

Error estimates and variance reduction for nonequilibrium stochastic dynamics

Gabriel STOLTZ

(CERMICS, UMR 9032, Ecole des Ponts & MATHERIALS team, Inria Paris)

Project funded by ANR SINEQ & ERC EMC2

Evry, February 2026

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**

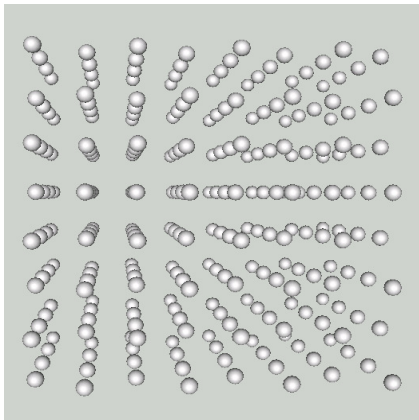
- distances $\sim 1 \text{ \AA} = 10^{-10} \text{ m}$
- energy per particle $\sim k_B T \sim 4 \times 10^{-21} \text{ J}$ at room temperature
- atomic masses $\sim 10^{-26} \text{ kg}$
- time $\sim 10^{-15} \text{ s}$
- number of particles $\sim \mathcal{N}_A = 6.02 \times 10^{23}$

- **“Standard” simulations**

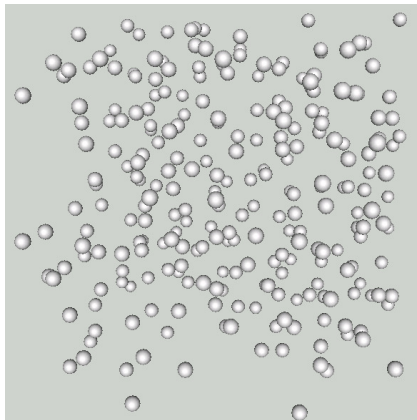
- 10^6 particles [“world records”: around 10^9 particles]
- integration time: (fraction of) ns [“world records”: (fraction of) μs]

General perspective (2)

What is the **melting temperature** of Argon?



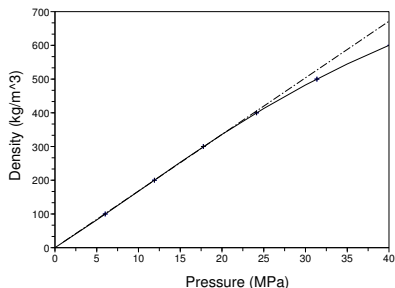
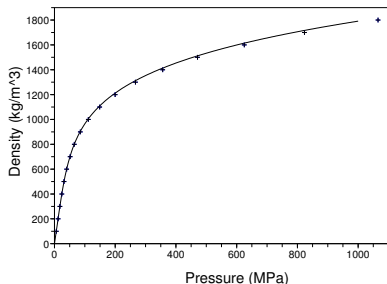
(a) Solid Argon (low temperature)



(b) Liquid Argon (high temperature)

General perspective (3)

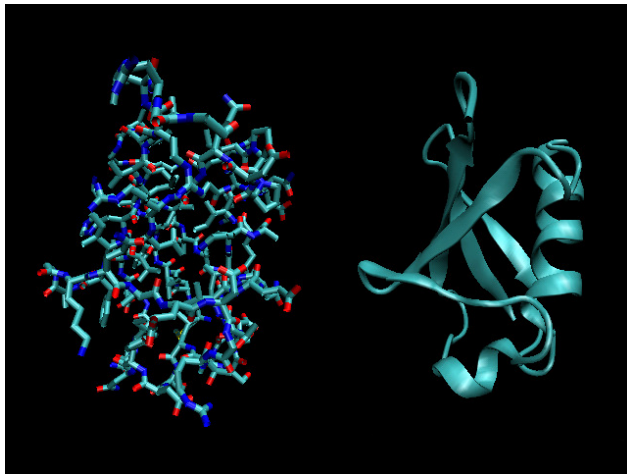
“Given the structure and the laws of interaction of the particles, what are the **macroscopic properties** of the matter composed of these particles?”



