Coupling Based Control Variate Strategies for Computing Transport Coefficients

Shiva Darshan

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

In collaboration with A. Eberle and G. Stoltz

based on arXiv:2409.15500

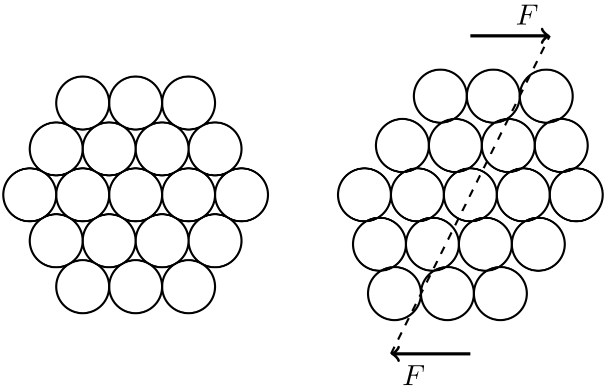
Project funded by ANR SINEQ

Outline

- **Linear response for stationary perturbed dynamics**
  - Transport Coefficients
  - Standard NEMD
  - Variance of NEMD estimator
- **Couplings based estimators**
  - Couplings based estimators
  - Synchronous coupling
  - Sticky coupling
- **Numerical Illustrations**
- **Extensions and perspectives**

Linear response for stationary perturbed dynamics

Transport Coefficients, Kézako[1]?



What is the first-order response of a system to an external forcing?

[1] French slang for "what's that"

Transport Coefficients, Kézako?

Our setup:

$$dX_t^\eta = (b(X_t^\eta) + \eta F(X_t^\eta))\, dt + \sqrt{\frac{2}{\beta}} dW_t, \tag{1}$$

with invariant probability measure ν_η and where $b, F : \mathbb{R}^d \to \mathbb{R}^d$, $\beta > 0$, and $\eta \in \mathbb{R}$.

Transport Coefficients quantify the first order response of the invariant probability measure with respect to the perturbation

$$\alpha_R := \frac{\mathrm{d}}{\mathrm{d}\eta} \nu_\eta(R) \Big|_{\eta=0} = \lim_{\eta \to 0} \frac{\nu_\eta(R) - \nu_0(R)}{\eta} \tag{2}$$

Some Assumptions

Technical Assumptions

- $F, b, R \in \mathscr{S}$, i.e. smooth, grow at most polynomially, and have derivatives that grow at most polynomially
- F, b Lipschitz
- F bounded

Important Assumption

Strong Contractivity at Infinity
There exists $M \geqslant 0$ and $m > 0$ such that

$$\langle x - y, b(x) - b(y) \rangle \leqslant -m\,|x - y|^2\,, \quad \text{if } |x - y| \geqslant M.$$

Example if $b(x) = -\nabla\left(V_1(x) + V_2(x)\right)$, where V_1 is a confining potential and V_2 is a compactly supported.

Under these assumptions, a unique invariant probability measure ν_η with smooth density w.r.t to Lebesgue exists of any $\eta \in \mathbb{R}$.

Estimating transport coefficients

Assume from here on that the observable R is such that $\nu_0(R) = 0$
The transport coefficient α_R is well-defined and

$$\alpha_R = \lim_{\eta \to 0} \frac{\nu_\eta(R)}{\eta} = \int_{\mathbb{R}^d} R\,\mathfrak{f}\,d\nu_0, \qquad \mathfrak{f} = -(\mathcal{L}_0^*)^{-1}\widetilde{\mathcal{L}}^*\mathbf{1},$$

Estimator of linear response:

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

Sources of error:

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

Couplings Based Estimators

Couplings Based Estimator

Definition

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

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

$$\widehat{\Psi}_{\eta,t} = \frac{1}{\eta t} \int_0^t \left[R\left(X_s^\eta\right) - R\left(Y_s^0\right) \right] ds, \tag{3}$$

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

By choosing $W = \widetilde{W}$, we *synchronously* couple the X^η and Y^0, giving

$$d\left(X_t^\eta - Y_t^0\right) = \left(b\left(X_t^\eta\right) - b\left(Y_t^0\right) + \eta F\left(X_t^\eta\right)\right) dt.$$

If the drift is strongly contractive everywhere, i.e.

$$\langle x - y, b(x) - b(y) \rangle \leqslant -m \left|x - y\right|^2, \quad \forall x, y \in \mathbb{R}^d, \tag{4}$$

then we have pointwise control over the distance between the coupled trajectories:

$$\left|X_t^\eta - Y_t^0\right| \leqslant \left(\left|X_0^\eta - Y_0^0\right| - \frac{\eta \left\|F\right\|_\infty}{2m}\right) \mathrm{e}^{-mt} + \frac{\eta \left\|F\right\|}{2m}.$$

As a consequence,

$$\mathbb{E}\left[\left|\widehat{\Psi}_{\eta,t}^{\mathrm{sync}}\right|^p\right] \leqslant C \left(\frac{\left|X_0^\eta - Y_0^0\right|^p}{\eta^p} \mathrm{e}^{-pmt} + \left(1 - e^{-p\lambda t}\right)^p \left(\frac{\eta \left\|F\right\|}{2m}\right)^p\right),$$

and a fortiori bounded variance and bias as $\eta \downarrow 0$ if $\left|X_0^\eta - Y_0^0\right|^p = O\left(\eta^p\right)$.

