KU LEUVEN

DTAI MACHINE LEARNING

AutoML in the Wild

LSB: Local Self-Balancing MCMC in Discrete Spaces

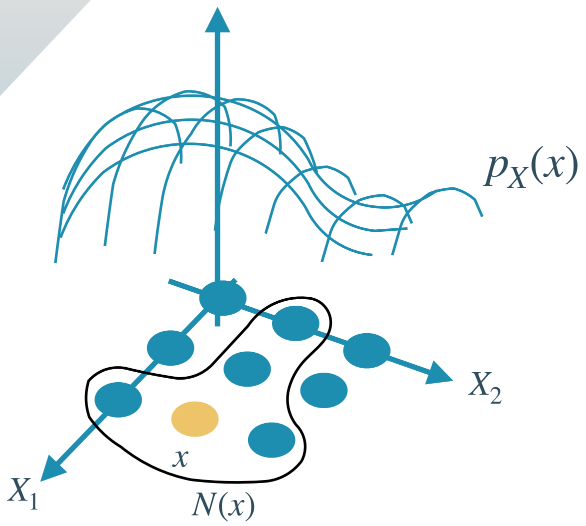
Emanuele Sansone

Problem

- Sampling in high dimensions $p_X(x) = \frac{\tilde{p}_X(x)}{Z}$
  - MCMC $T(x'|x) = A(x',x)Q(x'|x)$
    - Locally Balanced Proposal [1]

$$Q(x'|x) = \frac{g\left(\frac{\tilde{p}(x')}{\tilde{p}(x)}\right)1[x' \in N(x)]}{Z(x)}$$

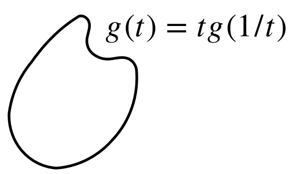
$$g(t) = tg(1/t)$$



Question: How to adapt the proposal . to target .
to improve sampling efficiency?

Solution

Parametrizations



Linear (LSB 1)

$$g_\theta(t) = \sum_{i=1}^{I} \theta_i g_i(t)$$

Nonlinear (LSB 2)

$$g_\theta(t) = \min\left\{\ell_\theta(t), t\ell\left(\frac{1}{t}\right)\right\}$$

