

Averaged LMC FI lower bound

August 10, 2026

1 Setting and main results

As usual, let $\pi \propto \exp(-V)$, where

$$Z_V := \int_{\mathbb{R}^d} \exp(-V(x)) \, dx < \infty \quad \text{and} \quad \sup_{x \in \mathbb{R}^d} \|\nabla^2 V(x)\|_{\text{op}} \leq \beta. \quad (1.1)$$

For a step size $h > 0$, Langevin Monte Carlo (LMC) is the recursion

$$\hat{X}_{(n+1)h} = \hat{X}_{nh} - h\nabla V(\hat{X}_{nh}) + \sqrt{2h} Z_{n+1}, \quad Z_{n+1} \sim \mathcal{N}(0, I_d). \quad (1.2)$$

Its standard continuous interpolation freezes the drift on each step: for $0 \leq u \leq h$,

$$\hat{X}_{nh+u} = \hat{X}_{nh} - u\nabla V(\hat{X}_{nh}) + \sqrt{2} (B_{nh+u} - B_{nh}). \quad (1.3)$$

Here the Brownian increments in successive steps are independent, and the value at $u = h$ is the next grid point in (1.2). Let $\hat{\mu}_t$ denote the law of $\hat{X}_t$, and define the continuously averaged output law by

$$\bar{\mu}_{N,h} := \frac{1}{Nh} \int_0^{Nh} \hat{\mu}_t \, dt. \quad (1.4)$$

Equivalently, one samples a time uniformly from $[0, Nh]$ and outputs the interpolated chain at that time.

The following theorem shows that the Fisher information analysis of averaged LMC is sharp. The two terms in the lower bound are obtained from different instances; for any fixed values of the parameters, one chooses the instance corresponding to the larger term.

Theorem 1.1 (Sharpness of FI bounds for averaged LMC). There are universal constants $c, c_0 > 0$ with the following property. For every $N, d \geq 1$, $K_0 > 0$, $\beta > 0$, and $0 < h \leq c_0/\beta$,

$$\sup_{\substack{V \text{ satisfies (1.1)} \\ \text{KL}(\mu_0 \parallel \pi) \leq K_0}} \text{FI}(\bar{\mu}_{N,h} \parallel \pi) \geq c \max \left\{ K_0 \min \left\{ \beta, \frac{1}{Nh} \right\}, \beta^2 dh \right\}. \quad (1.5)$$

Moreover, for every $h \geq c_0/\beta$,

$$\sup_{\substack{V \text{ satisfies (1.1)} \\ \text{KL}(\mu_0 \parallel \pi) \leq K_0}} \text{FI}(\bar{\mu}_{N,h} \parallel \pi) \geq c\beta d. \quad (1.6)$$

For comparison, Theorem 11.2.1 in [Che26], together with the convexity of Fisher information recalled immediately afterward, gives the following averaged-law upper bound.

Proposition 1.1 (Averaged LMC upper bound). If $\text{KL}(\mu_0 \parallel \pi) \leq K_0$ and $0 < h \leq 1/(4\beta)$, then

$$\text{FI}(\bar{\mu}_{N,h} \parallel \pi) \leq \frac{2K_0}{Nh} + 6\beta^2 dh. \quad (1.7)$$

Thus, whenever $Nh \geq \beta^{-1}$ and $h \leq c_0/\beta$, Theorem 1.1 and Proposition 1.1 give

$$\sup_{\substack{V \text{ satisfies (1.1)} \\ \text{KL}(\mu_0 \parallel \pi) \leq K_0}} \text{FI}(\bar{\mu}_{N,h} \parallel \pi) \asymp \max \left\{ \frac{K_0}{Nh}, \beta^2 dh \right\}. \quad (1.8)$$

The constants implicit in $\asymp$ are universal.