Synchronous Coupling

In fact, as long as we have sufficient contractivity, say due to sufficiently high temperature[2] or in the underdamped case[3], we can control the moments of the estimator as

$$\mathbb{E}\left[\left|\widehat{\Psi}^{\mathrm{sync}}_{\eta,t}\right|^p\right] \leqslant C\left(\frac{\left|X_0^\eta - Y_0^0\right|^p}{\eta^p}\mathrm{e}^{-pmt} + \left(1-e^{-p\lambda t}\right)^p\left(\frac{\eta\,\|F\|}{2m}\right)^p\right),$$

Moral: When there is enough strong contractivity, synchronous coupling is hard to beat.

What to do when we do not have enough strong contractivity?

[2] P. Monmarché (2022) *Wasserstein contraction and Poincaré inequalities for elliptic diffusions at high temperature*

[3] P. Monmarché (2023) *Almost sure contraction for diffusions on $\mathbb{R}^d$. Applications to generalized Langevin diffusions.*

Sticky Coupling

One can construct a coupling[4] such that $\left(X_t^\eta - Y_t^0\right)_{t\geqslant 0}$ is *sticky at 0* in the sense that its norm is controlled by a one-dimensional process $(r_t^\eta)_{t\geqslant 0}$ that spends a positive amount of time at 0

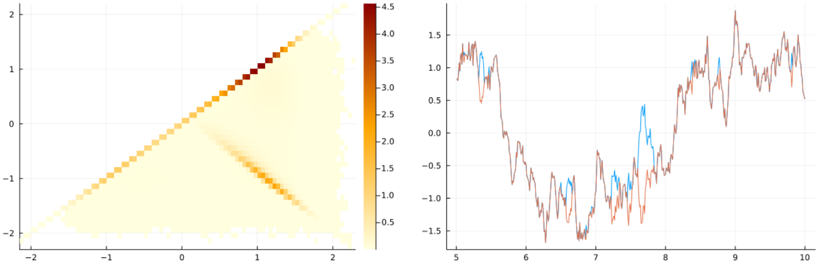


Figure: Sticky coupling of a 1D particle in a double well potential perturbed by a constant force to the right. **Left:** histogram of coupled process; **Right:** segment of trajectory of coupled process

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

Difficulties with Continuous-Time Sticky Coupling

- Non-explicit construction—constructed as the limit point of a tight family of processes
- Long-time properties of sticky coupled process are unclear. Unknown if it is ergodic, admits a unique invariant measure, etc.
- Convergence of discrete approximations also unclear

These difficulties arise because the limit object is highly degenerate. If it satisfied an SDE, the equation would have discontinuous coefficients and likely could not admit a strong solution.

The problem is that we have a "sticky" diffusion in $\mathbb{R}^d$

Discrete-Time Sticky Coupling

Work instead with discrete version of sticky coupling [5]. Consider the estimator

$$\widehat{\Psi}_{\eta,N}^{\Delta t} = \frac{1}{\eta N} \sum_{k=0}^{N-1} \left[R\left(X_n^{\eta,\Delta t}\right) - R\left(Y_n^{0,\Delta t}\right) \right]$$

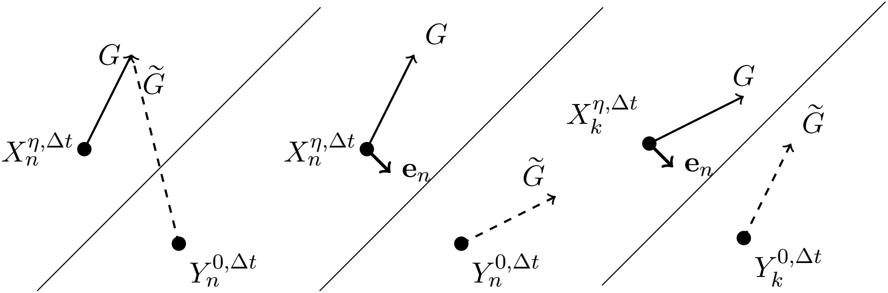
with $\left\{X_n^{\eta,\Delta t}, Y_n^{0,\Delta t}\right\}_{k\in\mathbb{N}}$ the discrete sticky coupling of the Euler-Maruyama discretizations of $\left(X_t^{\eta}\right)_{t\geqslant 0}$ and $\left(Y^0\right)_{t\geqslant 0}$.
Let $\{G_k\}_{k\geqslant 1}$ and $\{U_k\}_{k\geqslant 1}$ be i.i.d sequences of Gaussian and uniform random variables respectively.

