Mobility estimation for Langevin dynamics using control variates

MCQMC 2022

Urbain Vaes
urbain.vaes@inria.fr

MATHERIALS – Inria Paris & CERMICS – École des Ponts ParisTech

July 2022

Collaborators and outline

Grigorios Pavliotis
Imperial College London
Department of Mathematics

Gabriel Stoltz
École des Ponts ParisTech
CERMICS

Reference: G. A. Pavliotis, G. Stoltz, and U. Vaes. Mobility estimation for Langevin dynamics using control variates. arXiv preprint, 2022

Outline:

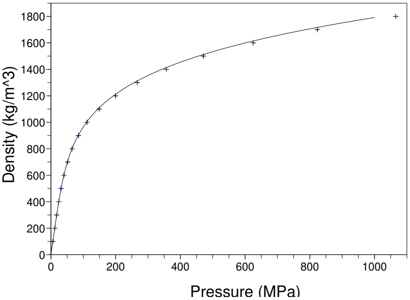
The big picture

Efficient mobility estimation

Numerical experiments

Goals of molecular dynamics

- Computation of macroscopic properties from Newtonians atomistic models:
  - Static properties, such as
    - the heat capacity and
    - the equations of state $P = P(\rho, T)$.
  - Dynamical properties, such as transport coefficients:
    - the viscosity;
    - the thermal conductivity;
    - the mobility of ions in solution.



Equation of state of argon at 300K.
- '+': molecular simulation;
- Solid line: experimental measurements[1].

- Numerical microscope: used in physics, biology, chemistry.

[1] `https://webbook.nist.gov/chemistry/fluid/`

Consider the (one-particle) Langevin equation

$$\begin{cases} d\mathbf{q}_t = \mathbf{p}_t\, dt, \\ d\mathbf{p}_t = -\nabla V(\mathbf{q}_t)\, dt - \gamma \mathbf{p}_t\, dt + \sqrt{2\gamma\beta^{-1}}\, d\mathbf{W}_t, \end{cases} \qquad (\mathbf{q}_0, \mathbf{p}_0) \sim \mu,$$

where γ is the friction, V is a periodic potential, and $\beta = \frac{1}{k_B T}$.

- The invariant probability measure is

$$\mu(\mathbf{q}, \mathbf{p}) = \frac{1}{Z}\, e^{-\beta H(\mathbf{q},\mathbf{p})} = \frac{1}{Z}\, e^{-\beta\left(V(\mathbf{q}) + \frac{|\mathbf{p}|^2}{2}\right)}, \quad \text{on } \mathbf{T}^d \times \mathbf{R}^d.$$

- The generator of the associated Markov semigroup

$$\left(e^{\mathcal{L}t}\varphi\right)(\mathbf{q}, \mathbf{p}) = \mathbf{E}\big(\varphi(\mathbf{q}_t, \mathbf{p}_t) \big| (\mathbf{q}_0, \mathbf{p}_0) = (\mathbf{q}, \mathbf{p})\big)$$

  is the following operator:

$$\mathcal{L} = (\mathbf{p} \cdot \nabla_{\mathbf{q}} - \nabla V(q) \cdot \nabla_{\mathbf{p}}) + \gamma \left(-\mathbf{p}\nabla_{\mathbf{p}} + \beta^{-1}\Delta_{\mathbf{p}}\right) =: \mathcal{L}_{\mathrm{ham}} + \gamma\, \mathcal{L}_{\mathrm{FD}}.$$

We denote by $\|\cdot\|$ and $\langle\cdot,\cdot\rangle$ the norm and inner product of $L^2(\mu)$, and

$$L^2_0(\mu) = \left\{\varphi \in L^2(\mu) : \langle\varphi, 1\rangle = \mathbf{E}_\mu \varphi = 0\right\}.$$

Definition of the mobility

Consider Langevin dynamics with additional forcing in a direction $\mathbf{e}$:

$$\begin{cases} d\mathbf{q}_t = \mathbf{p}_t \, dt, \\ d\mathbf{p}_t = -\nabla V(\mathbf{q}_t) \, dt + \eta \mathbf{e} \, dt - \gamma \mathbf{p}_t \, dt + \sqrt{2\gamma\beta^{-1}} \, d\mathbf{W}_t. \end{cases}$$

This dynamics admits a unique invariant probability distribution $\mu_\eta \in \mathcal{P}(\mathbf{T}^d \times \mathbf{R}^d)$.