Corollary 1.1. There are universal constants $c_1, c_2, c, C > 0$ such that, if

$$c_1 \frac{d}{N} \leq K_0 \leq c_2 dN, \quad (1.9)$$

then

$$c\beta \sqrt{\frac{dK_0}{N}} \leq \inf_{h>0} \sup_{\substack{V \text{ satisfies (1.1)} \\ \text{KL}(\mu_0 \parallel \pi) \leq K_0}} \text{FI}(\bar{\mu}_{N,h} \parallel \pi) \leq C\beta \sqrt{\frac{dK_0}{N}}. \quad (1.10)$$

Corollary 1.2. There is a universal constant $c > 0$ such that, in the regime

$$0 < \varepsilon^2 \leq c\beta \min\{K_0, d\}, \quad (1.11)$$

any fixed-step continuously averaged LMC guarantee which holds uniformly over the class in (1.1) and ensures

$$\text{FI}(\bar{\mu}_{N,h} \parallel \pi) \leq \varepsilon^2$$

must have

$$N \geq c \frac{\beta^2 d K_0}{\varepsilon^4}. \quad (1.12)$$

2 A Stein lower bound

We first record the Stein lower bound used for all three constructions.

Lemma 2.1 (Stein lower bound). Let $\pi \propto \exp(-V)$ and let μ have the integrability needed for the expressions below. If $\Phi: \mathbb{R}^d \rightarrow \mathbb{R}^d$ is a smooth vector field, then

$$\text{FI}(\mu \parallel \pi) \geq \frac{(\mathbb{E}_\mu[\text{div } \Phi - \langle \nabla V, \Phi \rangle])^2}{\mathbb{E}_\mu \|\Phi\|^2}. \quad (2.1)$$

The assertion is interpreted as trivial if the Fisher information is infinite or the denominator vanishes.

Proof. Suppose first that Φ is compactly supported and that the relative score s , defined by

$$s := \nabla \log \frac{d\mu}{d\pi},$$

belongs to $L^2(\mu)$. Integration by parts gives

$$\mathbb{E}_\mu[\text{div } \Phi - \langle \nabla V, \Phi \rangle] = -\mathbb{E}_\mu \langle \Phi, s \rangle.$$

The result follows from Cauchy–Schwarz. For the bounded periodic vector fields used below, and for the linear field used with Gaussian laws, apply this identity to smooth compactly supported cutoffs and let the cutoff radius tend to infinity. All the laws in question have finite moments of every required order, so dominated convergence applies. $\square$

3 The finite-time term

Fix $\alpha \in (0, \beta]$ and consider

$$\pi_\alpha = \mathcal{N}(0, \alpha^{-1}I_d) \quad \text{and} \quad \mu_0 = \mathcal{N}(m, \alpha^{-1}I_d),$$

where m is chosen so that

$$\frac{\alpha}{2} \|m\|^2 = K_0.$$

Then $\text{KL}(\mu_0 \parallel \pi_\alpha) = K_0$. We have

$$\begin{aligned} \mathbb{E} \hat{X}_{nh} &= (1 - \alpha h)^n m, \\ \mathbb{E} \hat{X}_{nh+u} &= (1 - \alpha u) (1 - \alpha h)^n m. \end{aligned}$$

It follows that the mean $\bar{m}$ of $\bar{\mu}_{N,h}$ is

$$\bar{m} = m \left(1 - \frac{\alpha h}{2}\right) \frac{1 - (1 - \alpha h)^N}{N\alpha h}. \quad (3.1)$$

Applying Lemma 2.1 with constant vector fields gives, for every μ with finite first moment,

$$\text{FI}(\mu \parallel \pi_\alpha) \geq \alpha^2 \left\| \int x \mu(dx) \right\|^2. \quad (3.2)$$

Therefore,

$$\text{FI}(\bar{\mu}_{N,h} \parallel \pi_\alpha) \geq 2\alpha K_0 \left[\left(1 - \frac{\alpha h}{2}\right) \frac{1 - (1 - \alpha h)^N}{N\alpha h} \right]^2. \quad (3.3)$$

Choose

$$\alpha = \min \left\{ \beta, \frac{1}{Nh} \right\}.$$

If $h \leq 1/(4\beta)$, then $0 \leq \alpha h \leq 1/4$ and $0 \leq N\alpha h \leq 1$. For $q \in [0, 1]$ and $z \in [0, 1]$,

$$1 - (1 - q)^N \geq 1 - e^{-Nq} \quad \text{and} \quad \frac{1 - e^{-z}}{z} \geq 1 - e^{-1}.$$

