Sticky Coupling as a Control Variate for Computing Transport Coefficients

Shiva Darshan

(CERMICS, Ecole des Ponts & MATHERIALS team, Inria Paris)

Project funded by ANR SINEQ

CECAM Mixed-Gen Session

Outline

- **Linear response for steady-state nonequilibrium dynamics**
  - Equilibrium dynamics and their perturbations
  - Definition of transport coefficients
  - Variance & bias of NEMD estimator
- **Couplings based estimators**
  - Couplings based estimators
  - Synchronous coupling
  - Sticky coupling
  - Variance & bias of sticky coupling based estimator
- **Extensions and perspectives**

Linear response for steady-state nonequilibrium dynamics

Physical context and motivations

Transport coefficients (*e.g.* thermal conductivity): quantitative estimates

$$J = -\kappa \nabla T \qquad \text{(Fourier's law)}$$

Slow convergence due to large noise to signal ratio
Long computational times to estimate κ (up to several weeks/months)

Nonequilibrium stochastic dynamics

Consider the following family of SDEs with values in $\mathbb{R}^d$ and additive noise:

$$dX_t^\eta = \left(b\left(X_t^\eta\right) + \eta F\left(X_t^\eta\right)\right) dt + \sqrt{\frac{2}{\beta}} dW_t,$$

where $b, F : \mathbb{R}^d \to \mathbb{R}^d$ are smooth and $\eta \in \mathbb{R}$. The above dynamics has generator

$$\mathcal{L}_\eta = \mathcal{L} + \eta\tilde{\mathcal{L}}, \quad \widetilde{\mathcal{L}} = F \cdot \nabla, \quad \mathcal{L} = b \cdot \nabla + \frac{1}{\beta}\Delta,$$

and we assume that for each η the above dynamics admits a unique invariant measure ν_η. We further assume that the drift is *contractive at infinity*, i.e.

Assumption

There exists $m, M > 0$ *such that*

$$\langle x - y, b(x) - b(y)\rangle \leqslant -m\left|x - y\right|^2, \qquad \textit{if } |x - y| \geqslant M.$$

Technical Assumptions

We need the following technical assumption to ensure that our arguments based on the solutions to the Poisson equation are justified. Let $\mathscr{S}$ be the space of smooth function who grow at most polynomially along with their derivatives. Denoting by $\partial^k = \partial_{x_1}^{k_1} \cdots \partial_{x_d}^{k_d}$ for $k = (k_1, \ldots, k_d) \in \mathbb{N}^d$ and $\mathcal{K}_n(x) = 1 + |x|^n$,

$$\mathscr{S} := \left\{ \varphi \in C^\infty(\mathbb{R}^d) \,\middle|\, \forall k \in \mathbb{N}^d \, \exists C > 0, \exists n \in \mathbb{N} \, \text{s.t.} \, \left| \partial^k \varphi \right| \leqslant C \mathcal{K}_n \right\}$$

Assumption

Assume that $b, F \in \mathscr{S}$ and for any $\eta_* > 0$ there exists $\lambda_{\eta_*} > 0$ such that

$$\nabla \left(b(x) + \eta F(x) \right) \cdot (h, h) \leqslant \lambda_{\eta_*} |h|^2 \,, \qquad \forall \eta \in [-\eta_*, \eta_*] \,, \forall x, h \in \mathbb{R}^d$$

The two assumption are satisfied if $b(x) = -V_1(x) - V_2(x)$, where V_1 is a strongly convex confining potential and V_2 is a compactly supported potential modeling the local interactions.

