IMPERIAL

Special Session 27 · Machine Learning for Numerical Methods in PDEs

Feedback Control for Sampling

A Feedback-Control Perspective on Sampling for Langevin and McKean–Vlasov Dynamics

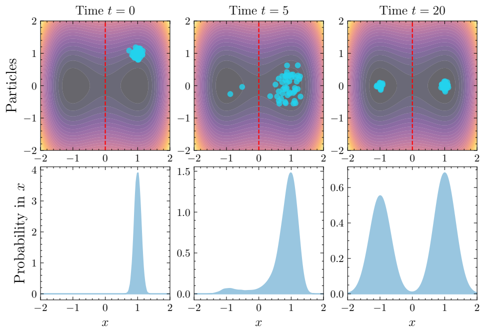
Dante Kalise, **Lucas M. Moschen**, and Grigorios A. Pavliotis

Department of Mathematics, Imperial College London

10 September 2026

Energy Barriers and Slow Convergence

A model problem for controlled Fokker–Planck equations



Particles in the energy landscape
$V(x, y) = (x^2 - 1)^2 + y^2$.

Target equilibrium: $\bar{\mu}(dx) \propto e^{-V(x)/\sigma}\, dx$

[*] Lelièvre, Nier and Pavliotis (2013); Pavliotis (2014).

- **Overdamped Langevin dynamics**

$$dX_t = -\nabla V(X_t)\, dt + \sqrt{2\sigma}\, dB_t, \qquad \mu_t = \mathrm{Law}(X_t).$$

- If $\bar{\mu}$ satisfies some regularity assumptions, then

$$\mathrm{dist}(\mu_t, \bar{\mu}) \leq e^{-\lambda t}\, \mathrm{dist}(\mu_0, \bar{\mu}).$$

  For nonconvex V, λ can be small $\Rightarrow$ **slow convergence**.

- **Goal:** construct a feedback control that improves the convergence rate towards $\bar{\mu}$.

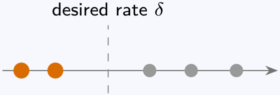
- **Applications:** molecular dynamics, equilibrium computation, and Langevin-based algorithms.

Can a Few Modes Accelerate a Particle Solver?

Langevin sampling

A few slowly relaxing modes can dominate the cost of reaching equilibrium.

1. Identify slow modes



Compute a few eigenfunctions of the linearised PDE

→

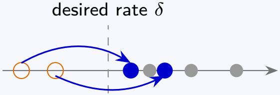
2. Measure their amplitudes

$$m_j^N = \frac{1}{N}\sum_{i=1}^{N} e_j(X_t^i)$$

Evaluate observables on the particle cloud

→

3. Apply feedback



Correct the selected modes through an added drift

Numerical question: does the reduction in particle steps pay for the feedback?

High-dimensional challenge: approximate or learn the slow modes efficiently.

Interacting Particles
McKean–Vlasov dynamics and the noisy Kuramoto model

Noisy Kuramoto model on $\mathbb{T}$

$$d\theta^i_t = -\frac{K}{N}\sum_{\ell=1}^{N} \sin(\theta^i_t - \theta^\ell_t)\, dt + \sqrt{2\sigma}\, dB^i_t.$$

Mean-field limit

Let $\mu^N_t = \frac{1}{N}\sum_{\ell=1}^N \delta_{\theta^\ell_t}$. Then, for i.i.d. initial data, $\mu^N_t \to \mu_t$, where μ_t solves

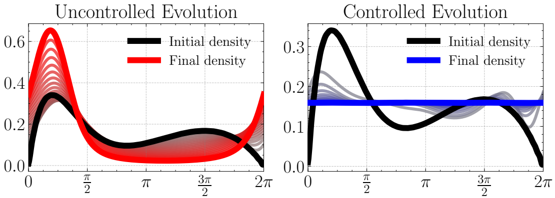
$$\partial_t \mu_t = \sigma\Delta\mu_t + \nabla\cdot(\mu_t \nabla(V + W * \mu_t))\,,$$

where $V = 0$ and $W(x) = -K\cos(x)$. This is the **McKean–Vlasov equation**.

