On Couplings for Kinetic Langevin Diffusions

Nawaf Bou-Rabee
Rutgers, Flatiron & Wharton

joint with Sonja Cox and Roy Schieven
(University of Amsterdam)

`arXiv:2605.31088`

Sampling on $\mathbb{R}^d$ via diffusions

- **Target.** $\mu_{\text{target}}(dx) \propto e^{-U(x)}\,dx$ on $\mathbb{R}^d$, with $U \in C^1$.

Sampling on $\mathbb{R}^d$ via diffusions

- **Target.** $\mu_{\mathrm{target}}(dx) \propto e^{-U(x)}\,dx$ on $\mathbb{R}^d$, with $U \in C^1$.
- **Overdamped Langevin.**

$$dX_t = -\nabla U(X_t)\,dt + \sqrt{2}\,dW_t \qquad \text{leaves } \mu_{\mathrm{target}} \text{ invariant.}$$

Sampling on $\mathbb{R}^d$ via diffusions

- **Target.** $\mu_{\text{target}}(dx) \propto e^{-U(x)}\,dx$ on $\mathbb{R}^d$, with $U \in C^1$.
- **Overdamped Langevin.**

$$dX_t = -\nabla U(X_t)\,dt + \sqrt{2}\,dW_t \qquad \text{leaves } \mu_{\text{target}} \text{ invariant.}$$

- **Kinetic Langevin.** Lift to $\mathbb{R}^{2d}$ by adding velocity:

$$\boxed{\begin{aligned} dX_t &= V_t\,dt, \\ dV_t &= -\nabla U(X_t)\,dt - \gamma V_t\,dt + \sqrt{2\gamma}\,dW_t \end{aligned}}$$

leaves the Boltzmann–Gibbs measure $\mu = \mu_{\text{target}} \otimes \mathcal{N}(0, \mathrm{Id}_d)$ invariant.

Sampling on $\mathbb{R}^d$ via diffusions

- **Target.** $\mu_{\text{target}}(dx) \propto e^{-U(x)}\,dx$ on $\mathbb{R}^d$, with $U \in C^1$.
- **Overdamped Langevin.**

$$dX_t = -\nabla U(X_t)\,dt + \sqrt{2}\,dW_t \qquad \text{leaves } \mu_{\text{target}} \text{ invariant.}$$

- **Kinetic Langevin.** Lift to $\mathbb{R}^{2d}$ by adding velocity:

$$\boxed{\begin{aligned} dX_t &= V_t\,dt, \\ dV_t &= -\nabla U(X_t)\,dt - \gamma V_t\,dt + \sqrt{2\gamma}\,dW_t \end{aligned}}$$

  leaves the Boltzmann–Gibbs measure $\mu = \mu_{\text{target}} \otimes \mathcal{N}(0, \mathrm{Id}_d)$ invariant.

- Noise enters *only the velocity* ⇒ hypoelliptic.

Diffusive-to-ballistic acceleration

- **Overdamped Langevin**[1]: under a Poincaré inequality with constant m,

$$T_{\mathrm{relax}} \asymp m^{-1} \quad \text{(diffusive)}.$$

---

[1] (Bakry, Gentil & Ledoux (2014), Thm. 4.2.5)

Diffusive-to-ballistic acceleration

- **Overdamped Langevin**[1]: under a Poincaré inequality with constant m,

$$T_{\mathrm{relax}} \asymp m^{-1} \quad \text{(diffusive)}.$$

- **Critically-tuned kinetic Langevin**[2]:

$$T_{\mathrm{relax}} \asymp m^{-1/2} \quad \text{(ballistic)}.$$

---

[1] (Bakry, Gentil & Ledoux (2014), Thm. 4.2.5)

[2] (Cao, Lu & Wang (2023); Eberle & Lörler (2024))

Diffusive-to-ballistic acceleration

- **Overdamped Langevin**[1]: under a Poincaré inequality with constant m,

$$T_{\mathrm{relax}} \asymp m^{-1} \quad \text{(diffusive)}.$$

- **Critically-tuned kinetic Langevin**[2]:

$$T_{\mathrm{relax}} \asymp m^{-1/2} \quad \text{(ballistic)}.$$

- A *square-root speedup* in ill-conditioned problems; the motivation for studying kinetic Langevin and its splitting discretizations.

---

[1] (Bakry, Gentil & Ledoux (2014), Thm. 4.2.5)
[2] (Cao, Lu & Wang (2023); Eberle & Lörler (2024))

From SDE to MCMC: splitting schemes

- Decompose the generator into three solvable flows:

$$\underbrace{O_h}_{\text{OU step (exact)}}, \quad \underbrace{\theta_h^{(\mathrm{A})}}_{\text{kinematic}}, \quad \underbrace{\theta_h^{(\mathrm{B})}}_{\text{kick by } -\nabla U}.$$

From SDE to MCMC: splitting schemes

- Decompose the generator into three solvable flows:

$$\underbrace{O_h}_{\text{OU step (exact)}}, \quad \underbrace{\theta_h^{(\mathrm{A})}}_{\text{kinematic}}, \quad \underbrace{\theta_h^{(\mathrm{B})}}_{\text{kick by } -\nabla U}.$$

- **OBABO splitting**:

$$\boxed{\Phi_h^{\mathrm{OBABO}} = O_{h/2} \circ \theta_{h/2}^{(\mathrm{B})} \circ \theta_h^{(\mathrm{A})} \circ \theta_{h/2}^{(\mathrm{B})} \circ O_{h/2}}$$

From SDE to MCMC: splitting schemes

- Decompose the generator into three solvable flows:

$$\underbrace{O_h}_{\text{OU step (exact)}}, \quad \underbrace{\theta_h^{(\mathrm{A})}}_{\text{kinematic}}, \quad \underbrace{\theta_h^{(\mathrm{B})}}_{\text{kick by } -\nabla U}.$$

- **OBABO splitting**[3]:

$$\boxed{\Phi_h^{\mathrm{OBABO}} = O_{h/2} \circ \theta_{h/2}^{(\mathrm{B})} \circ \theta_h^{(\mathrm{A})} \circ \theta_{h/2}^{(\mathrm{B})} \circ O_{h/2}}$$

- Second-order weakly accurate.

---

[3] (Bussi & Parrinello (2007); Leimkuhler & Matthews (2013))

From SDE to MCMC: splitting schemes

- Decompose the generator into three solvable flows:

$$\underbrace{O_h}_{\text{OU step (exact)}}, \quad \underbrace{\theta_h^{(\mathrm{A})}}_{\text{kinematic}}, \quad \underbrace{\theta_h^{(\mathrm{B})}}_{\text{kick by } -\nabla U}.$$

- **OBABO splitting**[3]:

$$\boxed{\Phi_h^{\mathrm{OBABO}} = O_{h/2} \circ \theta_{h/2}^{(\mathrm{B})} \circ \theta_h^{(\mathrm{A})} \circ \theta_{h/2}^{(\mathrm{B})} \circ O_{h/2}}$$

- Second-order weakly accurate.
- One step is a deterministic map of two Gaussian increments $\xi_k = (\xi_k^{(1)}, \xi_k^{(2)}) \sim \mathcal{N}(0, \mathrm{Id}_{2d})$:

$$Z_{k+1}^h = \Psi_{Z_k^h}(\xi_{k+1}), \qquad Z_k^h = (X_k^h, V_k^h).$$

---

[3] (Bussi & Parrinello (2007); Leimkuhler & Matthews (2013))

Question

- (Z_n^h) on $\mathbb{R}^{2d}$: OBABO chain, step h, friction γ, invariant law $\mu = \mu_{\text{target}} \otimes \mathcal{N}(0, \mathrm{Id}_d)$.