Definition (Mobility)

The mobility in direction $\mathbf{e}$ is defined mathematically as

$$M_{\mathbf{e}} = \lim_{\eta \to 0} \frac{1}{\eta} \mathbf{E}_{\mu_\eta} [\mathbf{e}^\mathsf{T} \mathbf{p}]$$

$\approx$ factor relating the mean momentum to the strength of the inducing force.

- There is a symmetric mobility tensor $\mathbf{M}$ such that $M_{\mathbf{e}} = \mathbf{e}^\mathsf{T} \mathbf{M} \mathbf{e}$.
- **Einstein's relation:** $\mathbf{D} = \beta^{-1} \mathbf{M}$, with $\mathbf{D}$ the effective diffusion coefficient.

Effective diffusion

It is possible to show a functional central limit theorem for the Langevin dynamics[2]:

$$\varepsilon \mathbf{q}_{s/\varepsilon^2} \xrightarrow[\varepsilon \to 0]{} \sqrt{2\mathbf{D}}\,\mathbf{W}_s \qquad \text{weakly on } C([0,\infty)).$$

In particular, $\mathbf{q}_t/\sqrt{t} \xrightarrow[t\to\infty]{} \mathcal{N}(0, 2\mathbf{D})$ weakly.

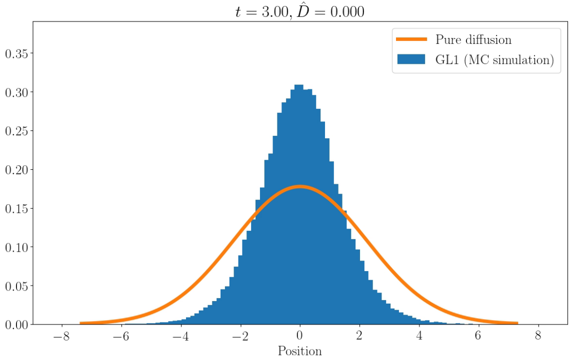


Figure: Histogram of $q_t/\sqrt{t}$. The potential $V(q) = -\cos(q)/2$ is illustrated in the background.

[2] R. N. Bhattacharya. On the functional central limit theorem and the law of the iterated logarithm

Mathematical expression for the effective diffusion (dimension 1)

Expression of D in terms of the solution to a Poisson equation

The effective diffusion coefficient is given by where $D = \langle \phi, p \rangle$ and ϕ is the solution to

$$-\mathcal{L}\phi = p, \qquad \phi \in L^2_0(\mu) := \left\{ u \in L^2(\mu) : \langle u, 1 \rangle = 0 \right\}.$$

Key idea of the proof: Apply Itô's formula to ϕ

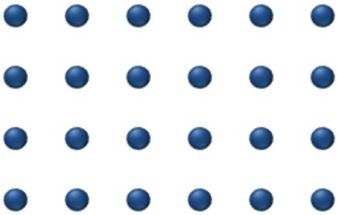
$$\mathrm{d}\phi(q_s, p_s) = -p_s \, \mathrm{d}s + \sqrt{2\gamma\beta^{-1}} \, \frac{\partial \phi}{\partial p}(q_s, p_s) \, \mathrm{d}W_s$$

and then rearrange:

$$\begin{aligned} \varepsilon(q_{t/\varepsilon^2} - q_0) &= \varepsilon \int_0^{t/\varepsilon^2} p_s \, \mathrm{d}s \\ &= \underbrace{\varepsilon\Big(\phi(q_0, p_0) - \phi(q_{t/\varepsilon^2}, p_{t/\varepsilon^2})\Big)}_{\to 0} + \underbrace{\sqrt{2\gamma\beta^{-1}}\varepsilon \int_0^{t/\varepsilon^2} \frac{\partial \phi}{\partial p}(q_s, p_s) \, \mathrm{d}W_s}_{\to \sqrt{2D} W_t \ \text{ weakly by MCLT}} . \end{aligned}$$