Definition of transport coefficients

Perturbative regime: invariant measure $\nu_\eta = f_\eta \nu_0$ with $f_\eta = 1 + \mathrm{O}(\eta)$

$$\forall \varphi, \qquad 0 = \int_{\mathbb{R}^d} \left[\left(\mathcal{L} + \eta \widetilde{\mathcal{L}} \right) \varphi \right] f_\eta \, d\nu_0 = \int_{\mathbb{R}^d} \varphi \left[\left(\mathcal{L} + \eta \widetilde{\mathcal{L}} \right)^* f_\eta \right] d\nu_0$$

* = adjoints on $L^2(\nu_0)$

Fokker–Planck equation

$$\left(\mathcal{L} + \eta \widetilde{\mathcal{L}} \right)^* f_\eta = 0$$

By identifying powers of η (and denoting by $\Pi_0 \varphi := \varphi - \nu_0(\varphi)$)

$$f_\eta = 1 + \eta \mathfrak{f}_1 + \eta^2 \mathfrak{f}_2 + \dots, \qquad \mathfrak{f}_1 = (-\mathcal{L}^*)^{-1} \widetilde{\mathcal{L}}^* \mathbf{1}$$

Response property $R \in L^2_0(\mu) = \Pi_0 L^2(\mu)$, the transport coefficient α_R satisfies:

$$\alpha_R = \lim_{\eta \to 0} \frac{\mathbb{E}_\eta(R)}{\eta} = \int_{\mathbb{R}^d} R \mathfrak{f}_1 \, d\nu_0$$

Error estimates for NEMD

Principle of nonequilibrium molecular dynamics

Estimator of linear response (observable R average 0 with respect to ν_0)

$$\widehat{\Phi}_{\eta,t} = \frac{1}{\eta t}\int_0^t R(X_t^\eta)\,ds \xrightarrow[t\to+\infty]{\text{a.s.}} \alpha_{R,\eta} := \frac{1}{\eta}\int_{\mathbb{R}^d} R\,d\nu_\eta = \alpha_R + \mathrm{O}(\eta)$$

Issues with linear response methods:

- Statistical error with asymptotic variance $\mathrm{O}(\eta^{-2})$
- Bias from finite integration time
- Bias $\mathrm{O}(\eta)$ due to $\eta \neq 0$
- Timestep discretization bias

Principle of nonequilibrium molecular dynamics

Estimator of linear response (observable R average 0 with respect to ν_0)

$$\widehat{\Phi}_{\eta,t} = \frac{1}{\eta t}\int_0^t R(X_t^\eta)\,ds \xrightarrow[t\to+\infty]{\text{a.s.}} \alpha_{R,\eta} := \frac{1}{\eta}\int_{\mathbb{R}^d} R\,d\nu_\eta = \alpha_R + \mathrm{O}(\eta)$$

Issues with linear response methods:

- Statistical error with asymptotic variance $\mathrm{O}(\eta^{-2})$
- Bias from finite integration time

- Bias $\mathrm{O}(\eta)$ due to $\eta \neq 0$
- Timestep discretization bias

Analysis of variance / finite integration time bias

- **Statistical error** dictated by Central Limit Theorem:

$$\sqrt{t}\left(\widehat{\Phi}_{\eta,t} - \alpha_\eta\right) \xrightarrow[t\to+\infty]{\text{law}} \mathcal{N}\left(0, \frac{\sigma^2_{R,\eta}}{\eta^2}\right), \qquad \sigma^2_{R,\eta} = \sigma^2_{R,0} + \mathrm{O}(\eta)$$

so $\widehat{\Phi}_{\eta,t} = \alpha_\eta + \mathrm{O}_{\mathrm{P}}\left(\frac{1}{\eta\sqrt{t}}\right)$ → requires long simulation times $t \sim \eta^{-2}$

- **Finite time integration bias**: $\left|\mathbb{E}\left(\widehat{\Phi}_{\eta,t}\right) - \alpha_\eta\right| \leqslant \frac{K}{\eta t}$

Bias due to $t < +\infty$ is $\mathrm{O}\left(\frac{1}{\eta t}\right)$ → typically smaller than statistical error

- Key equality for the proofs: introduce $-\mathcal{L}_\eta \widehat{R}_\eta = R - \int_{\mathbb{R}^d} R \, d\nu_\eta$

$$\widehat{\Phi}_{\eta,t} - \frac{1}{\eta}\int_{\mathbb{R}^d} R \, d\nu_\eta = \frac{\widehat{R}_\eta(X_0^\eta) - \widehat{R}_\eta(X_t^\eta)}{\eta t} + \frac{\sqrt{2}}{\eta t\sqrt{\beta}} \int_0^t \nabla \widehat{R}_\eta(X_s^\eta) \cdot dW_s$$