Hence the bracket in (3.3) is bounded below by a positive universal constant, proving

$$\text{FI}(\bar{\mu}_{N,h} \parallel \pi_\alpha) \gtrsim K_0 \min \left\{ \beta, \frac{1}{Nh} \right\}. \quad (3.4)$$

4 The discretization term

The hard instance has curvature oscillating on the Brownian scale $\sqrt{h}$. We first isolate the periodic calculation.

4.1 A periodic Fourier calculation

Fix $0 < \delta \leq 1$. On the circle $\mathbb{R}/(2\pi\mathbb{Z})$, let

$$p_\delta(d\theta) \propto \exp(\delta \cos \theta) d\theta. \quad (4.1)$$

Consider the chain

$$\Theta_{n+1} = \Theta_n - \delta \sin \Theta_n + \sqrt{2} Z_{n+1} \pmod{2\pi}, \quad \Theta_0 \sim p_\delta, \quad (4.2)$$

and its within-step interpolation

$$\Theta_{n,s} = \Theta_n - s\delta \sin \Theta_n + \sqrt{2s} Z \pmod{2\pi}, \quad 0 \leq s \leq 1. \quad (4.3)$$

Let $\bar{p}_{N,\delta}$ be the law obtained by averaging (4.3) uniformly over $n \in \{0, \dots, N-1\}$ and $s \in [0, 1]$.

Lemma 4.1 (Non-vanishing periodic Stein bias). There are universal constants $c, \delta_0 > 0$ such that, for every $N \geq 1$ and $0 < \delta \leq \delta_0$,

$$|\mathbb{E}_{\bar{p}_{N,\delta}} \cos \Theta - \delta \mathbb{E}_{\bar{p}_{N,\delta}} \sin^2 \Theta| \geq c\delta. \quad (4.4)$$

Proof. Set

$$c_{k,n} := \mathbb{E} \exp(ik\Theta_n).$$

The laws in (4.2) are invariant under reflection, so all the Fourier coefficients used below are real. Gaussian convolution gives

$$c_{k,n+1} = e^{-k^2} \mathbb{E} \exp(ik\Theta_n - ik\delta \sin \Theta_n). \quad (4.5)$$

We calculate the initial coefficients explicitly. Put

$$D_\delta := \frac{1}{2\pi} \int_0^{2\pi} e^{\delta \cos \theta} d\theta.$$

Taylor's theorem, with a remainder uniform over $0 \leq \delta \leq 1$ and $\theta \in [0, 2\pi]$, together with Fourier orthogonality, gives

$$\begin{aligned} D_\delta &= 1 + \frac{\delta^2}{4} + O(\delta^4), \\ \frac{1}{2\pi} \int_0^{2\pi} e^{i\theta} e^{\delta \cos \theta} d\theta &= \frac{\delta}{2} + O(\delta^3), \\ \frac{1}{2\pi} \int_0^{2\pi} e^{2i\theta} e^{\delta \cos \theta} d\theta &= \frac{\delta^2}{8} + O(\delta^4). \end{aligned}$$

Indeed, in the second line the only contribution through order two is

$$\frac{\delta}{2\pi} \int_0^{2\pi} e^{i\theta} \cos \theta d\theta = \frac{\delta}{2},$$

whereas in the third line the first non-zero contribution is

$$\frac{\delta^2}{4\pi} \int_0^{2\pi} e^{2i\theta} \cos^2 \theta d\theta = \frac{\delta^2}{8}.$$

Dividing the last two displays by D_δ therefore yields

$$c_{1,0} = \frac{\delta}{2} + O(\delta^3) \quad \text{and} \quad c_{2,0} = \frac{\delta^2}{8} + O(\delta^4) = O(\delta^2). \quad (4.6)$$

