European Research Council

Established by the European Commission

Reducing errors in molecular dynamics

Gabriel STOLTZ

gabriel.stoltz@enpc.fr

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

Work also supported by ANR Funding ANR-14-CE23-0012 ("COSMOS")

What can (and cannot) applied mathematics do?

We, applied mathematicians...

- work on simplified models (one dimensional reaction coordinates, overdamped Langevin dynamics, etc) $\rightarrow$ this make us look like fools...
- exaggerate sources of errors (for instance by considering situations which rarely happen in practice) $\rightarrow$ this makes us insufferable

...but the aim is to

- rigorously understand why some methods work and some don't (mathematical proofs)
- devise new numerical strategies based on this theoretical understanding
- validate them on toy examples
- transfer the knowledge to practitionners by helping them to implement the methods into their own codes

Outline

- Systematic errors for the **computation of average properties**
- Systematic errors for **transport coefficients**
- Reducing **statistical errors**

Main message

A careful construction of the numerical method can reduce both systematic and statistical errors

T. Lelièvre and G. Stoltz, Partial differential equations and stochastic methods in molecular dynamics, *Acta Numerica* **25**, 681-880 (2016)

Computational Statistical Physics

• *Predict macroscopic properties of materials from their microscopic description*

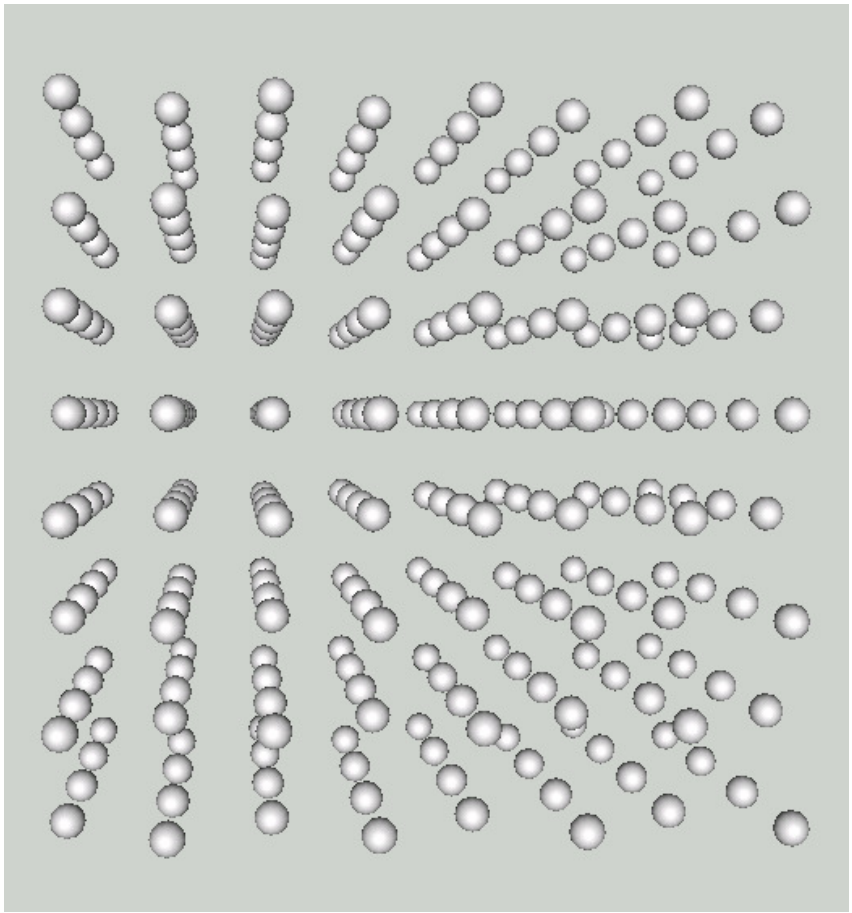
• **Microstate**
  - positions $q = (q_1, \ldots, q_N)$ and momenta $p = (p_1, \ldots, p_N)$
  - energy of the system $H(q,p) = V(q) + \sum_{i=1}^{N} \frac{p_i^2}{2m_i}$
  - (almost) all the physics is in the choice of V...

