Error estimates and variance reduction for transport coefficients

Gabriel STOLTZ

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

Research funded by ANR SINEQ and ERC Synergy EMC2

SCALES conference 2025, Mainz

Outline

Molecular dynamics

- Microscopic description of physical systems
- Macroscopic description: average properties

Practical computation of static properties

- Ergodic averages using Langevin dynamics
- Error estimates (statistical error, bias)

Estimation of transport coefficients

- Linear response of nonequilibrium dynamics
- Error estimates
- Variance reduction

Molecular dynamics

General perspective (1)

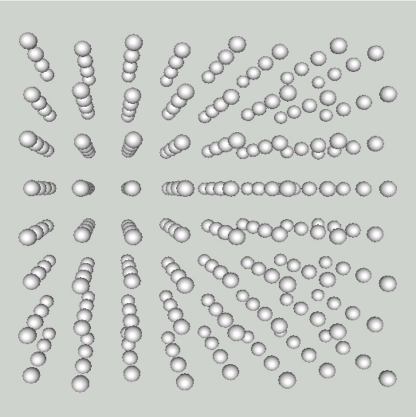
- **Aims of computational statistical physics:**
  - numerical microscope
  - computation of average properties, static or dynamic

- **Orders of magnitude**
  - distances $\sim 1\ \AA = 10^{-10}$ m
  - energy per particle $\sim k_{\rm B}T \sim 4 \times 10^{-21}$ J at room temperature
  - atomic masses $\sim 10^{-26}$ kg
  - time $\sim 10^{-15}$ s
  - number of particles $\sim \mathcal{N}_A = 6.02 \times 10^{23}$

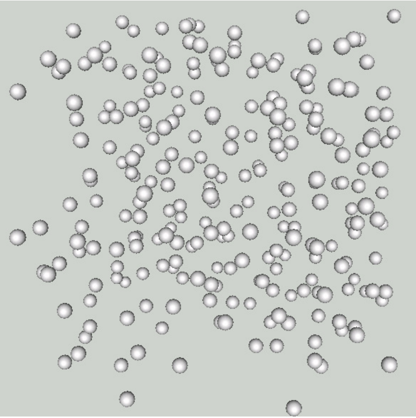
- **"Standard" simulations**
  - 10^6 particles ["world records": around 10^9 particles]
  - integration time: (fraction of) ns ["world records": (fraction of) μs]

General perspective (2)

What is the melting temperature of Argon?



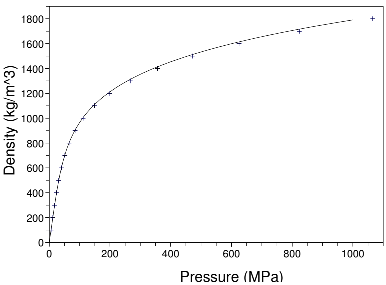
(a) Solid Argon (low temperature)

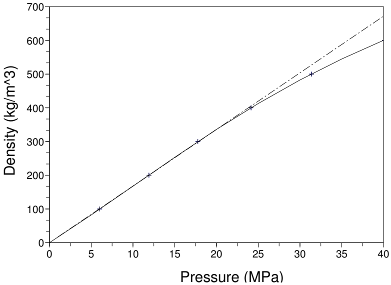


(b) Liquid Argon (high temperature)

General perspective (3)

"Given the structure and the laws of interaction of the particles, what are the macroscopic properties of the matter composed of these particles?"

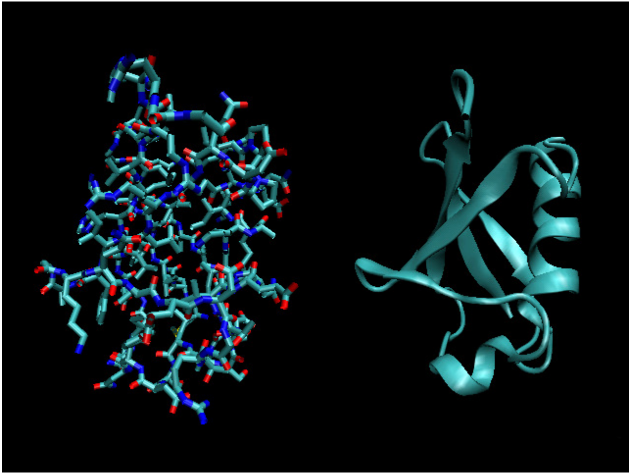




Equation of state (pressure/density diagram) for Argon at $T = 300$ K

General perspective (4)

What is the structure of the protein? What are its typical conformations, and what are the transition pathways from one conformation to another?