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:

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

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

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

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 $\varphi(q, p) = \frac{1}{3Nk_B} \sum_{i=1}^N \frac{p_i^2}{m_i}$

- **Canonical** ensemble = measure on (q, p) (average energy fixed)

$$\mu_{\text{NVT}}(dq dp) = Z_{\text{NVT}}^{-1} e^{-\beta H(q, p)} dq dp, \quad \beta = \frac{1}{k_B T}$$

Practical computation of average properties

Computing average properties

Main issue

Computation of **high-dimensional** integrals... **Ergodic** averages

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

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

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

Almost-sure convergence of ergodic averages¹

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

Hamiltonian and overdamped limits

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

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

Time $\sim \gamma^{-1}$ to change energy levels in this limit²

- Overdamped** limit $\gamma \rightarrow +\infty$ with $M = \text{Id}$: rescaling of time γt

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

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 $\min(\gamma, \gamma^{-1})$

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

Statistical error (1/2)

- **Evolution semigroup** $(e^{t\mathcal{L}}\varphi)(q, p) = \mathbb{E} \left[\varphi(q_t, p_t) \mid (q_0, p_0) = (q, p) \right]$,
generator $\mathcal{L} = \mathcal{L}_{\text{ham}} + \gamma \mathcal{L}_{\text{FD}}$

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

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

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

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 case³: $\widehat{\varphi}_t - \mathbb{E}_\mu(\varphi) \simeq \frac{\sigma_\varphi}{\sqrt{t}} \mathcal{G}$

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

Statistical error (2/2)

Prove **exponential convergence** of the semigroup $e^{t\mathcal{L}}$ on $E \subset L_0^2(\mu)$

- Lyapunov techniques⁴ $L_{\mathcal{X}}^\infty(\mathcal{E}) = \left\{ \varphi \text{ measurable, } \sup \left| \frac{\varphi}{\mathcal{X}} \right| < +\infty \right\}$
- “historic” hypocoercive⁵ setup $H^1(\mu)$
- $L^2(\mu)$ after hypoelliptic regularization⁶ from $H^1(\mu)$
- direct transfer from $H^1(\mu)$ to $L^2(\mu)$ by spectral argument⁷
- directly⁸ $L^2(\mu)$ (recently⁹ Poincaré using $\partial_t - \mathcal{L}_{\text{ham}}$)
- coupling arguments¹⁰
- direct estimates on the resolvent using Schur complements¹¹

Rate of convergence $\min(\gamma, \gamma^{-1})$ so **variance** $\sim \max(\gamma, \gamma^{-1})$

⁴Wu ('01); Mattingly/Stuart/Higham ('02); Rey-Bellet ('06); Hairer/Mattingly ('11)

⁵Villani (2009) and before Talay (2002), Eckmann/Hairer (2003), Hérau/Nier (2004),...

⁶Hérau, *J. Funct. Anal.* (2007)

⁷Deligiannidis/Paulin/Doucet, *Ann. Appl. Probab.* (2020)

⁸Hérau (2006), Dolbeaut/Mouhot/Schmeiser (2009, 2015)

⁹Albritton/Armstrong/Mourrat/Novack (2019), Cao/Lu/Wang (2019), Brigatti (2021), Dietert/Hérau/Hutridurga/Mouhot (2022), Brigati/Stoltz (2023)

¹⁰Eberle/Guillin/Zimmer, *Ann. Probab.* (2019)

¹¹Bernard/Fathi/Levitt/Stoltz, *Annales Henri Lebesgue* (2022)

Practical computation of average properties

- Numerical scheme = **Markov chain** characterized by evolution operator

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

- Discretization of the Langevin dynamics: **splitting** strategy¹²

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

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

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

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

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

¹²Leimkuhler/Matthews (2013, 2016); Leimkuhler/Matthews/Stoltz (2016)

Types of errors

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

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

Statistical error (variance of the estimator)

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

Bias (expectation of the estimator)

