Markov Chain Monte Carlo and Couplings
Part III

Pierre E. Jacob

ESSEC BUSINESS SCHOOL

Optimization-Conscious Econometrics Summer School

June 3-6th, 2024

Outline

Outline

Outline

Outline

Outline

Coupling from the past

Propp & Wilson, 1996, *Exact sampling with coupled Markov chains and applications to statistical mechanics*.

Consider a finite state space $\mathbb{X} = \{1, \ldots, n\}$, and a π-invariant, irreducible, aperiodic Markov chain $(X_t)_{t\geq 0}$ on $\mathbb{X}$.

Iterated random function representation:

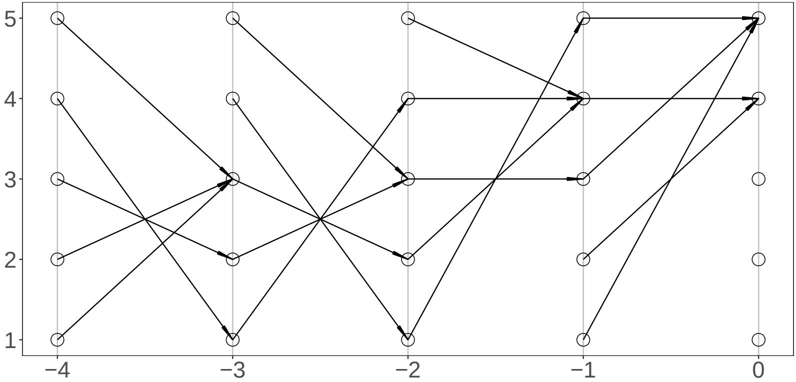
$$\begin{aligned} X_t = \psi(X_{t-1}, U_t) &= \Psi_t(X_{t-1}) \\ &= \Psi_t \circ \Psi_{t-1} \circ \ldots \circ \Psi_1(X_0). \end{aligned}$$

For any t, X_t has the same marginal distribution as

$$Y_t = \Psi_1 \circ \Psi_2 \circ \ldots \circ \Psi_t(X_0),$$

where the random functions $(\Psi_t)_{t\geq 0}$ are in reverse order.

Coupling from the past



Coupling from the past

The process (Y_t) might coalesce, i.e. reach some value $Y_{T^\star}$ at some coalescence time $T^\star$, such that

$$\forall x_0 \in \mathbb{X} \quad \Psi_1 \circ \Psi_2 \circ \ldots \circ \Psi_{T^\star}(x_0) = Y_{T^\star}.$$

Then $Y_t = Y_{T^\star}$ for all $t \geq T^\star$, so

$$Y_{T^\star} = \lim_{t\to\infty} Y_t.$$

But $\mathcal{L}(X_t) = \mathcal{L}(Y_t)$ and $\lim_{t\to\infty} \mathcal{L}(X_t) = \pi$.

So $Y_{T^\star}$, obtained in finite random time, is distributed as π.

Coupling from the past

Algorithm:

Set $t \leftarrow 0$.
Set $\Psi_0 = \Psi_{0\to 0} = \mathrm{Id}$, the identity function.

Repeat:

- $t \leftarrow t - 1$,
- draw U_t and set $\Psi_t = \psi(\cdot, U_t)$,
- set $\Psi_{t\to 0} = \Psi_{t+1\to 0} \circ \Psi_t$,

until $\Psi_{t\to 0}$ is a constant function, i.e. $|\Psi_{t\to 0}(\mathbb{X})| = 1$.

Output $T^\star = t$ and the unique value $Y_{T^\star}$ in the range of $\Psi_{T^\star\to 0}$.

Coupling from the past

It remains to see whether coalescence time $T^\star$ is a.s. finite.

Assumption of irreducibility: $K^k(x \to y) > 0$ for some $k \geq 1$ and all $x, y \in \mathbb{X}$.

Then there is a choice of random functions such that $|\Psi_1 \circ \ldots \circ \Psi_k(\mathbb{X})| = 1$ with positive probability.

For such functions the coalescence time behaves as a Geometric random variable.

This is an assumption on the random function representation: for the same kernel K some random function representation might lead to coalescence, and some might not.

Coupling from the past

Propp & Wilson, 1996, *Exact sampling with coupled Markov chains and applications to statistical mechanics*.

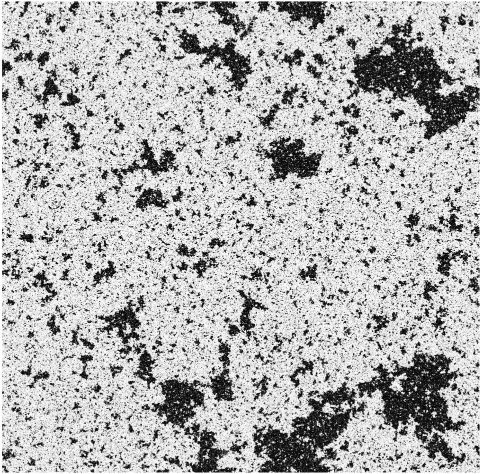


Figure 1: An equilibrated Ising state at the critical temperature on a 4200×4200 toroidal grid.