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

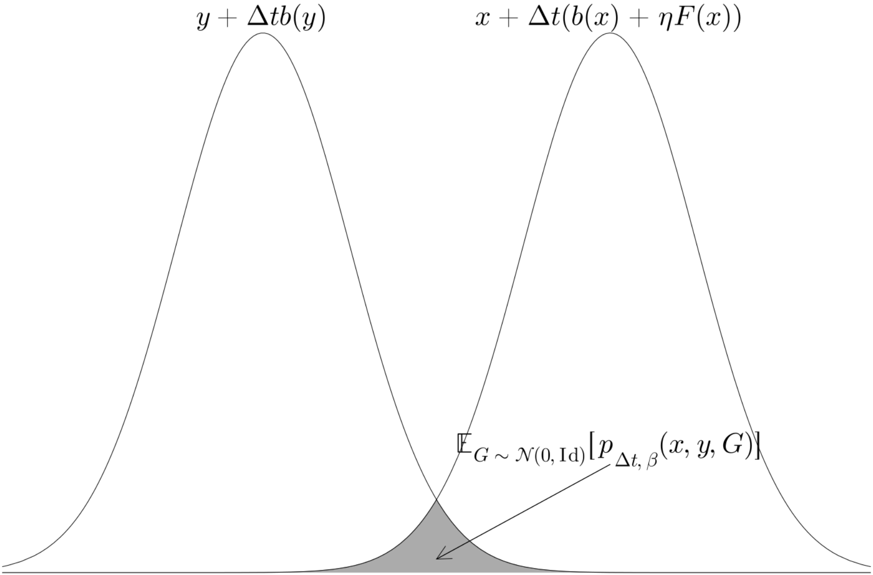
[5] A. Durmus, A. Eberle, A. Enfroy, A. Guillin, P. Monmarché (2024) *Discrete sticky couplings of functional autoregressive processes*

Discrete-Time Sticky Coupling

$$Y_{n+1}^{0,\Delta t} = \begin{cases} X_{n+1}^{\eta,\Delta t} & \text{if } p_{\Delta t,\beta}\left(X_n^{\eta,\Delta t}, Y_n^{0,\Delta t}, G_{n+1}\right) \geqslant U_{n+1} \\ Y_n^{0,\Delta t} + b\left(Y_n^{0,\Delta t}\right)\Delta t + \sqrt{\frac{2\Delta t}{\beta}}\left[\mathrm{Id} - 2\mathbf{e}_n\mathbf{e}_n^T\right] G_{n+1} & \text{otherwise} \end{cases}$$



(a) Collision (b) Reflection resulting in separation (c) Reflection resulting in contraction


$y + \Delta t b(y)$
$x + \Delta t(b(x) + \eta F(x))$
$\mathbb{E}_{G \sim \mathcal{N}(0, \mathrm{Id})}[p_{\Delta t, \beta}(x, y, G)]$

Discrete-Time Sticky Coupling

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

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

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

Discrete-Time Sticky Coupling

Proposition

For Δt sufficiently small, the discrete-time sticky coupled process $\left\{X_k^\eta, Y_k^0\right\}_{k\in\mathbb{N}}$ admits a unique invariant measure, $\mu_{\eta,\Delta t}$. Furthermore it is geometrically ergodic wrt to this measure.

Proof: Use Hairer & Mattingly strategy[6]

Denote by $T^{\eta,\Delta t}$ the Markov kernel of the coupled process

Contractivity implies $e^{c|x|^2} + e^{c|y|^2}$ is a Lyapunov function.

$p_{\Delta t,\beta}(x,y,z) > 0 \implies$ always strictly positive probability of returning to the diagonal

$\forall K > 0$ there exists $\rho_{K,\Delta t} \in (0,1)$ such that

$$\inf_{\max\{|x|,|y|\} \leqslant K} T^{\eta,\Delta t}\left((x,y),\cdot\right) \geqslant \rho_{K,\Delta t}\xi_K(\cdot)$$

with ξ_K the uniform probability on $\{x=y\} \cap \{\max\{|x|,|y|\} \leqslant K\}$

[6] M. Hairer and J. Mattingly *Yet another look at Harris's ergodic theorem for Markov chains*

Performance of the Sticky Coupling Based Estimator

The coupling based estimator improves by η^{-1} factor of :