Sznitman (1991); Bertini, Giacomin and Pakdaman (2010)

By linearising this equation at $\bar{\mu} = 1/(2\pi)$, the first Fourier mode satisfies

$$\frac{d}{dt}\widehat{\eta}_1 = \left(\frac{K}{2} - \sigma\right)\widehat{\eta}_1, \qquad \widehat{\eta}_1 = \int_{\mathbb{T}} e^{ix}\eta(x)\, dx.$$



For $K < 2\sigma$, convergence becomes arbitrarily slow as $K \uparrow 2\sigma$.

For $K > 2\sigma$, the uniform density is unstable and synchronisation emerges.

The McKean–Vlasov Equation as a Gradient Flow
The free energy determines both the equilibrium and the dynamics

- On $\mathbb{T}^d = (\mathbb{R}/\mathbb{Z})^d$, define the free energy (assuming W is even)

$$\mathcal{F}(\mu) = \int_{\mathbb{T}^d} V \, d\mu + \frac{1}{2} \iint_{\mathbb{T}^d \times \mathbb{T}^d} W(x-y)\, d\mu(x)\, d\mu(y) + \sigma \int_{\mathbb{T}^d} \mu \log \mu \, dx.$$

- The McKean–Vlasov equation is the W_2-gradient flow of $\mathcal{F}$:

$$\partial_t \mu_t = \nabla \cdot \left(\mu_t \nabla \frac{\delta \mathcal{F}}{\delta \mu}(\mu_t) \right) = \sigma \Delta \mu_t + \nabla \cdot (\mu_t \nabla V) + \nabla \cdot \big(\mu_t \nabla (W * \mu_t)\big), \qquad \frac{d}{dt}\mathcal{F}(\mu_t) \le 0.$$

  Here W_2 is the quadratic optimal-transport distance and $\delta \mathcal{F}/\delta \mu$ is the first variation.

- Equilibria satisfy

$$\bar{\mu} \propto \exp\left\{ -\frac{V + W * \bar{\mu}}{\sigma} \right\}.$$

* Jordan, Kinderlehrer and Otto (1998); Carrillo, Gvalani, Pavliotis and Schlichting (2020).

What Prevents Fast Convergence?
Looking at the linearised PDE

Fix one of the equilibria as the target $\bar{\mu}$.
To understand what prevents fast convergence, we linearise around $\bar{\mu}$.

Set
$$y = \mu - \bar{\mu} = \mathcal{I}_{\bar{\mu}}\xi, \quad \mathcal{I}_{\bar{\mu}}\xi := -\nabla \cdot (\bar{\mu}\nabla\xi), \quad \xi \in X := \dot{H}^1_{\bar{\mu}} \cong_{\nabla} T_{\bar{\mu}}\mathcal{P}(\mathbb{T}^d),$$
with inner product $\langle \xi, \eta \rangle_X = \int_{\mathbb{T}^d} \nabla\xi \cdot \nabla\eta \, d\bar{\mu}$.

Linearising the dynamics in X leads to
$$\partial_t \xi = -A\xi, \quad A = \mathrm{Hess}_{\mathrm{W}_2} \mathcal{F}(\bar{\mu}), \quad Ae_k = \lambda_k e_k$$
where A is self-adjoint, lower bounded, and has **compact resolvent**.

Mode e_k evolves as $e^{-\lambda_k t}$ and $\lambda_k \to \infty$. For a desired rate $\delta > 0$, define $m = \dim \mathcal{X}_\delta < \infty$, where
$$\mathcal{X}_\delta = \mathrm{span}\{e_k : \lambda_k \leq \delta\} \longrightarrow \textbf{Only modes that can obstruct decay at rate } \boldsymbol{\delta}\textbf{.}$$

* Kalise, Moschen, Pavliotis (2026, forthcoming).

Designing Feedback for Faster Convergence

We add a finite-dimensional control potential at the PDE level to **accelerate convergence** to $\bar{\mu}$:

$$\partial_t \mu_t = \nabla \cdot \left[\mu_t \nabla \left(\frac{\delta \mathcal{F}}{\delta \mu}(\mu_t) + \sum_{j=1}^{m} u_j(t)\alpha_j(x) \right) \right], \quad Bu := \sum_{j=1}^{m} u_j \alpha_j.$$

1. Act only where needed

Set $\alpha_j = e_j$. The controlled linearisation is

$$\partial_t \xi = -A\xi - Bu,$$

where $P_\delta := BB^*$ is the X-orthogonal projection onto X_δ.