Question

- (Z_n^h) on $\mathbb{R}^{2d}$: OBABO chain, step h, friction γ, invariant law $\mu = \mu_{\mathsf{target}} \otimes \mathcal{N}(0, \mathrm{Id}_d)$.
- **TV mixing time** from initial law ν:

$$t_{\mathrm{mix}}(\nu, \varepsilon) := \inf\left\{ n \geq 0 \ : \ d_{\mathsf{TV}}(\nu P^n, \mu) \leq \varepsilon \right\}.$$

Question

- (Z_n^h) on $\mathbb{R}^{2d}$: OBABO chain, step h, friction γ, invariant law $\mu = \mu_{\mathsf{target}} \otimes \mathcal{N}(0, \mathrm{Id}_d)$.
- **TV mixing time** from initial law ν:

$$t_{\mathrm{mix}}(\nu, \varepsilon) := \inf\{\, n \geq 0 \;:\; d_{\mathsf{TV}}(\nu P^n, \mu) \leq \varepsilon \,\}.$$

- Curvature of the target: $m\,\mathrm{Id} \preceq \nabla^2 U \preceq L\,\mathrm{Id}$, condition number $\kappa := L/m$.

Question

- (Z_n^h) on $\mathbb{R}^{2d}$: OBABO chain, step h, friction γ, invariant law $\mu = \mu_{\text{target}} \otimes \mathcal{N}(0, \mathrm{Id}_d)$.
- **TV mixing time** from initial law ν:

$$t_{\mathrm{mix}}(\nu, \varepsilon) := \inf\left\{ n \geq 0 \,:\, d_{\mathsf{TV}}(\nu P^n, \mu) \leq \varepsilon \right\}.$$

- Curvature of the target: $m\,\mathrm{Id} \preceq \nabla^2 U \preceq L\,\mathrm{Id}$, condition number $\kappa := L/m$.

For which initial laws ν does OBABO attain the **acceleration** $t_{\mathrm{mix}}(\nu, \varepsilon) \asymp \kappa^{1/2} \log(1/\varepsilon)$?

Our approach: couplings + Wasserstein-to-TV regularisation

Two pieces of standard machinery.

Our approach: couplings + Wasserstein-to-TV regularisation

Two pieces of standard machinery.

1. Coupling characterisation of TV.

$$d_{\mathsf{TV}}(\nu_1, \nu_2) \;=\; \inf_{(X,\tilde{X}):\, X\sim\nu_{\mathbf{1}},\, \tilde{X}\sim\nu_{\mathbf{2}}} \mathbb{P}(X \neq \tilde{X}).$$

Construct a coupling of the two chains and bound $\mathbb{P}$(disagree).

Our approach: couplings + Wasserstein-to-TV regularisation

Two pieces of standard machinery.

1. Coupling characterisation of TV.

$$d_{\mathsf{TV}}(\nu_1, \nu_2) \;=\; \inf_{(X,\tilde{X}):\, X\sim\nu_{\mathbf{1}},\, \tilde{X}\sim\nu_{\mathbf{2}}} \mathbb{P}(X \neq \tilde{X}).$$

Construct a coupling of the two chains and bound $\mathbb{P}(\text{disagree})$.

2. Wasserstein-to-TV regularisation.

$$d_{\mathsf{TV}}\big(\delta_z P^n, \delta_{\tilde{z}} P^n\big) \;\leq\; C_n\,|z-\tilde{z}| \;\Longrightarrow\; d_{\mathsf{TV}}(\nu P^n, \tilde{\nu} P^n) \;\leq\; C_n\, W_1(\nu, \tilde{\nu}),$$

via joint convexity of d_{TV} along a W_1-optimal coupling.

Our approach: couplings + Wasserstein-to-TV regularisation

Two pieces of standard machinery.

1. Coupling characterisation of TV.

$$d_{\mathrm{TV}}(\nu_1,\nu_2) \;=\; \inf_{(X,\tilde{X}):\,X\sim\nu_1,\,\tilde{X}\sim\nu_2} \mathbb{P}(X\neq\tilde{X}).$$

Construct a coupling of the two chains and bound $\mathbb{P}(\text{disagree})$.

2. Wasserstein-to-TV regularisation.

$$d_{\mathrm{TV}}\big(\delta_z P^n,\delta_{\tilde{z}} P^n\big) \;\le\; C_n\,|z-\tilde{z}| \;\Longrightarrow\; d_{\mathrm{TV}}(\nu P^n,\tilde{\nu} P^n) \;\le\; C_n\,W_1(\nu,\tilde{\nu}),$$

via joint convexity of d_{TV} along a W_1-optimal coupling.

High-acceptance localisation[4] transfers these bounds to the Metropolis-adjusted chain.

---

[4] (B.-R. & Oberdörster (2024), EJP)

Markovian vs. non-Markovian couplings

Let $\mathcal{F}_k = \sigma\big((Z_j^h, \tilde{Z}_j^h) : 0 \leq j \leq k\big)$.

- A coupling $(Z_k^h, \tilde{Z}_k^h)_{k\geq 0}$ is **Markovian** if the joint chain is Markov w.r.t. $(\mathcal{F}_k)$.

Markovian vs. non-Markovian couplings

Let $\mathcal{F}_k = \sigma\big((Z_j^h, \tilde{Z}_j^h) : 0 \leq j \leq k\big)$.

- A coupling $(Z_k^h, \tilde{Z}_k^h)_{k\geq 0}$ is **Markovian** if the joint chain is Markov w.r.t. $(\mathcal{F}_k)$.
- **Now-equals-forever**: $Z_s^h = \tilde{Z}_s^h \Rightarrow Z_t^h = \tilde{Z}_t^h$ for $t \geq s$, a.s.

Markovian vs. non-Markovian couplings

Let $\mathcal{F}_k = \sigma\big((Z_j^h, \tilde{Z}_j^h) : 0 \le j \le k\big)$.

- A coupling $(Z_k^h, \tilde{Z}_k^h)_{k \ge 0}$ is **Markovian** if the joint chain is Markov w.r.t. $(\mathcal{F}_k)$.
- **Now-equals-forever**: $Z_s^h = \tilde{Z}_s^h \Rightarrow Z_t^h = \tilde{Z}_t^h$ for $t \ge s$, a.s.
- **Non-Markovian**: the joint chain is not adapted as a Markov chain to $(\mathcal{F}_k)$; the driving noises $(\tilde{\xi}_k)$ may depend on the entire sequence (ξ_j) at once.

A brief history of couplings for *hypoelliptic* SDEs

- **Stochastic oscillator.**[5]
  First hypoelliptic coupling: switching between synchronous and antithetic drivers.

