

Hypocoercivity of Tempered Bouncy Particle Sampler for heavy-tailed targets

Vinayak Niraj

Department of Statistics



Joint work with **Aleksandar Mijatović**

October 12, 2026

EPSRC

Engineering and Physical Sciences
Research Council

EP/W006227/1 & EP/V009478/1

Problem: heavy tails defeat finite-speed BPS

Target: $\mu(dx) = Z_U^{-1} e^{-U(x)} dx$ on $\mathbb{R}^d$. Main example **(ME)**:

$$U(x) \asymp (d+r) \log |x|, \text{ for } r > 0 \text{ as } |x| \rightarrow \infty.$$

Bouncy Particle Sampler (BPS) with Hamiltonian $H(x, v) = U(x) + |v|^2/2$:

$$X_t = X_0 + \int_0^t V_s ds$$

$$V_t = V_0 + \int_0^t (R_{\nabla_x H(X_s, V_{s-})} V_{s-} - V_{s-}) dN_s^b + \int_0^t (\tilde{V}_s - V_{s-}) dN_s^r$$

Problem: heavy tails defeat finite-speed BPS

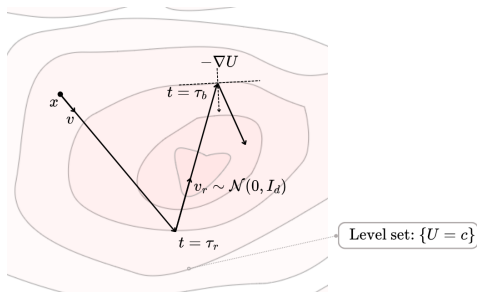
Target: $\mu(dx) = Z_U^{-1} e^{-U(x)} dx$ on $\mathbb{R}^d$. Main example **(ME)**:

$$U(x) \asymp (d+r) \log |x|, \text{ for } r > 0 \text{ as } |x| \rightarrow \infty.$$

Bouncy Particle Sampler (BPS) with Hamiltonian $H(x, v) = U(x) + |v|^2/2$:

$$X_t = X_0 + \int_0^t V_s ds$$

$$V_t = V_0 + \int_0^t (R_{\nabla_x H(X_s, V_{s-})} V_{s-} - V_{s-}) dN_s^b + \int_0^t (\tilde{V}_s - V_{s-}) dN_s^r$$



- Between events: $dX_t = V_t dt$ and $dV_t = 0$.
- Bounce rate*:
 $\lambda_b(x, v) = \langle v, \nabla U(x) \rangle_+$, bounce map $R_u v := v - 2|u|^{-2} \langle v, u \rangle u$.
- Refresh rate*: $\lambda_r > 0$ constant.
 $\tilde{V}_s \sim \mathcal{N}(0, I_d)$

Problem: heavy tails defeat finite-speed BPS

Target: $\mu(dx) = Z_U^{-1} e^{-U(x)} dx$ on $\mathbb{R}^d$. Main example **(ME)**:

$$U(x) \asymp (d+r) \log |x|, \text{ for } r > 0 \text{ as } |x| \rightarrow \infty.$$

Bouncy Particle Sampler (BPS) with Hamiltonian $H(x, v) = U(x) + |v|^2/2$:

$$X_t = X_0 + \int_0^t V_s ds$$

$$V_t = V_0 + \int_0^t (R_{\nabla_x H(X_s, V_{s-})} V_{s-} - V_{s-}) dN_s^b + \int_0^t (\tilde{V}_s - V_{s-}) dN_s^r$$

- Bounce rate $\lambda_b(x, v) = \langle v, \nabla U(x) \rangle_+$.
- In **(ME)**, $|\nabla U(x)| \rightarrow 0$ as $|x| \rightarrow \infty$ making *bounce clock* disappear in tails.
- Refresh $\tilde{V}_s \sim \mathcal{N}(0, I_d) = O(1)$. Therefore, V_t is bdd with high probability, implying $\lambda_b(x, V_t) \rightarrow 0$ as $|x| \rightarrow \infty$.