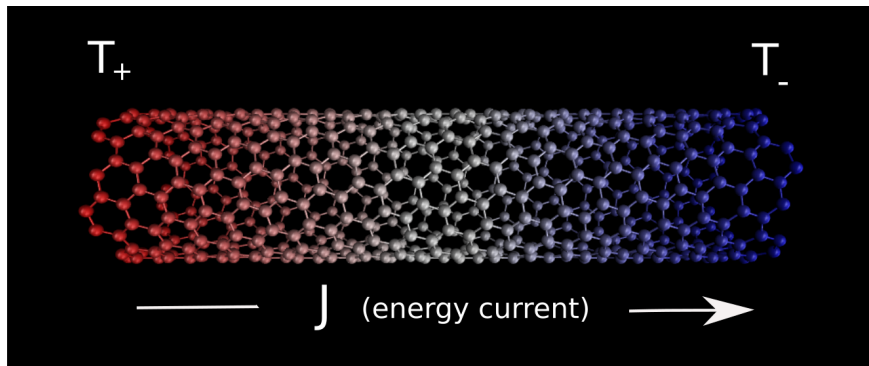
- **finite time** integration time $\rightarrow$ bias $O\left(\frac{1}{N\Delta t}\right)$
- **discretization** of the dynamics $\rightarrow$ bias $O(\Delta t^\alpha)$

Estimation of transport coefficients

Physical context and motivations

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

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



Slow convergence due to **large noise to signal ratio**

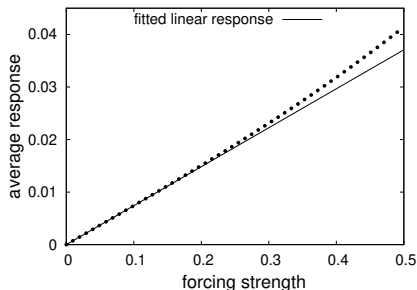
Long computational times to estimate κ (up to several weeks/months)

Linear response of nonequilibrium stochastic dynamics

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

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

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



Existence/uniqueness of invariant probability measure (Lyapunov)

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

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

α = **transport coefficient**

Definition of transport coefficients (1)

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

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

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

Fokker-Planck equation

$$(\mathcal{L} + \eta \tilde{\mathcal{L}})^* f_\eta = 0$$

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

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

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

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

Definition of transport coefficients (2)

Response property $R \in L^2_0(\mu) = \Pi L^2(\mu)$, conjugated response $S = \tilde{\mathcal{L}}^* \mathbf{1}$:

$$\begin{aligned}\alpha &= \lim_{\eta \rightarrow 0} \frac{\mathbb{E}_\eta(R)}{\eta} = \int_{\mathcal{E}} R \mathbf{f}_1 \, d\mu = \int_{\mathcal{E}} R \left[(-\mathcal{L}^*)^{-1} S \right] d\mu = \int_{\mathcal{E}} (-\mathcal{L}^{-1} R) S \, d\mu \\ &= \int_0^{+\infty} \left[\int_{\mathcal{E}} (e^{t\mathcal{L}} R) S \, d\mu \right] dt = \int_0^{+\infty} \mathbb{E}_0 \left(R(q_t, p_t) S(q_0, p_0) \right) dt\end{aligned}$$

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

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

Error estimates for nonequilibrium molecular dynamics

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

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

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

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

Issues with linear response methods:

- Statistical error with asymptotic variance $O(\eta^{-2}t^{-1})$
- Bias $O(\eta)$ due to $\eta \neq 0$
- Bias from finite integration time
- Timestep discretization bias¹³

¹³Leimkuhler/Matthews/Stoltz (2016)

Analysis of variance / finite integration time bias

Statistical error dictated by **Central Limit Theorem**:

$$\sqrt{t} \left(\hat{A}_{\eta,t} - \alpha_{\eta} \right) \xrightarrow[t \rightarrow +\infty]{\text{law}} \mathcal{N} \left(0, \frac{\sigma_{R,\eta}^2}{\eta^2} \right), \quad \sigma_{R,\eta}^2 = \sigma_{R,0}^2 + O(\eta)$$

so $\hat{A}_{\eta,t} = \alpha_{\eta} + O_P \left(\frac{1}{\eta\sqrt{t}} \right) \rightarrow$ requires **long simulation times** $t \sim \eta^{-2}$