• **Macrostate**
  - described by a probability measure μ
  - constraints fixed exactly or in average (number of particles, volume, energy)

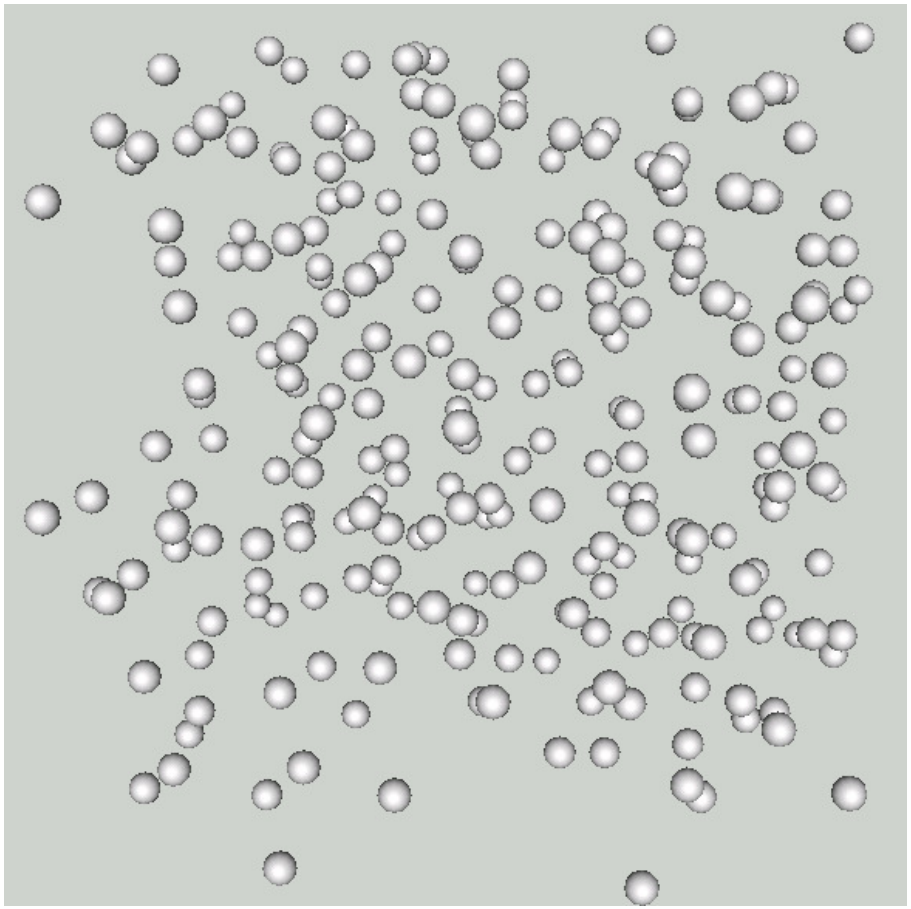
• **Properties:** static (equation of state, heat capacities, etc) and dynamical (transport coefficient, transition pathway, etc)

Example of molecular systems

What is the melting temperature of Argon?



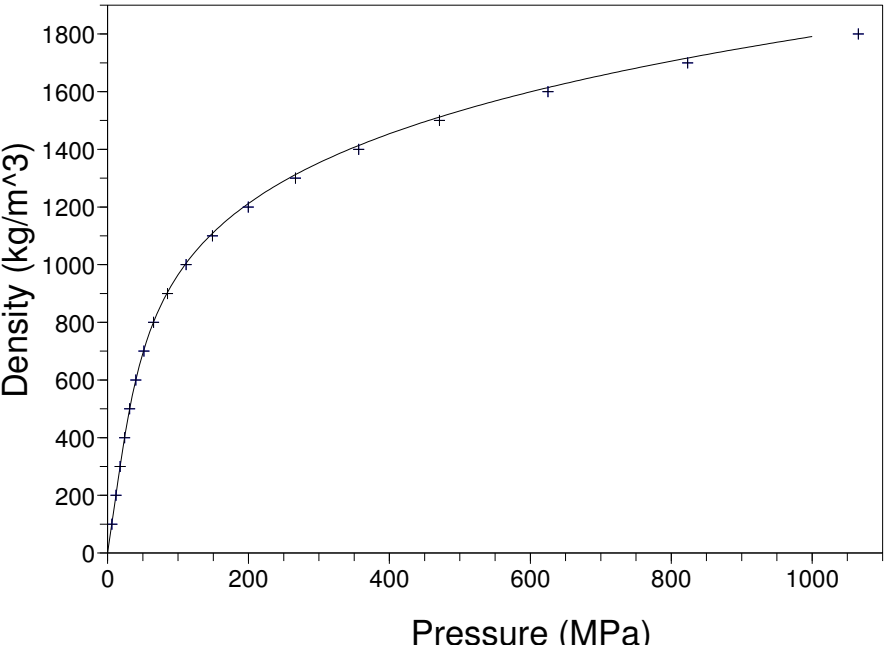
(a) Solid Argon (low temperature)