Problem: heavy tails defeat finite-speed BPS

Target: $\mu(dx) = Z_U^{-1} e^{-U(x)} dx$ on $\mathbb{R}^d$. Main example **(ME)**:

$$U(x) \asymp (d+r) \log |x|, \text{ for } r > 0 \text{ as } |x| \rightarrow \infty.$$

Bouncy Particle Sampler (BPS) with Hamiltonian $H(x, v) = U(x) + |v|^2/2$:

$$X_t = X_0 + \int_0^t V_s ds$$

$$V_t = V_0 + \int_0^t (R_{\nabla_x H(X_s, V_{s-})} V_{s-} - V_{s-}) dN_s^b + \int_0^t (\tilde{V}_s - V_{s-}) dN_s^r$$

- Bounce rate $\lambda_b(x, v) = \langle v, \nabla U(x) \rangle_+$.
- In **(ME)**, $|\nabla U(x)| \rightarrow 0$ as $|x| \rightarrow \infty$ making *bounce clock* disappear in tails.
- Refresh $\tilde{V}_s \sim \mathcal{N}(0, I_d) = O(1)$. Therefore, V_t is bdd with high probability, implying $\lambda_b(x, V_t) \rightarrow 0$ as $|x| \rightarrow \infty$.
- Hence $\|\text{Law}(X_t) - \mu\|_{\text{TV}} \rightarrow 0$ **polynomially** as $t \rightarrow \infty$ [RR23; ADW21].

Problem: heavy tails defeat finite-speed BPS

Target: $\mu(dx) = Z_U^{-1} e^{-U(x)} dx$ on $\mathbb{R}^d$. Main example **(ME)**:

$$U(x) \asymp (d+r) \log |x|, \text{ for } r > 0 \text{ as } |x| \rightarrow \infty.$$

Bouncy Particle Sampler (BPS) with Hamiltonian $H(x, v) = U(x) + |v|^2/2$:

$$X_t = X_0 + \int_0^t V_s ds$$

$$V_t = V_0 + \int_0^t (R_{\nabla_x H(X_s, V_{s-})} V_{s-} - V_{s-}) dN_s^b + \int_0^t (\tilde{V}_s - V_{s-}) dN_s^r$$

- Bounce rate $\lambda_b(x, v) = \langle v, \nabla U(x) \rangle_+$.
- In **(ME)**, $|\nabla U(x)| \rightarrow 0$ as $|x| \rightarrow \infty$ making *bounce clock* disappear in tails.
- Refresh $\tilde{V}_s \sim \mathcal{N}(0, I_d) = O(1)$. Therefore, V_t is bdd with high probability, implying $\lambda_b(x, V_t) \rightarrow 0$ as $|x| \rightarrow \infty$.
- Hence $\|\text{Law}(X_t) - \mu\|_{\text{TV}} \rightarrow 0$ **polynomially** as $t \rightarrow \infty$ [RR23; ADW21].

Goal: increase the tail velocity scale to obtain exponential convergence, while retaining piece-wise linear trajectories.

Tempered Bouncy Particle Sampler

Choose smooth $\sigma : \mathbb{R}^d \rightarrow [1, \infty)$ and define the Hamiltonian

$$H(x, v) = U(x) + \frac{|v|^2}{2\sigma^2(x)} + d \log \sigma(x).$$

$$\begin{aligned} X_t &= X_0 + \int_0^t V_s \, ds, \\ V_t &= V_0 + \int_0^t (R_{\nabla_x H(X_s, V_{s-})} V_{s-} - V_{s-}) \, dN_s^b + \int_0^t (\tilde{V}_s - V_{s-}) \, dN_s^r, \end{aligned} \quad (\text{T-BPS})$$

- As in BPS, between events: $dX_t = V_t \, dt$, $dV_t = 0$.
- Bounce rate: $\lambda_b(x, v) = \langle v, \nabla_x H(x, v) \rangle_+$ with jump $v \mapsto R_{\nabla_x H(x, v)} v$.
- Refresh rate λ_r : $\tilde{V} \sim \kappa_x = \mathcal{N}(0, \sigma^2(x) I_d) = O(\sigma(x))$ (position-dependent refresh distribution).