Outline

Unbiased estimators

Forsythe & Leibler, 1950, *Matrix inversion by a Monte Carlo method.*

How to obtain an unbiased estimator of a series $\sum_{k=0}^{\infty} a_k$?

Note: any limit $\lim_{t\to\infty} b_t$ can be written as a telescoping series.

Unbiased estimators

Forsythe & Leibler, 1950, *Matrix inversion by a Monte Carlo method.*
How to obtain an unbiased estimator of a series $\sum_{k=0}^{\infty} a_k$?
Note: any limit $\lim_{t\to\infty} b_t$ can be written as a telescoping series.

Let K be a random variable on $\{0, 1, \ldots\}$ with $p_k := \mathbb{P}(K = k)$.
Algorithm: sample K and compute:

$$G = \frac{a_K}{p_K}.$$

Unbiased estimators

Forsythe & Leibler, 1950, *Matrix inversion by a Monte Carlo method.*
How to obtain an unbiased estimator of a series $\sum_{k=0}^{\infty} a_k$?
Note: any limit $\lim_{t\to\infty} b_t$ can be written as a telescoping series.

Let K be a random variable on $\{0, 1, \ldots\}$ with $p_k := \mathbb{P}(K = k)$.
Algorithm: sample K and compute:

$$G = \frac{a_K}{p_K}.$$

Useful is

$$\mathbb{E}[K] = \sum_{k\geq 0} k p_k < \infty, \quad \mathbb{E}[G^2] = \sum_{k\geq 0} \frac{a_k^2}{p_k} < \infty.$$

So a_k^2 must decay faster than k^{-3}.

Unbiased estimators

Pioneering contribution in the context of Markov chains: Glynn & Rhee, 2014, *Exact estimation for Markov chain equilibrium expectations*.

In the context of MCMC more specifically:

McLeish, 2011, *A general method for debiasing a Monte Carlo estimator*,

Agapiou, Roberts & Vollmer, 2018, *Unbiased Monte Carlo: Posterior estimation for intractable/infinite-dimensional models*,

Jacob, O'Leary & Atchadé, 2020, *Unbiased Markov chain Monte Carlo with couplings*.

Douc, Jacob, Lee & Vats, 2024, *Solving the Poisson equation using coupled Markov chains*.

Coupled chains with a lag

Sample (X_0, Y_0) from $\bar{\pi}_0$,

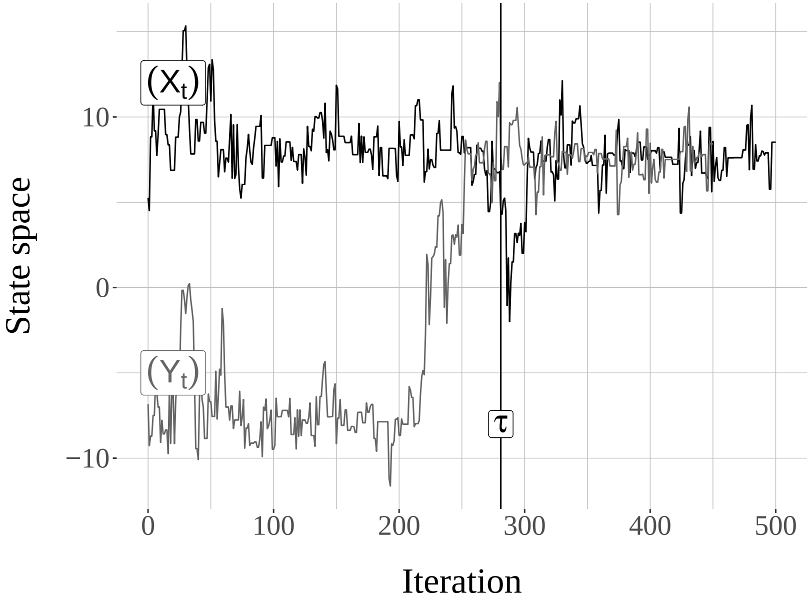
sample $X_t | X_{t-1}$ with transition P, for $t = 1, \ldots, L$,

for $t > L$, sample $(X_t, Y_{t-L}) | (X_{t-1}, Y_{t-L-1})$ with transition $\bar{P}$.