Theorem

Let $\eta_\star > 0$ and $R \in \mathscr{S}$ such that $\nu_0(R) = 0$. Assume that X^η and Y^0 have the same initial value. Assume above assumptions hold and Δt small enough, then the estimator $\left\{\widehat{\Psi}_{\eta,N}^{\Delta t}\right\}_{N\in\mathbb{N}}$ converges almost surely and satisfies a CLT with asymptotic variance $\sigma^2_{\mathrm{sticky},R,\eta,\Delta t}$. There exists $K > 0$ such that

$$\forall \eta \in [-\eta_\star, \eta_\star], \quad \left|\mathbb{E}\left[\widehat{\Psi}_{\eta,N}^{\Delta t}\right] - \alpha_{R,\eta}\right| \leqslant K\left(\frac{1}{N} + \Delta t\right), \tag{5}$$

and for any $n \in \mathbb{N}$ there exists $K_n > 0$ such that

$$\forall \eta \in [-\eta_\star, \eta_\star], \quad \sigma^2_{\mathrm{sticky},R,\eta,\Delta t} \leqslant K_n\left(\frac{1}{\eta} + \frac{\Delta t^{4n}}{\eta^2}\right). \tag{6}$$

Key Idea of Proof

Proposition

Under the same hypothesis as the theorem, there exists $c > 0$ such that

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \left(\mathrm{e}^{c|x|^2} + \mathrm{e}^{c|y|^2} \right) \mathbf{1}_{\{x \neq y\}} \, d\mu_{\eta,\Delta t}\left(dx\,dy\right) \leqslant C\eta \left(\nu_{\eta,\Delta t}\left(\mathrm{e}^{c|x|^2}\right) + \nu_{0,\Delta t}\left(\mathrm{e}^{c|y|^2}\right)\right)$$

Heuristic "proof" of proposition

$$\begin{aligned} &\int_{\mathbb{R}^d \times \mathbb{R}^d} \left(\mathrm{e}^{c|x|^2} + \mathrm{e}^{c|y|^2} \right) \mathbf{1}_{\{x \neq y\}} \, d\mu_{\eta,\Delta t}\left(dx\,dy\right) \\ &\leqslant \mu_{\eta,\Delta t}\left(\{x \neq y\}\right) \int_{\mathbb{R}^d \times \mathbb{R}^d} \left(\mathrm{e}^{c|x|^2} + \mathrm{e}^{c|y|^2} \right) d\mu_{\eta,\Delta t}\left(dx\,dy\right), \end{aligned} \tag{7}$$

The sticky coupled process spends an $\mathrm{O}(\eta)$ proportion of time off the diagonal. Furthermore $\mu_{\eta,\Delta t}$ is clearly a coupling of $\nu_{\eta,\Delta t}$ and $\nu_{0,\Delta t}$.

Numerical Illustrations

Numerical Illustrations: Strongly Convex Potential

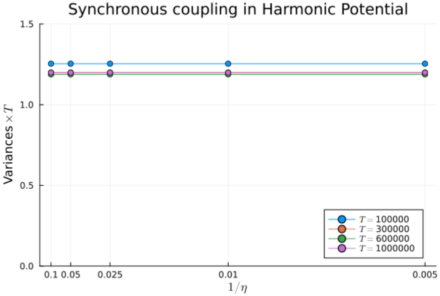
Consider a 2-dimensional Ornstein-Uhlenbeck process

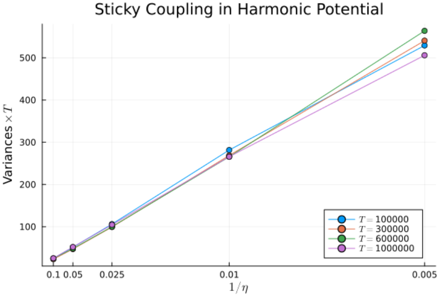
$$dX_t^\eta = -\begin{bmatrix} 1 & -\eta \\ 0 & 1 \end{bmatrix} X_t^\eta dt + \sqrt{\frac{2}{\beta}} dW_t;$$

here $b(x) = -\nabla U = -x$ and $F(x) = [x_2 \; 0]^T$. We choose as response function the covariance between the components. In this case, α_R can be computed analytically.