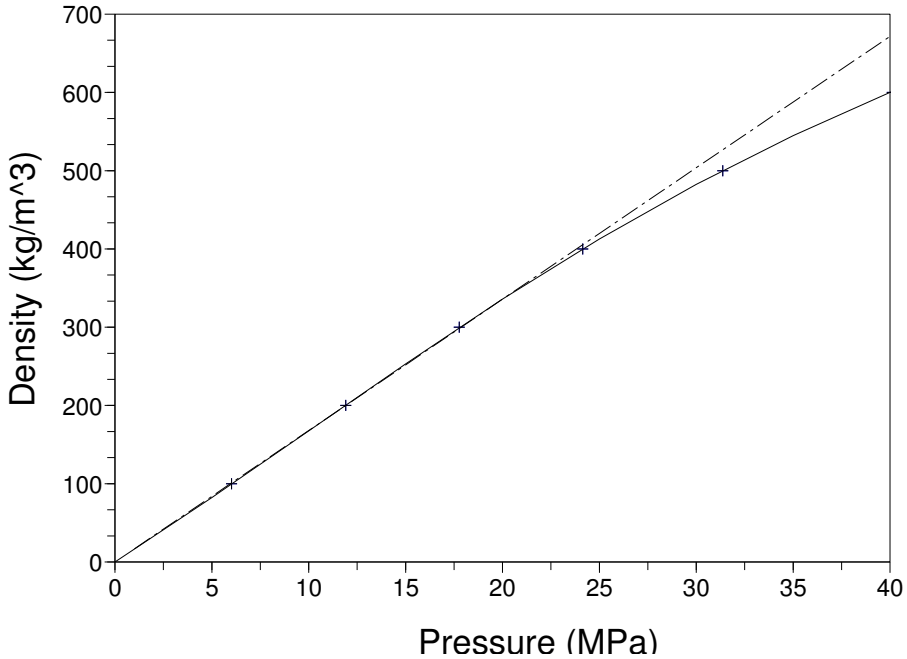


(b) Liquid Argon (high temperature)

Example of molecular systems (2)

Equation of state of Argon: density as a function of pressure, $T = 300$ K (comparison with data of *National Institute of Standards and Technology*)





Explore extreme conditions of matter...

Some orders of magnitude...

- **Physical quantities**
  - distances ~ 1 Å $= 10^{-10}$ m
  - energy per particle $\sim k_{\mathrm{B}} T \sim 4 \times 10^{-21}$ J at 300 K
  - atomic masses $\sim 10^{-26}$ kg
  - typical times $\sim 10^{-15}$ s
  - number of particles $\sim \mathcal{N}_A = 6.02 \times 10^{23}$

- **"Standard" simulations**
  - 10^6 particles ["heroic": 10^9 particles and more]
  - total time: (fraction of) ns ["heroic": (fraction of) μs]

- **Analogy** to understand what such large numbers represent...
  - about 10^{22} moles of water on Earth
  - 10^6 moles of water ~ 1 m^3

Aims of computational statistical physics

- **"Numerical microscope"**
  - gaining some insight into physical mechanisms at the atomic scale
  - From the press release for the Nobel prize in Chemistry 2013 (Karplus/Levitt/Warshel)

    Today the computer is just as important a tool for chemists as the test tube. Simulations are so realistic that they predict the outcome of traditional experiments.

- **Computation of average properties** (static)

$$\langle A \rangle = \int_{\mathcal{E}} A(q,p)\, \mu(dq\, dp)$$

Computation of high dimensional integrals

- μ is a probability measure
- A is the observable

Systematic errors for average properties

Langevin dynamics

- Positions $q \in \mathcal{D} = (L\mathbb{T})^d$ or $\mathbb{R}^d$ and momenta $p \in \mathbb{R}^d$

$\rightarrow$ phase-space $\mathcal{E} = \mathcal{D} \times \mathbb{R}^d$

- Hamiltonian (more general kinetic energies $U(p)$ can be considered[1])

$$H(q,p) = V(q) + \frac{1}{2} p^T M^{-1} p$$

Stochastic perturbation of the Hamiltonian dynamics

$$\begin{cases} dq_t = M^{-1} p_t \, dt \\ dp_t = -\nabla V(q_t) \, dt - \gamma M^{-1} p_t \, dt + \sqrt{\frac{2\gamma}{\beta}} \, dW_t \end{cases}$$

- Friction $\gamma > 0$ and inverse temperature $\beta = \frac{1}{k_{\mathrm{B}} T}$

[1] Redon, Stoltz, Trstanova, *J. Stat. Phys.* (2016)

