Finite particle limit in the Ensemble Kalman Sampler

Louis Carillo

PhD under the supervision of Tony Lelièvre, **Urbain Vaes** & Gabriel Stoltz

June 2026

ÉCOLE NATIONALE DES PONTS ET CHAUSSÉES | INSTITUT POLYTECHNIQUE DE PARIS

Inria

MATHerials

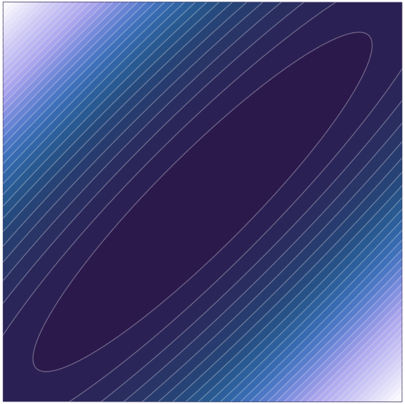
cnrs

In collaboration with:

Urbain Vaes

Nicolai Gerber

Julien Reygner



Sampling is ubiquitous in science nowadays

Today challenge: anisotropy in the target distribution

Take a Bolzmann distribution

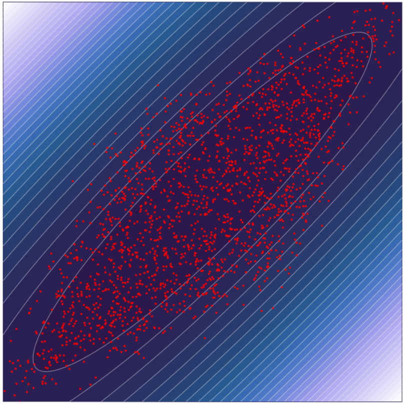
$$\mu_{\mathrm{B}}(x) = \frac{1}{Z}\mathrm{e}^{-V(x)}$$

with Z the normalization constant
and V an **anisotropic** potential

Exemple $V(x) = x^T A x$ with

$$\lambda_{\mathrm{max}}(A) \gg \lambda_{\mathrm{min}}(A)$$

Sampling **anisotropy**: Overdamped Langevin dynamics



Take N particles following

$$\mathrm{d}X_t = -\nabla V(X_t)\,\mathrm{d}t + \sqrt{2}\,\mathrm{d}W_t$$

Studied for decades [1]

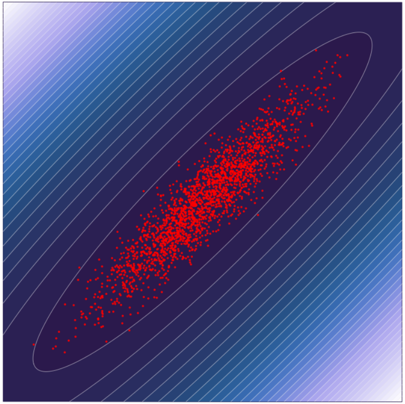
The invariant measure is $\mu_{\mathrm{B}}^{\otimes N}$

Problem:

convergence rate is constrained by $\lambda_{\min}(A)$

time step is constrained by $\lambda_{\max}(A)$

[1] Lelièvre, Stoltz *Acta Numerica* (2026)



Take N particles following

$$\mathrm{d}X_t = -C_\mu \nabla V(X_t)\,\mathrm{d}t + \sqrt{2C_\mu}\,\mathrm{d}W_t$$

C_μ is a preconditioning matrix

$\Rightarrow$ mitigate anisotropy[2]

Why?

$Y_t = C_\mu^{-\frac{1}{2}} X_t$ follows a well-conditioned Overdamped Langevin dynamics

[2] Goodman, Weare *Commun. Appl. Math. Comput. Sci.* (2010)

Ensemble Kalman Sampler (EKS/ALDI)

Consider the following particle system

$$\mathrm{d}X_t^n = -\mathcal{C}(\rho_{\mathcal{X}_t})\nabla V(X_t^n)\,\mathrm{d}t + \frac{d+1}{N}(X_t^n - \mathcal{M}_{\mathcal{X}_t}) + \sqrt{2\mathcal{C}(\rho_{\mathcal{X}_t})}\,\mathrm{d}W_t^n \qquad \text{(EKS)}$$

with $\mathcal{C}(\rho_{\mathcal{X}_t})$ the empirical covariance of the finite particle ensemble
and correction term for **finite number** of particles