---

[5] (Ben Arous, Cranston & Kendall (1995))

A brief history of couplings for *hypoelliptic* SDEs

- **Stochastic oscillator.**[5]
  First hypoelliptic coupling: switching between synchronous and antithetic drivers.
- **Kolmogorov diffusion** $\left(W_t, \int_0^t W_s \, ds\right)$.[6]
  No Markovian coupling matches the asymptotic TV rate.

---

[5] (Ben Arous, Cranston & Kendall (1995))

[6] (Banerjee & Kendall (2016))

A brief history of couplings for *hypoelliptic* SDEs

- **Stochastic oscillator.**[5]
  First hypoelliptic coupling: switching between synchronous and antithetic drivers.
- **Kolmogorov diffusion** $\left(W_t, \int_0^t W_s\, ds\right)$.[6]
  No Markovian coupling matches the asymptotic TV rate.
- **Wasserstein contraction for kinetic Langevin.**[7]
  Hybrid sync/reflection in phase space, non-convex U.
  - Discrete-time:[8] step-size-restricted analogues for OBABO and variants.

---

[5] (Ben Arous, Cranston & Kendall (1995))
[6] (Banerjee & Kendall (2016))
[7] (Eberle, Guillin & Zimmer (2019))
[8] (Cheng et al. (2018, 2020); Dalalyan & Riou-Durand (2020); Leimkuhler, Paulin & Whalley (2024); Schuh & Whalley (2025))

A brief history of couplings for *hypoelliptic* SDEs

- **Stochastic oscillator.**[5]
  First hypoelliptic coupling: switching between synchronous and antithetic drivers.
- **Kolmogorov diffusion** $\left(W_t, \int_0^t W_s\, ds\right)$.[6]
  No Markovian coupling matches the asymptotic TV rate.
- **Wasserstein contraction for kinetic Langevin.**[7]
  Hybrid sync/reflection in phase space, non-convex U.
  - Discrete-time:[8] step-size-restricted analogues for OBABO and variants.
- **TV mixing via Wasserstein-to-TV regularisation.**[9]
  Hypoelliptic smoothing transfers Wasserstein to TV control.

---

[5] (Ben Arous, Cranston & Kendall (1995))

[6] (Banerjee & Kendall (2016))

[7] (Eberle, Guillin & Zimmer (2019))

[8] (Cheng et al. (2018, 2020); Dalalyan & Riou-Durand (2020); Leimkuhler, Paulin & Whalley (2024); Schuh & Whalley (2025))

[9] (Roberts & Rosenthal (2002); Madras & Sezer (2010); Monmarché (2021); Gouraud et al. (2025); Chak & Monmarché (2025))

Prior work: Chak & Monmarché (2025)

Theorem (Chak & Monmarché, 2025, Thm. 2.2)

Let $U \in C^2(\mathbb{R}^d)$ *satisfy*

- *bounded Hessian:* $|\nabla^2 U(x)\, v| \leq L\, |v|$ *for all* v, x*,*
- *Hessian-Lipschitz (Frobenius):* $\|\nabla^2 U(x) - \nabla^2 U(y)\|_F \leq L_H\, |x - y|$*.*

There exist $\beta, h_0 > 0$ *depending on* L, L_H *such that for all* $h \leq h_0$*,* $n \geq 2$ *with* $hn \leq 1$*:*

$$d_{\mathrm{TV}}\big(\pi_n(\delta_z),\, \pi_n(\delta_{\tilde{z}})\big) \;\leq\; \beta\Big[\tfrac{1}{(hn)^{3/2}} + \tfrac{1}{(hn)^{1/2}} + (hn)^{1/2}\Big]\, |\tilde{z} - z|.$$

Prior work: Chak & Monmarché (2025)

Theorem (Chak & Monmarché, 2025, Thm. 2.2)

Let $U \in C^2(\mathbb{R}^d)$ *satisfy*

- *bounded Hessian:* $|\nabla^2 U(x)\, v| \le L\,|v|$ *for all* v, x,
- *Hessian-Lipschitz (Frobenius):* $\|\nabla^2 U(x) - \nabla^2 U(y)\|_F \le L_H\,|x-y|$.

There exist $\beta, h_0 > 0$ *depending on* L, L_H *such that for all* $h \le h_0$, $n \ge 2$ *with* $hn \le 1$:

$$d_{\mathrm{TV}}\big(\pi_n(\delta_z),\, \pi_n(\delta_{\tilde z})\big) \;\le\; \beta\Big[\tfrac{1}{(hn)^{3/2}} + \tfrac{1}{(hn)^{1/2}} + (hn)^{1/2}\Big]\; |\tilde z - z|.$$

- **Mechanism.** An explicit *coalescence map* $\Psi^n_{z,\tilde z}$ built from an ansatz, with bound $d_{\mathrm{TV}}(\pi_n(\delta_z), \pi_n(\delta_{\tilde z})) \le d_{\mathrm{TV}}(\mathrm{Law}(\xi), \mathrm{Law}(\Psi^n_{z,\tilde z}(\xi)))$ closed by [1].

---

[1] (B.-R. & Eberle, 2023, Lem. 15)

Prior work: Chak & Monmarché (2025)

Theorem (Chak & Monmarché, 2025, Thm. 2.2)

Let $U \in C^2(\mathbb{R}^d)$ *satisfy*

- *bounded Hessian:* $|\nabla^2 U(x)\, v| \le L\,|v|$ *for all* v, x,
- *Hessian-Lipschitz (Frobenius):* $\|\nabla^2 U(x) - \nabla^2 U(y)\|_F \le L_H\,|x-y|$.

There exist $\beta, h_0 > 0$ *depending on* L, L_H *such that for all* $h \le h_0$, $n \ge 2$ *with* $hn \le 1$:

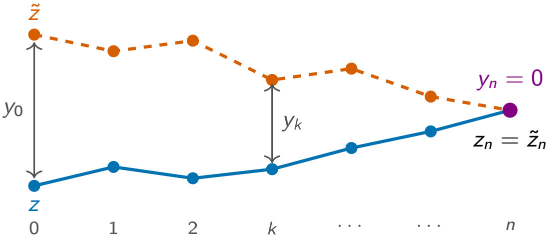
$$d_{\mathrm{TV}}\big(\pi_n(\delta_z),\, \pi_n(\delta_{\tilde z})\big) \;\le\; \beta\Big[\tfrac{1}{(hn)^{3/2}} + \tfrac{1}{(hn)^{1/2}} + (hn)^{1/2}\Big]\; |\tilde z - z|.$$

- **Mechanism.** An explicit *coalescence map* $\Psi^n_{z,\tilde z}$ built from an ansatz, with bound $d_{\mathrm{TV}}(\pi_n(\delta_z), \pi_n(\delta_{\tilde z})) \le d_{\mathrm{TV}}(\mathrm{Law}(\xi), \mathrm{Law}(\Psi^n_{z,\tilde z}(\xi)))$ closed by [1].
- **Our take.** Read $\Psi^n_{z,\tilde z}$ as the *noise transport* of a non-Markovian coupling; *optimise* the trajectory and *realise* the coupling.

