



European Research Council
Established by the European Commission



Error estimates on the computation of transport coefficients

Gabriel STOLTZ

`gabriel.stoltz@enpc.fr`

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

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

NESP 2015, ICTS (Bangalore), November 2015

Motivation

- **Transport coefficients** can be obtained from
 - nonequilibrium dynamics in the **linear response** regime
 - integrated **correlation functions** at equilibrium (Green-Kubo)
 - transient response to displacement from equilibrium

For concreteness: mobility/self-diffusion (direction $F \in \mathbb{R}^d$)

$$F^T DF = \lim_{t \rightarrow +\infty} \frac{\mathbb{E} \left[\left(F \cdot (Q_t - Q_0) \right)^2 \right]}{2t}$$

What is the numerical error arising from time discretization?

B. Leimkuhler, Ch. Matthews and G. Stoltz, The computation of averages from equilibrium and nonequilibrium Langevin molecular dynamics, *IMA J. Numer. Anal.* (2015)

M. Fathi, A.-A. Homman and G. Stoltz, Error analysis of the transport properties of Metropolized schemes, *ESAIM Proc.* (2015)

M. Fathi and G. Stoltz, Improving dynamical properties of stabilized discretizations of overdamped Langevin dynamics, *arXiv* **1505.04905** (2015)

Reminder: Error estimates in Monte Carlo simulations

- General SDE $dx_t = b(x_t) dt + \sigma(x_t) dW_t$, invariant measure π
- **Discretization** $x^n \simeq x_{n\Delta t}$, **invariant measure** $\pi_{\Delta t}$. For instance,

$$x^{n+1} = x^n + \Delta t b(x^n) + \sqrt{\Delta t} \sigma(x^n) G^n, \quad G^n \sim \mathcal{G}(0, \text{Id}) \text{ i.i.d.}$$

Error estimates for **finite** trajectory averages

$$\hat{A}_{N_{\text{iter}}} = \frac{1}{N_{\text{iter}}} \sum_{n=1}^{N_{\text{iter}}} A(x^n) = \mathbb{E}_{\pi}(A) + \underbrace{C\Delta t^\alpha}_{\text{bias}} + \underbrace{\frac{\sigma_{A,\Delta t}}{\sqrt{N_{\text{iter}}}} \mathcal{G}}_{\text{statistical error}}$$

- Bias $\mathbb{E}_{\pi_{\Delta t}}(A) - \mathbb{E}_{\pi}(A) \longrightarrow$ **Focus of this talk**

- Variance $\frac{\sigma_{A,\Delta t}^2}{N_{\text{iter}}} \simeq \frac{2 \int_0^{+\infty} \mathbb{E}[\delta A(x_t) \delta A(x_0)] dt}{T}$ where $T = N_{\text{iter}} \Delta t$

Definition of the self-diffusion

Definition of the self-diffusion: Langevin dynamics

- **Periodic potential** V , Langevin dynamics for $(q, p) \in \mathcal{E} = (L\mathbb{T})^d \times \mathbb{R}^d$

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

Mobility: averages in a nonequilibrium steady-state

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

Effective diffusion computed at equilibrium: $\eta = 0$

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

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

- Integrability by exp. convergence to $\mu(dq dp) = Z^{-1} e^{-\beta H(q,p)} dq dp$

Definition of the self-diffusion: overdamped Langevin

- Dynamics $dq_t = \left(-\nabla V(q_t) + \eta F \right) dt + \sqrt{\frac{2}{\beta}} dW_t$
- Invariant measure $\nu(dq) = Z^{-1} e^{-\beta V(q)} dq$

Mobility: averages in a nonequilibrium steady-state

$$\nu_F = \lim_{\eta \rightarrow 0} \frac{\mathbb{E}_\eta \left(-F^T \nabla V(q) \right)}{\eta} = \beta \int_{\mathcal{E}} -F^T \nabla V(q) g_{0,1}(q) \nu(dq) = \frac{1}{\beta} F^T (\text{Id} - D) F$$