In the multidimensional setting, $D_{\mathbf{e}} = \left\langle \phi_{\mathbf{e}}, \mathbf{e}^\mathsf{T} \mathbf{p} \right\rangle$ with $-\mathcal{L}\phi_{\mathbf{e}} = \mathbf{e}^\mathsf{T} \mathbf{p}$

Open question: surface diffusion when $\gamma \ll 1$[3]



Applications:

- integrated circuits;
- catalysis.

Open question: behavior of the effective diffusion coefficient when $\gamma \ll 1$?

$$D = \lim_{t \to \infty} \frac{\langle |\mathbf{q}(t)|^2 \rangle}{4t} \sim \gamma^{-\sigma}, \qquad \sigma = ???$$

---

[3] Source of the video: `https://en.wikipedia.org/wiki/Surface_diffusion`

Langevin dynamics: underdamped and overdamped regimes[4]

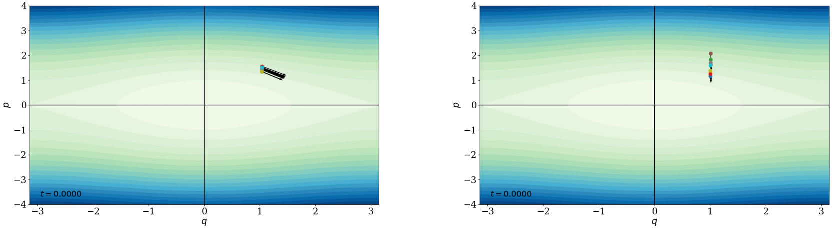


Figure: Langevin dynamics with friction $\gamma = 0.1$ (left) and $\gamma = 10$ (right)

- The underdamped limit as $\gamma \to 0$ is well understood in dimension 1 but not in the multi-dimensional setting.
- Overdamped limit: as $\gamma \to \infty$, the rescaled process $t \mapsto q_{\gamma t}$ converges weakly to the solution of the overdamped Langevin equation:

$$\dot{\mathbf{q}} = -\nabla V(q) + \sqrt{2\,\beta^{-1}}\,\dot{\mathbf{W}}.$$

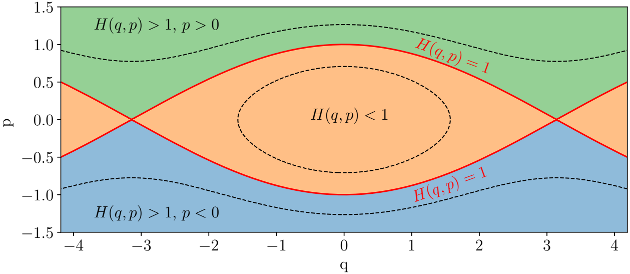
[4] M. Hairer and G. A. Pavliotis. From ballistic to diffusive behavior in periodic potentials. J. Stat. Phys., 2008.

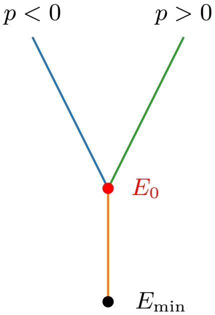
The underdamped limit in dimension 1

As $\gamma \to 0$, the Hamiltonian of the rescaled process

$$\begin{cases} q_\gamma(t) = q(t/\gamma), \\ p_\gamma(t) = p(t/\gamma), \end{cases}$$

converges weakly to a diffusion process on a graph.