---

[1] (B.-R. & Eberle, 2023, Lem. 15)

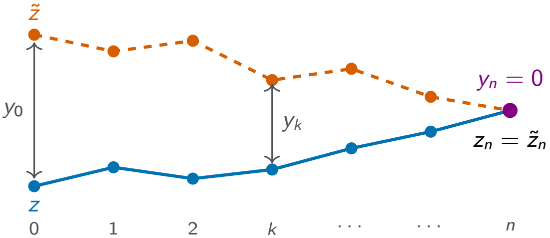
Coalescence map



Chain map. OBABO from z driven by noise $\xi = (\xi_1, \ldots, \xi_n)$:

$$z = z_0 \to z_1 \to \cdots \to z_n =: \Psi_z^n(\xi).$$

Coalescence map

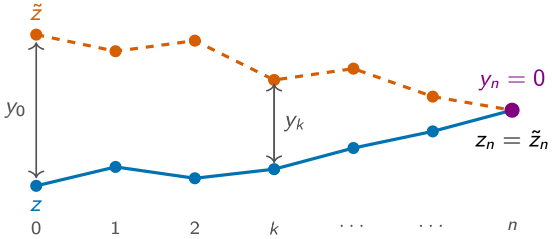


Chain map. OBABO from z driven by noise $\xi = (\xi_1, \ldots, \xi_n)$:

$$z = z_0 \to z_1 \to \cdots \to z_n =: \Psi_z^n(\xi).$$

Gap trajectory. $y_k := \tilde{z}_k - z_k$ (the bars in the picture), with $y_0 = \tilde{z} - z$ and $y_n = 0$.

Coalescence map



Chain map. OBABO from z driven by noise $\xi = (\xi_1, \dots, \xi_n)$:

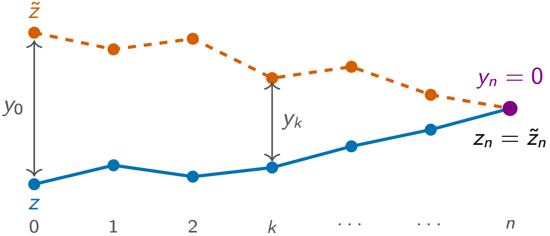
$$z = z_0 \to z_1 \to \cdots \to z_n =: \Psi_z^n(\xi).$$

Gap trajectory. $y_k := \tilde{z}_k - z_k$ (the bars in the picture), with $y_0 = \tilde{z} - z$ and $y_n = 0$.

Coalescence map. Given the noise ξ for the z-chain, $\tilde{\xi} := \Psi_{z,\tilde{z}}^n(\xi)$ is the noise that drives the $\tilde{z}$-chain s.t.:

$$\Psi_{\tilde{z}}^n\big(\Psi_{z,\tilde{z}}^n(\xi)\big) = \Psi_z^n(\xi).$$

Coalescence map



Chain map. OBABO from z driven by noise $\xi = (\xi_1, \ldots, \xi_n)$:

$$z = z_0 \to z_1 \to \cdots \to z_n =: \Psi_z^n(\xi).$$

Gap trajectory. $y_k := \tilde{z}_k - z_k$ (the bars in the picture), with $y_0 = \tilde{z} - z$ and $y_n = 0$.

Coalescence map. Given the noise ξ for the z-chain, $\tilde{\xi} := \Psi_{z,\tilde{z}}^n(\xi)$ is the noise that drives the $\tilde{z}$-chain s.t.:

$$\Psi_{\tilde{z}}^n\big(\Psi_{z,\tilde{z}}^n(\xi)\big) = \Psi_z^n(\xi).$$

Optimization. The interior $y_1, \ldots, y_{n-1}$ is free, and later picked by LQ optimization.

General lemma: TV bound between a reference and a perturbed Gaussian

Lemma (B.-R. & Eberle, 2023, Lem. 15)

Let $\xi \sim \mathcal{N}(0, \mathrm{Id}_d)$ *and* $\Phi\colon \mathbb{R}^d \to \mathbb{R}^d$ *be a* C^1 *diffeomorphism. Then*

$$d_{\mathsf{TV}}\big(\mathrm{Law}(\xi), \mathrm{Law}(\Phi(\xi))\big) \ \leq \ \sqrt{2\,\mathrm{KL}},$$

where $\mathrm{KL} \ = \ \mathbb{E}\big[\tfrac{1}{2}|\Phi(\xi) - \xi|^2 \ + \ \mathrm{tr}\big(D\Phi(\xi) - \mathrm{Id}\big) \ - \ \log|\det D\Phi(\xi)|\big]\,.$

General lemma: TV bound between a reference and a perturbed Gaussian

Lemma (B.-R. & Eberle, 2023, Lem. 15)

Let $\xi \sim \mathcal{N}(0, \mathrm{Id}_d)$ *and* $\Phi \colon \mathbb{R}^d \to \mathbb{R}^d$ *be a* C^1 *diffeomorphism. Then*

$$d_{\mathsf{TV}}\big(\mathrm{Law}(\xi), \mathrm{Law}(\Phi(\xi))\big) \;\le\; \sqrt{2\,\mathrm{KL}},$$

where $\mathrm{KL} \;=\; \mathbb{E}\big[\tfrac{1}{2}|\Phi(\xi) - \xi|^2 \;+\; \mathrm{tr}\big(D\Phi(\xi) - \mathrm{Id}\big) \;-\; \log|\det D\Phi(\xi)|\big]\,.$

Sketch.

- Pinsker gives the first inequality;
- Change of variables yields $\mathrm{KL} = \mathbb{E}[\frac{1}{2}|\Phi(\xi) - \xi|^2 + (\Phi(\xi) - \xi)\cdot\xi - \log|\det D\Phi(\xi)|]$;
- Gaussian IBP (Stein) replaces the middle term: $\mathbb{E}[(\Phi(\xi) - \xi)\cdot\xi] = \mathbb{E}\,\mathrm{tr}(D\Phi(\xi) - \mathrm{Id})$.

☐

General lemma: TV bound between a reference and a perturbed Gaussian

Lemma (B.-R. & Eberle, 2023, Lem. 15)

Let $\xi \sim \mathcal{N}(0, \mathrm{Id}_d)$ *and* $\Phi \colon \mathbb{R}^d \to \mathbb{R}^d$ *be a* C^1 *diffeomorphism. Then*

$$d_{\mathsf{TV}}\big(\mathrm{Law}(\xi), \mathrm{Law}(\Phi(\xi))\big) \;\le\; \sqrt{2\,\mathrm{KL}},$$

where $\mathrm{KL} \;=\; \mathbb{E}\big[\tfrac{1}{2}|\Phi(\xi) - \xi|^2 \;+\; \mathrm{tr}\big(D\Phi(\xi) - \mathrm{Id}\big) \;-\; \log|\det D\Phi(\xi)|\big]\,.$

Sketch.

