Estimating Convergence of Markov chains with L-Lag Couplings

Niloy Biswas[1], Pierre E. Jacob[1], Paul Vanetti[2]

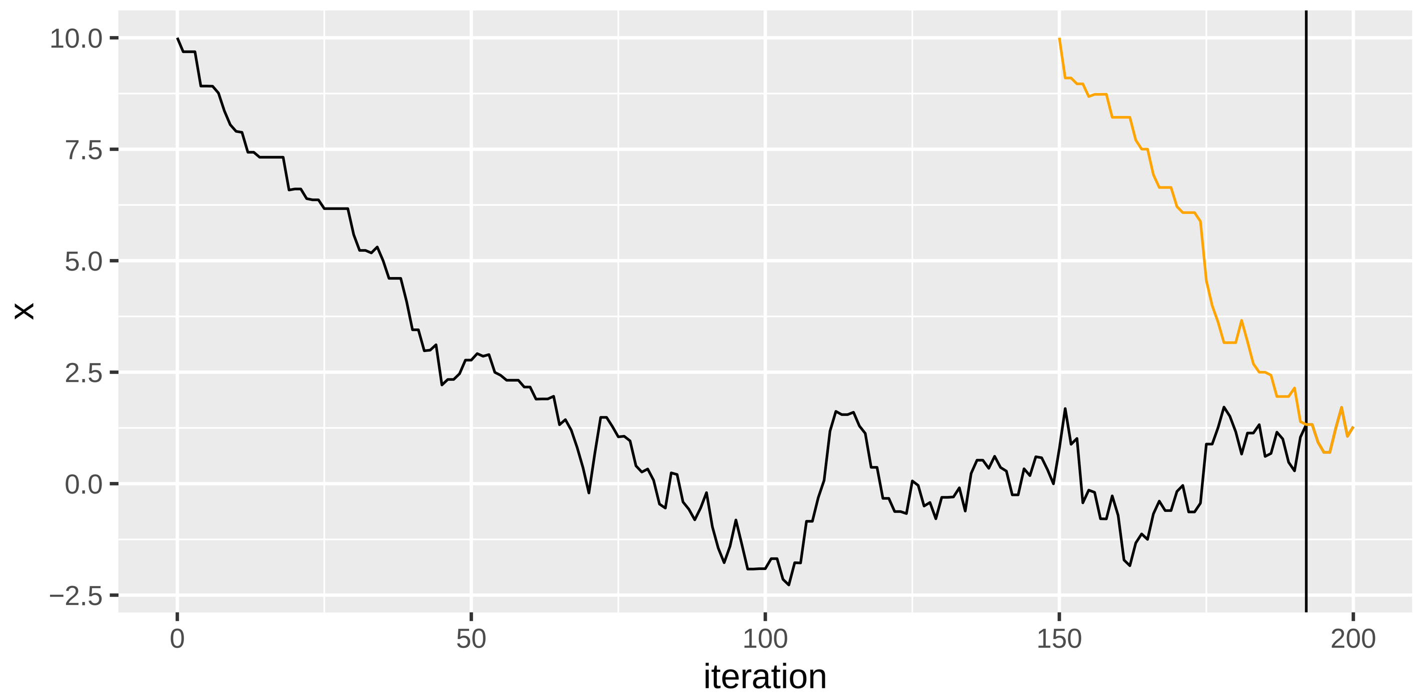
1: Department of Statistics, Harvard University 2: Department of Statistics, University of Oxford

Motivation: finite-time bias of MCMC

- MCMC methods have **non-asymptotic bias**: they only reach the target distribution as the number of iterations goes to infinity.
- We introduce L-**lag couplings** to generate computable, non-asymptotic upper bound estimates for the total variation or the Wasserstein distance of general Markov chains.
- We apply L-lag couplings to:
  - Determine MCMC burn-in
  - Compare different MCMC algorithms with the same target
  - Compare exact and approximate MCMC

What are L-Lag Couplings?

- A pair of Markov chains $(X_t, Y_t)_{t\geq 0}$ such that:
  - Same marginal distributions: $X_t \sim Y_t \sim \pi_t \; \forall t \geq 0$ with $\pi_t \overset{t\to\infty}{\Rightarrow} \pi$
  - X_t and Y_{t-L} meet *exactly* at time $\tau^{(L)} := inf\{t > L : X_t = Y_{t-L}\}$



Main Theorem

Consider an L-lag coupling of chains $(X_t, Y_t)_{t\geq 0}$, where $X_t \overset{t\to\infty}{\Rightarrow} \pi$ and meeting time $\tau^{(L)} := inf\{t > L : X_t = Y_{t-L}\}$ has sub-exponential tails. Then,

$$d_{\mathsf{TV}}(\pi_t, \pi) \leq \mathbb{E}\Big[\max(0, \lceil \frac{\tau^{(L)} - L - t}{L} \rceil)\Big] \tag{1}$$

Further assume for some $\eta > 0$, $(2+\eta)$-moments of chain $(X_t)_{t\geq 0}$ are uniformly bounded. Then,

$$d_{\mathsf{W}}(\pi_t, \pi) \leq \mathbb{E}\Big[\sum_{j=1}^{\lceil \frac{\tau^{(L)} - L - t}{L} \rceil} \|X_{t+jL} - Y_{t+(j-1)L}\|_1\Big] \tag{2}$$

Stylized Example

Section 2c

Ising Model: Single-site Gibbs vs. Parallel Tempering

Hamiltonian Monte Carlo

Code and References