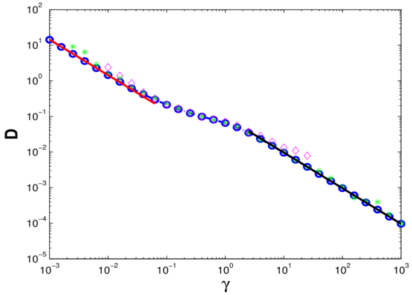


In this limit, it holds that

$$\phi = -\mathcal{L}^{-1} p = \gamma^{-1} \phi_{\text{und}} + \mathcal{O}(\gamma^{-1/2}).$$

Scaling of the effective diffusion coefficient for Langevin dynamics[5]

In dimension 1, $\lim_{\gamma\to 0} \gamma D^\gamma = D_{\text{und}}$ and $\lim_{\gamma\to\infty} \gamma D^\gamma = D_{\text{ovd}}$.



Our aims in this work:

- How can we efficiently estimate the effective diffusion coefficient when $\gamma \ll 1$?
- How does the mobility scale as $\gamma \to 0$ in the multidimensional setting?

[5] G. A. Pavliotis and A. Vogiannou. Diffusive transport in periodic potentials: underdamped dynamics. Fluct. Noise Lett., 2008.

Outline

The big picture

Efficient mobility estimation

Numerical experiments

Brief literature review

In dimension > 1, it does not hold that $\gamma D_{\mathbf{e}}^{\gamma} \xrightarrow[\gamma \to 0]{} D_{\mathrm{und}}$ when V is non separable, e.g.

$$V(\mathbf{q}) = -\frac{1}{2}\Big(\cos(q_1) + \cos(q_2)\Big) - \delta \cos(q_1)\cos(q_2)$$

Open question: how does $D_{\mathbf{e}}^{\gamma}$ behave when $\gamma \ll 1$ and $d = 2$?

Various answers are given in the literature:

- $D_{\mathbf{e}}^{\gamma} \propto \gamma^{-1/2}$ for specific potentials[6];
- $D_{\mathbf{e}}^{\gamma} \propto \gamma^{-1/3}$ for specific potentials[7];
- $D_{\mathbf{e}}^{\gamma} \propto \gamma^{-\sigma}$ with σ depending on the potential[8].

[6] L. Y. Chen, M. R. Baldan, and S. C. Ying. Surface diffusion in the low-friction limit: Occurrence of long jumps. Phys. Rev. B, 1996.

[7] O. M. Braun and R. Ferrando. Role of long jumps in surface diffusion. Phys. Rev. E, 2002.

[8] J. Roussel. Theoretical and Numerical Analysis of Non-Reversible Dynamics in Computational Statistical Physics. PhD thesis, Université Paris-Est, 2018.

Numerical approaches for calculating the effective diffusion coefficient

- Linear response approach:

$$D_{\mathbf{e}} = \lim_{\eta \to 0} \frac{1}{\beta \eta} \mathbf{E}_{\mu_\eta} \left(\mathbf{e}^\mathsf{T} \mathbf{p} \right).$$

where μ_η is the invariant distribution of the system with external forcing.

- Green–Kubo formula: Since $-\mathcal{L}^{-1} = \int_0^\infty \mathrm{e}^{t\mathcal{L}} \, \mathrm{d}t$,

$$\begin{aligned} D_{\mathbf{e}} &= \int -\mathcal{L}^{-1} (\mathbf{e}^\mathsf{T} \mathbf{p}) \, (\mathbf{e}^\mathsf{T} \mathbf{p}) \, \mathrm{d}\mu = \int_0^\infty \int \mathrm{e}^{t\mathcal{L}} (\mathbf{e}^\mathsf{T} \mathbf{p}) (\mathbf{e}^\mathsf{T} \mathbf{p}) \, \mathrm{d}\mu \, \mathrm{d}t \\ &= \int_0^\infty \mathbf{E}_\mu \big((\mathbf{e}^\mathsf{T} \mathbf{p}_0) (\mathbf{e}^\mathsf{T} \mathbf{p}_t) \big) \, \mathrm{d}t. \end{aligned}$$

- Einstein's relation:

$$D_{\mathbf{e}} = \lim_{t \to \infty} \frac{1}{2t} \mathbf{E}_\mu \left[\left| \mathbf{e}^\mathsf{T} (\mathbf{q}_t - \mathbf{q}_0) \right|^2 \right].$$