Ergodicity results

- Canonical measure: invariant probability measure

$$\mu(dq\,dp) = Z^{-1}\mathrm{e}^{-\beta H(q,p)}\,dq\,dp$$

- Almost-sure convergence[2] of ergodic averages

$$\widehat{\varphi}_t = \frac{1}{t}\int_0^t \varphi(q_s,p_s)\,ds \xrightarrow[t\to+\infty]{} \int_{\mathcal{E}} \varphi\,d\mu$$

- Asymptotic variance of ergodic averages (in fact, Central Limit Theorem)

$$\sigma_\varphi^2 = \lim_{t\to+\infty} t\mathbb{E}\left[\widehat{\varphi}_t^2\right] = 2\int_{\mathcal{E}} \left(-\mathcal{L}^{-1}\Pi\varphi\right)\Pi\varphi\,d\mu$$

where $\Pi\varphi = \varphi - \mathbb{E}_\mu(\varphi)$

[2] Kliemann, *Ann. Probab.* **15**(2), 690-707 (1987)

Numerical integration: splitting schemes

- Decompose the dynamics in three parts as

$$(A) \qquad \begin{cases} dq_t = M^{-1} p_t \, dt, \\ dp_t = 0, \end{cases}$$

$$(B) \qquad \begin{cases} dq_t = 0, \\ dp_t = -\nabla V(q_t) \, dt, \end{cases}$$

$$(C) \qquad \begin{cases} dq_t = 0, \\ dp_t = -M^{-1} p_t \, dt + \sqrt{\dfrac{2}{\beta}} \, dW_t. \end{cases}$$

- First order splitting schemes: Trotter splitting

$$P_{\Delta t}^{ZYX} = \mathrm{e}^{\Delta t Z} \mathrm{e}^{\Delta t Y} \mathrm{e}^{\Delta t X} \simeq \mathrm{e}^{\Delta t (A + B + \gamma C)}$$

- Second order schemes: Strang splitting

$$P_{\Delta t}^{ZYXYZ} = \mathrm{e}^{\Delta t Z/2} \mathrm{e}^{\Delta t Y/2} \mathrm{e}^{\Delta t X} \mathrm{e}^{\Delta t Y/2} \mathrm{e}^{\Delta t Z/2}$$

- Other category: Geometric Langevin algorithms, *e.g.* $P_{\Delta t}^{\gamma C, A, B, A}$

Examples of splitting schemes for Langevin dynamics (2)

- $P_{\Delta t}^{B,A,\gamma C}$ corresponds to $\begin{cases} \widetilde{p}^{n+1} = p^n - \nabla V(q^n)\,\Delta t, \\ q^{n+1} = q^n + \Delta t\, M^{-1}\widetilde{p}^{n+1}, \\ p^{n+1} = \alpha_{\Delta t}\widetilde{p}^{n+1} + \sqrt{\dfrac{1-\alpha_{\Delta t}^2}{\beta}}\, M\, G^n \end{cases}$

where G^n are i.i.d. Gaussian and $\alpha_{\Delta t} = \exp(-\gamma M^{-1}\Delta t)$

- $P_{\Delta t}^{\gamma C,B_\eta,A,B_\eta,\gamma C}$ for $\begin{cases} \widetilde{p}^{n+1/2} = \alpha_{\Delta t/2} p^n + \sqrt{\dfrac{1-\alpha_{\Delta t}}{\beta}}\, M\, G^n, \\ p^{n+1/2} = \widetilde{p}^{n+1/2} - \dfrac{\Delta t}{2}\nabla V(q^n), \\ q^{n+1} = q^n + \Delta t\, M^{-1} p^{n+1/2}, \\ \widetilde{p}^{n+1} = p^{n+1/2} - \dfrac{\Delta t}{2}\nabla V(q^{n+1}), \\ p^{n+1} = \alpha_{\Delta t/2}\widetilde{p}^{n+1} + \sqrt{\dfrac{1-\alpha_{\Delta t}}{\beta}}\, M\, G^{n+1/2} \end{cases}$

Error estimates (systematic bias)

- Invariant measure $\mu_{\gamma,\Delta t}$ different from μ

Error estimates on the bias

$$\int_{\mathcal{E}} \varphi \, d\mu_{\gamma,\Delta t} = \int_{\mathcal{E}} \varphi \Big(1 + \Delta t^p f_\gamma\Big) d\mu + \mathrm{O}(\Delta t^{p+1})$$

with

- $p = 1$ for Trotter splitting
- $p = 2$ for Strang splitting and GLA

- **Various asymptotic results** in the Hamiltonian limit $\gamma \to 0$ or in the overdamped limit $\gamma \to +\infty$

$\to$ e.g. understanding why the BAOAB scheme is much better than other discretizations of Brownian dynamics

B. Leimkuhler, Ch. Matthews and G. Stoltz, The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics, *IMA J. Numer. Anal.* **36**(1), 13-79 (2016)