Microscopic description of physical systems

• Microstate of a classical system of N particles:

$$(q,p) = (q_1,\ldots,q_N,\ p_1,\ldots,p_N) \in \mathcal{E}$$

Positions q (configuration), momenta p (to be thought of as $M\dot{q}$)

• Here, periodic boundary conditions: $\mathcal{E} = \mathcal{D} \times \mathbb{R}^{3N}$ with $\mathcal{D} = (L\mathbb{T})^{3N}$

• More complicated situations can be considered: molecular constraints defining submanifolds of the phase space

• Hamiltonian $H(q,p) = E_{\mathrm{kin}}(p) + V(q)$, where the kinetic energy is

$$E_{\mathrm{kin}}(p) = \frac{1}{2} p^\top M^{-1} p, \qquad M = \begin{pmatrix} m_1 \, \mathrm{Id}_3 & & 0 \\ & \ddots & \\ 0 & & m_N \, \mathrm{Id}_3 \end{pmatrix}.$$

Average properties

- Macrostate of the system described by a probability measure

Equilibrium thermodynamic properties (pressure,...)

$$\mathbb{E}_\mu(\varphi) = \int_{\mathcal{E}} \varphi(q,p)\,\mu(dq\,dp)$$

- Examples of observables:
  - Pressure $\varphi(q,p) = \frac{1}{3|\mathcal{D}|}\sum_{i=1}^N \left(\frac{p_i^2}{m_i} - q_i \cdot \nabla_{q_i} V(q)\right)$
  - Kinetic temperature $\varphi(q,p) = \frac{1}{3Nk_\mathrm{B}}\sum_{i=1}^N \frac{p_i^2}{m_i}$
- Canonical ensemble = measure on (q,p) (average energy fixed)

$$\mu_\mathrm{NVT}(dq\,dp) = Z_\mathrm{NVT}^{-1}\,\mathrm{e}^{-\beta H(q,p)}\,dq\,dp, \qquad \beta = \frac{1}{k_\mathrm{B}T}$$

Practical computation of average properties

Computing average properties

Main issue

Computation of high-dimensional integrals... Ergodic averages

$$\mathbb{E}_\mu(\varphi) = \lim_{t \to +\infty} \widehat{\varphi}_t, \qquad \widehat{\varphi}_t = \frac{1}{t}\int_0^t \varphi(q_s, p_s)\,ds$$

• One possible choice: Langevin dynamics with friction parameter $\gamma > 0$
= Stochastic perturbation of the Hamiltonian dynamics

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

Almost-sure convergence of ergodic averages[1]

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

Hamiltonian and overdamped limits

- As $\gamma \to 0$, the Hamiltonian dynamics is recovered

$$\frac{d}{dt}\mathbb{E}\left[H(q_t,p_t)\right] = -\gamma\left(\mathbb{E}\left[p_t^\top M^{-2} p_t\right] - \frac{1}{\beta}\mathrm{Tr}(M^{-1})\right)dt$$

Time $\sim \gamma^{-1}$ to change energy levels in this limit[2]

- Overdamped limit $\gamma \to +\infty$ with $M = \mathrm{Id}$: rescaling of time γt

$$\begin{aligned} q_{\gamma t} - q_0 &= -\frac{1}{\gamma}\int_0^{\gamma t} \nabla V(q_s)\,ds + \sqrt{\frac{2}{\gamma\beta}} W_{\gamma t} - \frac{1}{\gamma}\left(p_{\gamma t} - p_0\right) \\ &= -\int_0^t \nabla V(q_{\gamma s})\,ds + \sqrt{2\beta^{-1}} B_t - \frac{1}{\gamma}\left(p_{\gamma t} - p_0\right) \end{aligned}$$

which converges to the solution of $dQ_t = -\nabla V(Q_t)\,dt + \sqrt{2\beta^{-1}}\,dB_t$

- In both cases, slow convergence, with rate scaling as $\min\left(\gamma, \gamma^{-1}\right)$

[2] Hairer and Pavliotis, *J. Stat. Phys.*, **131**(1), 175-202 (2008)