Couplings Based Estimators

Couplings Based Estimator

Definition

A coupling of two random variables X and Y is a couple $\left(\widetilde{X}, \widetilde{Y}\right)$ of random variables such that $\widetilde{X} \stackrel{\text{Law}}{=} X$ and $\widetilde{Y} \stackrel{\text{Law}}{=} Y$

Idea: Use the reference dynamics to reduce the variance and bias of the estimator:

$$\widehat{\Psi}_{\eta,t} = \frac{1}{\eta t} \int_0^t \left[R\left(X_s^\eta\right) - R\left(Y_s^0\right)\right] ds, \tag{1}$$

with $(X_t^\eta, Y_t^\eta)_{t\geqslant 0}$ the solution of

$$dX_t^\eta = \left(b\left(X_t^\eta\right) + \eta F\left(X_t^\eta\right)\right) dt + \sqrt{\frac{2}{\beta}} dW_t,$$

$$dY_t^0 = b\left(Y_t^0\right) dt + \sqrt{\frac{2}{\beta}} d\widetilde{W}_t,$$

where the driving noises $\left(W_t, \widetilde{W}_t\right)_{t\geqslant 0}$ are cleverly coupled.

Synchronous Coupling

A simple method coupling is to take $W = \widetilde{W}$, i.e. use the same driving Brownian motion. This is called synchronous coupling.

If the drift is everywhere contractive, i.e

$$\langle x - y, b(x) - b(y) \rangle \leqslant -m \left| x - y \right|^2, \quad \forall x, y \in \mathbb{R}^d,$$

then this method works quite well. One can show that the coupling distance, $|X_t^\eta - Y_t^\eta|$, is bounded by $C\eta$ and terms that decays exponential fast time.

This implies that $\left(\widehat{\Psi}_{\eta,t}^{\mathrm{sync}} - \alpha_{R,\eta} \right)$ has bounded moments.

One can constructed a coupling[1] such that $\left(X_t^\eta - Y_t^0\right)_{t\geqslant 0}$ is *sticky at 0* in the sense that they the difference is controlled by a one-dimensional process $(r_t^\eta)_{t\geqslant 0}$ that spends a positive amount of time at 0.

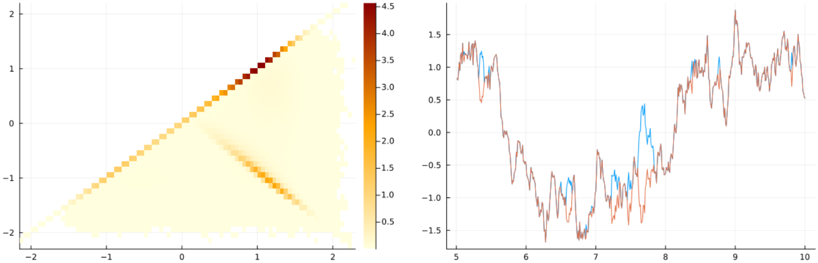


Figure: Sticky coupling of a 1D particle in a double well potential perturbed by a constant force to the right. **Left:** histogram of coupled process; **Right:** segment of trajectory of coupled process

[1] A. Eberle, R. Zimmer (2019) *Sticky couplings of multidimensional diffusions with different drifts*

Performance of the Sticky Coupling Based Estimator

The coupling based estimator improves upon the bias and variance of the NEMD estimator by a factor of η^{-1}:

Theorem

Let $\eta_* > 0$ *and* $R \in \mathscr{S}$ *such that* $\nu_0(R) = 0$*. Assume that* X^η *and* Y^0 *have the same initial value. If the two previously stated assumptions hold, then there exists* $K_1, K_2 > 0$ *such that*

$$\forall \eta \in [-\eta_*, \eta_*], \quad \lim_{t \to \infty} t \mathrm{Var}\left(\widehat{\Psi}_{\eta,t}^{\mathrm{sticky}}\right) \leqslant \frac{K_1}{\eta}, \tag{2}$$

and

$$\left| \mathbb{E}\left[\widehat{\Psi}_{\eta,t}^{\mathrm{sticky}}\right] - \alpha_{R,\eta} \right| \leqslant \frac{K_2}{t}. \tag{3}$$