Systematic errors for transport coefficients

Transport coefficients

- "Steady-state" dynamical information
  - mobility
  - shear viscosity
  - thermal conductivity

- **Bias** due to timestep for
  - Green-Kubo formulas
  - linear response approaches

Linear response and Green–Kubo formula

$$\lim_{\eta \to 0} \frac{\mathbb{E}_\eta(R)}{\eta} = \int_0^{+\infty} \mathbb{E}_0(R(x_t)S(x_0))dt$$

A paradigmatic example

Langevin dynamics perturbed by a constant force term

$$\begin{cases} dq_t = M^{-1} p_t \, dt, \\ dp_t = (-\nabla V(q_t) + \eta F) \, dt - \gamma M^{-1} p_t \, dt + \sqrt{\dfrac{2\gamma}{\beta}} \, dW_t, \end{cases} \tag{1}$$

where

- $F \in \mathbb{R}^d$ with $|F| = 1$ is a given direction
- $\eta \in \mathbb{R}$ determines the strength of the external forcing

- Non-zero velocity in the direction F is expected in the steady-state

- F **does not derive from the gradient of a periodic function**
  - of course, $F = -\nabla W_F(q)$ with $W_F(q) = -F^T q$
  - ...but W_F is not periodic!

Mobility and self-diffusion

- **Linear response approaches:** averages in a nonequilibrium steady-state

Mobility

$$\nu_F = \lim_{\eta \to 0} \frac{\mathbb{E}_\eta \left(F^T M^{-1} p \right)}{\eta} = \beta \int_{\mathcal{E}} F^T M^{-1} p \, f_{0,1}(q,p) \, \mu(dq\, dp) = \beta \, F^T D F$$

- **Green-Kubo formulas:** integrated correlation functions

Effective diffusion at equilibrium ($\eta = 0$)

Unperiodized displacement $Q_t - Q_0 = \displaystyle\int_0^t M^{-1} p_s \, ds$

$$F^T D F = \int_0^{+\infty} \mathbb{E}_0 \left[\left(F^T M^{-1} p_t \right) \left(F^T M^{-1} p_0 \right) \right] dt$$

Error estimates on linear response

- Numerical schemes as for equilibrium dynamics with force $-\nabla V(q) + \eta F$

Error estimates for nonequilibrium dynamics

There exists a function $f_{\alpha,1,\gamma} \in H^1(\mu)$ such that

$$\int_{\mathcal{E}} \psi \, d\mu_{\gamma,\eta,\Delta t} = \int_{\mathcal{E}} \psi \Big(1 + \eta f_{0,1,\gamma} + \Delta t^\alpha f_{\alpha,0,\gamma} + \eta \Delta t^\alpha f_{\alpha,1,\gamma} \Big) d\mu + r_{\psi,\gamma,\eta,\Delta t},$$

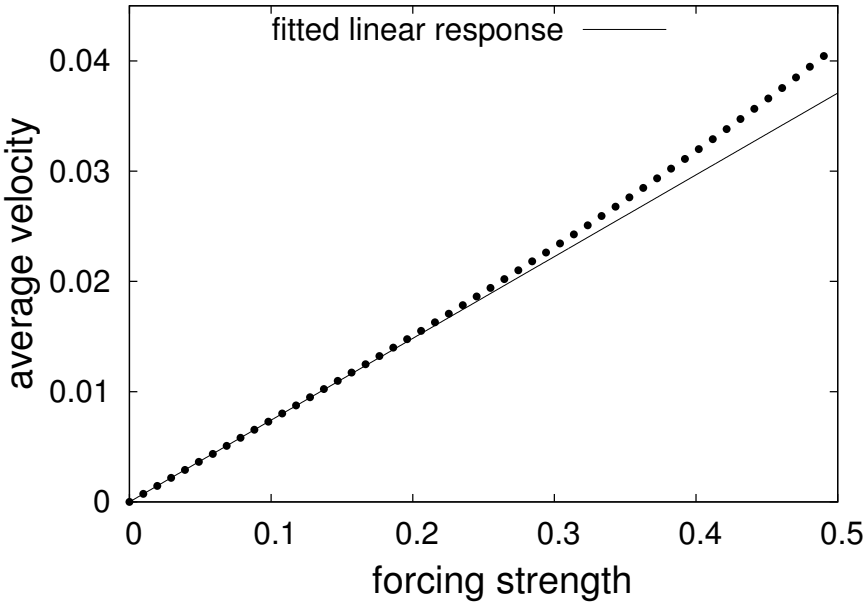
where the remainder is compatible with linear response

