightarrow \text{until they coalesce!} \]
Monotony for the ferromagnetic Ising model

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

- 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**
  → independent of attachment to data term !!

- **example**: \( L^2 + (\text{anisotropic}) \text{ TV} \)

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

- **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 → done
- efficient hardware/software implementation
  - in progress ...
Recalls: Markov Chain Monte Carlo (MCMC)

- **Doeblin coupling**

\[
\tilde{Q}((a, a'), (b, b')) = Q(a, b) \cdot Q(a', b') \quad \text{if } a \neq b
\]

\[
= Q(a, b) \quad \text{if } a = a', b = b'
\]

\[
= 0 \quad \text{else}
\]

- **marginally each Markov chain follows kernel** \(Q\)

Recalls: Gibbs Sampler

- **classical Gibbs sampler**
  
  \[
  x^{(n)} \quad \text{current configuration} \\
  s \quad \text{current site} \\
  \mathcal{N}_s^{(n)} \quad \text{current neighborhood}
  \]

  \[
  x_s^{(n+1)} \leftarrow \pi(X_s = . \mid \mathcal{N}_s^{(n)})
  \]

- **coupled Gibbs sampler (Markov(t))**
Recalls: Gibbs Sampler (followed)

- a remark

\[ \mathcal{N}_s', \mathcal{N}_s \]

\[ u_t \]

\[ x_s, x'_s \]
References


http://hal.archives-ouvertes.fr/hal-00764175/.

http://www.maths.bristol.ac.uk/~mapjg/papers/MurdochGreenSJS.pdf.