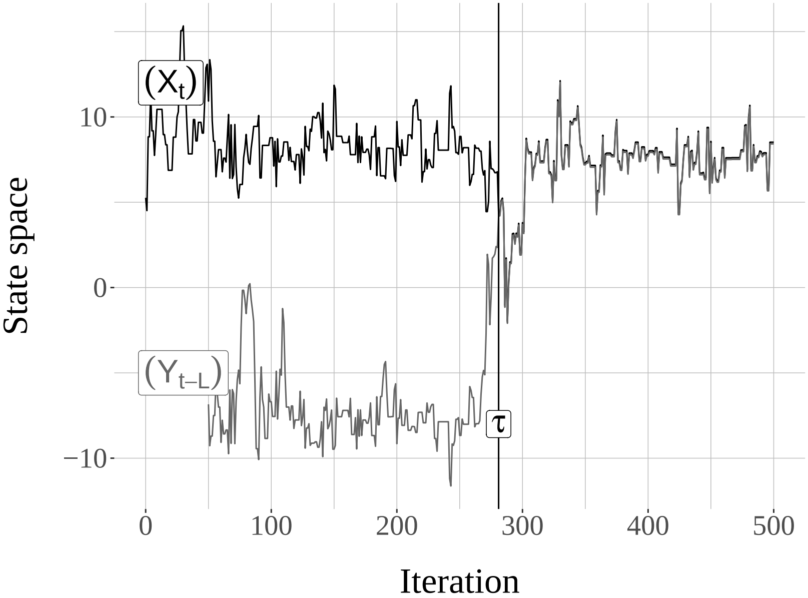
$\bar{\pi}_0$ and $\bar{P}$ must be such that

- X_t and Y_t have the same marginal distribution for all $t \geq 0$,
- $\exists$ finite meeting time τ such that $X_t = Y_{t-L}$ for $t \geq \tau$.

Coupled chains with a lag



Coupled chains with a lag



Unbiased estimators

Limit as a telescopic sum, for all $k \geq 0$,

$$\begin{aligned}\mathbb{E}_\pi[h(X)] &= \lim_{t\to\infty} \mathbb{E}[h(X_t)] \\ &= \mathbb{E}[h(X_k)] + \sum_{j=1}^{\infty} \mathbb{E}[h(X_{k+jL}) - h(X_{k+(j-1)L})].\end{aligned}$$

Unbiased estimators

Limit as a telescopic sum, for all $k \geq 0$,

$$\begin{aligned}\mathbb{E}_\pi[h(X)] &= \lim_{t\to\infty} \mathbb{E}[h(X_t)] \\ &= \mathbb{E}[h(X_k)] + \sum_{j=1}^{\infty} \mathbb{E}[h(X_{k+jL}) - h(X_{k+(j-1)L})].\end{aligned}$$

Since for all $t \geq 0$, X_t and Y_t have the same distribution,

$$= \mathbb{E}[h(X_k)] + \sum_{j=1}^{\infty} \mathbb{E}[h(X_{k+jL}) - h(Y_{k+(j-1)L})].$$

Unbiased estimators

Limit as a telescopic sum, for all $k \geq 0$,

$$\begin{aligned}\mathbb{E}_\pi[h(X)] &= \lim_{t\to\infty} \mathbb{E}[h(X_t)] \\ &= \mathbb{E}[h(X_k)] + \sum_{j=1}^{\infty} \mathbb{E}[h(X_{k+jL}) - h(X_{k+(j-1)L})].\end{aligned}$$

Since for all $t \geq 0$, X_t and Y_t have the same distribution,

$$= \mathbb{E}[h(X_k)] + \sum_{j=1}^{\infty} \mathbb{E}[h(X_{k+jL}) - h(Y_{k+(j-1)L})].$$

If we can swap expectation and limit,

$$= \mathbb{E}[h(X_k) + \sum_{j=1}^{\infty}(h(X_{k+jL}) - h(Y_{k+(j-1)L}))].$$

A first unbiased signed measure

- Choose integers k (warmup) and L (lag).
- Sample (X_0, Y_0) from $\pi_0 \otimes \pi_0$,
- sample $X_t | X_{t-1}$ with transition P, for $t = 1, \ldots, L$,
- then sample $(X_t, Y_{t-L}) | (X_{t-1}, Y_{t-L-1})$ with transition $\bar{P}$ until $t \geq k$ and the chains meet: $X_t = Y_{t-L}$.

A first unbiased signed measure

- Choose integers k (warmup) and L (lag).
- Sample (X_0, Y_0) from $\pi_0 \otimes \pi_0$,
- sample $X_t | X_{t-1}$ with transition P, for $t = 1, \ldots, L$,
- then sample $(X_t, Y_{t-L})|(X_{t-1}, Y_{t-L-1})$ with transition $\bar{P}$ until $t \geq k$ and the chains meet: $X_t = Y_{t-L}$.

The signed measure constructed as

$$\hat{\pi} = \delta_{X_k} + \sum_{j=1}^{\infty} (\delta_{X_{k+jL}} - \delta_{Y_{k+(j-1)L}})$$

is unbiased, in the sense that $\mathbb{E}[\hat{\pi}(h)] = \pi(h)$.

Unbiased MCMC via the Poisson equation

Poisson equation $g - Pg = h - \pi(h)$. Rearranged:

$$\pi(h) = h(y) + Pg(y) - g(y), \quad \forall y.$$

Idea: estimate $\pi(h)$ by estimating $h(y) + Pg(y) - g(y)$ for any y.

We can evaluate h, but
how do we estimate $Pg - g$ point-wise without bias?

Unbiased estimation via the Poisson equation

The function $g_\star : x \mapsto \sum_{t\geq 0} P^t\{h(x) - \pi(h)\}$ is fishy, and we can add any constant.

