Sticky Coupling as a Control Variate for Computing Transport Coefficients

Shiva Darshan

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

In collaboration with A. Eberle and G. Stoltz

Project funded by ANR SINEQ

GAMM: Modeling, analysis and simulation of molecular systems

Outline

- **Linear response for steady-state nonequilibrium dynamics**
  - Perturbations of equilibrium dynamics
  - Definition of transport coefficients
  - Variance & bias of NEMD estimator
- **Couplings based estimators**
  - Couplings based estimators
  - Synchronous coupling
  - Sticky coupling
- **Numerical Illustrations**
- **Extensions and perspectives**

Linear response for steady-state nonequilibrium dynamics

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 $b, F : \mathbb{R}^d \to \mathbb{R}^d$ Lipschitz, with F bounded, and $\eta \in \mathbb{R}$.

Assumption

There exists $M \geqslant 0$ and $m > 0$ such that

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

and that b, F are smooth with derivatives growing polynomially and for any $\eta_* > 0$ there exists $\lambda_{\eta_*} > 0$ such that

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

These two assumption are satisfied if $b(x) = -V_1(x) - V_2(x)$, where V_1 is a confining potential and V_2 is a compactly supported.

Estimating transport coefficients

Response property $R \in L^2(\nu_0)$, s.t. $\nu_0(R) = 0$, the transport coefficient α_R satisfies:

$$\alpha_R = \lim_{\eta \to 0} \frac{\mathbb{E}_\eta(R)}{\eta}$$

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

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

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 $\left(\widetilde{X}, \widetilde{Y}\right)$ of random variables such that $\widetilde{X} \stackrel{\mathrm{Law}}{=} X$ and $\widetilde{Y} \stackrel{\mathrm{Law}}{=} Y$

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

By choosing $W = \widetilde{W}$, we can *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.

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

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 long as we have sufficient contractivity, say due to sufficiently high temperature[1] or in the underdamped case[2], we can control the moments of the estimator as

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

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

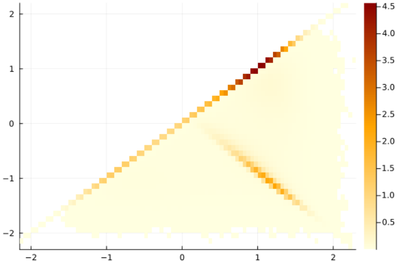
What to do when we do not have enough strong contractivity?

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

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

Sticky Coupling

One can constructed a coupling[3] 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



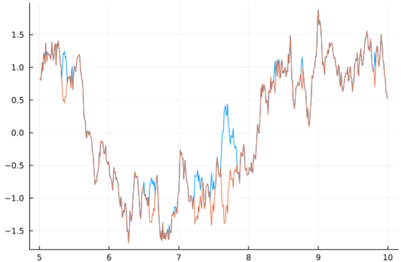


Figure: Sticky coupling of a 1D particle in a double well potential perturbed by a constant force to the right. **Left:** histogram of coupled process; **Right:** segment of trajectory of coupled process

[3] 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. Furthermore $\left\{t \geqslant 0 : X_t^\eta = Y_t^0\right\}$ is random fat Cantor set: for any $T > 0$

$$\mathbb{P}\left(\left|\left\{t \in [0,T] : X_t^\eta = Y_t^0\right\}\right| > 0\right) > 0,$$

but

$$\mathbb{P}\left(\exists a < b, \text{ s.t. } [a,b] \subset \left\{t \in [0,T] : X_t^\eta = Y_t^0\right\}\right) = 0.$$

Discrete-Time Sticky Coupling

Lets work with the discrete version of sticky coupling [4] instead. Consider the estimator