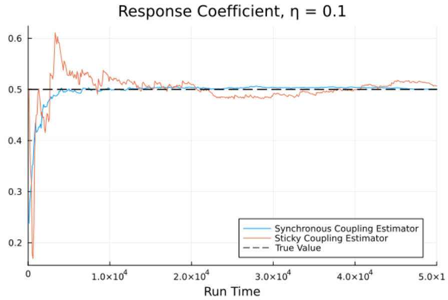
$$R(x) = x_1 x_2, \qquad \alpha_R = \frac{1}{2\beta}$$

Numerical Illustrations: Strongly Convex Potential





Numerical Illustrations: Strongly Convex Potential



Numerical Illustrations: Lennard-Jones Fluid

For less trivial example, we consider an 18 particles 2-D Lennard-Jones fluid. For $x = \left(x_1^1, x_2^1, x_1^2, x_2^2, \dots, x_1^{18}, x_2^{18}\right)$, the interaction is given by

$$U_1(x) = \sum_{i \geqslant j} \left[\left(\frac{1}{|r_{ij}|}\right)^{12} - 2\left(\frac{1}{|r_{ij}|}\right)^{6} \right],$$

with $r_{ij} = |x^i - x^j|$ if $i < j$ and $r_{ii} = |x^i|$. The confinement is give by

$$U_2(x) = \sum_{i=1}^{18} \left[\max\left\{|x_1^i| - 5, 0\right\}^2 + \max\left\{|x_2^i| - 5, 0\right\}^2 \right].$$

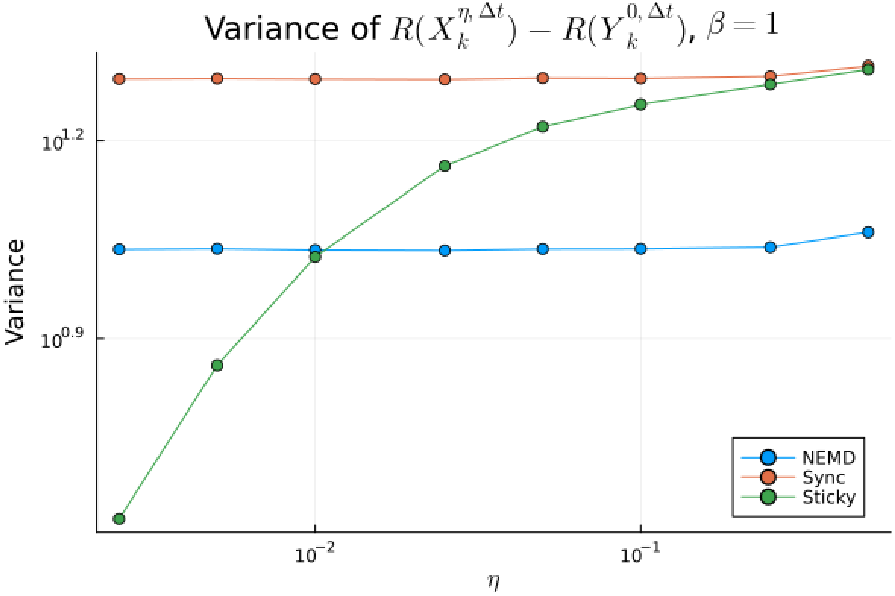
Thus $b(x) = -\nabla U = -\nabla(U_1 + U_2)$. For F we use sine shear

$$(F(x))_i = \begin{cases} \sin(\pi x_2^k/5) & \text{if } i = 2k-1 \\ 0 & \text{otherwise} \end{cases}$$

