NeurIPS 2026

Preconditioned Implicit Midpoint Langevin Sampling for Non-Smooth Bayesian Imaging

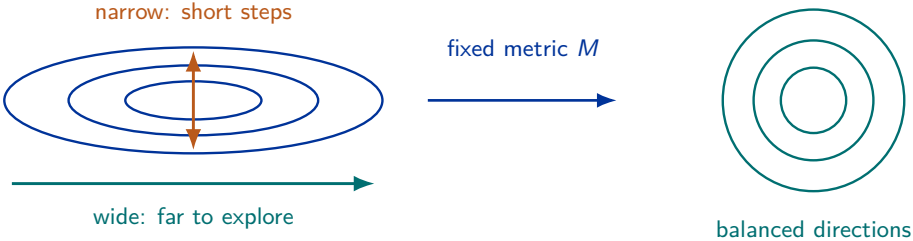
Can ideas from optimization make image sampling practical?

Sichen Wang[1] Zhipeng Lu[1,2]

[1]Department of Engineering, Shenzhen MSU-BIT University
[2]Guangdong Laboratory of Machine Perception and Intelligent Computing

The starting point: geometry in optimization

Earlier geometry experiments led us toward proximal Langevin sampling.



Optimization: find one solution.
$\min_x f(x)$

Sampling: explore plausible solutions.
$\pi(x) \propto e^{-f(x)}$

Smooth / Gaussian intuition. Exact circularization uses $M = \nabla^2 f$ for a quadratic f; TV posteriors need not be elliptical.

Large steps also need the right update

x: current state x': next state δ: step size $\xi \sim \mathcal{N}(0, I)$: random noise

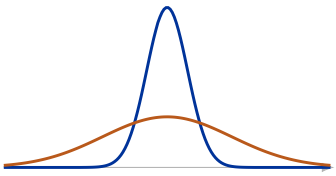
Euler–Maruyama: use the current slope

$$x' = x - \delta M^{-1} \nabla f(x) + \sqrt{2\delta}\, M^{-1/2} \xi$$

P-IMLA: use the midpoint slope

$$x' = x - \delta M^{-1} \nabla f\left(\frac{x + x'}{2}\right) + \sqrt{2\delta}\, M^{-1/2} \xi$$

Target $\mathcal{N}(0, 1)$, $\delta = 1.8$



Euler: variance 10

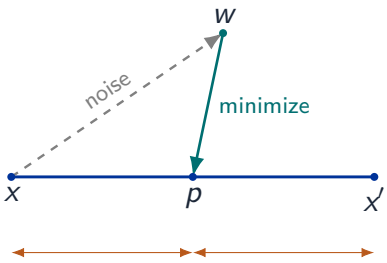
Target / midpoint: variance 1

M rescales both drift and noise. Midpoint preserves the Gaussian spread.

Exact midpoint solution; general targets can have discretization bias. Midpoint without preconditioning is IMLA (Klatzer et al., 2024).

A sampling step becomes an optimization problem

The midpoint equation has a useful form: add noise, solve for p, then reflect.



$p = (x + x')/2$: the midpoint.

$$w = x + \sqrt{\delta/2}\, M^{-1/2}\xi$$

$$p = \arg\min_u \left\{ f(u) + \frac{\|u - w\|_M^2}{\delta} \right\}$$

$$x' = 2p - x \quad \text{reflect } x \text{ about } p$$

For images: use $U_\lambda = f + g^{\lambda,M}$ in place of f.
g: edge-preserving TV prior; $g^{\lambda,M}$: its smoothed version in the same metric M.

Section 3. $\|v\|_M^2 = v^\top M v$; $M \succ 0$ is fixed. The diagram is a construction, not a measured trajectory.

JP-IMLA: making each image update cheaper

The image results were promising. Solving the inner problem was expensive.

Canonical P-IMLA$_\lambda$

$p = \mathrm{prox}^M_{(\delta/2)(f+g^{\lambda,M})}(w)$

Picard iterations with a TV solve inside

30×50 inner iterations

Joint-proximal variant: JP-IMLA

$p = \mathrm{prox}^M_{(\delta/2)(f+g)}(w)$

One joint image optimization

12 ADMM iterations

What the analysis resolves

Changing the proximal problem and stopping the solver early create distinct errors.
We bound them separately and establish a unique invariant law for exact JP.*

Section 3 and Appendix B. *Strongly convex quadratic f + finite TV penalties. The error comparison is to the canonical invariant law. Shown solver budgets are not matched accuracy.

Fewer steps, then less time

256×256 geometric image, blurred and noisy. Lower reconstruction error is better.

Method	Steps	SSE ↓	Minutes / chain
SP-MYULA	200,000	3432.0 ± 1.0	74
P-IMLA$_\lambda$	2,000	855.22 ± 0.66	70
JP-IMLA	**2,000**	$\mathbf{852.72 \pm 0.65}$	**1.2**

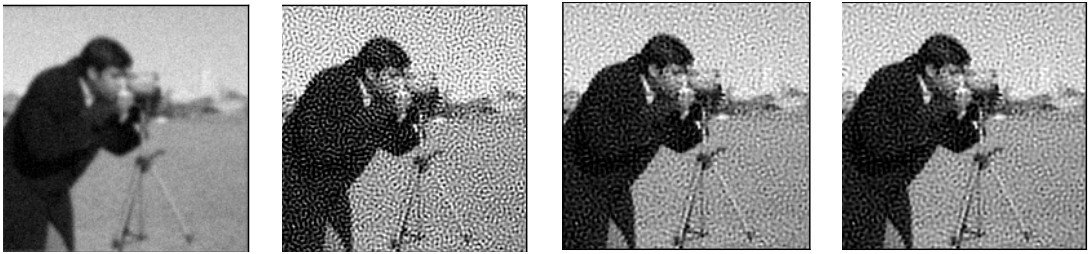
P-IMLA$_\lambda$
100× fewer steps; similar runtime.

JP-IMLA versus SP-MYULA
About 4× lower SSE in 1/62 the time.

Table 3. SSE of the estimated mean; mean ± standard error over 8 zero-start chains, no burn-in. RTX PRO 6000 Blackwell, 4-way parallel. Finite-budget comparison.

The reconstruction gain extends across images

Observation	SP-MYULA 200,000 steps	P-IMLA$_\lambda$ 2,000 steps	JP-IMLA 2,000 steps



SP-MYULA SSE / P-IMLA SSE

Geometric 4.01× Shepp–Logan 3.48× Cameraman 3.89×

JP-IMLA stays within 0.7% of canonical P-IMLA's SSE on all three images.

Figure 3: original cameraman panels (pooled means). Ratios: per-chain SSE in Tables 3–4; JP comparison: Table 8. SSE agreement does not imply identical distributions.

What we learned

Geometry makes exploration easier.
Use one metric for the drift, noise and prior smoothing.

Midpoint uses a proximal solve.
Optimize for the midpoint, then reflect to obtain the next sample.

JP makes the construction practical.
A cheaper solve, with its approximation made explicit.

Paper (arXiv)

Code

Thank you!

Sichen Wang and Zhipeng Lu