Linear flow with paths that are continuous in position and locally bounded rate λ_b imply non-explosion.

Invariant law of TBPS and the role of tempering

With $H(x, v) = U(x) + \frac{|v|^2}{2\sigma^2(x)} + d \log \sigma(x)$, invariant law of TBPS is

$$\mu_\sigma(dx, dv) = \mu(dx) \kappa_x(dv) \propto e^{-H(x,v)} dx dv.$$

- The position marginal remains exactly μ .

Invariant law of TBPS and the role of tempering

With $H(x, v) = U(x) + \frac{|v|^2}{2\sigma^2(x)} + d \log \sigma(x)$, invariant law of TBPS is

$$\mu_\sigma(dx, dv) = \mu(dx) \kappa_x(dv) \propto e^{-H(x, v)} dx dv.$$

- The position marginal remains exactly μ .
- Refresh at $X_t = x$ follows $\tilde{V}_t \sim \kappa_x = \mathcal{N}(0, \sigma^2(x)I_d)$. Thus typical magnitude

$$|V_t| \asymp \sigma(x) \quad \text{and} \quad \lambda_b(x, V_t) \asymp \sigma(x) |\nabla U(x)|.$$

Hence, σ should be chosen to keep $\lambda_b(x, V_t) \asymp 1$.

Invariant law of TBPS and the role of tempering

With $H(x, v) = U(x) + \frac{|v|^2}{2\sigma^2(x)} + d \log \sigma(x)$, invariant law of TBPS is

$$\mu_\sigma(dx, dv) = \mu(dx) \kappa_x(dv) \propto e^{-H(x, v)} dx dv.$$

- The position marginal remains exactly μ .
- Refresh at $X_t = x$ follows $\tilde{V}_t \sim \kappa_x = \mathcal{N}(0, \sigma^2(x)I_d)$. Thus typical magnitude

$$|V_t| \asymp \sigma(x) \quad \text{and} \quad \lambda_b(x, V_t) \asymp \sigma(x) |\nabla U(x)|.$$

Hence, σ should be chosen to keep $\lambda_b(x, V_t) \asymp 1$.

Tempering changes the velocity in the tails, not the position marginal distribution μ .

Main theorem: exponential L^2 -convergence

(A0) Weighted Poincaré inequality: $m\|g - \mu(g)\|_{L^2(\mu)}^2 \leq \int \sigma^2 |\nabla g|^2 d\mu$

(A1) Tempering bounds: $|\nabla \sigma|, \sigma |\nabla^2 \sigma|_F, \sigma |\nabla U|, \sigma^2 |\nabla^2 U|_F \leq M < \infty$

(A2) Weighted curvature lower bound: $\sigma^2 \nabla^2 U \succeq -K I_d$

Theorem ([MN26])

Let $(\mathcal{P}_t)$ be the TBPS semigroup. Under (A0)–(A2),

$$\|\mathcal{P}_t f\|_{L^2(\mu_\sigma)}^2 \leq 3e^{-\nu t} \|f\|_{L^2(\mu_\sigma)}^2, \quad f \in L_0^2(\mu_\sigma).$$

For $K_0 \lesssim d^3 M^2 m^{-1} K$, the explicit rate is

$$\nu^{-1} = \frac{3}{\lambda_r} \left[4 + \left(2\sqrt{K_0} + \frac{\lambda_r}{2\sqrt{m}} \right)^2 \right].$$

Recall that the BPS has polynomial decay for heavytails [ADW21; RR23].

Examples of heavy-tailed targets

Set $\langle x \rangle := \sqrt{1 + |x|^2}$ for $x \in \mathbb{R}^d$.