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

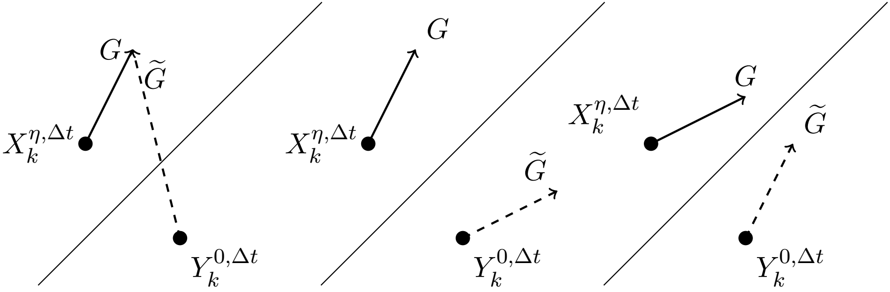
with $\left\{X_k^{\eta,\Delta t}, Y_k^{0,\Delta t}\right\}_{k\in\mathbb{N}}$ the discrete sticky coupling of the Euler-Maruyama discretizations of $(X_t^\eta)_{t\geqslant 0}$ and $\left(Y^0\right)_{t\geqslant 0}$.
Let $\{G_k\}_{k\geqslant 1}$ and $\{U_k\}_{k\geqslant 1}$ be i.i.d sequences of Gaussian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

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

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

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

Discrete-Time Sticky Coupling



(a) Collision $B_{k+1} = 1$

(b) Reflection, $B_{k+1} = 0$, resulting in separation

(c) Reflection, $B_{k+1} = 1$, resulting in contraction

Discrete-Time Sticky Coupling

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

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

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

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

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

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

Discrete-Time Sticky Coupling

Proposition

If b is strongly contractive at infinity and 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[5]
Strong contractivity implies that $e^{c|x|^2} + e^{c|y|^2}$ is a Lyapunov function.
Furthermore $p_{\Delta t,\beta}(x,y,z) > 0$ implies that there is always strictly positive probability of the process returning to the diagonal. Thus for any $K > 0$ there exists $\rho_{K,\Delta t} \in (0,1)$ such that

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

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

[5] 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 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 and Δt small enough, then $\left\{X_k^{\eta,\Delta t}, Y_k^{0,\Delta t}\right\}_{k\in\mathbb{N}}$ satisfies a CLT and there exists K_1, K_2 such that

$$\forall \eta \in [-\eta_*, \eta_*], \quad \lim_{N\to\infty} N \mathrm{Var}\left(\widehat{\Psi}_{\eta,N}^{\Delta t}\right) \leqslant \frac{K_1}{\eta}, \tag{3}$$

and

$$\left|\mathbb{E}\left[\widehat{\Psi}_{\eta,N}^{\Delta t}\right] - \alpha_{R,\eta}\right| \leqslant K_2 \left(\frac{1}{N} + \Delta t\right). \tag{4}$$

Key Idea of Proof

Proposition

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

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

Heuristic "proof" of proposition

$$\begin{aligned} &\int_{\mathbb{R}^d \times \mathbb{R}^d} \left(\mathrm{e}^{c|x|} + \mathrm{e}^{c|y|}\right) \mathbf{1}_{\{x \neq y\}} \, d\mu_{\eta,\Delta t}\,(dx\,dy) \\ &\leqslant \mu_{\eta,\Delta t}\,(\{x \neq y\}) \int_{\mathbb{R}^d \times \mathbb{R}^d} \left(\mathrm{e}^{c|x|} + \mathrm{e}^{c|y|}\right) \, d\mu_{\eta,\Delta t}\,(dx\,dy)\,, \end{aligned} \tag{5}$$

The sticky coupled process spends an $O(\eta)$ 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

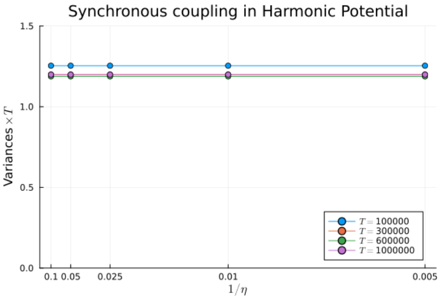
Consider a 2-dimensional Ornstein-Uhlenbeck process