- Pinsker gives the first inequality;
- Change of variables yields $\mathrm{KL} = \mathbb{E}[\frac{1}{2}|\Phi(\xi) - \xi|^2 + (\Phi(\xi) - \xi)\cdot\xi - \log|\det D\Phi(\xi)|]$;
- Gaussian IBP (Stein) replaces the middle term: $\mathbb{E}[(\Phi(\xi) - \xi)\cdot\xi] = \mathbb{E}\,\mathrm{tr}(D\Phi(\xi) - \mathrm{Id})$.

□

Remark (OBABO). For the OBABO coalescence map (next slide), $D\Psi^n_{z,\tilde z}$ is block lower-triangular with identity diagonal: $\mathrm{tr}(D\Psi - \mathrm{Id}) \equiv 0$, $\det D\Psi \equiv 1$, and $d_{\mathsf{TV}} \le \mathbb{E}[|\Psi^n_{z,\tilde z}(\xi) - \xi|^2]^{1/2}$.

Why OBABO has a lower-triangular coalescence Jacobian

Lemma (B.-R.–Cox–Schieven)

$D\Psi^n_{z,\tilde{z}}$ is block lower-triangular with identity diagonal, so $\det D\Psi^n_{z,\tilde{z}} \equiv 1$.

Why OBABO has a lower-triangular coalescence Jacobian

Lemma (B.-R.–Cox–Schieven)

$D\Psi^n_{z,\tilde{z}}$ *is block lower-triangular with identity diagonal, so* det $D\Psi^n_{z,\tilde{z}} \equiv 1$.

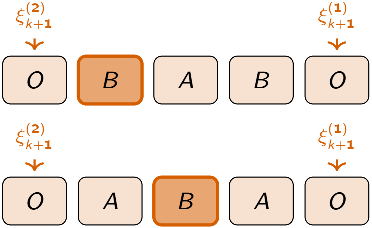
Why? Inductively, $\tilde{\xi}_i = \xi_i + g_i(\xi_1, \dots, \xi_{i-1})$. Hence $\dfrac{\partial \tilde{\xi}_i}{\partial \xi_j} = \begin{cases} I, & j = i, \\ *, & j < i, \\ 0, & j > i, \end{cases}.$

Why OBABO has a lower-triangular coalescence Jacobian

Lemma (B.-R.–Cox–Schieven)

$D\Psi^n_{z,\tilde z}$ is block lower-triangular with identity diagonal, so det $D\Psi^n_{z,\tilde z} \equiv 1$.

Why? Inductively, $\tilde\xi_i = \xi_i + g_i(\xi_1, \dots, \xi_{i-1})$. Hence $\frac{\partial \tilde\xi_i}{\partial \xi_j} = \begin{cases} I, & j = i, \\ *, & j < i, \\ 0, & j > i, \end{cases}$.



$\xi^{(2)}_{k+1}$ $\xi^{(1)}_{k+1}$

OBABO O B A B O q_{k+1} is *affine* in $\xi^{(1)}$

$\xi^{(2)}_{k+1}$ $\xi^{(1)}_{k+1}$

OABAO O A B A O q_{k+1} is *nonlinear* in $\xi^{(1)}$

Quantitative TV bound for OBABO

Theorem (B.-R.–Cox–Schieven)

Let $\gamma, h > 0$, $n \in \mathbb{N}$, *and* $U \in C^2(\mathbb{R}^d)$ *with* ∇U *L-Lipschitz. Then for all* $z, \tilde{z} \in \mathbb{R}^{2d}$*:*

$$d_{\mathrm{TV}}(\pi_n(\delta_z), \pi_n(\delta_{\tilde{z}})) \leq \frac{1}{\gamma^{1/2}}\left[\frac{5}{(hn)^{3/2}} + \frac{12+5\gamma}{(hn)^{1/2}} + \frac{L(\gamma h)^{1/2}}{(1-e^{-\gamma h})^{1/2}}\left(1+\frac{hn}{1+\gamma hn}\right)(hn)^{1/2}\right]|\tilde{z}-z|.$$

Quantitative TV bound for OBABO

Theorem (B.-R.–Cox–Schieven)

Let $\gamma, h > 0$, $n \in \mathbb{N}$, *and* $U \in C^2(\mathbb{R}^d)$ *with* ∇U *L-Lipschitz. Then for all* $z, \tilde{z} \in \mathbb{R}^{2d}$*:*

$$d_{\mathrm{TV}}(\pi_n(\delta_z), \pi_n(\delta_{\tilde{z}})) \leq \frac{1}{\gamma^{1/2}} \left[\frac{5}{(hn)^{3/2}} + \frac{12 + 5\gamma}{(hn)^{1/2}} + \frac{L(\gamma h)^{1/2}}{(1 - e^{-\gamma h})^{1/2}} \left(1 + \frac{hn}{1 + \gamma hn} \right) (hn)^{1/2} \right] |\tilde{z} - z|.$$

- Holds for **all** $h > 0$ and $n \in \mathbb{N}$ (no $hn \leq 1$, no $h \leq h_0$).

Quantitative TV bound for OBABO

Theorem (B.-R.–Cox–Schieven)

Let $\gamma, h > 0$, $n \in \mathbb{N}$, *and* $U \in C^2(\mathbb{R}^d)$ *with* ∇U *L-Lipschitz. Then for all* $z, \tilde{z} \in \mathbb{R}^{2d}$*:*

$$d_{\mathrm{TV}}\big(\pi_n(\delta_z), \pi_n(\delta_{\tilde{z}})\big) \leq \frac{1}{\gamma^{1/2}}\left[\frac{5}{(hn)^{3/2}} + \frac{12+5\gamma}{(hn)^{1/2}} + \frac{L(\gamma h)^{1/2}}{(1-e^{-\gamma h})^{1/2}}\left(1+\frac{hn}{1+\gamma hn}\right)(hn)^{1/2}\right]|\tilde{z}-z|.$$

- Holds for **all** $h > 0$ and $n \in \mathbb{N}$ (no $hn \leq 1$, no $h \leq h_0$).
- Only requires ∇U Lipschitz (no Lipschitz Hessian).

Quantitative TV bound for OBABO

Theorem (B.-R.–Cox–Schieven)

Let $\gamma, h > 0$, $n \in \mathbb{N}$, *and* $U \in C^2(\mathbb{R}^d)$ *with* ∇U *L-Lipschitz. Then for all* $z, \tilde{z} \in \mathbb{R}^{2d}$*:*

$$d_{\mathrm{TV}}(\pi_n(\delta_z), \pi_n(\delta_{\tilde{z}})) \leq \frac{1}{\gamma^{1/2}}\left[\frac{5}{(hn)^{3/2}} + \frac{12+5\gamma}{(hn)^{1/2}} + \frac{L(\gamma h)^{1/2}}{(1-e^{-\gamma h})^{1/2}}\left(1+\frac{hn}{1+\gamma hn}\right)(hn)^{1/2}\right]|\tilde{z}-z|.$$

- Holds for **all** $h > 0$ and $n \in \mathbb{N}$ (no $hn \leq 1$, no $h \leq h_0$).
- Only requires ∇U Lipschitz (no Lipschitz Hessian).
- Yields a Wasserstein-to-TV regularization for any initial laws $\nu, \tilde{\nu}$.