Target potential $U(x)$	$\sigma(x)$	Weighted gap
Polynomial $(d + r) \log \langle x \rangle$,	$\langle x \rangle$	$m^{-1} = d + r - 2$ for $r > d$ [BL09]
Sub-exponential $\langle x \rangle^p$, $0 < p < 1$	$\langle x \rangle^{1-p}$	$m^{-1} = e\{12d/p^3 + (d + p)/p^4\}$ [Cat+10]
Exponential or thinner	1	ordinary Poincaré gap

The same σ remains valid under bounded perturbations of U . Hence multimodality and anisotropy are allowed away from the tails.

From L^2 -decay to **mixing time** of TBPS

Mixing time $t_{\text{mix}}(\varepsilon, \mu_0) := \inf\{t \geq 0 : \chi^2(\mu_0 \mathcal{P}_t \| \mu_\sigma) < \varepsilon\}$ for $\varepsilon > 0$.

For $\chi^2(\mu_0 \| \mu_\sigma) < \infty$, using Theorem[MN26]

$$t_{\text{mix}}(\varepsilon, \mu_0) \leq \nu^{-1} \left[\log \frac{3\chi^2(\mu_0 \| \mu_\sigma)}{\varepsilon} \right]_+.$$

From L^2 -decay to **mixing time** of TBPS

Mixing time $t_{\text{mix}}(\varepsilon, \mu_0) := \inf\{t \geq 0 : \chi^2(\mu_0 \mathcal{P}_t \| \mu_\sigma) < \varepsilon\}$ for $\varepsilon > 0$.

For $\chi^2(\mu_0 \| \mu_\sigma) < \infty$, using Theorem[MN26]

$$t_{\text{mix}}(\varepsilon, \mu_0) \leq \nu^{-1} \left[\log \frac{3\chi^2(\mu_0 \| \mu_\sigma)}{\varepsilon} \right]_+.$$

Feasible start: $\log \chi^2(\mu_0^{(d)} \| \mu_\sigma^{(d)}) = O(d)$ (initial distribution may have lighter tails than target, e.g. $\mu_0^{(d)}$ Gaussian for $\mu_\sigma^{(d)}$ polynomial in x).

Theorem[MN26] implies:

- If $m^{-1} = O(d^\ell)$, then the mixing time is polynomial in dimension:

$$t_{\text{mix}}(1, \mu_0^{(d)}) = O(d^{\ell+5}).$$

- Polynomial and stretched-exponential targets have $\ell = 1$: $O(d^6)$.
- Pinsker gives the corresponding TV mixing-time bounds.

Previously $O(\exp(d))$ mixing time has been found in [And+26] for RWM on heavytails.

Future work: Adaptive tempering

Over-tempering: If σ too large, WPI (A0) holds but curvature assumption (A2) does not.

Conjecture. The heuristic aim that *bounce* rate $\lambda_b(X_t, V_t)$ should be $O(1)$ in tails, motivates the choice

$$\sigma(x) \asymp 1 + \frac{1}{|\nabla U(x)|}.$$

- It remains to prove that this choice is admissible for a large class of targets, and in practical scenarios where tail behaviour is anisotropic and/or unknown.

Future work: Adaptive tempering

Over-tempering: If σ too large, WPI (A0) holds but curvature assumption (A2) does not.

Conjecture. The heuristic aim that *bounce* rate $\lambda_b(X_t, V_t)$ should be $O(1)$ in tails, motivates the choice

$$\sigma(x) \asymp 1 + \frac{1}{|\nabla U(x)|}.$$

- It remains to prove that this choice is admissible for a large class of targets, and in practical scenarios where tail behaviour is anisotropic and/or unknown.

Boundary case:

- **If μ has no polynomial moments**, we prove WPI (A0) fails for every σ with bounded derivative. Hence, (A1) must be weakened to deal with this case.