- Deterministic method, e.g. Fourier/Hermite Galerkin, for the Poisson equation

$$-\mathcal{L} \phi_{\mathbf{e}} = \mathbf{e}^\mathsf{T} \mathbf{p}, \qquad D_{\mathbf{e}} = \langle \phi_{\mathbf{e}}, p \rangle \, .$$

Estimation of the effective diffusion coefficient from Einstein's relation

Consider the following estimator of the effective diffusion coefficient $D_{\mathbf{e}}$:

$$u(T) = \frac{\left|\mathbf{e}^\mathsf{T}(\mathbf{q}_T - \mathbf{q}_0)\right|^2}{2T}, \qquad (\mathbf{q}_0, \mathbf{p}_0) \sim \mu.$$

Bias of this estimator:

$$\mathbf{E}\big[u(T)\big] = D_{\mathbf{e}} - \int_0^\infty \left\langle \mathrm{e}^{t\mathcal{L}}(\mathbf{e}^\mathsf{T}\mathbf{p}), \mathbf{e}^\mathsf{T}\mathbf{p} \right\rangle \min\left\{1, \frac{t}{T}\right\} \mathrm{d}t.$$

Using the decay estimate for the semigroup[9]

$$\left\| \mathrm{e}^{t\mathcal{L}} \right\|_{\mathcal{B}\left(L_0^2(\mu)\right)} \leq L\, \mathrm{e}^{-\ell \min\{\gamma, \gamma^{-1}\} t},$$

we deduce

$$|\mathbf{E}[u(T)] - D_{\mathbf{e}}| \leq \frac{C \max\{\gamma^2, \gamma^{-2}\}}{T}.$$

[9] J. Roussel and G. Stoltz. Spectral methods for Langevin dynamics and associated error estimates. ESAIM: Math. Model. Numer. Anal., 2018.

Variance of the estimator $u(T)$ for large T

For $T \gg 1$, it holds approximately that

$$\frac{\mathbf{e}^\mathsf{T}(\mathbf{q}_T - \mathbf{q}_0)}{\sqrt{2T}} \sim \mathcal{N}(0, D_\mathbf{e}) \qquad \rightsquigarrow \qquad u(T) = \frac{\left|\mathbf{e}^\mathsf{T}(\mathbf{q}_T - \mathbf{q}_0)\right|^2}{2D_\mathbf{e}T} \sim \chi^2(1).$$

Therefore, we deduce

$$\lim_{T\to\infty} \mathbf{V}\left[u(T)\right] = 2D_\mathbf{e}^2.$$

The relative standard deviation (asymptotically as $T \to \infty$) is therefore

$$\lim_{T\to\infty} \frac{\sqrt{\mathbf{V}\left[u(T)\right]}}{\mathbf{E}\left[u(T)\right]} = \sqrt{2} \qquad \rightsquigarrow \text{independent of } \gamma.$$

Scaling of the mean square error when using J realizations

Assuming an asymptotic scaling as $\gamma^{-\sigma}$ of $D_\mathbf{e}$, we have

$$\forall \gamma \in (0,1), \qquad \frac{\mathrm{MSE}}{D_\mathbf{e}^2} \leq \frac{C}{\gamma^{4-2\sigma}T^2} + \frac{2}{J}$$

Variance reduction using control variates

Let $\phi_{\mathbf{e}}$ denote the solution to the Poisson equation

$$-\mathcal{L}\phi_{\mathbf{e}}(\mathbf{q}, \mathbf{p}) = \mathbf{e}^{\mathsf{T}}\mathbf{p}, \qquad \phi_{\mathbf{e}} \in L^2_0(\mu).$$

and let $\psi_{\mathbf{e}}$ denote an approximation of $\phi_{\mathbf{e}}$. By Itô's formula, we obtain

$$\phi_{\mathbf{e}}(\mathbf{q}_T, \mathbf{p}_T) - \phi_{\mathbf{e}}(\mathbf{q}_0, \mathbf{p}_0) = -\int_0^T \mathbf{e}^{\mathsf{T}}\mathbf{p}_t \, \mathrm{d}t + \sqrt{2\gamma\beta^{-1}} \int_0^T \nabla_{\mathbf{p}}\phi_{\mathbf{e}}(\mathbf{q}_t, \mathbf{p}_t) \cdot \mathrm{d}\mathbf{W}_t.$$