Unbiased estimation via the Poisson equation

The function $g_\star : x \mapsto \sum_{t\geq 0} P^t\{h(x) - \pi(h)\}$ is fishy, and we can add any constant.

Thus $g_y : x \mapsto g_\star(x) - g_\star(y) = \sum_{t\geq 0}\{P^t h(x) - P^t h(y)\}$ is also fishy for any fixed y.

Unbiased estimation via the Poisson equation

The function $g_\star : x \mapsto \sum_{t\geq 0} P^t\{h(x) - \pi(h)\}$ is fishy, and we can add any constant.

Thus $g_y : x \mapsto g_\star(x) - g_\star(y) = \sum_{t\geq 0}\{P^t h(x) - P^t h(y)\}$ is also fishy for any fixed y.

Define the following estimator of $g_y(x)$:

$$G_y(x) := \sum_{t=0}^{\tau-1}\{h(X_t) - h(Y_t)\},$$

where $X_0 = x$, and $Y_0 = y$, $\tau = \inf\{t \geq 1 : X_t = Y_t\}$.
Under some conditions, we get $\mathbb{E}[G_y(x)] = g_y(x)$ for π-a.e. y.

But we wanted to estimate $Pg - g$.

Unbiased estimation via the Poisson equation

For $y \in \mathbb{X}$, sample $X \sim P(y, \cdot)$ and compute $G_y(X)$. Then

$$\mathbb{E}_y[G_y(X)] = \mathbb{E}_y[g_y(X)] = \mathbb{E}_y[g_\star(X) - g_\star(y)] = Pg_\star(y) - g_\star(y).$$

Thus $h(y) + g_y(X)$ is an unbiased estimator of $\pi(h)$.

Turns out to be exactly equivalent to Glynn & Rhee's unbiased estimator, with $\pi_0 = \delta_y$, $L = 1$.

Unbiased MCMC: a more efficient approximation

- Choose integers k (warmup), ℓ (length) and L (lag).
- Sample (X_0, Y_0) from $\pi_0 \otimes \pi_0$,
- sample $X_t | X_{t-1}$ with transition P, for $t = 1, \ldots, L$,
- then sample $(X_t, Y_{t-L}) | (X_{t-1}, Y_{t-L-1})$ with transition $\bar{P}$ until $t \geq \ell$ and the chains meet: $X_t = Y_{t-L}$.

Unbiased MCMC: a more efficient approximation

- Choose integers k (warmup), ℓ (length) and L (lag).
- Sample (X_0, Y_0) from $\pi_0 \otimes \pi_0$,
- sample $X_t | X_{t-1}$ with transition P, for $t = 1, \ldots, L$,
- then sample $(X_t, Y_{t-L}) | (X_{t-1}, Y_{t-L-1})$ with transition $\bar{P}$ until $t \geq \ell$ and the chains meet: $X_t = Y_{t-L}$.

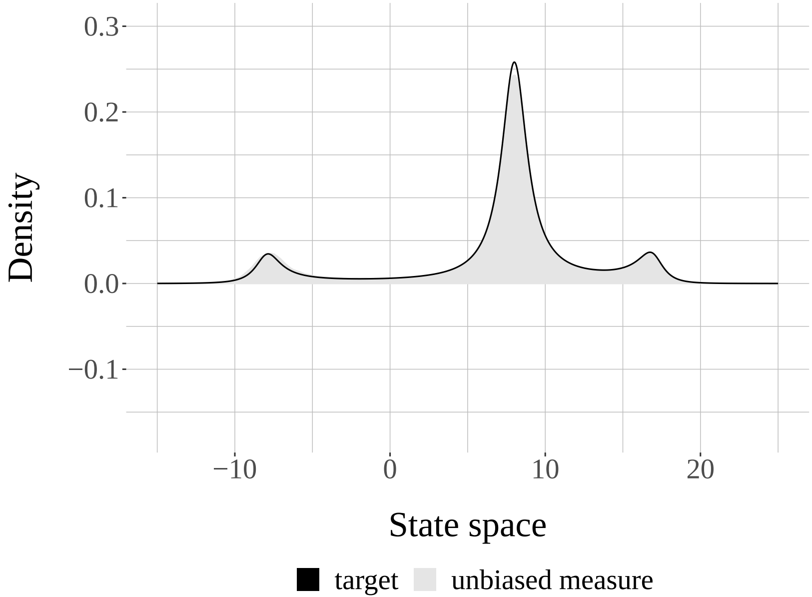
Unbiased signed measure is constructed as

$$\hat{\pi} = \frac{1}{\ell - k + 1} \sum_{t=k}^{\ell} \delta_{X_t} + \sum_{t=k+L}^{\tau-1} \frac{v_t}{\ell - k + 1} \left(\delta_{X_t} - \delta_{Y_{t-L}}\right)$$