Here and below the constants in $O(\cdot)$ are universal. The bound $|e^{-2i\delta \sin \theta} - 1| \leq 2\delta$ and (4.5) imply

$$|c_{2,n+1}| \leq e^{-4} |c_{2,n}| + 2e^{-4}\delta.$$

Iteration and (4.6) show that

$$c_{2,n} = O(\delta) \quad \text{uniformly in } n. \quad (4.7)$$

Since

$$-ie^{i\theta} \sin \theta = \frac{1 - e^{2i\theta}}{2}$$

and the remainder in $e^{-i\delta \sin \theta} = 1 - i\delta \sin \theta + O(\delta^2)$ is uniform in θ , equations (4.5) and (4.7) yield

$$c_{1,n+1} = e^{-1} c_{1,n} + \frac{\delta e^{-1}}{2} + O(\delta^2).$$

Together with (4.6), this recursion implies directly that there is a universal constant C such that

$$c_{1,n} \leq \frac{\delta}{2} + C\delta^2 \quad \text{uniformly in } n.$$

Indeed, choose C large enough to dominate the initial error and the uniform remainder in the recursion. If the bound holds at time n , then its leading-order part at time $n+1$ is at most $\delta/e < \delta/2$, while the error is at most $(C/e + O(1))\delta^2$; enlarging C closes the induction.

The same uniform Taylor expansion in (4.3) gives

$$\mathbb{E} e^{i\Theta_{n,s}} = e^{-s} \left[c_{1,n} + \frac{s\delta}{2} (1 - c_{2,n}) \right] + O(\delta^2).$$

All remainders here are uniform in n and $s \in [0, 1]$. Similarly,

$$|\mathbb{E} e^{2i\Theta_{n,s}}| \leq e^{-4s} (|c_{2,n}| + 2s\delta) = O(\delta).$$

Hence $\mathbb{E} \sin^2 \Theta_{n,s} = 1/2 + O(\delta)$, and therefore, uniformly in n and s ,

$$\mathbb{E} [\cos \Theta_{n,s} - \delta \sin^2 \Theta_{n,s}] \leq \frac{\delta}{2} [e^{-s} (1 + s) - 1] + O(\delta^2).$$

For every n , integration over the fractional time gives

$$\frac{1}{2} \int_0^1 [e^{-s} (1 + s) - 1] ds = -\frac{3 - e}{2e}.$$

Averaging over n now gives

$$\mathbb{E}_{\tilde{p}_{N,\delta}} [\cos \Theta - \delta \sin^2 \Theta] \leq -\frac{3 - e}{2e} \delta + O(\delta^2).$$

The remainder is uniform in N . Since $3 - e > 0$, the claim follows after decreasing δ_0 by a universal amount. $\square$

4.2 A normalizable target on Euclidean space

Fix a sufficiently small universal constant $\eta > 0$ and set

$$\delta := \eta\beta h. \tag{4.8}$$

For $\rho > 0$, define the product potential

$$V_{\rho,h}(x) := \sum_{i=1}^d \left[\frac{\rho}{2h} x_i^2 - \delta \cos \frac{x_i}{\sqrt{h}} \right]. \quad (4.9)$$

Its Hessian is diagonal, with entries

$$\frac{\rho}{h} + \eta\beta \cos \frac{x_i}{\sqrt{h}}.$$

In particular, if $0 < \eta < 1$ and $0 < \rho \leq (1 - \eta)\beta h$, then

$$\|\nabla^2 V_{\rho,h}(x)\|_{\text{op}} \leq \beta$$

for every x . The quadratic term makes the target normalizable. Initialize LMC at the target:

$$\mu_0 = \pi_{\rho,h} \propto \exp(-V_{\rho,h}), \quad \text{KL}(\mu_0 \parallel \pi_{\rho,h}) = 0. \quad (4.10)$$

The initialization is stationary only for the diffusion, not for the Euler chain.