Therefore

$$\begin{aligned} \mathbf{e}^{\mathsf{T}}(\mathbf{q}_T - \mathbf{q}_0) &= \int_0^T \mathbf{e}^{\mathsf{T}}\mathbf{p}_t \, \mathrm{d}t \\ &\approx -\psi_{\mathbf{e}}(\mathbf{q}_T, \mathbf{p}_T) + \psi_{\mathbf{e}}(\mathbf{q}_0, \mathbf{p}_0) + \sqrt{2\gamma\beta^{-1}} \int_0^T \nabla_{\mathbf{p}}\psi_{\mathbf{e}}(\mathbf{q}_t, \mathbf{p}_t) \cdot \mathrm{d}\mathbf{W}_t =: \xi_T. \end{aligned}$$

which suggests the improved estimator

$$v(T) = \frac{\left|\mathbf{e}^{\mathsf{T}}(\mathbf{q}_T - \mathbf{q}_0)\right|^2}{2T} - \left(\frac{|\xi_T|^2}{2T} - \lim_{T \to \infty} \mathbf{E}\left[\frac{|\xi_T|^2}{2T} \right] \right).$$

Properties of the improved estimator

Smaller bias if $-\mathcal{L}\psi_{\mathbf{e}} \approx \mathbf{e}^\mathsf{T}\mathbf{p}$:

$$\left|\mathbf{E}\left[v(T)\right] - D^{\gamma}_{\mathbf{e}}\right| \leq \frac{L\max\{\gamma^2, \gamma^{-2}\}}{T\ell^2} \left\|\mathbf{e}^\mathsf{T}\mathbf{p} + \mathcal{L}\psi_{\mathbf{e}}\right\| \left(\beta^{-1/2} + \|\mathcal{L}\psi_{\mathbf{e}}\|\right).$$

Smaller variance:

$$\mathbf{V}\left[v(T)\right] \leq C\left(T^{-1}\|\phi_{\mathbf{e}} - \psi_{\mathbf{e}}\|_{L^4(\mu)}{}^2 + \gamma\|\nabla_{\mathbf{p}}\phi_{\mathbf{e}} - \nabla_{\mathbf{p}}\psi_{\mathbf{e}}\|_{L^4(\mu)}{}^2\right) \\ \times\left(T^{-1}\|\phi_{\mathbf{e}} + \psi_{\mathbf{e}}\|^2_{L^4(\mu)} + \gamma\|\nabla_{\mathbf{p}}\phi_{\mathbf{e}} + \nabla_{\mathbf{p}}\psi_{\mathbf{e}}\|^2_{L^4(\mu)}\right).$$

Construction of $\psi_{\mathbf{e}}$ in the one-dimensional setting. We consider two approaches:

- Approximate the solution to the Poisson equation by a Galerkin method.
- Use asymptotic result for the Poisson equation:

$$\gamma\phi \xrightarrow[\gamma\to 0]{L^2(\mu)} \phi_{\mathrm{und}},$$

which suggests letting $\psi = \phi_{\mathrm{und}}/\gamma$.

Construction of the approximate solution $\psi_{\mathbf{e}}$ in dimension 2

We consider the potential

$$V(\mathbf{q}) = -\frac{1}{2}\Big(\cos(q_1) + \cos(q_2)\Big) - \delta \cos(q_1)\cos(q_2).$$

- For this potential, $\mathbf{D}$ is isotropic $\rightsquigarrow$ sufficient to consider $\mathbf{e} = (1,0)$,

$$D_{(1,0)} = \left\langle \phi_{(1,0)}, p_1 \right\rangle, \qquad -\mathcal{L}\phi_{(1,0)}(\mathbf{q},\mathbf{p}) = p_1.$$