Effective diffusion computed at equilibrium: $\eta = 0$

Unperiodized displacement $Q_t - Q_0 = - \int_0^t \nabla V(q_s) ds + \sqrt{\frac{2}{\beta}} W_t$

$$F^T D F = |F|^2 - \beta^2 \int_0^{+\infty} \mathbb{E}_0 \left[\left(F^T \nabla V(q_t) \right) \left(F^T \nabla V(q_0) \right) \right] dt$$

General definition of transport coefficients

- **Ergodic** stochastic dynamics $dX_t = b(X_t) dt + \sigma dW_t$
 - Invariant measure $\pi(dx)$
 - Generator $\mathcal{L} = b \cdot \nabla + \frac{\sigma^2}{2} \Delta$
- **Drift perturbation** $b + \eta \tilde{b} \rightarrow$ perturbation of generator $\tilde{\mathcal{L}} = \tilde{b} \cdot \nabla$

Linear response and Green-Kubo type formulas

With adjoints taken on $L^2(\pi)$ and when $\mathbb{E}_0(R) = 0$,

$$\alpha = \lim_{\eta \rightarrow 0} \frac{\mathbb{E}_\eta(R)}{\eta} = - \int_{\mathcal{E}} [\mathcal{L}^{-1} R] [\tilde{\mathcal{L}}^* \mathbf{1}] d\pi = \int_0^{+\infty} \mathbb{E}_0(R(x_t) S(x_0)) dt$$

- Conjugated response function $S = \tilde{\mathcal{L}}^* \mathbf{1}$
- Relies on the equality $-\mathcal{L}^{-1} = \int_0^{+\infty} e^{t\mathcal{L}} dt$ (**functional analysis**)

Error estimates on α resulting from finite $\Delta t > 0$?

Error estimates for equilibrium dynamics

Weak type expansions

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

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

- (Infinitely) Many possibilities! Numerical analysis allows to **discriminate**

Δt -expansion of the evolution operator

$$P_{\Delta t}\varphi = \varphi + \Delta t \mathcal{A}_1\varphi + \Delta t^2 \mathcal{A}_2\varphi + \cdots + \Delta t^{p+1} \mathcal{A}_{p+1}\varphi + \Delta t^{p+2} r_{\varphi, \Delta t}$$

- Weak order** p when $\mathcal{A}_k = \mathcal{L}^k/k!$ for $1 \leq k \leq p$, namely

$$\sup_{0 \leq n \leq T/\Delta t} \left| \mathbb{E}[\varphi(x^n)] - \mathbb{E}[\varphi(x_{n\Delta t})] \right| \leq C \Delta t^p$$

- Ergodicity** of the numerical scheme with invariant measure $\pi_{\Delta t}$

$$\frac{1}{N_{\text{iter}}} \sum_{n=1}^{N_{\text{iter}}} A(x^n) \xrightarrow{N_{\text{iter}} \rightarrow +\infty} \int_{\mathcal{X}} A(x) \pi_{\Delta t}(dx)$$

Example: Euler-Maruyama, weak order 1

- Scheme $x^{n+1} = \Phi_{\Delta t}(x^n, G^n) = x^n + \Delta t b(x^n) + \sqrt{\Delta t} \sigma(x^n) G^n$

- Note that $P_{\Delta t} \varphi(x) = \mathbb{E}_G [\varphi(\Phi_{\Delta t}(x, G))]$

- Technical tool: **Taylor expansion**

$$\varphi(x + \delta) = \varphi(x) + \delta^T \nabla \varphi(x) + \frac{1}{2} \delta^T \nabla^2 \varphi(x) \delta + \frac{1}{6} D^3 \varphi(x) : \delta^{\otimes 3} + \dots$$

- Replace δ with $\sqrt{\Delta t} \sigma(x) G + \Delta t b(x)$ and **gather in powers of Δt**

$$\begin{aligned} \varphi(\Phi_{\Delta t}(x, G)) &= \varphi(x) + \sqrt{\Delta t} \sigma(x) G \cdot \nabla \varphi(x) \\ &\quad + \Delta t \left(\frac{\sigma(x)^2}{2} G^T [\nabla^2 \varphi(x)] G + b(x) \cdot \nabla \varphi(x) \right) + \dots \end{aligned}$$