Quantitative TV bound for OBABO

Theorem (B.-R.–Cox–Schieven)

Let $\gamma, h > 0$, $n \in \mathbb{N}$, *and* $U \in C^2(\mathbb{R}^d)$ *with* ∇U *L-Lipschitz. Then for all* $z, \tilde{z} \in \mathbb{R}^{2d}$*:*

$$d_{\mathrm{TV}}(\pi_n(\delta_z), \pi_n(\delta_{\tilde{z}})) \leq \frac{1}{\gamma^{1/2}}\left[\frac{5}{(hn)^{3/2}} + \frac{12+5\gamma}{(hn)^{1/2}} + \frac{L(\gamma h)^{1/2}}{(1-e^{-\gamma h})^{1/2}}\left(1+\frac{hn}{1+\gamma hn}\right)(hn)^{1/2}\right]|\tilde{z}-z|.$$

- Holds for **all** $h > 0$ and $n \in \mathbb{N}$ (no $hn \leq 1$, no $h \leq h_0$).
- Only requires ∇U Lipschitz (no Lipschitz Hessian).
- Yields a Wasserstein-to-TV regularization for any initial laws $\nu, \tilde{\nu}$.
- **Realized by an explicit non-Markovian coupling**.

The optimal noise-transport trajectory (force-free)

Force-free bound. For $\nabla U \equiv 0$, Lemma 15 reduces to a controlled L^2 cost on the gap trajectory y_k:

$$d_{\mathsf{TV}} \;\leq\; \tfrac{1}{2}\left(\sum_{k=1}^{n} |E_k|^2\right)^{1/2}, \qquad E_{k+1} = -L_h^{-1}(y_{k+1} - A_h y_k).$$

The optimal noise-transport trajectory (force-free)

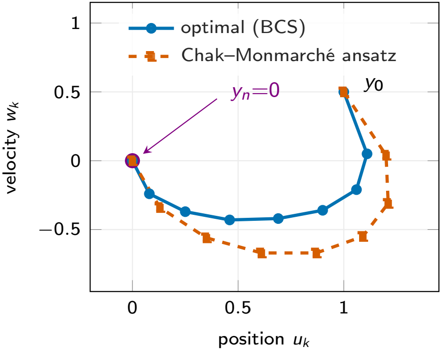
Force-free bound. For $\nabla U \equiv 0$, Lemma 15 reduces to a controlled L^2 cost on the gap trajectory y_k:

$$d_{\mathrm{TV}} \;\le\; \tfrac{1}{2}\Big(\sum_{k=1}^{n} |E_k|^2\Big)^{1/2}, \qquad E_{k+1} = -L_h^{-1}(y_{k+1} - A_h y_k).$$

LQ-optimal trajectory. Minimum-energy control $\min_E \|E\|^2$ s.t. $y_0{=}\Delta z$, $y_n{=}0$. Closed form via controllability Gramian:

$$\sum_k |E_k|^2 \;=\; |\Sigma_{h,n}^{-1/2} A_h^n \Delta z|^2.$$

General U: use this y and treat ∇U as a Lipschitz perturbation.



The optimal noise-transport trajectory (force-free)

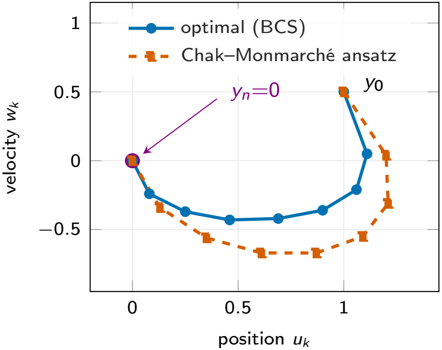
Force-free bound. For $\nabla U \equiv 0$, Lemma 15 reduces to a controlled L^2 cost on the gap trajectory y_k:

$$d_{\mathrm{TV}} \le \frac{1}{2}\left(\sum_{k=1}^{n} |E_k|^2\right)^{1/2}, \qquad E_{k+1} = -L_h^{-1}(y_{k+1} - A_h y_k).$$

LQ-optimal trajectory. Minimum-energy control $\min_E \|E\|^2$ s.t. $y_0 = \Delta z$, $y_n = 0$. Closed form via controllability Gramian:

$$\sum_k |E_k|^2 = |\Sigma_{h,n}^{-1/2} A_h^n \Delta z|^2.$$

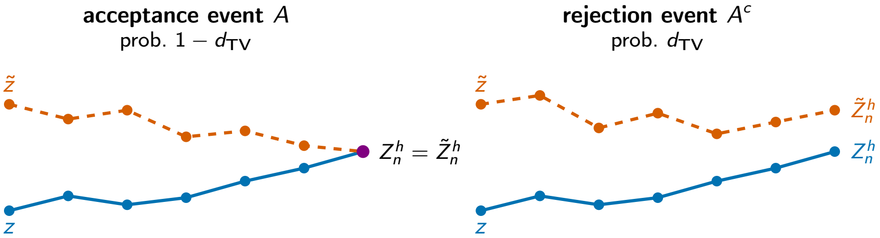
General U: use this y and treat ∇U as a Lipschitz perturbation.



Design principle (EGZ & BCS). Construct for $\nabla U \equiv 0$, then treat ∇U as a Lipschitz perturbation.

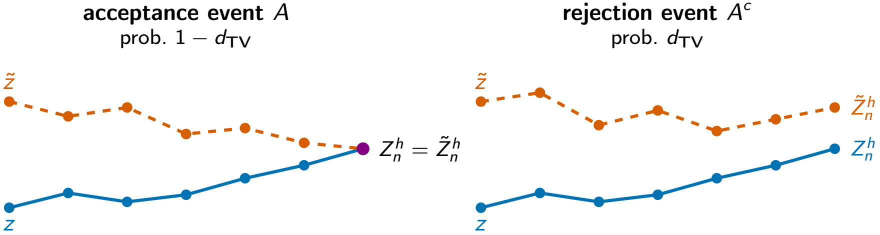
The non-Markovian coupling

- Sample $\xi \sim \mathcal{N}(0, \mathrm{Id}_{2dn})$; proposal $\tilde{\xi}_\star := \Psi^n_{z,\tilde{z}}(\xi)$; maximally couple ξ to $\tilde{\xi}_\star$.
- On accept: $\tilde{\xi} = \tilde{\xi}_\star$ — chains meet at time n. On reject: $\tilde{\xi}$ drawn independently — chains evolve independently.
- $\mathbb{P}(Z^h_n \neq \tilde{Z}^h_n) = d_{\mathsf{TV}}\big(\mathrm{Law}(\xi), \mathrm{Law}(\Psi^n_{z,\tilde{z}}(\xi))\big)$.



The non-Markovian coupling

- Sample $\xi \sim \mathcal{N}(0, \mathrm{Id}_{2dn})$; proposal $\tilde{\xi}_\star := \Psi^n_{z,\tilde{z}}(\xi)$; maximally couple ξ to $\tilde{\xi}_\star$.
- On accept: $\tilde{\xi} = \tilde{\xi}_\star$ — chains meet at time n. On reject: $\tilde{\xi}$ drawn independently — chains evolve independently.
- $\mathbb{P}(Z^h_n \neq \tilde{Z}^h_n) = d_{\mathsf{TV}}\big(\mathrm{Law}(\xi), \mathrm{Law}(\Psi^n_{z,\tilde{z}}(\xi))\big)$.