Finite time integration bias: $\left| \mathbb{E} \left(\hat{A}_{\eta,t} \right) - \alpha_{\eta} \right| \leq \frac{K}{\eta t}$

Bias due to $t < +\infty$ is $O \left(\frac{1}{\eta t} \right) \rightarrow$ typically **smaller than statistical error**

Key equality for the proofs: introduce $-\left(\mathcal{L} + \eta \tilde{\mathcal{L}} \right) \mathcal{R}_{\eta} = R - \int_{\mathcal{E}} R f_{\eta} d\mu$

$$\hat{A}_{\eta,t} - \frac{1}{\eta} \int_{\mathcal{E}} R f_{\eta} d\mu = \frac{\mathcal{R}_{\eta}(q_0^{\eta}, p_0^{\eta}) - \mathcal{R}_{\eta}(q_t^{\eta}, p_t^{\eta})}{\eta t} + \frac{\sqrt{2\gamma}}{\eta t \sqrt{\beta}} \int_0^t \nabla_p \mathcal{R}_{\eta}(q_s^{\eta}, p_s^{\eta})^T dW_s$$

Error estimates on the Green–Kubo formula

- Aim: approximate $\alpha = \int_0^{+\infty} \mathbb{E}_0 \left(R(q_t, p_t) S(q_0, p_0) \right) dt$
- **Estimator** (continuous time) $\hat{\alpha}_{K,T} = \frac{1}{K} \sum_{k=1}^K \int_0^T R(q_t, p_t) S(q_0, p_0) dt$
- **Sources of error** : trade off in truncation time T
 - Time truncation bias $|\mathbb{E}_\pi[\hat{\alpha}_{K,T}] - \alpha| \leq \frac{C}{\kappa} \|R\|_{L^2(\pi)} \|S\|_{L^2(\pi)} e^{-\kappa T}$
 - Statistical error¹⁴ $\text{Var}[\hat{\alpha}_{K,T}] \sim \frac{2T}{K} \|S\|_{L^2(\pi)}^2 \langle R, -\mathcal{L}^{-1} R \rangle$
 - Timestep bias and quadrature formula when numerically integrating¹⁵

¹⁴Pavliotis/Spacek/Stoltz/Vaes (2025); Gastaldello/Stoltz/Vaes (2025)

¹⁵Leimkuhler/Matthews/Stoltz (2016)

Study of alternative approaches: several year workplan!

Alternatives to direct NEMD/GK, possibly with some **blending**

- Alternative fluctuation formulas¹⁶
- Control variate approaches¹⁷ (better solutions to Poisson equation needed...)
- Use coupling methods between X_t^η and X_t^0 , e.g. sticky coupling¹⁸
- Rely on tangent dynamics¹⁹ for $T_t = \lim_{\eta \rightarrow 0} (X_t^\eta - X_t^0)/\eta$
- Optimize synthetic forcings²⁰
- Large deviation techniques to estimate second order cumulants²¹
- Norton dynamics²² (dual approach where the flux is fixed)
- Transient methods: subtraction²³ and TTCF²⁴

Quantify variance and bias and apply to physical systems

¹⁶ Plechac/Stoltz/Wang (2021, 2023)

¹⁷ Mangaud/Rotenberg (2020); Roussel/Stoltz (2019), Pavliotis/Stoltz/Vaes (2022), Pavliotis/Spacek/Stoltz/Vaes (2024)

¹⁸ Eberle/Zimmer (2019); Durmus/Eberle/Enfroy/Guillin/Monmarché (2021); Darshan/Eberle/Stoltz (2024)

¹⁹ Assaraf/Jourdain/Lelièvre/Roux, *Stoch. Partial Differ. Equ. Anal. Comput.* (2018)

²⁰ Evans/Morriss (2008); Spacek/Stoltz (2023)

²¹ Limmer/Gao/Poggioli (2021); currently Guyader/Gastaldello/Stoltz/Vaes

²² Evans/Morriss (2008); Blassel/Stoltz (2023) and now Darshan/Stoltz

²³ Ciccotti/Jacucci (1975); Monmarché/Spacek/Stoltz (2025)

²⁴ Evans/Morriss (1987, 1988)