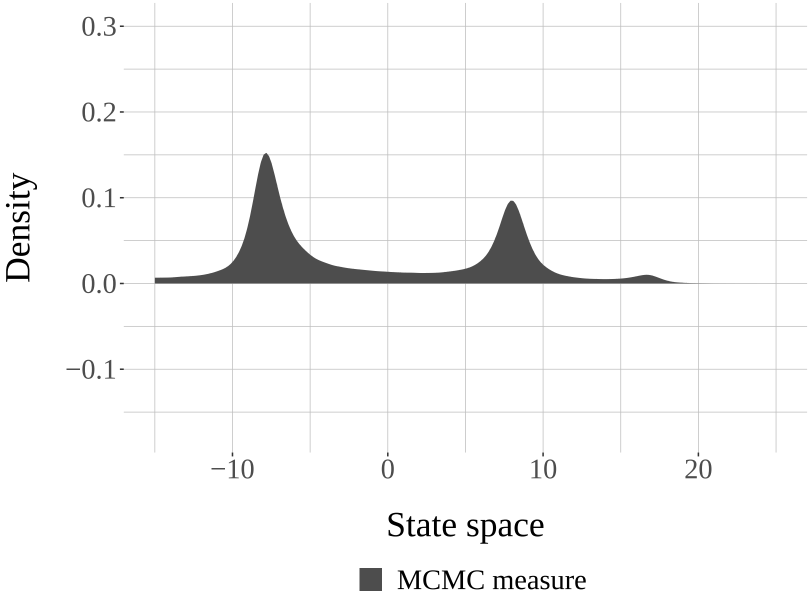
with

$$v_t = \lfloor (t-k)/L \rfloor - \lceil \max(L, t-\ell)/L \rceil + 1.$$

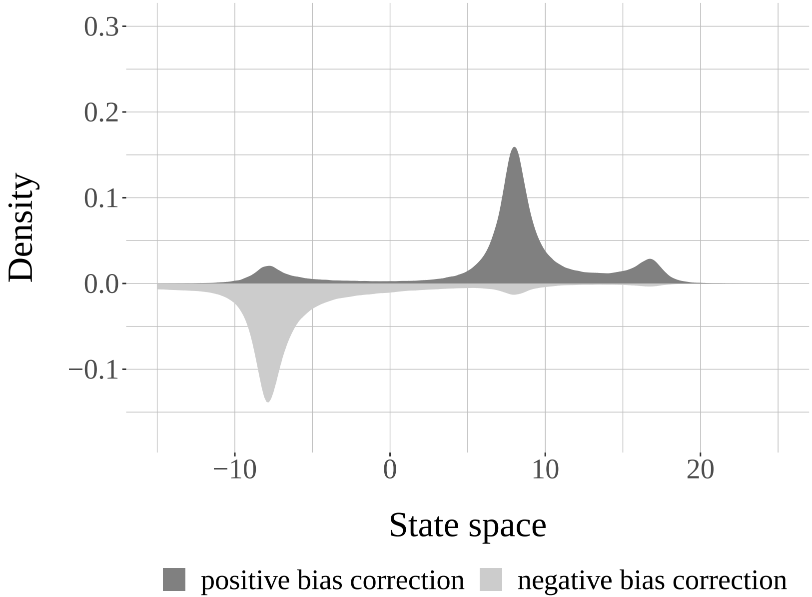
Signed measure



Signed measure



Signed measure



Assumption on meeting time

Main assumption. For some $\kappa > 1$, $\mathbb{E}_{\pi\otimes\pi}[\tau^\kappa] < \infty$.

Equivalent to $\mathbb{P}(\tau > t)$ being at most $t^{-\kappa}$ as $t \to \infty$.

Holds for all $\kappa > 1$ if tails of τ are Geometric.

Finite moments

Let $h \in L^m(\pi)$ (i.e. $\int |h(x)|^m d\pi(x) < \infty$) for some $m > \kappa/(\kappa - 1)$.

Denote by $\hat{\pi}(h)$ the unbiased MCMC estimator of $\pi(h)$, for any integers k, ℓ, L.

Then $\mathbb{E}[\hat{\pi}(h)] = \pi(h)$,
and for $p \geq 1$ such that $\frac{1}{p} > \frac{1}{m} + \frac{1}{\kappa}$, $\mathbb{E}[|\hat{\pi}(h)|^p] < \infty$.

Efficiency

If $\mathbb{E}[|\hat{\pi}(h)|^2] < \infty$, then one can estimate $\pi(h)$ by averaging R independent copies of $\hat{\pi}(h)$, as well as its variance.

The performance of unbiased estimators can be measured with

$$\text{asymptotic inefficiency} = \mathbb{E}[\text{cost}] \times \text{variance}.$$

The tuning parameters k, ℓ and L affect cost and efficiency.