- **Non-Markovian**, designed for horizon n: $\tilde{\xi}$ depends on the whole ξ at once; intermediate chains generally disagree.

Are Markovian couplings enough?

- Test problem: **quadratic potential** $U(x) = \alpha|x|^2$, $\alpha \geq 0$, overdamped regime $\gamma^2 > 4\alpha$.

$$dX_t = V_t\,dt, \qquad dV_t = -\alpha X_t\,dt - \gamma V_t\,dt + \sqrt{2\gamma}\,dW_t.$$

Are Markovian couplings enough?

- Test problem: **quadratic potential** $U(x) = \alpha|x|^2$, $\alpha \geq 0$, overdamped regime $\gamma^2 > 4\alpha$.

$$dX_t = V_t\,dt, \qquad dV_t = -\alpha X_t\,dt - \gamma V_t\,dt + \sqrt{2\gamma}\,dW_t.$$

- Drift matrix has eigenvalues

$$\lambda_\pm = -\tfrac{\gamma}{2} \pm \tfrac{1}{2}\sqrt{\gamma^2 - 4\alpha}, \qquad \lambda_- < \lambda_+ \leq 0.$$

Are Markovian couplings enough?

- Test problem: **quadratic potential** $U(x) = \alpha|x|^2$, $\alpha \geq 0$, overdamped regime $\gamma^2 > 4\alpha$.

$$dX_t = V_t\, dt, \qquad dV_t = -\alpha X_t\, dt - \gamma V_t\, dt + \sqrt{2\gamma}\, dW_t.$$

- Drift matrix has eigenvalues

$$\lambda_\pm = -\tfrac{\gamma}{2} \pm \tfrac{1}{2}\sqrt{\gamma^2 - 4\alpha}, \qquad \lambda_- < \lambda_+ \leq 0.$$

- If $\Delta z = (\Delta x, \Delta v)$ lies in the λ_--eigenspace $\big(\lambda_- \Delta x = \Delta v\big)$:

$$d_{\mathsf{TV}}(\mathrm{Law}(Z_t), \mathrm{Law}(\tilde{Z}_t)) \;\lesssim\; e^{\lambda_- t}|\Delta z|.$$

Are Markovian couplings enough?

- Test problem: **quadratic potential** $U(x) = \alpha|x|^2$, $\alpha \geq 0$, overdamped regime $\gamma^2 > 4\alpha$.

$$dX_t = V_t\,dt, \qquad dV_t = -\alpha X_t\,dt - \gamma V_t\,dt + \sqrt{2\gamma}\,dW_t.$$

- Drift matrix has eigenvalues

$$\lambda_\pm = -\tfrac{\gamma}{2} \pm \tfrac{1}{2}\sqrt{\gamma^2 - 4\alpha}, \qquad \lambda_- < \lambda_+ \leq 0.$$

- If $\Delta z = (\Delta x, \Delta v)$ lies in the λ_--eigenspace $\big(\lambda_- \Delta x = \Delta v\big)$:

$$d_{\mathsf{TV}}(\mathrm{Law}(Z_t), \mathrm{Law}(\tilde{Z}_t)) \;\lesssim\; e^{\lambda_- t}|\Delta z|.$$

- Question: can a *Markovian* coupling match this rate?

No Markovian coupling matches the TV decay

Theorem (B.-R.–Cox–Schieven, continuous time)

Under $\gamma^2{>}4\alpha{\geq}0$*: (upper) if* $\lambda_-\Delta x{=}\Delta v$*,* $\Delta x{\neq}0$*, then* $d_{\mathsf{TV}} \leq Ce^{\lambda_- t}|\Delta z|$*; (lower) for every Markovian coupling* μ *and every* Δz*,* $\mu(Z_t \neq \tilde{Z}_t) \geq c_\mu \min\big(t^{-1/2}, e^{\lambda_+ t}\big)$ *for* $t \geq t_\mu$*.*

No Markovian coupling matches the TV decay

Theorem (B.-R.–Cox–Schieven, continuous time)

Under $\gamma^2{>}4\alpha{\geq}0$*: (upper) if* $\lambda_-\Delta x{=}\Delta v$, $\Delta x{\neq}0$*, then* $d_{\mathrm{TV}} \leq Ce^{\lambda_- t}|\Delta z|$*; (lower) for every Markovian coupling* μ *and every* Δz, $\mu(Z_t \neq \tilde{Z}_t) \geq c_\mu \min\big(t^{-1/2}, e^{\lambda_+ t}\big)$ *for* $t \geq t_\mu$*.*

Theorem (B.-R.–Cox–Schieven, discrete time)

Same setting. For every $h > 0$ *and Markovian* μ_h *(*c *indep. of* μ_h*):*

$$\mu_h(Z_k^h \neq \tilde{Z}_k^h) \;\geq\; c\min\Big(c_{\mu_h}(hk{+}1)^{-1/2},\; c_{\mu_h}e^{\lambda_+ hk},\; h^{-1}e^{\lambda_- hk}\Big).$$

No Markovian coupling matches the TV decay

Theorem (B.-R.–Cox–Schieven, continuous time)

Under $\gamma^2>4\alpha\geq 0$*: (upper) if* $\lambda_-\Delta x=\Delta v$, $\Delta x\neq 0$*, then* $d_{\mathrm{TV}}\leq Ce^{\lambda_- t}|\Delta z|$*; (lower) for every Markovian coupling* μ *and every* Δz, $\mu(Z_t\neq\tilde{Z}_t)\geq c_\mu\min\big(t^{-1/2}, e^{\lambda_+ t}\big)$ *for* $t\geq t_\mu$*.*

Theorem (B.-R.–Cox–Schieven, discrete time)

Same setting. For every $h>0$ *and Markovian* μ_h *(c indep. of* μ_h*):*

$$\mu_h(Z_k^h\neq\tilde{Z}_k^h)\ \geq\ c\min\Big(c_{\mu_h}(hk+1)^{-1/2},\ c_{\mu_h}e^{\lambda_+ hk},\ h^{-1}e^{\lambda_- hk}\Big).$$

- Extends (Banerjee & Kendall (2016)) (Kolmogorov diffusion) to kinetic Langevin.

The canonical Markovian candidate: iterated one-shot

- For linear Markov chains $Z_{k+1} = A_k Z_k + B_k \xi_{k+1}$:
  - Greedy: maximize the meeting probability *at the next step* by reflection coupling.
  - Satisfies the *now-equals-forever* property: $Z_k^h = \tilde{Z}_k^h \Rightarrow Z_j^h = \tilde{Z}_j^h$ for $j \geq k$.

The canonical Markovian candidate: iterated one-shot

- For linear Markov chains $Z_{k+1} = A_k Z_k + B_k \xi_{k+1}$:
  - Greedy: maximize the meeting probability *at the next step* by reflection coupling.
  - Satisfies the *now-equals-forever* property: $Z_k^h = \tilde{Z}_k^h \Rightarrow Z_j^h = \tilde{Z}_j^h$ for $j \geq k$.
- For overdamped Euler–Maruyama: asymptotically optimal[1].