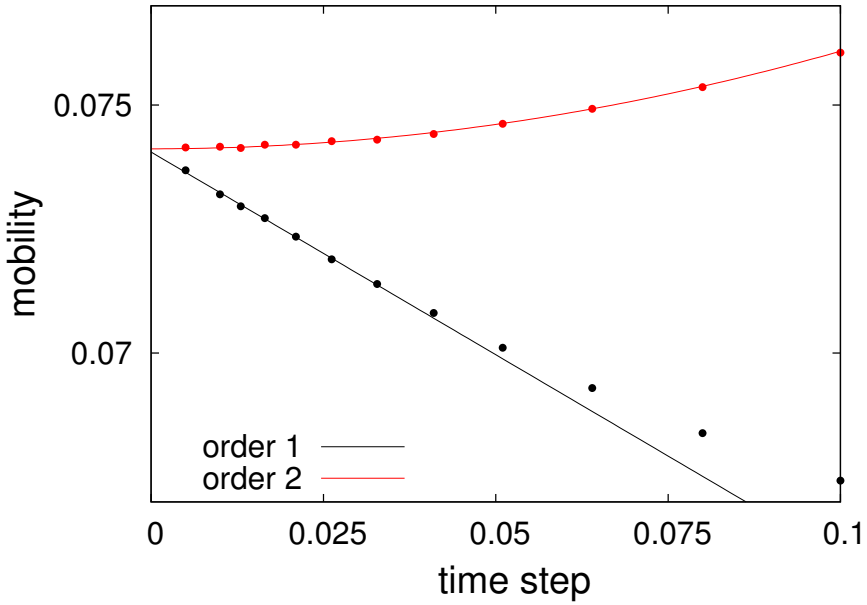
- Corollary: error estimates on the numerically computed mobility

$$\begin{aligned} \nu_{F,\gamma,\Delta t} &= \lim_{\eta \to 0} \frac{1}{\eta} \left(\int_{\mathcal{E}} F^T M^{-1} p \, \mu_{\gamma,\eta,\Delta t}(dq\, dp) - \int_{\mathcal{E}} F^T M^{-1} p \, \mu_{\gamma,0,\Delta t}(dq\, dp) \right) \\ &= \nu_{F,\gamma} + \Delta t^\alpha \int_{\mathcal{E}} F^T M^{-1} p \, f_{\alpha,1,\gamma} \, d\mu + \Delta t^{\alpha+1} r_{\gamma,\Delta t} \end{aligned}$$

- Results in the overdamped limit

Numerical results





Left: Linear response of the average velocity as a function of η for the scheme associated with $P_{\Delta t}^{\gamma C, B_\eta, A, B_\eta, \gamma C}$ and $\Delta t = 0.01, \gamma = 1$.

Right: Scaling of the mobility $\nu_{F,\gamma,\Delta t}$ for the first order scheme $P_{\Delta t}^{A, B_\eta, \gamma C}$ and the second order scheme $P_{\Delta t}^{\gamma C, B_\eta, A, B_\eta, \gamma C}$.