- Taking **expectations w.r.t. G** leads to

$$P_{\Delta t} \varphi(x) = \varphi(x) + \underbrace{\Delta t \left(\frac{\sigma(x)^2}{2} \Delta \varphi(x) + b(x) \cdot \nabla \varphi(x) \right)}_{=\mathcal{L}\varphi(x)} + O(\Delta t^2)$$

Error estimates on the invariant measure

- **Assumptions** on the operators in the weak-type expansion
 - invariance of π by \mathcal{A}_k for $1 \leq k \leq p$, namely $\int_{\mathcal{X}} \mathcal{A}_k \varphi d\pi = 0$
 - $\int_{\mathcal{X}} \mathcal{A}_{p+1} \varphi d\pi = \int_{\mathcal{X}} g_{p+1} \varphi d\pi$ (i.e. $g_{p+1} = \mathcal{A}_{p+1}^* \mathbf{1}$)

Error estimates on $\pi_{\Delta t}$

$$\int_{\mathcal{X}} \varphi d\pi_{\Delta t} = \int_{\mathcal{X}} \varphi (1 + \Delta t^p f_{p+1}) d\pi + \Delta t^{p+1} R_{\varphi, \Delta t}$$

- In fact, $f_{p+1} = -(\mathcal{A}_1^*)^{-1} g_{p+1}$
 - when $\mathcal{A}_1 = \mathcal{L}$, the first order correction can be **estimated** by some integrated correlation function as $\int_0^{+\infty} \mathbb{E}(\varphi(x_t) g_{p+1}(x_0)) dt$
 - in general, first order term can be removed by Romberg extrapolation
- Error on invariant measure can be **(much) smaller** than the weak error

Sketch of proof (1)

Step 1: Establish the error estimate for $\varphi \in \text{Ran}(P_{\Delta t} - \text{Id})$

- Idea: $\pi_{\Delta t} = \pi(1 + \Delta t^p f_{p+1} + \dots)$
- by definition of $\pi_{\Delta t}$

$$\int_{\mathcal{X}} \left[\left(\frac{P_{\Delta t} - \text{Id}}{\Delta t} \right) \psi \right] d\pi_{\Delta t} = 0$$

- compare to first order correction to the invariant measure

$$\begin{aligned} & \int_{\mathcal{X}} \left[\left(\frac{P_{\Delta t} - \text{Id}}{\Delta t} \right) \psi \right] (1 + \Delta t^p f_{p+1}) d\pi \\ &= \Delta t^p \int_{\mathcal{X}} \left(\mathcal{A}_{p+1} \psi + (\mathcal{A}_1 \psi) f_{p+1} \right) d\pi + O(\Delta t^{p+1}) \\ &= \Delta t^p \int_{\mathcal{X}} \left(g_{p+1} + \mathcal{A}_1^* f_{p+1} \right) \psi d\pi + O(\Delta t^{p+1}) \end{aligned}$$

Suggests $f_{p+1} = -(\mathcal{A}_1^*)^{-1} g_{p+1}$

Sketch of proof (2)

Step 2: Define an approximate inverse

- Issue: derivatives of $(\text{Id} - P_{\Delta t})^{-1}\varphi$ are not controlled

- Consider $\left(\Pi \frac{P_{\Delta t} - \text{Id}}{\Delta t} \Pi\right) Q_{\Delta t} \psi = \psi + \Delta t^{p+1} \tilde{r}_{\psi, \Delta t}$ where

$$\Pi \varphi = \varphi - \int_{\mathcal{X}} \varphi d\pi$$

- Idea of the construction: truncate the formal series expression

$$(A + \Delta t B)^{-1} = A^{-1} - \Delta t A^{-1} B A^{-1} + \Delta t^2 A^{-1} B A^{-1} B A^{-1} + \dots$$

Step 3: Conclusion

- Write the invariances with $\Pi \left(\frac{P_{\Delta t} - \text{Id}}{\Delta t}\right) \Pi \psi$ instead of $\left(\frac{P_{\Delta t} - \text{Id}}{\Delta t}\right) \psi$
- Replace ψ by $Q_{\Delta t} \varphi$, and gather in $R_{\varphi, \Delta t}$ all the higher order terms

Error estimates
on the linear response
of nonequilibrium dynamics

Examples of splitting schemes for Langevin dynamics (1)