---

[1] (Durmus & Moulines (2019))

The canonical Markovian candidate: iterated one-shot

- For linear Markov chains $Z_{k+1} = A_k Z_k + B_k \xi_{k+1}$:
    - Greedy: maximize the meeting probability *at the next step* by reflection coupling.
    - Satisfies the *now-equals-forever* property: $Z_k^h = \tilde{Z}_k^h \Rightarrow Z_j^h = \tilde{Z}_j^h$ for $j \geq k$.
- For overdamped Euler–Maruyama: asymptotically optimal[1].
- For kinetic Langevin: necessarily suboptimal (by Theorem above). By how much?

---

[1] (Durmus & Moulines (2019))

Exact meeting probability for iterated one-shot

Theorem (B.-R.–Cox–Schieven)

Let $(Z_k, \tilde{Z}_k)$ be the iterated one-shot coupling of the linear chain $Z_{k+1} = A_{k+1} Z_k + B_{k+1} \xi_{k+1}$ $(A_k, B_k$ non-singular). Then

$$\mathbb{P}(Z_n \neq \tilde{Z}_n) = 2\,\Phi\left(\frac{1}{2\,\Theta_n^{1/2}}\right) - 1, \qquad \Theta_n = \sum_{k=1}^{n} \frac{1}{|B_k^{-1}\,\Pi_k\,\Delta z|^2}, \quad \Pi_k = A_k A_{k-1} \cdots A_1 .$$

Exact meeting probability for iterated one-shot

Theorem (B.-R.–Cox–Schieven)

Let $(Z_k, \tilde{Z}_k)$ *be the iterated one-shot coupling of the linear chain* $Z_{k+1} = A_{k+1} Z_k + B_{k+1} \xi_{k+1}$ *(*A_k, B_k *non-singular). Then*

$$\mathbb{P}(Z_n \neq \tilde{Z}_n) = 2\,\Phi\left(\frac{1}{2\,\Theta_n^{1/2}}\right) - 1, \qquad \Theta_n = \sum_{k=1}^{n} \frac{1}{|B_k^{-1}\,\Pi_k\,\Delta z|^2}, \quad \Pi_k = A_k A_{k-1} \cdots A_1.$$

- **Linear-drift companion** to (Durmus–Moulines (2019), Thm. 19); linearity makes it an equality.

Exact meeting probability for iterated one-shot

Theorem (B.-R.–Cox–Schieven)

Let $(Z_k, \tilde{Z}_k)$ *be the iterated one-shot coupling of the linear chain* $Z_{k+1} = A_{k+1} Z_k + B_{k+1} \xi_{k+1}$ $(A_k, B_k$ *non-singular). Then*

$$\mathbb{P}(Z_n \neq \tilde{Z}_n) = 2\,\Phi\left(\frac{1}{2\,\Theta_n^{1/2}}\right) - 1, \qquad \Theta_n = \sum_{k=1}^{n} \frac{1}{|B_k^{-1}\,\Pi_k\,\Delta z|^2}, \quad \Pi_k = A_k A_{k-1} \cdots A_1.$$

- **Linear-drift companion** to (Durmus–Moulines (2019), Thm. 19); linearity makes it an equality.
- Used as a sharp *lower bound* on $\mathbb{P}(Z_n \neq \tilde{Z}_n)$, not a TV upper bound.

Exact meeting probability for iterated one-shot

Theorem (B.-R.–Cox–Schieven)

Let $(Z_k, \tilde{Z}_k)$ *be the iterated one-shot coupling of the linear chain* $Z_{k+1} = A_{k+1} Z_k + B_{k+1} \xi_{k+1}$ *(*A_k, B_k *non-singular). Then*

$$\mathbb{P}(Z_n \neq \tilde{Z}_n) = 2\,\Phi\left(\frac{1}{2\,\Theta_n^{1/2}}\right) - 1, \qquad \Theta_n = \sum_{k=1}^{n} \frac{1}{|B_k^{-1}\,\Pi_k\,\Delta z|^2}, \quad \Pi_k = A_k A_{k-1} \cdots A_1.$$

- **Linear-drift companion** to (Durmus–Moulines (2019), Thm. 19); linearity makes it an equality.
- Used as a sharp *lower bound* on $\mathbb{P}(Z_n \neq \tilde{Z}_n)$, not a TV upper bound.
- Free kinetic Langevin ($\alpha = 0$, $\Delta z = (\Delta x, -\gamma \Delta x)$): $\mathbb{P}(Z_k^h \neq \tilde{Z}_k^h) \geq c\, h^{-1} e^{\lambda_- hk}$ — *saturates* the $h^{-1} e^{\lambda_- hk}$ term in the impossibility bound.

Summary and outlook

- Sharp TV bound for OBABO via an LQ-optimal coalescence map; realised by an explicit non-Markovian rejection sampler. Drops Hessian-Lipschitz and $hn \leq 1$ of (Chak & Monmarché (2025)).
- No Markovian coupling matches the TV decay for $U(x)=\alpha|x|^2$, $\gamma^2 > 4\alpha$; extends (Banerjee & Kendall (2016)).
- Exact meeting formula for iterated one-shot on linear chains; companion to (Durmus & Moulines (2019), Thm. 19).

Summary and outlook

- Sharp TV bound for OBABO via an LQ-optimal coalescence map; realised by an explicit non-Markovian rejection sampler. Drops Hessian-Lipschitz and $hn \leq 1$ of (Chak & Monmarché (2025)).
- No Markovian coupling matches the TV decay for $U(x){=}\alpha|x|^2$, $\gamma^2 > 4\alpha$; extends (Banerjee & Kendall (2016)).
- Exact meeting formula for iterated one-shot on linear chains; companion to (Durmus & Moulines (2019), Thm. 19).

Summary and outlook

- Sharp TV bound for OBABO via an LQ-optimal coalescence map; realised by an explicit non-Markovian rejection sampler. Drops Hessian-Lipschitz and $hn \leq 1$ of (Chak & Monmarché (2025)).
- No Markovian coupling matches the TV decay for $U(x) = \alpha|x|^2$, $\gamma^2 > 4\alpha$; extends (Banerjee & Kendall (2016)).
- Exact meeting formula for iterated one-shot on linear chains; companion to (Durmus & Moulines (2019), Thm. 19).

Open.

- ⋆ **Acceleration** at rate $\kappa^{-1/2}$: *reduced* to proving W_1 contraction at this rate.
- Underdamped regime $\gamma^2 < 4\alpha$ (rigidity uses real eigenvalues).

Related directions.

- KL / Rényi divergences[1]; other splittings[2]; Metropolis adjustment[3].

---

[1] (B.-R., Mitra & Wibisono (2026))
[2] (Schuh & Whalley (2025))
[3] (B.-R. & Oberdörster (2024))

Thank you!

On Couplings for Kinetic Langevin Diffusions

joint work with **Sonja Cox** and **Roy Schieven**

Questions?

Thanks to **Pierre Monmarché**, **Andreas Eberle**, and **Stefan Oberdörster**
for fruitful discussions.