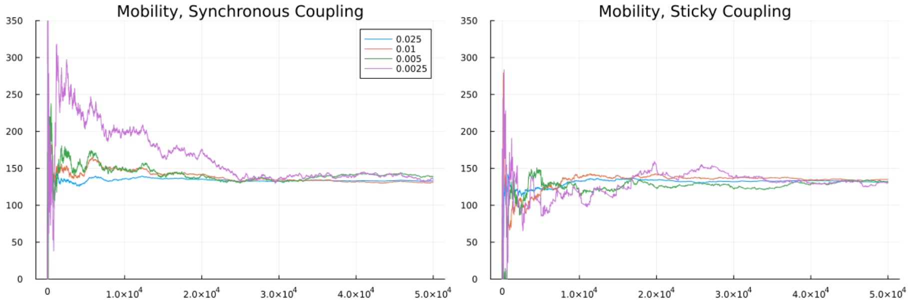
and we measure the mobility response

$$R(x) = F(x)^T \nabla U(x)$$

Numerical Illustrations: Lennard-Jones Fluid Sine Shear



Numerical Illustrations: Lennard-Jones Fluid Sine Shear

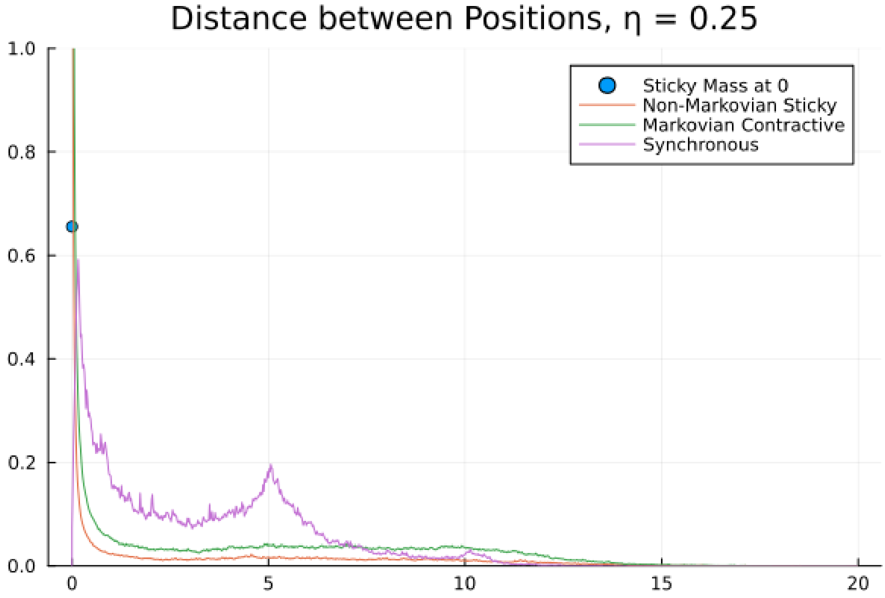


Extension to Hypoelliptic Dynamics

$$\begin{aligned} dq_t^\eta &= p_t^\eta dt \\ dp_t^\eta &= \left(b\left(q_t^\eta\right) + \eta F\left(q_t^\eta\right)\right) dt - \gamma p_t^\eta dt + \sqrt{\frac{2\gamma}{\beta}} dW_t \end{aligned} \quad (8)$$

One can construct a *non-Markovian* coupling of the Euler-Maruyamma discretization of perturbed $\eta \neq 0$ and reference dynamics where the noise at $n+1$ step is used to force a collision of the positions at the $n+2$ step and the noise at $n+2$ step is used to force a collision of the momenta at the $n+2$ step.

Extension to Hypoelliptic Dynamics



Some Extensions and Perspectives

- Componentwise and particle system coupling: Prefactors likely behave badly as $d \to \infty$. Idea: For particle clusters, couple each particle to either its same number particle or nearest particle in the other cluster[7]
- Hybrid coupling: Reflective part gives sticky coupling a long tail, while synchronous is unbeatable when there's contractivity. This suggests a hybrid approach of mixing sticky and synchronous couplings.
- Extension to Riemann manifolds: adapt reflection coupling part to geometry of the manifold via Kendall-Cranston coupling [8]
- Extension to kinetic Langevin dynamics [9] [10] [11]

[7] see works by A. Eberle, K. Schuh, R. Zimmer

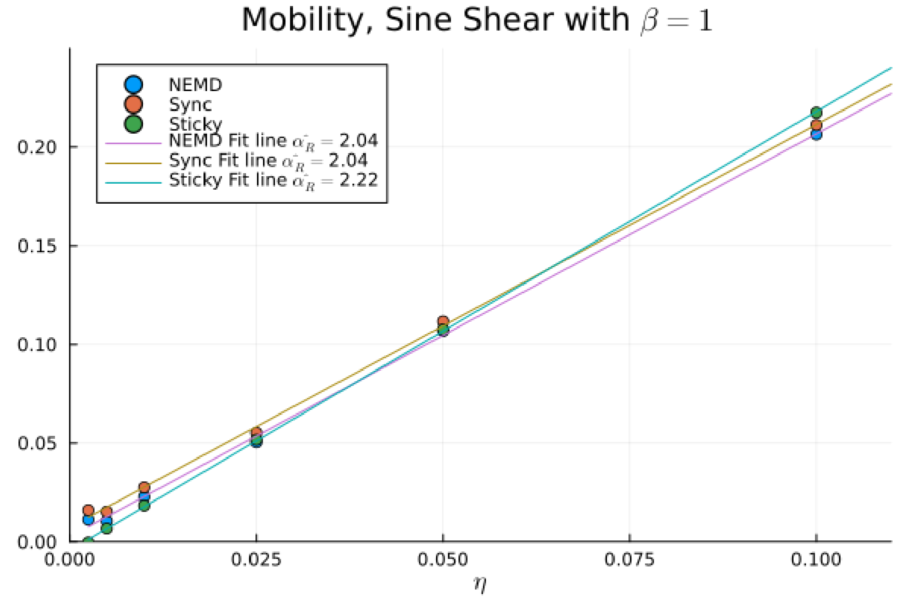
[8] A. Eberle (2016) *Reflection couplings and contraction rates for diffusions*