Proof: Work in Progress

Ideas of Proof (1)

Pass to discrete version of sticky coupling—maximal–reflection coupling[2]. Let $(G_k)_{k \geqslant 1}$ and $(U_k)_{k \geqslant 1}$ be i.i.d sequences of Gaussian and uniform random variables respectively. The coupled SDE discretizes as

$$X_{k+1}^{\eta,\Delta t} = X_k^{\eta,\Delta t} + \Delta t \left[b\left(X_k^{\eta,\Delta t}\right) + \eta F\left(X_k^{\eta,\Delta t}\right)\right] + \sqrt{\frac{2\Delta t}{\beta}} G_{k+1},$$

$$Y_{k+1}^{0,\Delta t} = X_k^{\eta,\Delta t} B_{k+1} + (1 - B_{k+1}) H_{\Delta t}\left(X_k^{\eta,\Delta t}, Y_k^{0,\Delta t}, G_{k+1}\right),$$

[2] A. Durmus, A. Eberle, A. Enfroy, A. Guillin, P. Monmarché (2021) *Discrete sticky couplings of functional autoregressive processes*

Ideas of Proof (2)

with $B_{k+1} = \mathbf{1}_{[0,1]}\left(p_{\Delta t,\beta}\left(X_k^{\eta,\Delta t}, Y_k^{0,\Delta t}, G_{k+1}\right) - U_{k+1}\right)$ and

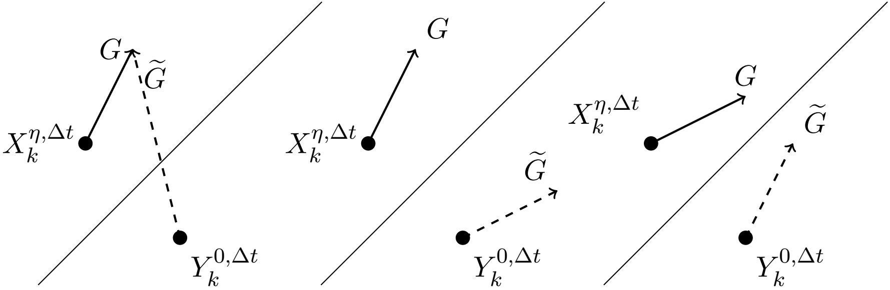
$$H_{\Delta t}(x,y,z) = y + \Delta t b(y) + \sqrt{\frac{2\Delta t}{\beta}}\left[\mathrm{Id} - 2\mathbf{e}(x,y)\,\mathbf{e}(x,y)^T\right] z,$$

$$\mathbf{E}(x,y) = y - x + \Delta t\left[b(y) - b(x) - \eta F(x)\right],$$

$$\mathbf{e}(x,y) = \begin{cases} \dfrac{\mathbf{E}(x,y)}{|\mathbf{E}(x,y)|} & \text{if } \mathbf{E}(x,y) \neq 0 \\ \mathbf{e}_0 & \text{otherwise,} \end{cases}$$

$$p_{\Delta t,\beta}(x,y,z) = \min\left\{1, \frac{\varphi\left(\sqrt{\frac{\beta}{2\Delta t}}\,|\mathbf{E}(x,y)| - \langle \mathbf{e}(x,y), z\rangle\right)}{\varphi\left(\langle \mathbf{e}(x,y), z\rangle\right)}\right\},$$

Ideas of Proof (3)



(a) Collision (b) Separation (c) Contraction

Ideas of Proof (4)

Denote by $\mu_{\eta,\Delta t}$, $\nu_{\eta,\Delta t}$, and $\nu_{0,\Delta t}$ the invariant measures of the discrete coupled and discrete marginal processes respectively. We use MR coupling to prove a weighted TV norm type estimate of the form

$$\int_{\mathbb{R}^d\times\mathbb{R}^d}(\mathcal{K}_n(x)+\mathcal{K}_n(y))\,\mathbf{1}_{\{x\neq y\}}\,d\mu_{\eta,\Delta t}\,(dx\,dy)\leqslant C_n\eta\left(\nu_{\eta,\Delta t}\left(\mathcal{K}_n\right)+\nu_{0,\Delta t}\left(\mathcal{K}_n\right)\right).$$

Question:

Can we pass to the limit $\Delta t\downarrow 0$ to recover a similar estimate for the continuous time processes?