Asymptotic variance

Let $h \in L^m(\pi)$ for some $m > 2\kappa/(\kappa-1)$. Then for any $k \in \mathbb{N}$,

$$\sqrt{\ell - k + 1}\,(\hat{\pi}(h) - \pi(h)) \xrightarrow[\ell\to\infty]{d} \text{Normal}(0, v(P,h)),$$

where $v(P,h)$ is the same asymptotic variance as in

$$\sqrt{t}(t^{-1}\sum_{s=0}^{t-1} h(X_s) - \pi(h)) \xrightarrow[t\to\infty]{d} \text{Normal}(0, v(P,h)).$$

Asymptotic variance

Let $h \in L^m(\pi)$ for some $m > 2\kappa/(\kappa-1)$. Then for any $k \in \mathbb{N}$,

$$\sqrt{\ell - k + 1}\,(\hat{\pi}(h) - \pi(h)) \xrightarrow[\ell\to\infty]{d} \text{Normal}(0, v(P,h)),$$

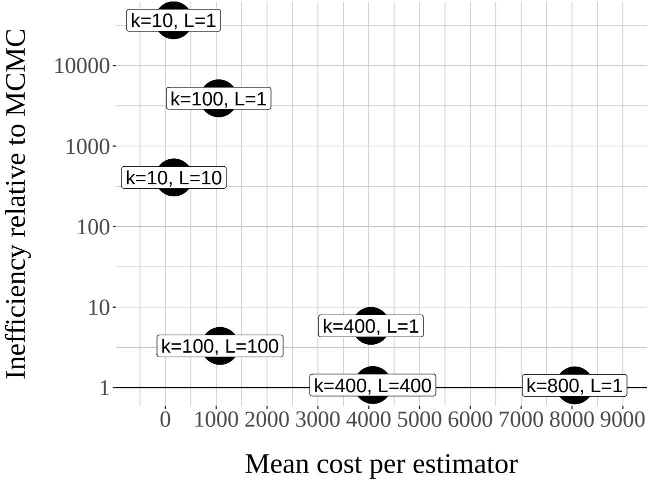
where $v(P,h)$ is the same asymptotic variance as in

$$\sqrt{t}(t^{-1}\sum_{s=0}^{t-1} h(X_s) - \pi(h)) \xrightarrow[t\to\infty]{d} \text{Normal}(0, v(P,h)).$$

By choosing ℓ large enough, we can make sure that the variance of unbiased MCMC is nearly that of standard MCMC.

The cost of unbiased MCMC is $\max(L, \ell - (\tau - L)) + 2(\tau - L)$. This is equivalent to ℓ as $\ell \to \infty$.

Performance for different (k, L) with $\ell = 10k$



Outline

Upper bounds using couplings

Johnson, 1996, *Studying convergence of Markov chain Monte Carlo algorithms using coupled sample paths.*
Johnson, 1998, *A coupling-regeneration scheme for diagnosing convergence in Markov chain Monte Carlo algorithms.*

Generate $C \gg 1$ chains until they all coincide.

Requires some knowledge about distance between π_0 and π, namely a value r such that, in a rejection sampler, a draw X from π_0 would be accepted as a draw from π with probability at least $1 - r$.

Then

$$|\pi_t - \pi|_{\text{TV}} \leq \mathbb{P}(\tau_C > t)(1 - r^C)^{-1}.$$

Upper bounds using couplings

Distributions π_{t+jL} with $j \geq 1$ between π_t and $\pi = \pi_\infty$
+ triangle inequalities + coupling representation of TV:

$$
\begin{aligned}
|\pi_t - \pi|_{\mathrm{TV}} &\leq \sum_{j=1}^{\infty} |\pi_{t+jL} - \pi_{t+(j-1)L}|_{\mathrm{TV}} \\
&\leq \sum_{j=1}^{\infty} \mathbb{P}(X_{t+jL} \neq Y_{t+(j-1)L}) \\
&= \sum_{j=1}^{\infty} \mathbb{E}[\mathbb{1}(\tau > t + jL)].
\end{aligned}
$$

Upper bounds using couplings

Distributions π_{t+jL} with $j \geq 1$ between π_t and $\pi = \pi_\infty$
+ triangle inequalities + coupling representation of TV:

$$\begin{aligned} |\pi_t - \pi|_{\mathrm{TV}} &\leq \sum_{j=1}^{\infty} |\pi_{t+jL} - \pi_{t+(j-1)L}|_{\mathrm{TV}} \\ &\leq \sum_{j=1}^{\infty} \mathbb{P}(X_{t+jL} \neq Y_{t+(j-1)L}) \\ &= \sum_{j=1}^{\infty} \mathbb{E}[\mathbb{1}(\tau > t + jL)]. \end{aligned}$$

Swapping infinite sum and expectation,

$$|\pi_t - \pi|_{\mathrm{TV}} \leq \mathbb{E}[\max(0, \lceil (\tau - L - t)/L \rceil)].$$

Upper bounds using couplings

Upper bounds for all $t \geq 0$:

$$|\pi_t - \pi|_{\mathrm{TV}} \leq \mathbb{E}[\max(0, \lceil(\tau - L - t)/L\rceil)].$$

Zero assumptions on π_0.

We can then estimate the right-hand side for all t, by generating $\tau^{(1)}, \ldots, \tau^{(R)}$ independently and averaging.