Statistical error

- **Evolution semigroup** $\left(\mathrm{e}^{t\mathcal{L}}\varphi\right)(q,p) = \mathbb{E}\left[\varphi(q_t,p_t)\,\middle|\,(q_0,p_0)=(q,p)\right]$,

generator $\mathcal{L} = \mathcal{L}_{\mathrm{ham}} + \gamma \mathcal{L}_{\mathrm{FD}}$

$$\mathcal{L}_{\mathrm{ham}} = p^\top M^{-1}\nabla_q - \nabla V^\top \nabla_p, \qquad \mathcal{L}_{\mathrm{FD}} = -p^\top M^{-1}\nabla_p + \frac{1}{\beta}\Delta_p$$

- **Asymptotic variance** $\sigma_\varphi^2 = \lim_{t\to+\infty} t\,\mathrm{Var}_\mu(\widehat{\varphi}_t)$: with $\Pi\varphi = \varphi - \int_{\mathcal{E}} \varphi\, d\mu$,

$$\begin{aligned}\sigma_\varphi^2 &= \lim_{t\to+\infty} \int_0^t \left(1-\frac{s}{t}\right)\mathbb{E}_\mu\left[\Pi\varphi(q_t,p_t)\Pi\varphi(q_0,p_0)\right] ds \\ &= 2\int_0^{+\infty}\int_{\mathcal{E}} (\mathrm{e}^{s\mathcal{L}}\Pi\varphi)\Pi\varphi\, d\mu\, ds = 2\int_{\mathcal{E}} (-\mathcal{L}^{-1}\Pi\varphi)\Pi\varphi\, d\mu\end{aligned}$$

Well-defined provided $-\mathcal{L}\Phi = \Pi\varphi$ has a solution in $L_0^2(\mu) = \Pi L^2(\mu)$

A Central Limit Theorem holds in this case[3]: $\boxed{\widehat{\varphi}_t - \mathbb{E}_\mu(\varphi) \simeq \frac{\sigma_\varphi}{\sqrt{t}}\mathcal{G}}$

[3] R. N. Bhattacharya, *Z. Wahrsch. Verw. Gebiete* (1982)

Practical computation of average properties

- Numerical scheme = Markov chain characterized by evolution operator

$$P_{\Delta t}\varphi(q,p) = \mathbb{E}\Big(\varphi\left(q^{n+1},p^{n+1}\right)\,\Big|(q^n,p^n)=(q,p)\Big)$$

- Discretization of the Langevin dynamics: splitting strategy[4]

$$A = M^{-1}p\cdot\nabla_q, \qquad B = -\nabla V(q)\cdot\nabla_p, \qquad C = -M^{-1}p\cdot\nabla_p + \frac{1}{\beta}\Delta_p$$

- Example: $P_{\Delta t}^{B,A,\gamma C}$ corresponds to (with $\alpha_{\Delta t} = \exp(-\gamma M^{-1}\Delta t)$)

$$\begin{cases} \widetilde{p}^{n+1} = p^n - \Delta t\,\nabla V(q^n), \\ q^{n+1} = q^n + \Delta t\, M^{-1}\widetilde{p}^{n+1}, \\ p^{n+1} = \alpha_{\Delta t}\widetilde{p}^{n+1} + \sqrt{(1-\alpha_{\Delta t}^2)\beta^{-1}M}\, G^n, \end{cases}$$

where G^n are i.i.d. standard Gaussian random variables

- Second order splittings schemes (Verlet + FD or BAOAB)

[4] Leimkuhler/Matthews (2013, 2016); Leimkuhler/Matthews/Stoltz (2016)

Types of errors

Estimators of $\mathbb{E}_\mu(\varphi)$

$$\widehat{\varphi}_t = \frac{1}{t}\int_0^t \varphi(q_s,p_s)\,ds, \qquad \widehat{\varphi}_{\Delta t}^N = \frac{1}{N}\sum_{n=1}^N \varphi(q^n,p^n)$$

Statistical error (variance of the estimator)

- $\mathrm{O}\left(\dfrac{\sigma_\varphi}{\sqrt{N\Delta t}}\right)$ from central limit theorem for continuous dynamics
- discrete dynamics: asymptotic variance coincides at order Δt^α

Bias (expectation of the estimator)