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

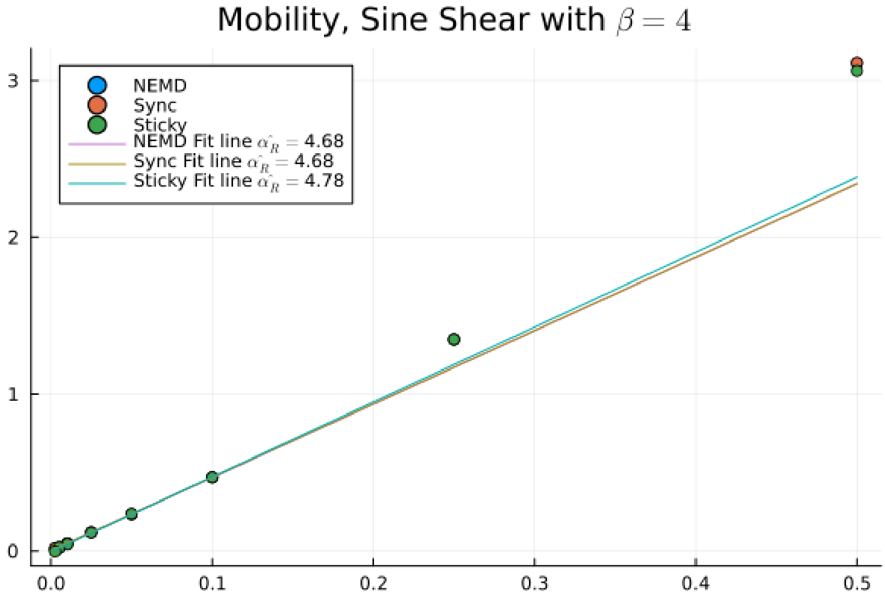
[10] N. Bou-Rabee, A. Eberle, R. Zimmer (2020) *Coupling and Convergence for Hamiltonian Monte Carlo*

[11] M. Chak and P. Monmarché (2024) *Reflection coupling for unadjusted generalized Hamiltonian Monte Carlo in the nonconvex stochastic gradient case*

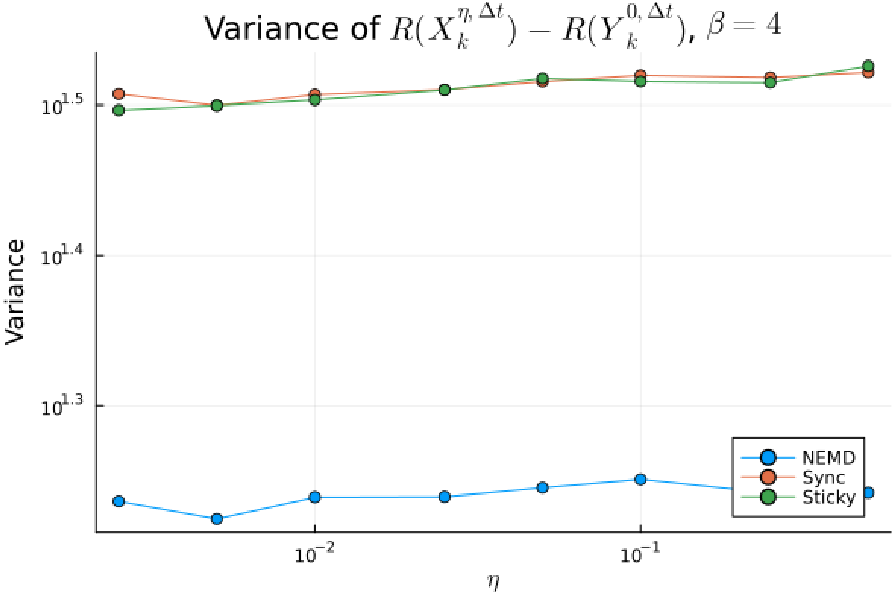
Numerical Illustrations: Lennard-Jones Fluid Sine Shear



Numerical Illustrations: Lennard-Jones Fluid Sine Shear



Numerical Illustrations: Lennard-Jones Fluid Sine Shear



Analysis of Variance/Finite-Time Bias of Standard Estimator

- **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 $-\mathcal{L}_\eta \widetilde{R}_\eta = R - \int_{\mathbb{R}^d} R\, d\nu_\eta$