If so, then we can use this estimate to control the differences $\widehat{R}_\eta\left(X_t^\eta\right)-\widehat{R}_0\left(Y_t^0\right)$ and $R_\eta\left(X_t^\eta\right)-R_0\left(Y_t^0\right)$ and thereby the bias and variance and prove the theorem.

Numerical Illustrations (1)

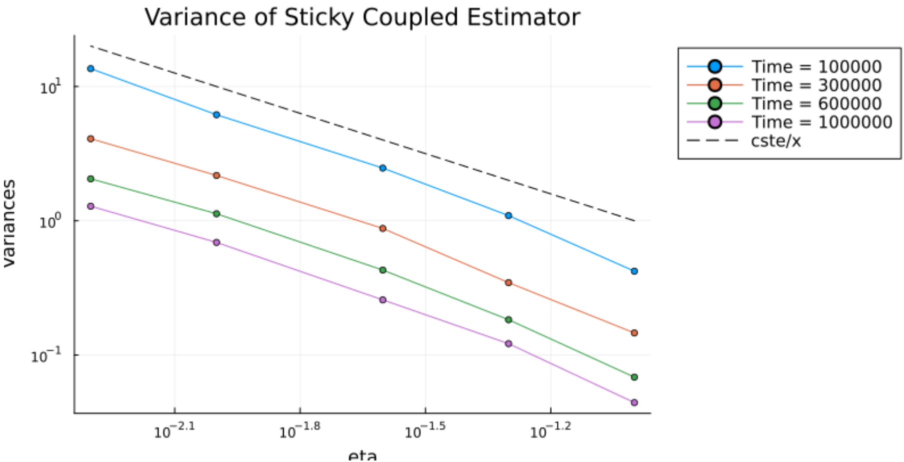


Figure: Variance of Sticky Coupled Estimator applied to a 2D particle in locally nonconvex potential experiencing shear forcing

Numerical Illustrations (2)

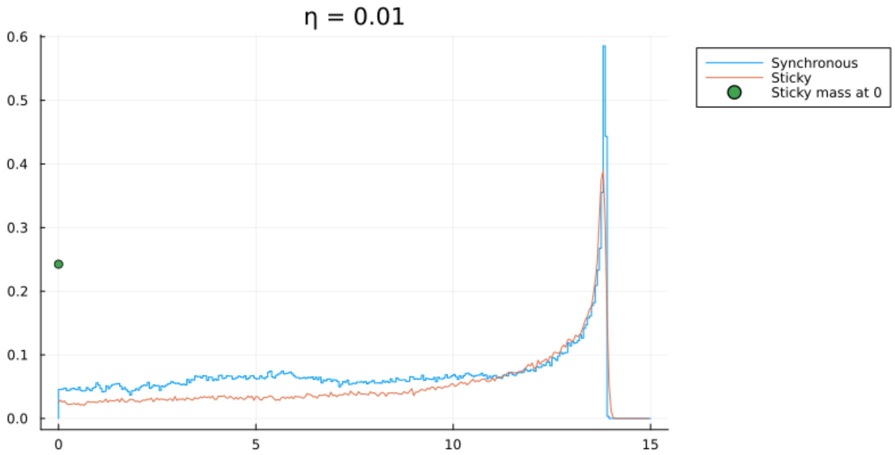


Figure: Histogram of distance between perturbed trajectory and reference trajectory of 2D overdamped Lennard-Jones cluster of 18 particles at $\beta = 16$ confined to $[-5, 5]^2$ by a quadratic potential and perturbed by shear forcing.