- Example: Langevin dynamics, discretized using a **splitting** strategy

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

- First order splitting schemes: Trotter splitting

$$P_{\Delta t}^{ZYX} = e^{\Delta t Z} e^{\Delta t Y} e^{\Delta t X} \simeq e^{\Delta t A}$$

- **Second order** schemes: Strang splitting

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

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

Examples of splitting schemes for Langevin dynamics (2)

- $P_{\Delta t}^{B_\eta, A, \gamma C}$ corresponds to

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

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

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

Error estimates on linear response

Error estimates for nonequilibrium dynamics

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

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

where the remainder is compatible with linear response

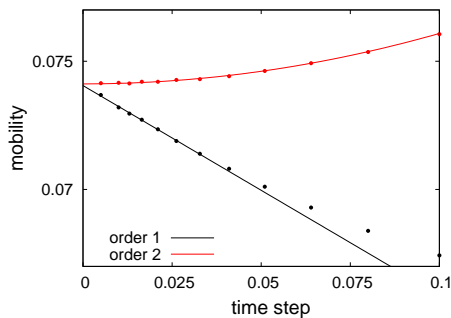
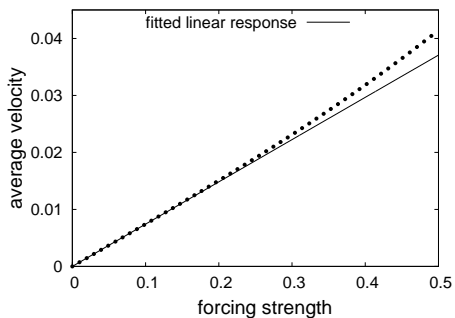
$$|r_{\psi,\gamma,\eta,\Delta t}| \leq K(\eta^2 + \Delta t^{\alpha+1}), \quad |r_{\psi,\gamma,\eta,\Delta t} - r_{\psi,\gamma,0,\Delta t}| \leq K\eta(\eta + \Delta t^{\alpha+1})$$

- Corollary: error estimates on the **numerically computed mobility**

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

- 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 Green-Kubo formulas (1)

- Error of **order α on invariant measure**: $\int_{\mathcal{X}} \psi d\pi_{\Delta t} = \int_{\mathcal{X}} \psi d\pi + O(\Delta t^\alpha)$
- Expansion of the evolution operator ($p + 1 \geq \alpha$ and $\mathcal{A}_1 = \mathcal{L}$)

$$P_{\Delta t}\varphi = \varphi + \Delta t \mathcal{L}\varphi + \Delta t^2 \mathcal{A}_2\varphi + \cdots + \Delta t^{p+1} \mathcal{A}_{p+1}\varphi + \Delta t^{p+2} r_{\varphi, \Delta t}$$

Ergodicity of the numerical scheme

$$\forall n \in \mathbb{N}, \quad \|P_{\Delta t}^n\|_{\mathcal{B}(L_{\mathcal{K}_s, \Delta t}^\infty)} \leq C_s e^{-\lambda_s n \Delta t}$$

where \mathcal{K}_s is a Lyapunov function ($1 + |\rho|^{2s}$ for Langevin) and

$$L_{\mathcal{K}_s, \Delta t}^\infty = \left\{ \frac{\varphi}{\mathcal{K}_s} \in L^\infty(\mathcal{X}), \int_{\mathcal{X}} \varphi d\pi_{\Delta t} = 0 \right\}$$

- Proof: Lyapunov condition + uniform-in- Δt minorization condition¹

¹M. Hairer and J. Mattingly, *Progr. Probab.* (2011)

Error estimates on Green-Kubo formulas (2)

Error estimates on integrated correlation functions

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

$$\int_0^{+\infty} \mathbb{E}(\psi(x_t)\varphi(x_0)) dt = \Delta t \sum_{n=0}^{+\infty} \mathbb{E}_{\Delta t}(\tilde{\psi}_{\Delta t, \alpha}(x^n)\varphi(x^0)) + \Delta t^\alpha r_{\Delta t}^{\psi, \varphi},$$

where $\mathbb{E}_{\Delta t}$ denotes expectations w.r.t. initial conditions $x_0 \sim \pi_{\Delta t}$ and over all realizations of the Markov chain (x^n) , and

$$\tilde{\psi}_{\Delta t, \alpha} = \psi_{\Delta t, \alpha} - \int_{\mathcal{X}} \psi_{\Delta t, \alpha} d\pi_{\Delta t}$$

with $\psi_{\Delta t, \alpha} = (\text{Id} + \Delta t \mathcal{A}_2 \mathcal{L}^{-1} + \dots + \Delta t^{\alpha-1} \mathcal{A}_\alpha \mathcal{L}^{-1})\psi$