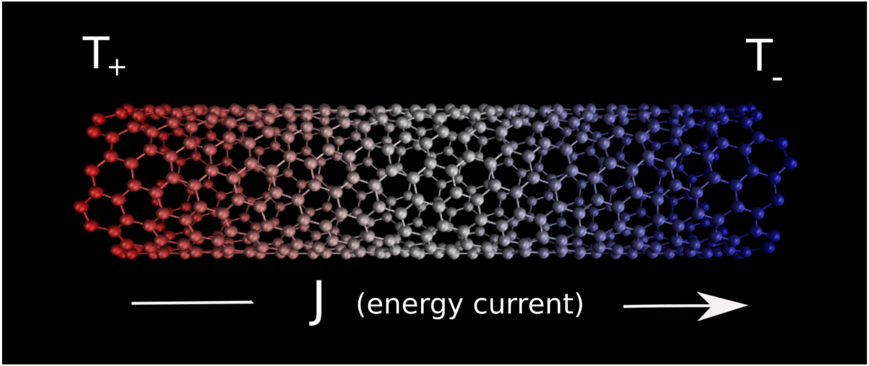
- finite time integration time $\to$ bias $\mathrm{O}\left(\dfrac{1}{N\Delta t}\right)$
- discretization of the dynamics $\to$ bias $\mathrm{O}(\Delta t^\alpha)$

Estimation of transport coefficients

Physical context and motivations

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

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



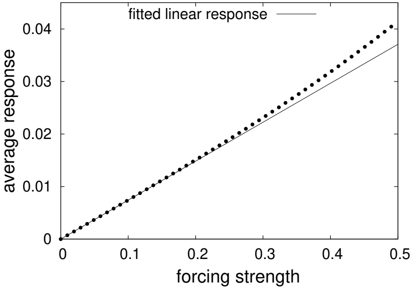
Slow convergence due to large noise to signal ratio

Long computational times to estimate κ (up to several weeks/months)

Example: $\mathcal{D} = (L\mathbb{T})^d$, non-gradient force $F \in \mathbb{R}^d$

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

Response function $R(q,p) = F^\top M^{-1}p =$ velocity in direction F



Existence/uniqueness of invariant probability measure (Lyapunov)

Generator $\mathcal{L} + \eta\widetilde{\mathcal{L}}$ with $\widetilde{\mathcal{L}} = F^\top \nabla_p$

$$\mathbb{E}_\eta(R) = \int_{\mathcal{E}} R\,\psi_\eta \approx \alpha\eta$$

$\alpha =$ transport coefficient

Definition of transport coefficients (1)

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

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

* = adjoints on $L^2(\mu)$ ($\partial_{q_i}^* = -\partial_{q_i} + \beta \partial_{q_i} V$ and $\partial_{p_i}^* = -\partial_{p_i} + \beta (M^{-1}p)_i$)

Fokker–Planck equation

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

By identifying powers of η (recalling $\Pi \varphi = \varphi - \mu(\varphi)$)

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

Running example: $\mathcal{L}^* = -\mathcal{L}_{\mathrm{ham}} + \gamma \mathcal{L}_{\mathrm{FD}}$ and $\widetilde{\mathcal{L}}^* = -\widetilde{\mathcal{L}} + \beta F^\top M^{-1} p$

$$S(q,p) = \beta F^\top M^{-1} p$$

Definition of transport coefficients (2)

Response property $R \in L_0^2(\mu) = \Pi L^2(\mu)$, conjugated response $S = \widetilde{\mathcal{L}}^* \mathbf{1}$:

$$\alpha = \lim_{\eta \to 0} \frac{\mathbb{E}_\eta(R)}{\eta} = \int_{\mathcal{E}} R \mathfrak{f}_1 \, d\mu = \int_{\mathcal{E}} R \left[(-\mathcal{L}^*)^{-1} S \right] d\mu = \int_{\mathcal{E}} \left(-\mathcal{L}^{-1} R \right) S \, d\mu$$

$$= \int_0^{+\infty} \left[\int_{\mathcal{E}} \left(\mathrm{e}^{t\mathcal{L}} R \right) S \, d\mu \right] dt = \int_0^{+\infty} \mathbb{E}_0 \Big(R(q_t, p_t) S(q_0, p_0) \Big) dt$$

In practice:

- Identify the response function and the reference dynamics
- Construct a physically meaningful perturbation (bulk or boundary driven)
- Obtain the transport coefficient α (thermal cond., shear viscosity,...)