Numerical Illustrations (3)

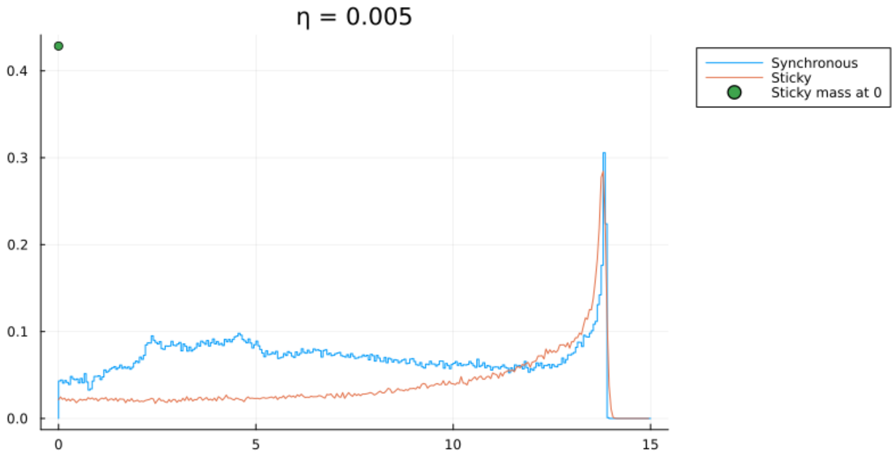


Figure: Histogram of distance between perturbed trajectory and reference trajectory of 2D overdamped Lennard-Jones cluster of 18 particles at $\beta = 16$ confined to $[-5, 5]^2$ by a quadratic potential and perturbed by shear forcing.

Extensions and perspectives

Extension to Kinetic Langevin Dynamics

For kinetic Langevin Dynamics the noise only effects the momentum.

$$dq_t^\eta = M^{-1} p_t^\eta dt,$$
$$dp_t^\eta = \left(-\nabla U\left(q_t^\eta\right) + \eta F\left(q_t^\eta\right)\right) dt - \gamma M^{-1} p_t^\eta dt + \sqrt{\frac{2\gamma}{\beta}} dW_t.$$

The coordinate change[3][4] hints at what the coupling should do: let $\left(Z_t^\eta, Q_t^\eta\right) = \left(q_t^\eta - q_t^0, q_t^\eta - q_t^0 + \gamma^{-1}\left(p_t^\eta - p_t^0\right)\right)$, then

$$dZ_t^\eta = -\gamma M^{-1} Z_t^\eta dt + \gamma M^{-1} Q_t^\eta dt,$$
$$dQ_t^\eta = -\gamma^{-1}\left(\nabla U\left(q_t^\eta\right) - \nabla U\left(q_t^0\right)\right) dt + \gamma^{-1} \eta F\left(q_t^\eta\right) dt + \sqrt{\frac{2}{\gamma\beta}} d(W - \widetilde{W})_t.$$

Z^η is contractive whenever $\|Z_t^\eta\|_\infty \geqslant \|Q_t^\eta\|_\infty$.

[3] A. Eberle, A. Guillin, R. Zimmer (2019) *Couplings and quantitative contraction rates for Langevin dynamics*

[4] N. Bou-Rabee, A. Eberle, R. Zimmer (2020) *Coupling and Convergence for Hamiltonian Monte Carlo*

Sticky Coupling on a Manifold

Morally one should be able to extend sticky coupling to processes that take values on a manifold. The reflection part of sticky coupling can be extended to manifold-valued processes using the Kendall-Cranston[5]

If the manifold is compact can we make the coupling work for an arbitrary potential? Example: Lennard-Jones fluids with periodic boundary conditions.

[5] A. Eberle (2016) Reflection couplings and contraction rates for diffusions

Hybrid Coupling

In contractive regions of the phase space, synchronous coupling (i.e choosing $\widetilde{W} = W$) is more effective at bring the coupled trajectories together than reflection coupling.

On the other hand, reflective coupling can separate trajectories just as easily as it can bring them together—MR coupling has a long "tail".

This suggests a hybrid approach of mixing MR coupling and synchronous coupling.