2. Compute the feedback

Solving a linear-quadratic problem with state-cost Q, we get $u = B^*\Pi\xi$, for Π solving the **Riccati equation**

$$-(A - \delta I)\Pi - \Pi(A - \delta I) - \Pi BB^*\Pi + Q = 0.$$

Then $\Pi e_k = \pi_k e_k$ if Q is diagonal in this eigenbasis and $\partial_t \xi = -A_\Pi \xi$ with $A_\Pi := A + P_\delta \Pi \succeq \delta I$.

Local exponential stabilisation

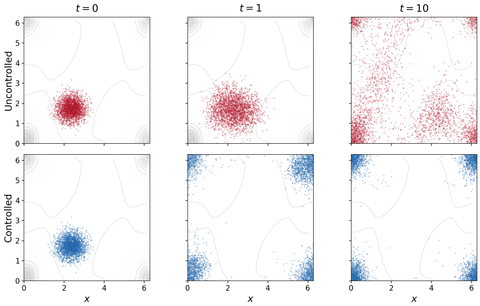
$$\|\mu_t - \bar{\mu}\|_{H^{-1}_{\bar{\mu}}} \leq C_\delta e^{-\delta t} \|\mu_0 - \bar{\mu}\|_{L^2} \quad \text{for } \|\mu_0 - \bar{\mu}\|_{L^2} \text{ small.}$$

From the PDE to the Particle System
How do we apply the PDE control in the SDEs?

With $\int_{\mathbb{T}^d} e_j \, d\bar{\mu} = 0$, the feedback depends only on expectations:

$$u_j(t) = \pi_j \int_{\mathbb{T}^d} e_j \, d(\mu_t - \bar{\mu}) \longrightarrow u_j^N(t) = \frac{\pi_j}{N} \sum_{\ell=1}^{N} e_j(X_t^\ell),$$

so no density reconstruction. In this framework, the sampling accelerator is constructed from the eigenmodes of the Wasserstein Hessian on X_δ.



Uncontrolled (top) vs. feedback-controlled (bottom).

Controlled N-particle system

$$dX_t^i = -\left(\nabla V(X_t^i) + \frac{1}{N} \sum_{k=1}^{N} \nabla W(X_t^i - X_t^k) + \sum_{j=1}^{m} u_j^N(t) \nabla e_j(X_t^i) \right) dt + \sqrt{2\sigma} \, dB_t^i.$$

Controlling the Noisy Kuramoto Model

Stabilising the incoherent state with feedback

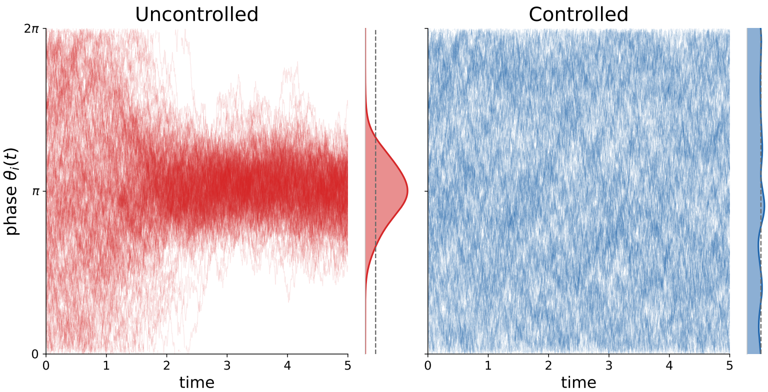


Figure: Particles following a noisy Kuramoto model on the circle. Without control, the particles **synchronise**. With control, they **remains dispersed**.