Upper bounds using couplings

Upper bounds for all $t \geq 0$:

$$|\pi_t - \pi|_{\mathrm{TV}} \leq \mathbb{E}[\max(0, \lceil(\tau - L - t)/L\rceil)].$$

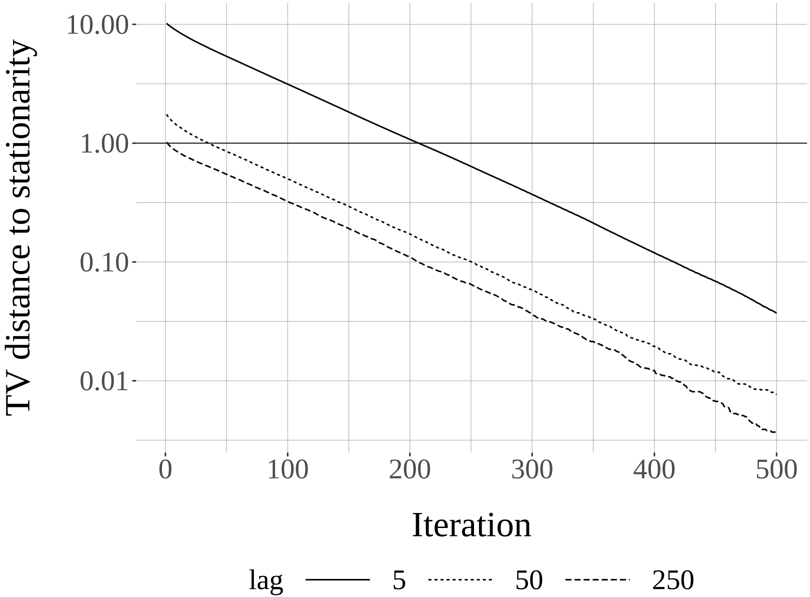
Zero assumptions on π_0.

We can then estimate the right-hand side for all t, by generating $\tau^{(1)}, \ldots, \tau^{(R)}$ independently and averaging.

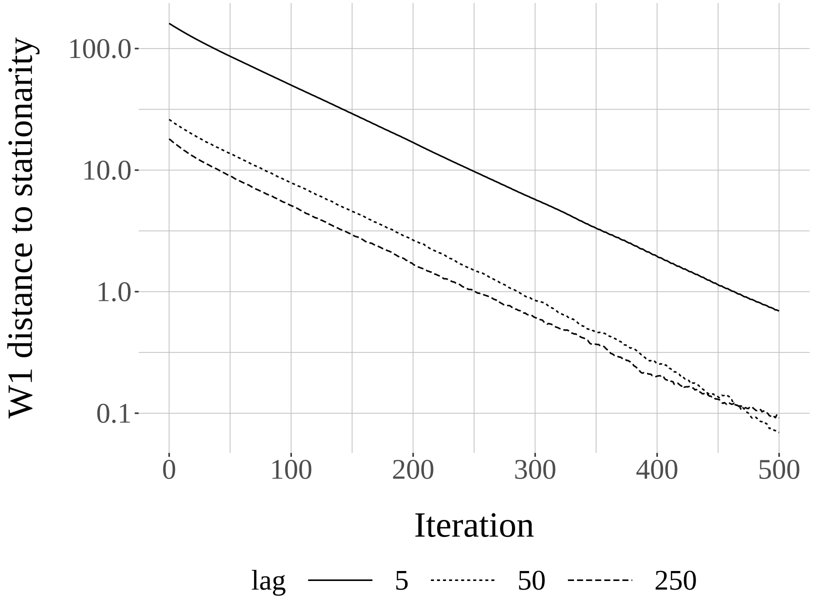
Similarly, can provide upper bounds on Wasserstein distances.

Biswas, Jacob & Vanetti, 2019, *Estimating Convergence of Markov chains with L-Lag Couplings.*
Craiu & Meng, 2020, *Double happiness: Enhancing the coupled gains of L-lag coupling via control variates.*
Papp & Sherlock, 2022, *Bounds on Wasserstein distances between continuous distributions using independent samples.*

Upper bounds using couplings



Upper bounds using couplings



Asymptotic variance and fishy functions

Using the Poisson equation to establish a CLT for Markov chain ergodic averages leads to the following expression

$$v(P,h) = \mathbb{E}_{\pi}[\{g(X_1) - Pg(X_0)\}^2].$$

Asymptotic variance and fishy functions

Using the Poisson equation to establish a CLT for Markov chain ergodic averages leads to the following expression

$$v(P,h) = \mathbb{E}_\pi[\{g(X_1) - Pg(X_0)\}^2].$$

Algebraic manipulations (expand square, note $\mathbb{E}[g(X_1)Pg(X_0)] = \mathbb{E}[\{Pg(X_0)\}^2]$, replace Pg by $g - (h - \pi(h))$) obtain

$$v(P,h) = 2\pi(\{h - \pi(h)\}g) - (\pi(h^2) - \pi(h)^2).$$

Asymptotic variance and fishy functions

Asymptotic variance:

$$v(P, h) = 2\pi(\{h - \pi(h)\}g) - (\pi(h^2) - \pi(h)^2).$$

Asymptotic variance and fishy functions

Asymptotic variance:

$$v(P,h) = 2\pi(\{h - \pi(h)\}g) - (\pi(h^2) - \pi(h)^2).$$

Using successful coupled chains, we can obtain:

1. unbiased MCMC approximation $\hat{\pi} = \sum_{n=1}^{N} \omega_n \delta_{Z_n}$ of π,

Asymptotic variance and fishy functions

Asymptotic variance:

$$v(P, h) = 2\pi(\{h - \pi(h)\}g) - (\pi(h^2) - \pi(h)^2).$$

Using successful coupled chains, we can obtain:

1. unbiased MCMC approximation $\hat{\pi} = \sum_{n=1}^{N} \omega_n \delta_{Z_n}$ of π,
2. unbiased estimator of g, point-wise:

$$\text{fishy evaluation:} \quad g_y(x) = \sum_{t=0}^{\infty} \{P^t h(x) - P^t h(y)\},$$

$$\text{its estimator:} \quad G_y(x) = \sum_{t=0}^{\infty} \{h(X_t) - h(Y_t)\}.$$

Asymptotic variance estimator

To estimate $$v(P,h) = 2\,\underbrace{\pi(\{h-\pi(h)\}g)}_{(a)} - \underbrace{(\pi(h^2)-\pi(h)^2)}_{(b)}.$$

1. Obtain $\hat{\pi}^{(1)}$ and $\hat{\pi}^{(2)}$, two independent approximations of π.
2. Compute (B) $= \frac{1}{2}(\hat{\pi}^{(1)}(h^2)+\hat{\pi}^{(2)}(h^2)) - \hat{\pi}^{(1)}(h)\times\hat{\pi}^{(2)}(h)$.

Asymptotic variance estimator

To estimate $$v(P,h) = 2\,\underbrace{\pi(\{h-\pi(h)\}g)}_{(a)} - \underbrace{(\pi(h^2)-\pi(h)^2)}_{(b)}.$$

1. Obtain $\hat{\pi}^{(1)}$ and $\hat{\pi}^{(2)}$, two independent approximations of π.
2. Compute (B) $= \frac{1}{2}(\hat{\pi}^{(1)}(h^2) + \hat{\pi}^{(2)}(h^2)) - \hat{\pi}^{(1)}(h) \times \hat{\pi}^{(2)}(h)$.
3. Write $\hat{\pi}^{(1)}(\cdot) = \sum_{n=1}^{N} \omega_n \delta_{Z_n}$.
4. Sample I uniformly in $\{1,\ldots,N\}$.
5. Generate $G(Z_I)$ with expectation $g(Z_I)$.
6. Compute (A) $= N\omega_I(h(Z_I) - \hat{\pi}^{(2)}(h))G(Z_I)$.

Asymptotic variance estimator

To estimate $$v(P,h) = 2\,\underbrace{\pi(\{h-\pi(h)\}g)}_{(a)} - \underbrace{(\pi(h^2)-\pi(h)^2)}_{(b)}.$$

1. Obtain $\hat{\pi}^{(1)}$ and $\hat{\pi}^{(2)}$, two independent approximations of π.
2. Compute (B) $= \frac{1}{2}(\hat{\pi}^{(1)}(h^2) + \hat{\pi}^{(2)}(h^2)) - \hat{\pi}^{(1)}(h) \times \hat{\pi}^{(2)}(h)$.
3. Write $\hat{\pi}^{(1)}(\cdot) = \sum_{n=1}^{N} \omega_n \delta_{Z_n}$.
4. Sample I uniformly in $\{1,\ldots,N\}$.
5. Generate $G(Z_I)$ with expectation $g(Z_I)$.
6. Compute (A) $= N\omega_I(h(Z_I) - \hat{\pi}^{(2)}(h))G(Z_I)$.
7. Output $\hat{v}(P,h) = 2(\text{A}) - (\text{B})$.

Asymptotic variance estimator

Assume that $\mathbb{E}_{\pi\otimes\pi}[\tau^\kappa] < \infty$ for some $\kappa > 2$.

Let $h \in L^m(\pi)$ for some $m > 2\kappa/(\kappa - 2)$.

Then $\mathbb{E}\left[\hat{v}(P, h)\right] = v(P, h)$,

and for $p \geq 1$ such that $\frac{1}{p} > \frac{2}{m} + \frac{2}{\kappa}$, $\mathbb{E}\left[|\hat{v}(P, h)|^p\right] < \infty$.

Douc, Jacob, Lee & Vats, 2024, *Solving the Poisson equation using coupled Markov chains.*
Atchadé & Jacob, 2024, *Unbiased Markov chain Monte Carlo: what, why and how.*

Coming soon...

- If we can design couplings of MCMC algorithms such that pairs of chains *meet* exactly in finite time, then we can...

  - obtain unbiased estimators of $\pi(h)$,

  - obtain upper bounds on the distance to stationarity,

  - obtain unbiased estimators of the asymptotic variance in the CLT for MCMC averages.

- It remains to see how to implement such couplings!

Outline