Error estimates on Green-Kubo formulas

Error estimates on integrated correlation functions

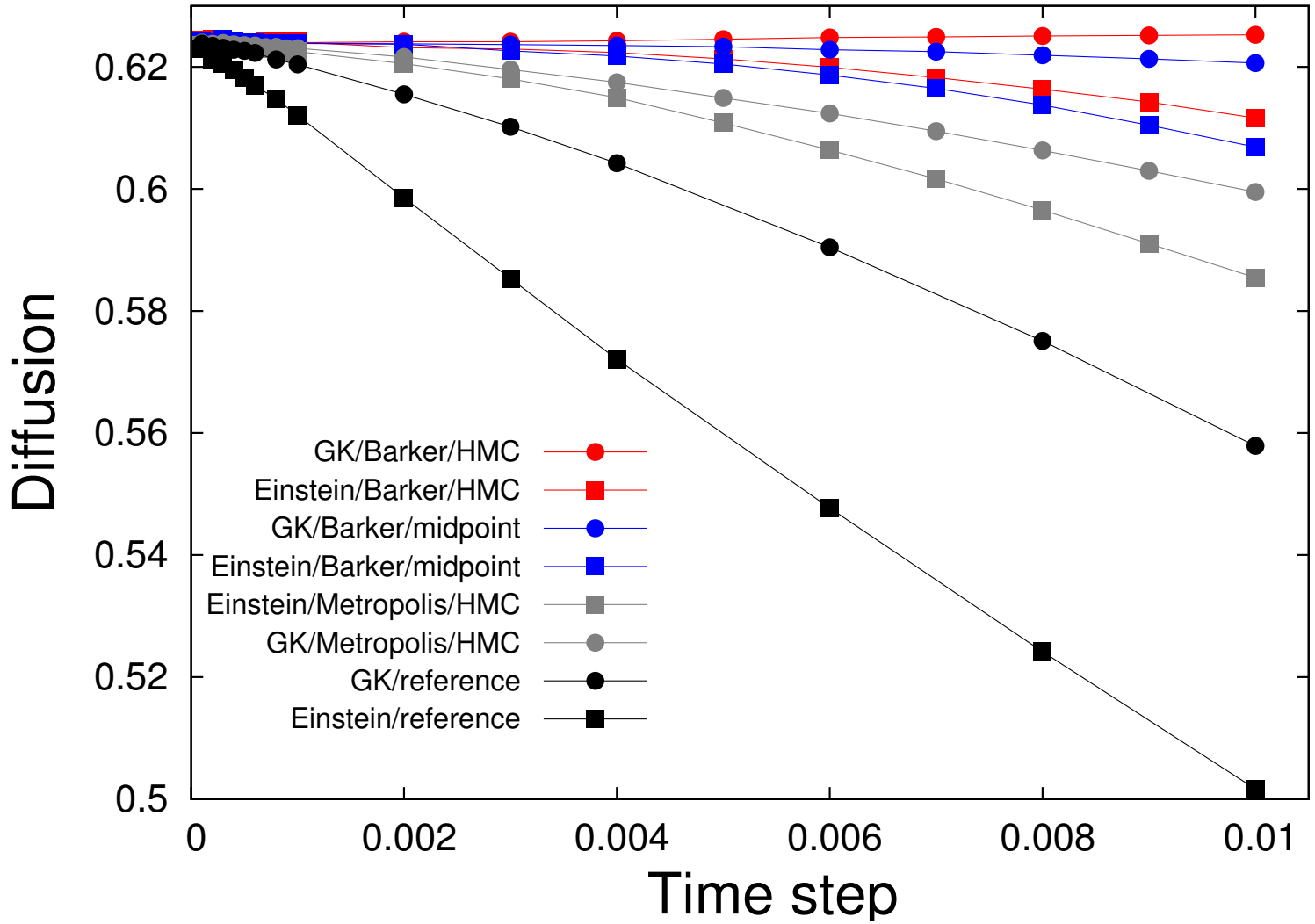
Observables φ, ψ with average 0 w.r.t. invariant measure μ

$$\int_0^{+\infty} \mathbb{E}\Big(\psi(q_t, p_t)\varphi(q_0, p_0)\Big)\, dt$$
$$= \Delta t \sum_{n=0}^{+\infty} \mathbb{E}_{\Delta t}\left(\widetilde{\psi}_{\Delta t,\alpha}\left(q^n, p^n\right)\varphi\left(q^0, p^0\right)\right) + \Delta t^{\alpha} r_{\Delta t}^{\psi,\varphi},$$

where $\widetilde{\psi}_{\Delta t,\alpha}$ is a corrected observable

- Useful when the corrected observable can be computed, e.g. schemes of weak order k
- Reduces to trapezoidal rule for second order schemes

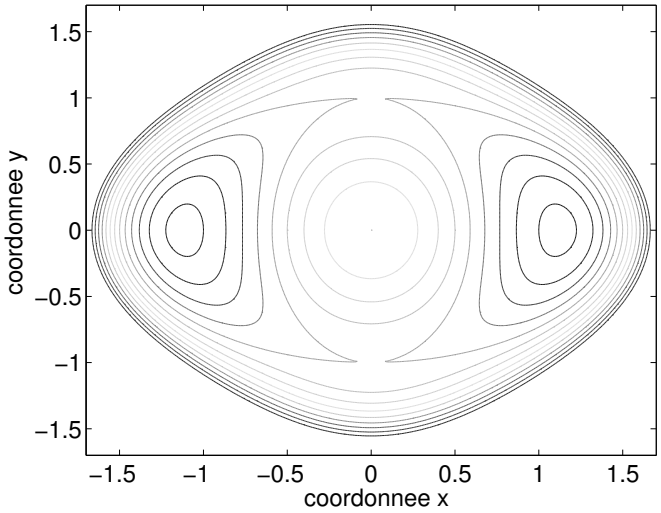
1D case, $\varphi = \psi = V'$, cosine potential, overdamped