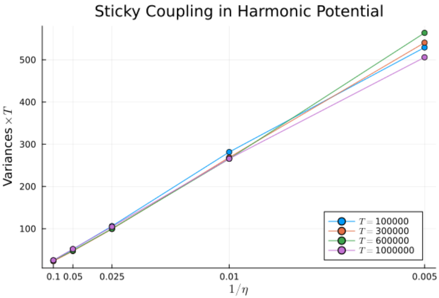
$$dX_t^\eta = -\begin{bmatrix} 1 & -\eta \\ 0 & 1 \end{bmatrix} X_t^\eta dt + \sqrt{\frac{2}{\beta}} dW_t;$$

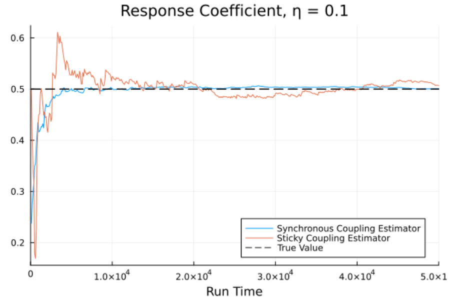
here $b\left(x\right) = -\nabla U = -x$ and $F(x) = \left[x_2 \; 0\right]^T$. We choose as response function the covariance between the components. In this case α_R is explicitly calculable.

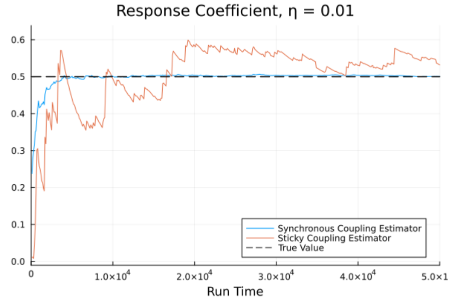
$$R(x) = x_1 x_2, \qquad \alpha_R = \frac{1}{2\beta}$$

Numerical Illustrations: Strongly Convex Potential






Response Coefficient, η = 0.1
0.6
0.5
0.4
0.3
0.2
0
1.0×10^4
2.0×10^4
3.0×10^4
4.0×10^4
5.0×1
Run Time
Synchronous Coupling Estimator
Sticky Coupling Estimator
True Value


Response Coefficient, η = 0.01
0.6
0.5
0.4
0.3
0.2
0.1
0.0
0
1.0×10^4
2.0×10^4
3.0×10^4
4.0×10^4
5.0×1
Run Time
Synchronous Coupling Estimator
Sticky Coupling Estimator
True Value

Numerical Illustrations: Lennard-Jones Fluid

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

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

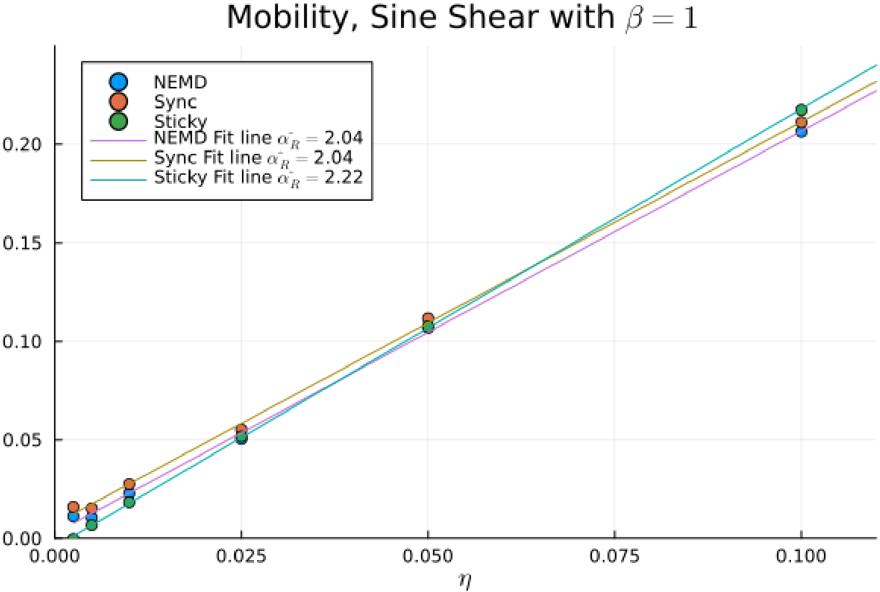
Thus $b(x) = -\nabla U = -\nabla(U_1 + U_2)$. For F we use sine shear