Sampling Without the Normalising Constant

The spectral construction only needs the unnormalised target

For an interacting model, we compute an equilibrium $\bar{\mu}$ numerically (e.g., by Newton's method) and then assemble the Hessian eigenproblem. For Gibbs sampling, however,

$$\nu(x) := e^{-V(x)/\sigma}, \qquad \bar{\mu}(x) = \frac{\nu(x)}{Z}, \qquad Z = \int_{\mathbb{T}^d} \nu(x)\, dx \quad \text{is unknown.}$$

An equivalent eigenvalue problem

When $W = 0$, instead of assembling A using integrals against $\bar{\mu}$, we solve

$$Ae_j^\nu = \lambda_j e_j^\nu \qquad \Longleftrightarrow \qquad -\mathcal{L}_\nu e_j^\nu = \lambda_j e_j^\nu, \qquad \mathcal{L}_\nu f := \sigma \Delta f - \nabla V \cdot \nabla f = \frac{\sigma}{\nu} \nabla \cdot (\nu \nabla f).$$

Here e_j^ν is orthonormal with respect to $\langle \varphi, \psi \rangle_\nu := \int_{\mathbb{T}^d} \nabla \varphi \cdot \nabla \psi \, d\nu$, so $e_j = \sqrt{Z} e_j^\nu$ is X-orthonormal.

$$\gamma_j = \overline{Z} \pi_j, \quad \overline{Z} \geq Z \quad \longrightarrow \quad \text{spectral shift} = \frac{\gamma_j}{Z} = \frac{\overline{Z}}{Z} \pi_j \geq \pi_j.$$

Hence we keep the prescribed decay rate δ.

* Kalise, Moschen and Pavliotis (work in progress).

Computational Cost of the Controlled Particle Method

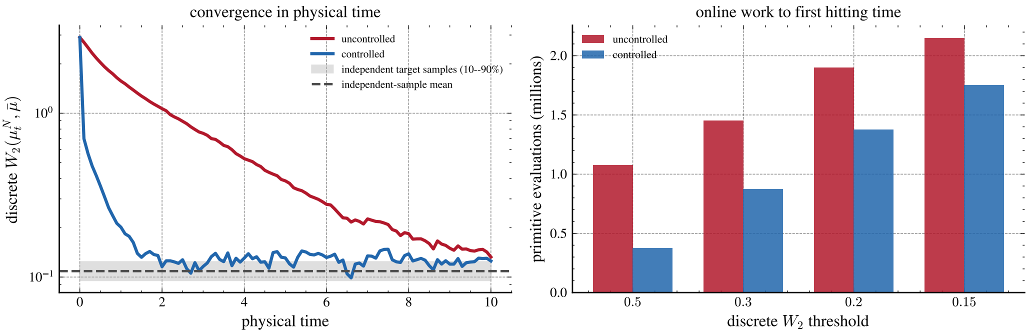


Figure: Two-dimensional four-well Gibbs target with $N = 2500$ particles and two controlled modes. Left: exact discrete W_2 on an 18×18 histogram. Right: equal unit cost for each evaluation of ∇V, e_j and ∇e_j; the offline eigensolve is excluded.

Long-Time Confinement

PDE contraction is stable under empirical approximation

Theorem

For N sufficiently large, if $\mathrm{W}_2(\mu_0^N, \bar{\mu}) < R$, then, for every $t \geq 0$,

$$\mathbb{E}\mathrm{W}_2^2(\mu_t^N, \bar{\mu}) \leq \underbrace{Ce^{-2\delta t}\mathrm{W}_2^2(\mu_0^N, \bar{\mu})}_{\textit{controlled contraction}} + \underbrace{C\beta_d(N)}_{\textit{empirical floor}} + \underbrace{C\min\{1, te^{-cN}\}}_{\textit{rare exits}}.$$

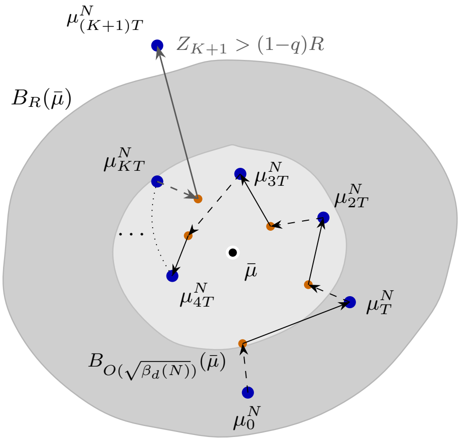
Here $\beta_d(N) \to 0$ as $N \to \infty$, with slower convergence in higher dimensions.