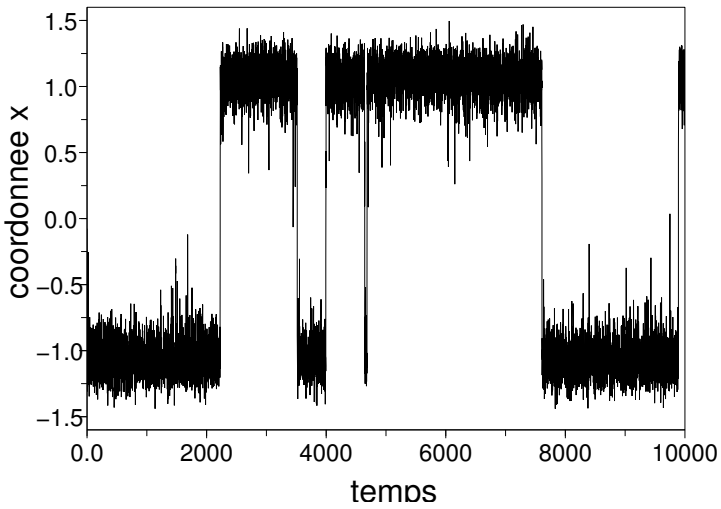


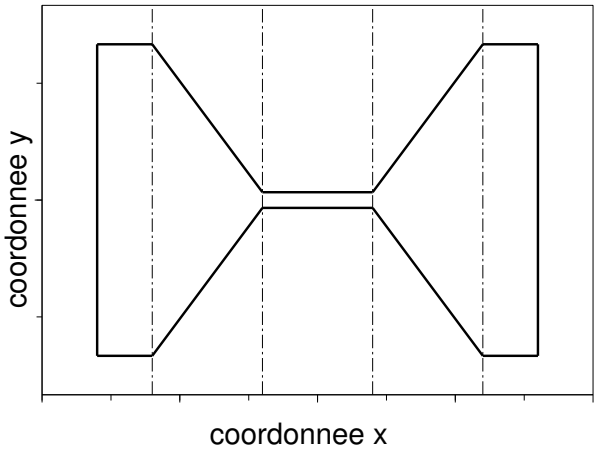
M. Fathi and G. Stoltz, *Numer. Math.* **136**(2), 545-602 (2017)

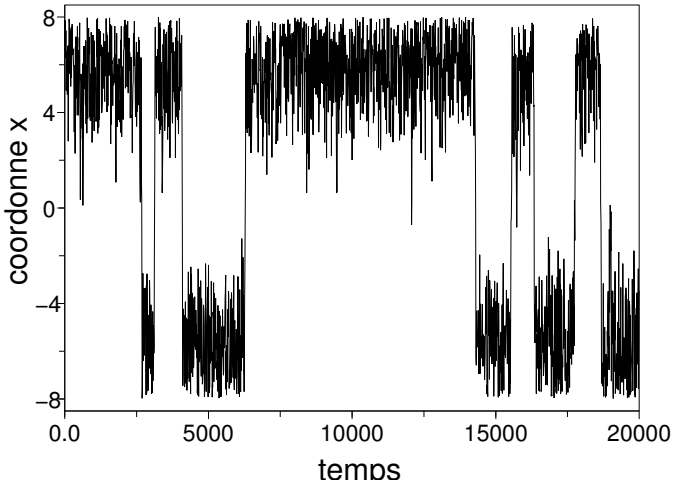
Reducing the statistical error

The actual mathematical challenge: metastability









Energetic vs. entropic barriers $\rightarrow$ Variance reduction techniques, e.g. modifying $-\nabla V$

Variance reduction techniques for equilibrium systems

The statistical error often dominates in simulations due to metastability issues

- **Stratification** (thermodynamic integration, MBAR/WHAM)
- **Importance sampling**
  - biasing by the free energy
  - adding non reversible perturbations preserving the invariant measure
- Other general techniques, not so used in molecular dynamics:
  - antithetic variables
  - coupling
  - control variates

Control variate approach for nonequilibrium systems

- Standard approach: compute steady-state average of response function as

$$\frac{1}{t}\int_0^t R(q_s,p_s)\,ds \xrightarrow[t\to+\infty]{} \int_{\mathcal{E}} R\,\psi_\eta = \mathrm{O}(\eta)$$

when R vanishes at equilibrium

- Variance of order 1 $\to$ relative error of order $1/\eta$!

- Need for variance reduction... but
  - no straightforward importance sampling
  - no easy stratification

Control variate approach

$$\frac{\mathbb{E}_\eta(R)}{\eta} = \frac{\mathbb{E}_\eta(R-\mathcal{L}_\eta\Phi)}{\eta} \quad \text{with} \quad \mathrm{Var}_\eta(R-\mathcal{L}_\eta\Phi) \ll \mathrm{Var}_\eta(R)$$

J. Roussel and G. Stoltz, A perturbative approach to control variates in molecular dynamics, *Multiscale Model. Simul.* (2019)