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

More Ideas of Proof of Theorem

Denote by $\nu_{\eta,\Delta t}$, and $\nu_{0,\Delta t}$ the invariant measures of the respective discrete marginal processes and let $\Pi_{\eta,\Delta t}$ and $\Pi_{0,\Delta t}$ be the operators that center function with respect to these measures. Denote by $P^{\eta,\Delta t}$ and $P^{0,\Delta}$ their Markov kernels.
The CLT from follows ergodicity, constructing an explicit solution to the discrete Poisson equation

$$\Delta t^{-1}\left(\mathrm{Id} - T^{\eta,\Delta t}\right) u(x,y) = \Pi_{\eta,\Delta t} R(x) - \Pi_{0,\Delta t} R(y),$$

and a CLT for Markov chains[12]. This further gives an expression for the asymptotic variance, $\sigma^2_{R,\eta,\Delta t}$ in terms of the

$$\widehat{R}_{\eta,\Delta t} = \Delta t \left(\mathrm{Id} - P^{\eta,\Delta t}\right)^{-1} \Pi_{\eta,\Delta t} R,$$

and

$$\widehat{R}_{0,\Delta t} = \Delta t \left(\mathrm{Id} - P^{0,\Delta t}\right)^{-1} \Pi_{0,\Delta t} R.$$

[12] R. Douc et. al (2018) *Markov Chains*

Variance of Coupling Based Estimator

μ_η invariant measure of the coupled process. $\widetilde{R}_\eta$ and $\widetilde{R}_0$ are solutions of the Poisson equation $\mathcal{L}_\eta \widetilde{R}_\eta = \Pi_\eta R$ and $\mathcal{L}_0 \widetilde{R}_0 = \Pi_0 R$. The asymptotic variance is given by

$$
\begin{aligned}
\widetilde{\sigma}^2_{R,\eta} &= \frac{2}{\eta^2} \int_{\mathbb{R}^d \times \mathbb{R}^d} \left(\widetilde{R}_\eta(x) - \widetilde{R}_0(y)\right) \left(\Pi_\eta R(x) - \Pi_0 R(y)\right) \mu_\eta\,(dx\,dy) \\
&\leqslant \frac{2}{\eta^2} \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} \left(\widetilde{R}_\eta(x) - \widetilde{R}_0(y)\right)^2 \mu_\eta\,(dx\,dy) \right)^{1/2} \\
&\quad \times \left(\int_{\mathbb{R}^d \times \mathbb{R}^d} \left(\Pi_\eta R(x) - \Pi_0 R(y)\right)^2 \mu_\eta\,(dx\,dy) \right)^{1/2}
\end{aligned}
$$

$$
\begin{aligned}
&\int_{\mathbb{R}^d \times \mathbb{R}^d} \left(\widetilde{R}_\eta(x) - \widetilde{R}_0(y)\right)^2 \mu_\eta\,(dx\,dy) \\
&\leqslant 2 \int_{\mathbb{R}^d \times \mathbb{R}^d} \left[\left(\widetilde{R}_\eta(x) - \widetilde{R}_0(x)\right)^2 + \left(\widetilde{R}_0(x) - \widetilde{R}_0(y)\right)^2 \right] \mu_\eta\,(dx\,dy)
\end{aligned}
$$

Variance of Coupling Based Estimator in Discrete-time

$\mu_{\eta,\Delta t}$ invariant measure of the discrete-time coupled process. The asymptotic variance can be bounded as

$$\sigma^2_{\mathrm{sticky},R,\eta,\Delta t} \leqslant \frac{8}{\eta^2}\int_{\mathbb{R}^d\times\mathbb{R}^d}\left(\widehat{R}_{\eta,\Delta t}(x)-\widehat{R}_{0,\Delta t}(x)\right)^2\mu_{\eta,\Delta t}(dx\,dy) + \frac{8}{\eta^2}\int_{\mathbb{R}^d\times\mathbb{R}^d}\left(\widehat{R}_{0,\Delta t}(x)-\widehat{R}_{0,\Delta t}(y)\right)^2\mu_{\eta,\Delta t}(dx\,dy)\,.$$

Second integral can be controlled using our proposition on how much mass $\mu_{\eta,\Delta t}$ puts off the diagonal. Adapting the strategy of Leimkuhler et al (2015)[13], we have

$$\left\|\widehat{R}_{\eta,\Delta t}-\widehat{R}_{0,\Delta t}\right\|_{\widetilde{V}_c} = \mathrm{O}(\eta)+\mathrm{O}\left(\Delta t^{2n}\right)$$

[13] B. Leimkuhler, C. Matthews, and G. Stoltz *The computation of averages from equilibrium and non-equilibrium Langevin molecular dynamics*

More Ideas of Proof of Theorem

A long computation adapting the strategies of Leimkuhler, et. al (2015)[14] and Plechac, et. al (2021)[15] lets us bound the bias and variance with terms of the form

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

and higher order terms. (Recall $\mathcal{K}_n = 1 + |x|^n$). It only remains to control this integral.

[14] B. Leimkuhler, C. Matthews, and G. Stoltz *The computation of averages from equilibrium and non-equilibrium Langevin molecular dynamics*

[15] P. Plechac, G. Stoltz, and T. Wang *Convergence of the likelihood ratio method for linear response of non-equilibrium stationary states*