Idea introduced in [3]

Consider the following particle system

$$\mathrm{d}X_t^n = -\mathcal{C}(\rho_{\mathcal{X}_t})\nabla V(X_t^n)\,\mathrm{d}t + \frac{d+1}{N}\left(X_t^n - \mathcal{M}_{\mathcal{X}_t}\right) + \sqrt{2\mathcal{C}(\rho_{\mathcal{X}_t})}\,\mathrm{d}W_t^n \qquad \text{(EKS)}$$

with $\mathcal{C}(\rho_{\mathcal{X}_t})$ the empirical covariance of the finite particle ensemble
and correction term for **finite number** of particles

Idea introduced in [3]

- **Affine invariant** method, always well-conditioned
- **Gradient-free** approximation can be used
- **Square root approximation** of covariance for high-dimension

[3] Garbuno-Inigo, Nüsken, Reich *SIAM J. Appl. Dyn. Syst.* (2020)

Finite particle and mean-field limit: Invariant measure

Theorem (Garbuno-Inigo, Nüsken, Reich 2020)

For μ_{B} with gaussian tails, (EKS) is ergodic with a unique invariant measure $\mu_{\mathrm{B}}^{\otimes N}$.

The particles are **interacting** though $C(\rho_{\mathcal{X}_t})$ yet they are as if they were **independent at equilibrium** as the invariant measure is a product measure

Open question:

What is the convergence rate?

Propagation of chaos has been established in [4] but only on finite time intervals

[4] Ding, Li *SIAM J. Math. Anal.* (2021)

Finite particle and mean-field limit: Main results

Theorem (Uniform-in-time propagation of chaos) [L.C., Gerber, Reygner, Vaes]

For $V(x) = \frac{1}{2}x^\top x$, consider $(X_t^n)_{n=1}^N$ following the (EKS) with initial condition $\mathcal{X}_0 \overset{i.i.d.}{\sim} \rho_0$. Then, there exits $C > 0$ such that for any $p \geq 2$,

$$\sup_{t \geq 0,\, n \in \{1,\ldots,N\}} W_p\left(\mathrm{Law}(X_t^n), \mathrm{Law}(\overline{X}_t)\right) \leq \frac{C}{\sqrt{N}}$$

with $\overline{X}_t$ following the mean-field limit of (EKS) with initial condition $\overline{X}_0 \sim \rho_0$.

Theorem (Convergence to equilibrium) [L.C., Gerber, Reygner, Vaes 2026+]

In dimension 1, under the same assumptions, there exists $C > 0$ such that

$$W_2\left(\rho_{\mathcal{X}_t}, \mu_{\mathrm{B}}^{\otimes N}\right) \leq C\,\mathrm{e}^{-(1+\varepsilon(N))t}$$

with $\varepsilon(N) \to 0$ as $N \to +\infty$.

Theorem (Convergence to equilibrium) [L.C., Gerber, Reygner, Vaes 2026+]

In dimension 1, under the same assumptions, there exists $C > 0$ such that

$$W_2\left(\rho_{\mathcal{X}_t}, \mu_{\mathrm{B}}^{\otimes N}\right) \leq C\,\mathrm{e}^{-(1+\varepsilon(N))t}$$

with $\varepsilon(N) \to 0$ as $N \to +\infty$.

We will use a coupling between two system, $(X_t^n)_{n=1}^N$ and $(Y_t^n)_{n=1}^N$ following (EKS) with initial condition $\mathcal{X}_0 \overset{i.i.d.}{\sim} \rho_0$ and $\mathcal{Y}_0 \overset{i.i.d.}{\sim} \mu_{\mathrm{B}}$ respectively
Then

$$W_2\left(\rho_{\mathcal{X}_t}, \mu_{\mathrm{B}}^{\otimes N}\right) \leq \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N} |X_t^n - Y_t^n|^2\right]^{\frac{1}{2}}$$