The particle system inherits the prescribed decay rate δ until it reaches the empirical scale. More precisely, after the corresponding burn-in time t_{burn}, for every $0 < c' < c$,

$$\sup_{t_{\mathrm{burn}} \leq t \leq e^{c'N}} \mathbb{E}\mathrm{W}_2^2(\mu_t^N, \bar{\mu}) \leq C\beta_d(N) + Ce^{-(c-c')N}.$$

[*] Fournier and Guillin (2015)

Proof Strategy



Blue: particles   Orange: PDE solution
Dashed: contraction   Solid: random particle error

$$D_k := \mathrm{W}_2(\mu^N_{kT}, \bar\mu), \quad Z_{k+1} := \mathrm{W}_2(\mu^N_{(k+1)T}, S^{\Pi}_T \mu^N_{kT}).$$

One iteration: $D_{k+1} \le qD_k + Z_{k+1}, \quad q < 1.$

Probability of exit:

$$\mathbb{P}\left(Z_{k+1} \ge \sqrt{C_T \beta_d(N)} + r \mid X^1_{kT}, \dots, X^N_{kT}\right) \le e^{-c_T N r^2}.$$

First exit: $\kappa^N_R := \inf\{k \ge 0 : D_k \ge R\}$.

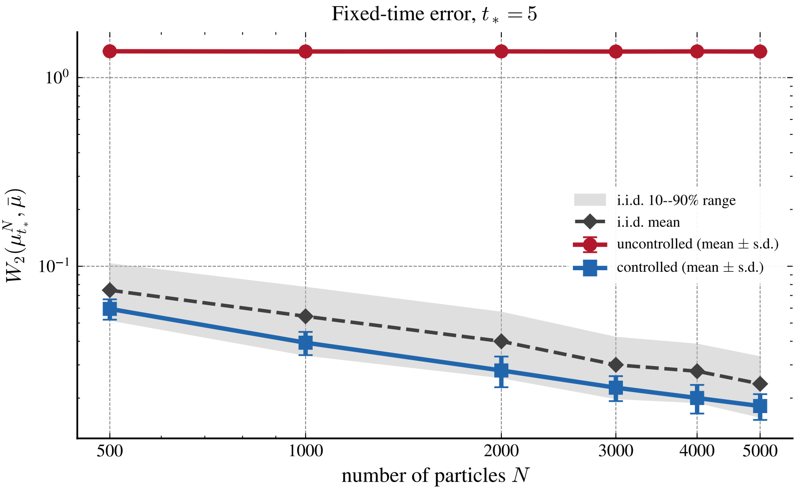
$$\kappa^N_R = k+1 \quad \Longrightarrow \quad Z_{k+1} > (1-q)R.$$

Hence, by a union bound,

$$\mathbb{P}(\kappa^N_R \le K) \le Ke^{-cN}.$$

Controlled Particles Reach the Empirical Error Scale

Finite-N scaling for the noisy Kuramoto model.



Summary and Open Questions

Takeaways

- A feedback control stabilises unstable targets and accelerates convergence toward stable ones at a chosen rate $\delta > 0$.
- The particle system inherits the PDE contraction for exponentially long times.
- For Gibbs measures, the controller can be constructed using only the unnormalised density.

Open questions

- Can the feedback modes be learned efficiently in high dimensions?
- When does the online acceleration compensate for the offline spectral computation?
- Extension to kinetic models?
- Can the local result be made global?

Summary and Open Questions

Takeaways

- A feedback control stabilises unstable targets and accelerates convergence toward stable ones at a chosen rate $\delta > 0$.
- The particle system inherits the PDE contraction for exponentially long times.
- For Gibbs measures, the controller can be constructed using only the unnormalised density.

Open questions

- Can the feedback modes be learned efficiently in high dimensions?
- When does the online acceleration compensate for the offline spectral computation?
- Extension to kinetic models?
- Can the local result be made global?

Happy to discuss any of these!

Papers:

2026: Linearization-Based Feedback Stabilization of McKean–Vlasov PDEs (to appear in SICON)

2026: Feedback Control and Local Convexification of Wasserstein Gradient Flows

lucas.moschen22@imperial.ac.uk