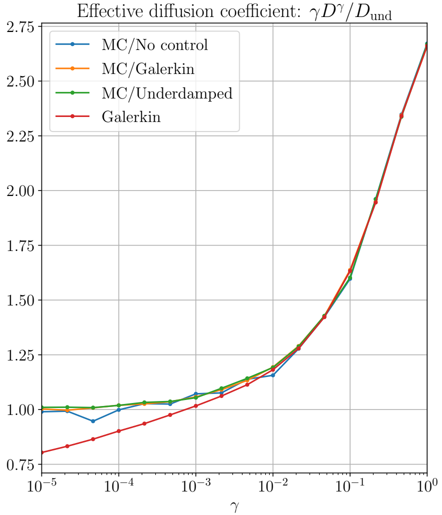
- If $\delta = 0$, then the solution is $\phi_{(1,0)}(\mathbf{q},\mathbf{p}) = \phi_{1\mathrm{D}}(q_1,p_1)$, where $\phi_{1\mathrm{D}}$ solves

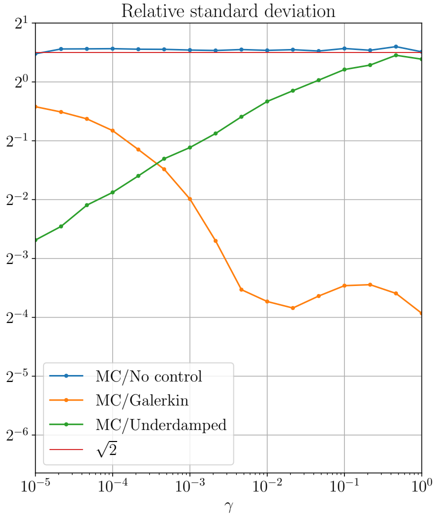
$$-\mathcal{L}_{1\mathrm{D}}\phi_{1\mathrm{D}}(q,p) = p, \qquad V_{1\mathrm{D}}(q) = \frac{1}{2}\cos(q).$$

- We take $\psi_{(1,0)}(\mathbf{q},\mathbf{p}) = \psi_{1\mathrm{D}}(q_1,p_1)$, where $\psi_{1\mathrm{D}} \approx \phi_{1\mathrm{D}}$.

Outline

Numerical experiments for the one-dimensional case (1/2)





Numerical experiments for the one-dimensional case (2/2)

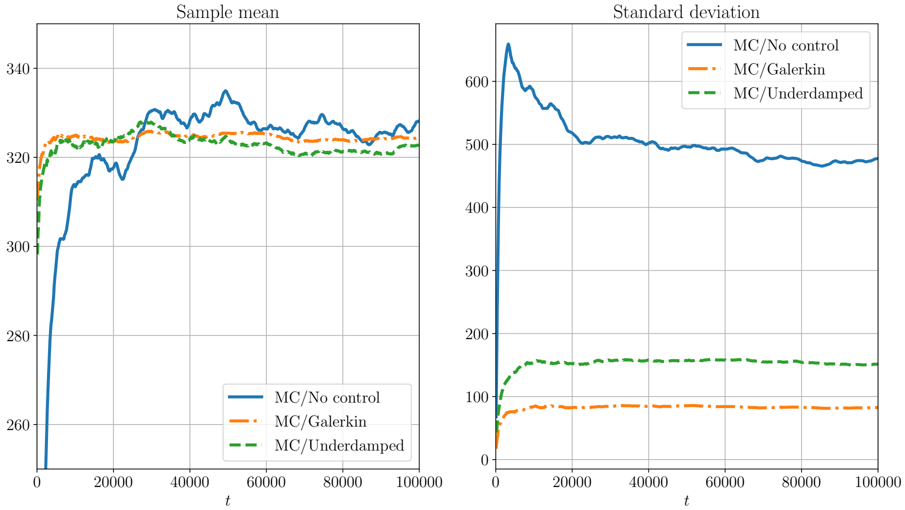
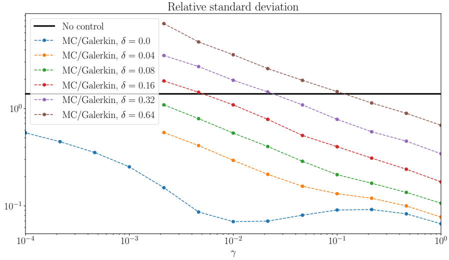
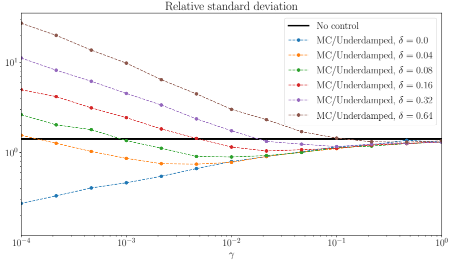