Any non-negative real function $\ell_\theta(t)$

Objective

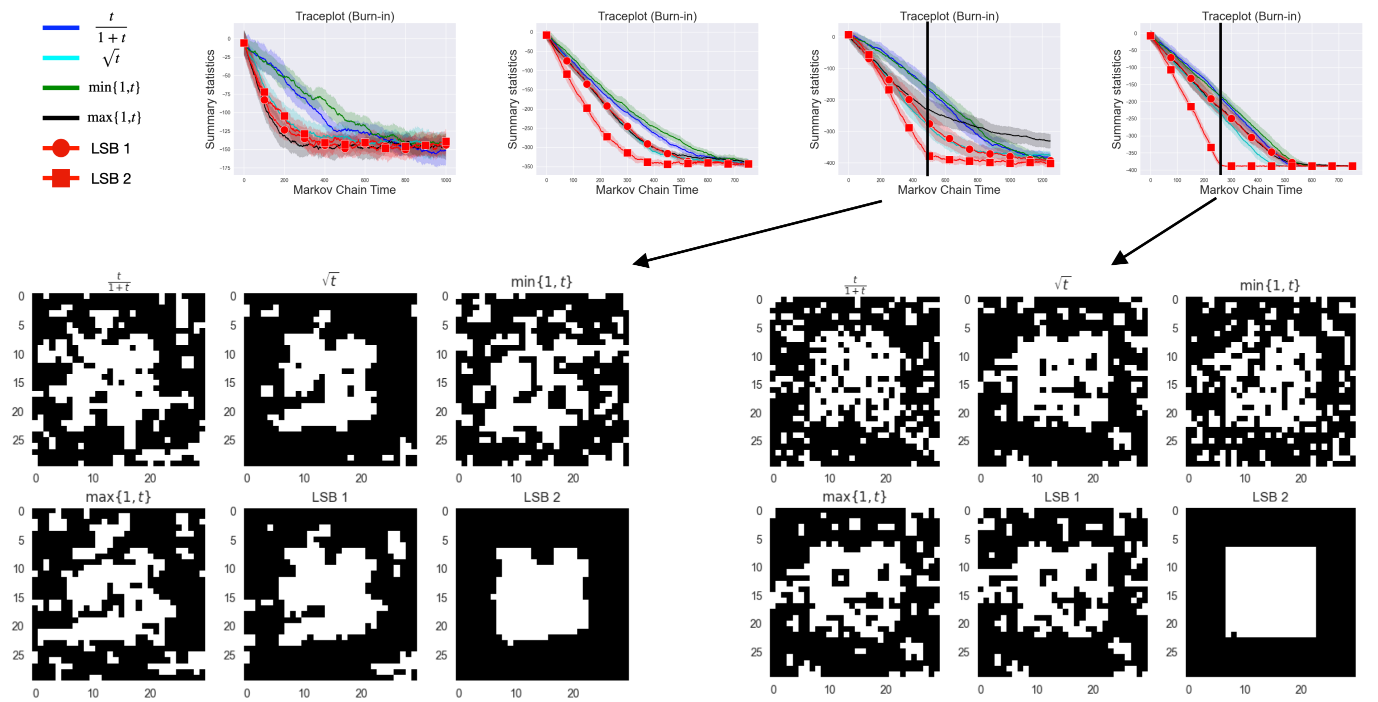
$$I_\theta = KL\{p_X(x)T(x'|x) \| p_X(x)p_X(x')\}$$

Learning procedure

Use historical samples to estimate the objective and update theta at each sampling iteration (during burn-in phase)

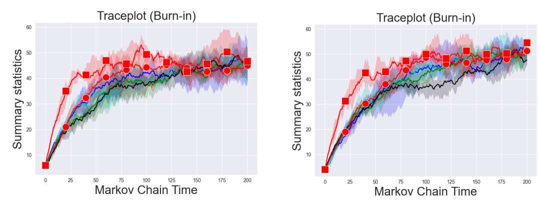
Experiments

2D Ising



Setting	$\frac{t}{1+t}$	$\sqrt{t}$	$\min\{1,t\}$	$\max\{1,t\}$	LSB 1	LSB 2
Case 1	**2.48 ± 0.21**	**2.30 ± 0.22**	**2.42 ± 0.19**	1.75 ± 0.17	**2.50 ± 0.28**	**2.46 ± 0.28**
Case 2	**3.33 ± 0.32**	**2.94 ± 0.36**	**3.33 ± 0.33**	1.72 ± 0.18	**2.98 ± 0.24**	**3.33 ± 0.43**
Case 3	**2.58 ± 0.73**	**1.99 ± 0.43**	**2.56 ± 0.62**	1.26 ± 0.12	**2.48 ± 0.61**	**2.67 ± 0.84**
Case 4	**32.8 ± 9.2**	**18.5 ± 6.8**	**31.8 ± 10.0**	2.60 ± 1.46	**18.4 ± 8.0**	**30.8 ± 9.2**

Bayesian Networks



Dataset	$\frac{t}{1+t}$	$\sqrt{t}$	$\min\{1,t\}$	$\max\{1,t\}$	LSB 1	LSB 2
BN 1	2.90 ± 0.76	3.41 ± 0.77	2.54 ± 0.32	2.70 ± 0.63	3.19 ± 0.46	3.22 ± 0.38
BN 2	3.43 ± 0.75	3.92 ± 0.94	3.78 ± 0.50	3.63 ± 0.67	3.52 ± 0.42	3.44 ± 0.44

References

[1] Zanella. Informed Proposals for Local MCMC in Discrete Spaces. Journal of the American Statistical Association 2020

Contact: emanuele.sansone@kuleuven.be