It is enough to analyze one coordinate. Under the scaling $Y = \hat{X}/\sqrt{h}$, its target density is

$$q_\rho(y) \propto \exp\left(-\frac{\rho}{2} y^2 + \delta \cos y\right), \quad (4.11)$$

and the LMC recursion becomes

$$Y_{n+1} = (1 - \rho)Y_n - \delta \sin Y_n + \sqrt{2} Z_{n+1}. \quad (4.12)$$

For $0 \leq s \leq 1$, its interpolation is

$$Y_{n,s} = (1 - s\rho)Y_n - s\delta \sin Y_n + \sqrt{2s} Z_{n+1}. \quad (4.13)$$

We now justify carefully the limit from (4.12) to the periodic chain.

Lemma 4.2. Fix N and δ . As $\rho \searrow 0$, the phase of the initial law $Y_0 \bmod 2\pi$ converges in distribution to p_δ . Moreover, for every $0 \leq n < N$ and $s \in [0, 1]$,

$$Y_{n,s} \bmod 2\pi \xrightarrow{d} \Theta_{n,s},$$

where $\Theta_{n,s}$ is defined in (4.3). Finally,

$$\max_{0 \leq n < N} \sup_{0 \leq s \leq 1} \rho \mathbb{E} |Y_{n,s}| \longrightarrow 0. \quad (4.14)$$

Proof. The density of $Y_0 \bmod 2\pi$ is proportional on $[0, 2\pi)$ to

$$e^{\delta \cos \theta} \sum_{k \in \mathbb{Z}} \exp\left(-\frac{\rho}{2} (\theta + 2\pi k)^2\right).$$

By the Poisson summation formula,

$$\sqrt{2\pi\rho} \sum_{k \in \mathbb{Z}} \exp\left(-\frac{\rho}{2} (\theta + 2\pi k)^2\right) = \sum_{m \in \mathbb{Z}} e^{-m^2/(2\rho)} e^{im\theta}.$$

The right-hand side converges uniformly in θ to 1, proving the initial phase convergence.

Since $e^{-\delta} \leq e^{\delta \cos y} \leq e^{\delta}$, comparison with a centered Gaussian of variance ρ^{-1} gives

$$\mathbb{E} |Y_0| = O(\rho^{-1/2}). \quad (4.15)$$

For $0 < \rho \leq 1$, recursion (4.12) and its interpolation then imply, for fixed N ,

$$\max_{0 \leq n < N} \sup_{0 \leq s \leq 1} \mathbb{E} |Y_{n,s}| = O(\rho^{-1/2}) + O_N(1).$$

This proves (4.14). In particular, $\rho Y_n \rightarrow 0$ in probability. Induction in (4.12), followed by the continuous mapping theorem and Slutsky's lemma, now proves convergence of the phases at grid points and at every fractional time. Since periodic test functions are bounded, dominated convergence also permits averaging over the finitely many steps and over $s \in [0, 1]$. $\square$

Take the one-dimensional test function

$$\phi_h(x) = \sin(x/\sqrt{h}).$$

Its Stein expectation under the one-dimensional marginal of $\bar{\mu}_{N,h}$ is

$$\begin{aligned} S_\rho &:= \mathbb{E}_{\bar{\mu}_{N,h}} [\phi'_h(\hat{X}) - V'_{\rho,h}(\hat{X}) \phi_h(\hat{X})] \\ &= \frac{1}{\sqrt{h}} \mathbb{E} [\cos Y - \delta \sin^2 Y] - \frac{\rho}{\sqrt{h}} \mathbb{E} [Y \sin Y]. \end{aligned} \quad (4.16)$$

Here every expectation is also averaged uniformly over the N steps and over the fractional time. Lemma 4.2 shows that the last term tends to zero and that the first term converges to the periodic expectation in Lemma 4.1. Consequently, for $\delta \leq \delta_0$ and all sufficiently small positive ρ ,

$$|S_\rho| \gtrsim \frac{\delta}{\sqrt{h}} = \eta \beta \sqrt{h} \asymp \beta \sqrt{h}. \quad (4.17)$$

The required choice of ρ may depend on N, h , which is harmless because the theorem is an existence statement for each fixed collection of parameters.