- Useful when $\mathcal{A}_k \mathcal{L}^{-1}$ can be computed, e.g. $\mathcal{A}_k = a_k \mathcal{L}^k$
- Reduces to trapezoidal rule for second order schemes

Extension to Metropolized overdamped Langevin dynamics

- Superimpose Metropolis-Hastings correction to discretization of SDE

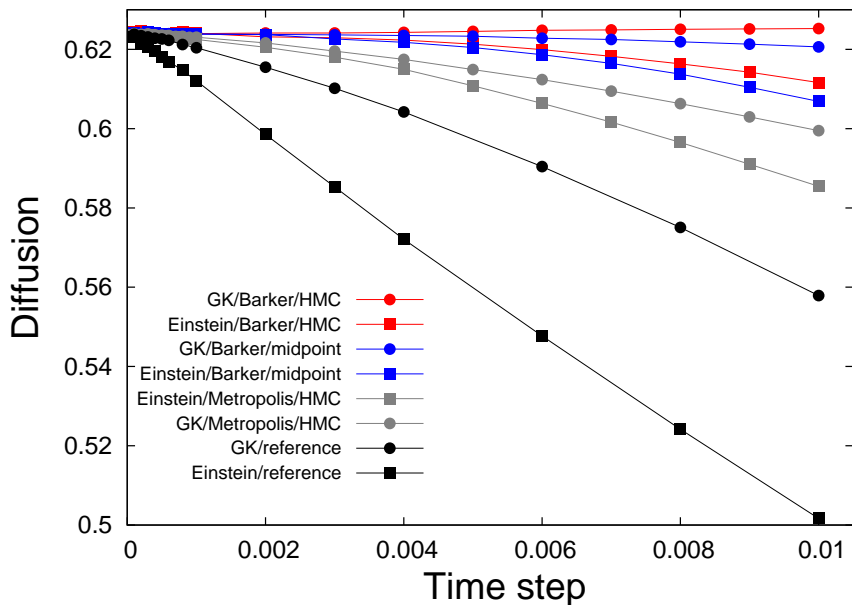
$$\tilde{q}^{n+1} = q^n + \Delta t \nabla V(q^n) + \sqrt{\frac{2\Delta t}{\beta}} G^n$$

- **no bias** on invariant measure / **stabilization** for singular potentials
- error estimates on the diffusion of order Δt

- **HMC-like** scheme $\tilde{q}^{n+1} = q^n + \Delta t \nabla V \left(q^n + \sqrt{\frac{\Delta t}{2\beta}} G^n \right) + \sqrt{\frac{2\Delta t}{\beta}} G^n$

- Error of order $\Delta t^{3/2}$ when the Metropolis-Hastings rule is used
 - Reduced to Δt^2 when a **Barker rule** is used (replace $\min(1, r)$ by $r/(r+1)$)
 - Requires some **time renormalization** since rejection rate $\simeq 1/2$
 - Trade-off between **increased variance** (factor 2) and reduced bias
- Extension to diffusions with multiplicative noise

Numerical results



Conclusion and perspectives

Conclusion and perspectives

- **Systematic errors in Green-Kubo formulas are larger or equal than for linear response results** – but beware prefactors!

- **equilibrium** invariant measures $\pi_{\Delta t} = \pi + O(\Delta t^\alpha)$
- linear response of **nonequilibrium** perturbations \rightarrow errors Δt^α
- Green-Kubo formula \rightarrow error limited by **weak order type** conditions
- Numerical analysis allows to derive numerical schemes to decrease the bias

- **Reduction of statistical error?**

- Driving model: thermal transport in 1D systems \rightarrow **I am fed up of waiting months for simulation results to converge!**
- Issues: non-trivial perturbations of the invariant measure when the dynamics is (locally) modified
- PhD work of Julien Roussel