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

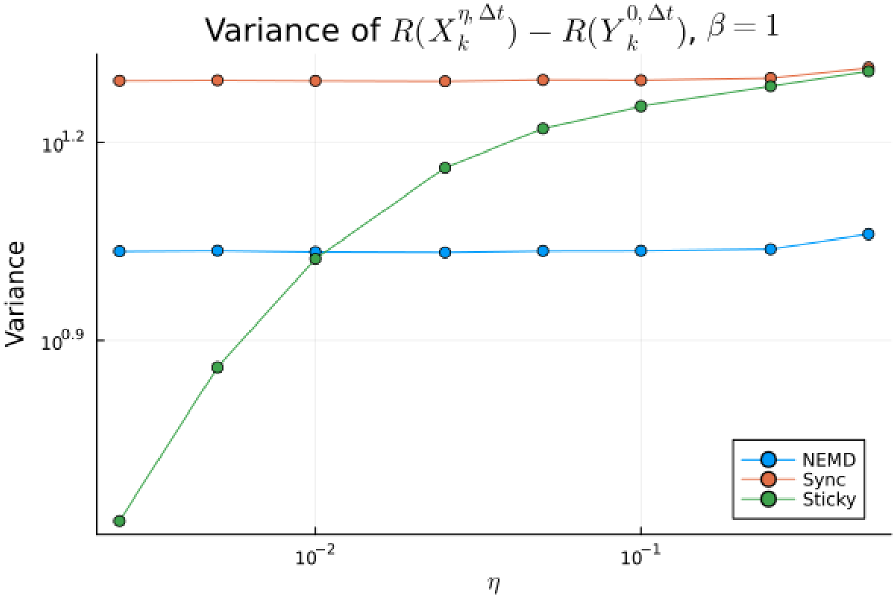
and we measure the mobility response

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

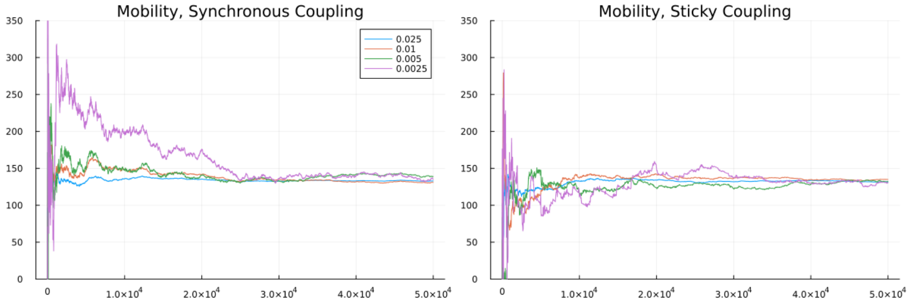
Numerical Illustrations: Lennard-Jones Fluid Sine Shear