Idea of proof: difficulty with covariance moments

Consider the following coupling of two systems of particles following (EKS)

$$\mathrm{d}X_t^n = -\mathcal{C}(\rho_{\mathcal{X}_t})X_t^n\,\mathrm{d}t + \frac{d+1}{N}\left(X_t^n - \mathcal{M}_{\mathcal{X}_t}\right)\mathrm{d}t + \sqrt{2\mathcal{C}(\rho_{\mathcal{X}_t})}\,\mathrm{d}W_t^n$$
$$\mathrm{d}Y_t^n = -\mathcal{C}(\rho_{\mathcal{Y}_t})Y_t^n\,\mathrm{d}t + \frac{d+1}{N}\left(Y_t^n - \mathcal{M}_{\mathcal{Y}_t}\right)\mathrm{d}t + \sqrt{2\mathcal{C}(\rho_{\mathcal{Y}_t})}\,\mathrm{d}W_t^n$$

Key idea: If the covariances matrices $\mathcal{C}(\rho_{\mathcal{X}_t})$ and $\mathcal{C}(\rho_{\mathcal{Y}_t})$ are close, then the two systems would be close too

Idea of proof: difficulty with covariance moments

Consider the following coupling of two systems of particles following (EKS)

$$\mathrm{d}X_t^n = -\mathcal{C}(\rho_{\mathcal{X}_t})X_t^n\,\mathrm{d}t + \frac{d+1}{N}\left(X_t^n - \mathcal{M}_{\mathcal{X}_t}\right)\mathrm{d}t + \sqrt{2\mathcal{C}(\rho_{\mathcal{X}_t})}\,\mathrm{d}W_t^n$$
$$\mathrm{d}Y_t^n = -\mathcal{C}(\rho_{\mathcal{Y}_t})Y_t^n\,\mathrm{d}t + \frac{d+1}{N}\left(Y_t^n - \mathcal{M}_{\mathcal{Y}_t}\right)\mathrm{d}t + \sqrt{2\mathcal{C}(\rho_{\mathcal{Y}_t})}\,\mathrm{d}W_t^n$$

Key idea: If the covariances matrices $\mathcal{C}(\rho_{\mathcal{X}_t})$ and $\mathcal{C}(\rho_{\mathcal{Y}_t})$ are close, then the two systems would be close too

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{E}_t^p \leq -2\mathcal{E}_t^p + \frac{C_p}{\sqrt{N}}\left(\mathcal{E}_t^{p+1} + \mathcal{F}_t^{p+1}\right) \\ \frac{\mathrm{d}}{\mathrm{d}t}\mathcal{F}_t^p \leq -2\mathcal{F}_t^p + \frac{C_p}{N}\left(\mathcal{E}_t^{p+1} + \mathcal{F}_t^{p+1}\right) \end{cases} \quad \text{with} \quad \begin{cases} \mathcal{E}_t^p = \mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}\left|X_t^n - Y_t^n\right|^{2p}\right]^{\frac{1}{2p}} \\ \mathcal{F}_t^p = \mathbb{E}\left[\left\|C_{\rho_{\mathcal{X}_t}}^{-1} - C_{\rho_{\mathcal{Y}_t}}^{-1}\right\|_{\mathrm{F}}^{2p}\right]^{\frac{1}{2p}} \end{cases}$$

This leads to a hierarchy of moments that is difficult to close

Idea of proof: difficulty with covariance moments

The hierarchy of moments leads to results with a remainders:

$$\mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}|X_t^n - Y_t^n|^2\right] \leq C\,\mathrm{e}^{-(2+\varepsilon(N))t} + \frac{C_{\mathbf{p}}}{N^{\mathbf{p}}}$$

for some **p** arbitrarily large but fixed.

Idea of proof: difficulty with covariance moments

The hierarchy of moments leads to results with a remainders:

$$\mathbb{E}\left[\frac{1}{N}\sum_{n=1}^{N}|X_t^n - Y_t^n|^2\right] \leq C\,\mathrm{e}^{-(2+\varepsilon(N))t} + \frac{C_{\mathbf{p}}}{N^{\mathbf{p}}}$$

for some **p** arbitrarily large but fixed.

Better idea: before doing sticky coupling, recenter the system:

$$U_t^n = C(\rho_{\mathcal{X}_t})^{-\frac{1}{2}}(X_t^n - \mathcal{M}_{\mathcal{X}_t})$$

It is easy to do in $d = 1$.

In larger dimensions, we need the Gradient and Hessian of $\varphi(A) = \sqrt{A}$ for A a positive definite symetric matrix