In d dimensions, use the vector field

$$\Phi_h(x) = \left(\sin \frac{x_1}{\sqrt{h}}, \dots, \sin \frac{x_d}{\sqrt{h}} \right).$$

The numerator in (2.1) equals dS_ρ , while

$$\mathbb{E} \|\Phi_h\|^2 \leq d.$$

It follows that

$$\text{FI}(\bar{\mu}_{N,h} \parallel \pi_{\rho,h}) \geq dS_\rho^2 \gtrsim \beta^2 dh. \quad (4.18)$$

Choose η and then $c_0 \leq 1/4$ sufficiently small that $\delta = \eta\beta h \leq \delta_0$ whenever $h \leq c_0/\beta$. Together with (3.4), this proves (1.5).

5 A large-step obstruction

The periodic construction was only needed for small steps. A Gaussian rules out all larger steps in the high-accuracy regime.

Proposition 5.1 (Large Gaussian steps). For every $h > 0$, take

$$\pi = \mathcal{N}(0, \beta^{-1}I_d) \quad \text{and} \quad \mu_0 = \pi.$$

Then, uniformly in $N \geq 1$,

$$\text{FI}(\bar{\mu}_{N,h} \parallel \pi) \geq \beta d \frac{(\beta h)^4/9}{1 + (\beta h)^2/3}. \quad (5.1)$$

Proof. Every interpolated law is centered and isotropic. Let e_1 denote the first standard basis vector in $\mathbb{R}^d$, and set

$$r_n := \beta \mathbb{E}[\langle \hat{X}_{nh}, e_1 \rangle^2] \quad \text{and} \quad r_{n,s} := \beta \mathbb{E}[\langle \hat{X}_{nh+sh}, e_1 \rangle^2].$$

Since $r_0 = 1$, the Gaussian LMC recursion gives

$$r_{n,s} = (1 - s\beta h)^2 r_n + 2s\beta h,$$

$$r_{n,s} - 1 = (1 - s\beta h)^2 (r_n - 1) + s^2 (\beta h)^2.$$

Taking $s = 1$ inductively shows that $r_n \geq 1$ for all n . If $\bar{r}$ denotes the variance parameter averaged over n and s , then

$$\bar{r} - 1 \geq \int_0^1 s^2 (\beta h)^2 ds = \frac{(\beta h)^2}{3}.$$

Apply Lemma 2.1 with $\Phi(x) = x$. Its numerator is $d(1 - \bar{r})$ and its denominator is $d\bar{r}/\beta$, so

$$\text{FI}(\bar{\mu}_{N,h} \parallel \pi) \geq \beta d \frac{(\bar{r} - 1)^2}{\bar{r}}.$$

The function $(r-1)^2/r$ is increasing for $r \geq 1$. Substituting $\bar{r} \geq 1 + (\beta h)^2/3$ proves (5.1). $\square$

If $h \geq c_0/\beta$, then $\beta h \geq c_0$, so (5.1) proves (1.6) and completes the proof of Theorem 1.1.

6 Consequences

Proof of Corollary 1.1. Proposition 1.1 with $h = (2\beta)^{-1} \sqrt{K_0/(dN)}$ gives the upper bound. For the lower bound, fix h . If $h \geq c_0/\beta$, use (1.6); if instead $Nh < \beta^{-1}$, use the βK_0 term in (1.5); otherwise, use $\max\{K_0/(Nh), \beta^2 dh\} \geq \beta \sqrt{dK_0/N}$. The assumptions $d/N \lesssim K_0 \lesssim dN$ make the first two cases at least as large as the last. $\square$

Proof of Corollary 1.2. The assumption on ε and (1.6) force $h < c_0/\beta$. Equation (1.5) then gives $h \lesssim \varepsilon^2/(\beta^2 d)$ and $Nh \gtrsim K_0/\varepsilon^2$, where the latter uses $\varepsilon^2 \ll \beta K_0$. Thus $N = (Nh)/h \gtrsim \beta^2 d K_0/\varepsilon^4$. $\square$

References

- [Che26] S. Chewi. *Log-concave sampling*. Book draft. 2026. URL: <https://chewisinho.github.io/main.pdf>.