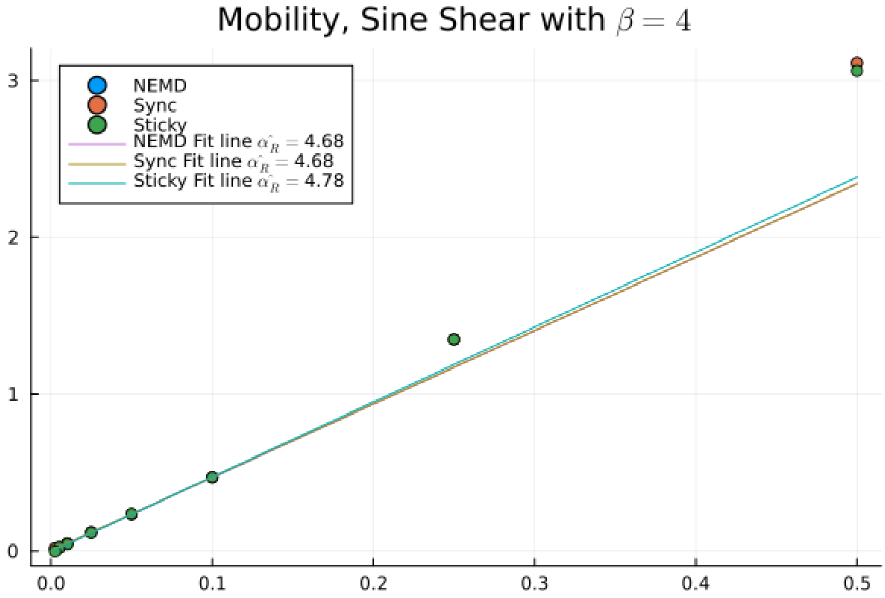
Numerical Illustrations: Lennard-Jones Fluid Sine Shear



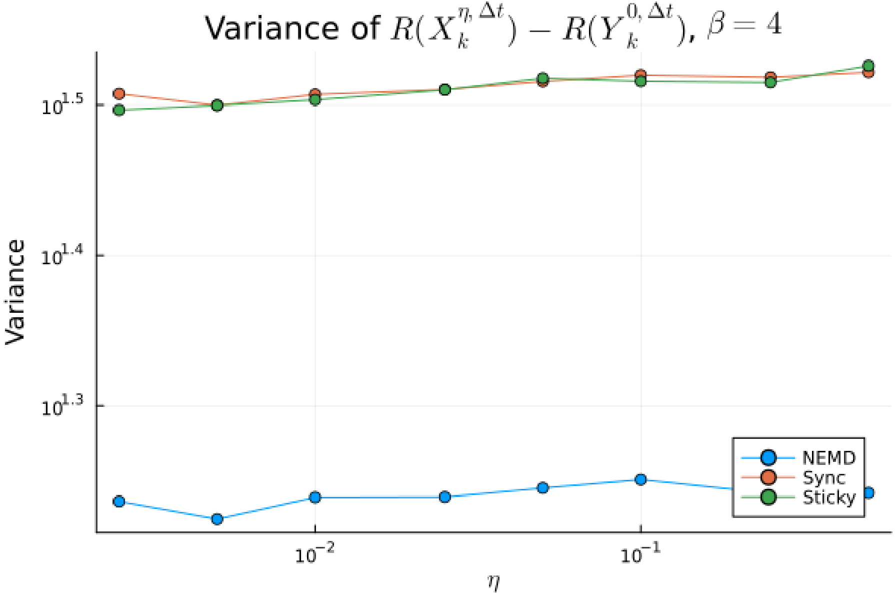
Numerical Illustrations: Lennard-Jones Fluid Sine Shear



Numerical Illustrations: Lennard-Jones Fluid Sine Shear



Numerical Illustrations: Lennard-Jones Fluid Sine Shear



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[6]
- 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 [7]
- Extension to kinetic Langevin dynamics [8] [9]

Vielen Dank für Ihre Aufmerksamkeit!!

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

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

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

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

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

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

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 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[10]. 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

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

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

More Ideas of Proof of Theorem

A long computation adapting the strategies of Leimkuhler, et. al (2015)[11] and Plechac, et. al (2021)[12] 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.

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

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