Figure: Evolution of the sample mean and standard deviation, estimated from $J = 5000$ realizations for $\gamma = 10^{-3}$.

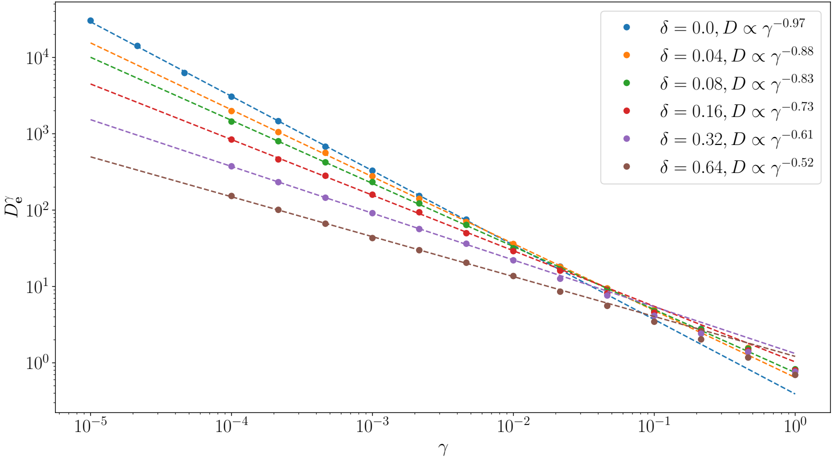
Performance of the control variates approach in dimension 2





- Variance reduction is possible if $|\delta|/\gamma \ll 1$;
- Control variates are not very useful when $\gamma \ll 1$ and δ is fixed.

Scaling of the mobility in dimension 2



Summary and perspectives for future work

In this talk, we presented

- a variance reduction approach for efficiently estimating the mobility;
- numerical results showing that the scaling of the mobility is not universal.

Perspectives for future work:

- Use alternative methods (PINNs, Gaussian processes) to solve the Poisson equation;
- Improve and study variance reduction approaches for other transport coefficients.

Thank you for your attention!

Connection with the asymptotic variance of MCMC estimators

Ergodic theorem[10]: for an observable $\varphi \in L^1(\mu)$,

$$\widehat{\varphi}_t = \frac{1}{t}\int_0^t \varphi(\mathbf{q}_s, \mathbf{p}_s)\,\mathrm{d}s \xrightarrow[t\to\infty]{a.s.} \mathbf{E}_\mu \varphi.$$

Central limit theorem[11]: If the following Poisson equation has a solution $\phi \in L^2(\mu)$,

$$-\mathcal{L}\phi = \varphi - \mathbf{E}_\mu \varphi,$$

then a central limit theorem holds:

$$\sqrt{t}\big(\widehat{\varphi}_t - \mathbf{E}_\mu \varphi\big) \xrightarrow[t\to\infty]{\mathrm{Law}} \mathcal{N}(0, \sigma_\varphi^2), \qquad \sigma_\varphi^2 = \langle \phi, \varphi - \mathbf{E}_\mu \varphi \rangle\,.$$

Connection with effective diffusion: Apply this result with $\varphi(\mathbf{q}, \mathbf{p}) = \mathbf{e}^\mathsf{T}\mathbf{p}$.

[10] W. Kliemann. Recurrence and invariant measures for degenerate diffusions. Ann. Probab., 1987.
[11] R. N. Bhattacharya. On the functional central limit theorem and the law of the iterated logarithm for Markov processes. Z. Wahrsch. Verw. Gebiete, 1982.