Nonreversible dynamics related to TBPS

TBPS is a *second-order lift* [EL26, Def. 1]) of the **Tempered Langevin dynamics** (TLD) process:

$$Y_t = Y_0 + \int_0^t (2\sigma(Y_s)\nabla_x\sigma(Y_s) - \sigma^2(Y_s)\nabla_x U(Y_s)) \, ds + \sqrt{2} \int_0^t \sigma(Y_s) \, dB_s$$

Nonreversible dynamics related to TBPS

TBPS is a *second-order lift* [EL26, Def. 1]) of the **Tempered Langevin dynamics** (TLD) process:

$$Y_t = Y_0 + \int_0^t (2\sigma(Y_s)\nabla_x\sigma(Y_s) - \sigma^2(Y_s)\nabla_x U(Y_s)) ds + \sqrt{2} \int_0^t \sigma(Y_s) dB_s$$

Tempered Underdamped Langevin dynamics (TULD) is also a second-order lift of TLD:

$$dQ_t = P_t dt,$$

$$dP_t = \left[-\sigma^2(Q_t)\nabla_x U(Q_t) - \left(d - 2 - \frac{|P_t|^2}{\sigma^2(Q_t)} \right) \sigma(Q_t)\nabla_x\sigma(Q_t) \right] dt \\ - \frac{\lambda}{\sigma^2(Q_t)} P_t dt + \sqrt{2\lambda} dW_t$$

Nonreversible dynamics related to TBPS

TBPS is a *second-order lift* [EL26, Def. 1]) of the **Tempered Langevin dynamics** (TLD) process:

$$Y_t = Y_0 + \int_0^t (2\sigma(Y_s)\nabla_x\sigma(Y_s) - \sigma^2(Y_s)\nabla_x U(Y_s)) ds + \sqrt{2} \int_0^t \sigma(Y_s) dB_s$$

Tempered Underdamped Langevin dynamics (TULD) is also a second-order lift of TLD:

$$dQ_t = P_t dt,$$

$$dP_t = \left[-\sigma^2(Q_t)\nabla_x U(Q_t) - \left(d - 2 - \frac{|P_t|^2}{\sigma^2(Q_t)} \right) \sigma(Q_t)\nabla_x\sigma(Q_t) \right] dt \\ - \frac{\lambda}{\sigma^2(Q_t)} P_t dt + \sqrt{2\lambda} dW_t$$

TULD should be useful in practice as an algorithm:

- it has additive noise, allowing higher-order discretisations.
- if $\frac{|P_t|}{\sigma(Q_t)} = O(1)$ as in TBPS case, drift of TULD is at most linear.

References

- [ADW21] Christophe Andrieu, Paul Dobson, and Andi Q. Wang. Subgeometric hypocoercivity for piecewise-deterministic Markov process Monte Carlo methods. *Electron. J. Probab.* 2021.
- [And+26] Christophe Andrieu, Anthony Lee, Sam Power, and Andi Q. Wang. Weak Poincaré inequalities for Markov chains: Theory and applications. *The Ann. Appl. Probab.* 2026.
- [BL09] Sergey G. Bobkov and Michel Ledoux. Weighted Poincaré-type inequalities for Cauchy and other convex measures. *The Ann. Probab.* 2009.
- [Cat+10] Patrick Cattiaux, Nathael Gozlan, Arnaud Guillin, and Cyril Roberto. Functional Inequalities for Heavy Tailed Distributions and Application to Isoperimetry. *Electron. J. Probab.* 2010.
- [EL26] Andreas Eberle and Francis Lörler. Non-reversible lifts of reversible diffusion processes and relaxation times. *Probab. Theory Relat. Fields*, 2026.
- [MN26] Aleksandar Mijatović and Vinayak Niraj. Hypocoercivity of Tempered Bouncy Particle Samplers for Heavy-Tailed Targets. *arXiv: 2608.29657*, 2026.
- [RR23] Gareth O. Roberts and Jeffrey S. Rosenthal. Polynomial Convergence Rates of Piecewise Deterministic Markov Processes. *Methodol. Comput. Appl. Probab.* 2023.