For the running example, definition of mobility with $R(q,p) = F^\top M^{-1} p$

$$\lim_{\eta \to 0} \frac{\mathbb{E}_\eta \left(F^\top M^{-1} p \right)}{\eta} = \beta \, F^\top D F, \quad D = \int_0^{+\infty} \mathbb{E}_0 \Big((M^{-1} p_t) \otimes (M^{-1} p_0) \Big) \, dt$$

Error estimates for nonequilibrium molecular dynamics

Example: $\mathcal{D} = (L\mathbb{T})^d$, non-gradient force $F \in \mathbb{R}^{3N}$

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

Estimator of linear response (observable R with equilibrium average 0)

$$\widehat{A}_{\eta,t} = \frac{1}{\eta t} \int_0^t R(q_s^\eta, p_s^\eta) \, ds \xrightarrow[t \to +\infty]{\text{a.s.}} \alpha_\eta := \frac{1}{\eta} \int_{\mathcal{E}} R \, f_\eta \, d\mu = \alpha + \mathrm{O}(\eta)$$

Issues with linear response methods:

- Statistical error with asymptotic variance $\mathrm{O}(\eta^{-2} t^{-1})$
- Bias $\mathrm{O}(\eta)$ due to $\eta \neq 0$
- Bias from finite integration time
- Timestep discretization bias[5]

[5] Leimkuhler/Matthews/Stoltz (2016)

Error estimates on the Green–Kubo formula

- Aim: approximate $\alpha = \displaystyle\int_0^{+\infty} \mathbb{E}_0\Big(R(q_t,p_t)S(q_0,p_0)\Big)dt$
- **Estimator** (continuous time) $\widehat{\alpha}_{K,T} = \dfrac{1}{K}\displaystyle\sum_{k=1}^{K}\int_0^T R(q_t,p_t)S(q_0,p_0)\,dt$
- **Sources of error** : trade off in truncation time T
  - Time truncation bias $\left|\mathbb{E}_\pi[\widehat{\alpha}_{K,T}]-\alpha\right| \leqslant \dfrac{C}{\kappa}\|R\|_{L^2(\pi)}\|S\|_{L^2(\pi)}\mathrm{e}^{-\kappa T}$
  - Statistical error[6] $\mathrm{Var}[\widehat{\alpha}_{K,T}] \sim \dfrac{2T}{K}\|S\|^2_{L^2(\pi)}\left\langle R, -\mathcal{L}^{-1}R\right\rangle$
  - Timestep bias and quadrature formula when numerically integrating[7]

[6] Pavliotis/Spacek/Stoltz/Vaes (2025); Gastaldello/Stoltz/Vaes (2025)
[7] Leimkuhler/Matthews/Stoltz (2016)

Study of alternative approaches: several year workplan!

Alternatives to direct NEMD/GK, possibly with some blending

- Alternative fluctuation formulas[8]
- Control variate approaches[9] (better solutions to Poisson equation needed...)
- Use coupling methods between X_t^η and X_t^0, *e.g.* sticky coupling[10]
- Rely on tangent dynamics[11] for $T_t = \lim_{\eta \to 0}(X_t^\eta - X_t^0)/\eta$
- Optimize synthetic forcings[12]
- Large deviation techniques to estimate second order cumulants[13]
- Norton dynamics[14] (dual approach where the flux is fixed)
- Transient methods: subtraction[15] and TTCF[16]

Quantify variance and bias and apply to physical systems

[8] Plechac/Stoltz/Wang (2021, 2023)

[9] Mangaud/Rotenberg (2020); Roussel/Stoltz (2019), Pavliotis/Stoltz/Vaes (2022), Pavliotis/Spacek/Stoltz/Vaes (2024)

[10] Eberle/Zimmer (2019); Durmus/Eberle/Enfroy/Guillin/Monmarché (2021); Darshan/Eberle/Stoltz (2024)

[11] Assaraf/Jourdain/Lelièvre/Roux, *Stoch. Partial Differ. Equ. Anal. Comput.* (2018)

[12] Evans/Morriss (2008); Spacek/Stoltz (2023)

[13] Limmer/Gao/Poggioli (2021); currently Guyader/Gastaldello/Stoltz/Vaes

[14] Evans/Morriss (2008); Blassel/Stoltz (2023) and now Darshan/Stoltz

[15] Ciccotti/Jacucci (1975); Monmarché/Spacek/Stoltz (2025)

[16] Evans/Morriss (1987, 1988)