Multiscale stochastic dynamics: effective dynamics and parareal computations

Frédéric Legoll

Ecole des Ponts & Inria Paris

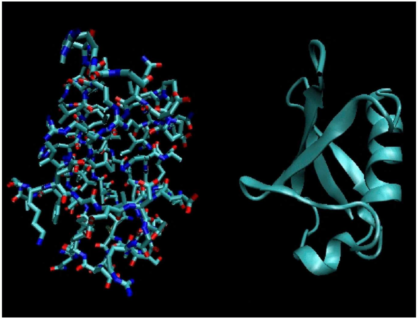
Joint with T. Lelièvre and U. Sharma (Ecole des Ponts),
S. Olla (Paris Dauphine), K. Myerscough and G. Samaey (KU Leuven)

Workshop on Multiscale methods for deterministic and stochastic dynamics

Genève, 29-30 January 2020

Motivation

This work is motivated by molecular simulation.



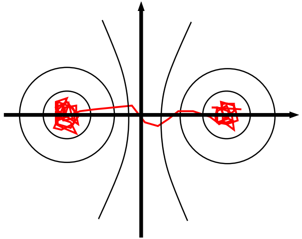
Typical dynamics: the overdamped Langevin equation

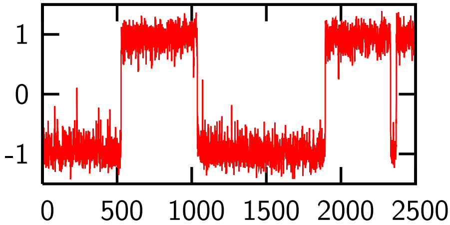
$$dX_t = -\nabla V(X_t)\,dt + \sqrt{2\beta^{-1}}\,dW_t \quad X_t \equiv \text{position of all atoms}$$

Metastability and reaction coordinate

$$dX_t = -\nabla V(X_t)\,dt + \sqrt{2\beta^{-1}}\,dW_t$$

- the dynamics is often metastable:





- we assume that wells are fully described through a well-chosen reaction coordinate

$$\xi : \mathbb{R}^n \mapsto \mathbb{R}$$

Quantity of interest: path $t \mapsto \xi(X_t)$.

Our aim

- Reversible case:

$$dX_t = -\nabla V(X_t)\, dt + \sqrt{2\beta^{-1}}\, dW_t \quad \text{in } \mathbb{R}^n$$

Given a reaction coordinate $\xi : \mathbb{R}^n \mapsto \mathbb{R}$,

and under some time-scale separation assumptions,

propose a dynamics z_t that approximates $\xi(X_t)$.

Our aim

- Reversible case:

$$dX_t = -\nabla V(X_t)\,dt + \sqrt{2\beta^{-1}}\,dW_t \quad \text{in } \mathbb{R}^n$$

Given a reaction coordinate $\xi : \mathbb{R}^n \mapsto \mathbb{R}$,
and under some time-scale separation assumptions,
propose a dynamics z_t that approximates $\xi(X_t)$.

- Non-reversible case:

$$dX_t = \mathcal{F}(X_t)\,dt + \sqrt{2\beta^{-1}}\,\sigma(X_t)\,dW_t \qquad \text{non degenerate } \sigma$$

This case is encountered in practice:

- as a mathematical tool to reduce the variance when computing $\mathbb{E}_{\rm Gibbs}\left[\Phi(X)\right]$, e.g. [Duncan et al 2016]:

$$dX_t = -(I + \alpha J_{\rm skew})\nabla V(X_t)\,dt + \sqrt{2\beta^{-1}}\,dW_t$$

- modelization of polymer chains in a macroscopic flow: micro SDEs coupled with macro Navier-Stokes

Remarks

- Alternative and/or related approaches by Schuette, Pavliotis and Stuart, Hartmann, Papanicolaou, E & Vanden-Eijnden, Hudson & Li, ..., Mori-Zwanzig approaches, ...

- Examples of applications:
  - Molecular Dynamics problems
  - Classical slow / fast SDEs such as

$$
\begin{aligned}
dX_t^{\varepsilon,1} &= -\partial_1 V(X_t^\varepsilon)\, dt + \sqrt{2\beta^{-1}}\, dW_t^1 \\
dX_t^{\varepsilon,i} &= -\varepsilon^{-1}\, \partial_i V(X_t^\varepsilon)\, dt + \sqrt{2\beta^{-1}\varepsilon^{-1}}\, dW_t^i \qquad \text{for } i = 2, \dots, n
\end{aligned}
$$

  - ...

Construction of an effective dynamics

FL, T. Lelièvre, Nonlinearity 2010

Effective dynamics using conditional expectations - 1

$$dX_t = -\nabla V(X_t)\,dt + \sqrt{2\beta^{-1}}\,dW_t, \qquad \xi : \mathbb{R}^n \to \mathbb{R}$$

From the dynamics on X_t, we obtain (chain rule)

$$d\left[\xi(X_t)\right] = \left(-\nabla V \cdot \nabla\xi + \beta^{-1}\Delta\xi\right)(X_t)\,dt + \sqrt{2\beta^{-1}}\ \ |\nabla\xi|(X_t)\ \ dB_t$$

where B_t is a 1D brownian motion.

Effective dynamics using conditional expectations - 1

$$dX_t = -\nabla V(X_t)\, dt + \sqrt{2\beta^{-1}}\, dW_t, \qquad \xi : \mathbb{R}^n \to \mathbb{R}$$

From the dynamics on X_t, we obtain (chain rule)

$$d\left[\xi(X_t)\right] = \left(-\nabla V \cdot \nabla \xi + \beta^{-1} \Delta \xi\right)(X_t)\, dt + \sqrt{2\beta^{-1}} \;\; |\nabla \xi|(X_t) \;\; dB_t$$

where B_t is a 1D brownian motion.

Introduce the average of the drift term:

$$\begin{aligned} b(z) &:= \int \left(-\nabla V \cdot \nabla \xi + \beta^{-1} \Delta \xi\right)(X) \; \psi_{\rm Gibbs}(X)\, \delta_{\xi(X)-z}\, dX \\ &= \mathbb{E}_{\rm Gibbs}\left[\left(-\nabla V \cdot \nabla \xi + \beta^{-1} \Delta \xi\right)(X) \;\middle|\; \xi(X) = z\right] \end{aligned}$$

Effective dynamics using conditional expectations - 1

$$dX_t = -\nabla V(X_t)\,dt + \sqrt{2\beta^{-1}}\,dW_t, \qquad \xi : \mathbb{R}^n \to \mathbb{R}$$

From the dynamics on X_t, we obtain (chain rule)

$$d\left[\xi(X_t)\right] = \left(-\nabla V \cdot \nabla \xi + \beta^{-1}\Delta\xi\right)(X_t)\,dt + \sqrt{2\beta^{-1}}\ \ |\nabla\xi|(X_t)\ \ dB_t$$

where B_t is a 1D brownian motion.

Introduce the average of the drift term:

$$\begin{aligned} b(z) \ &:= \ \int \left(-\nabla V \cdot \nabla \xi + \beta^{-1}\Delta\xi\right)(X)\ \psi_{\rm Gibbs}(X)\,\delta_{\xi(X)-z}\,dX \\ &= \ \mathbb{E}_{\rm Gibbs}\left[\left(-\nabla V \cdot \nabla \xi + \beta^{-1}\Delta\xi\right)(X)\ \Big|\ \xi(X) = z\right] \end{aligned}$$

and likewise for the diffusion term:

$$\sigma^2(z) := \int |\nabla\xi(X)|^2\ \ \psi_{\rm Gibbs}(X)\,\delta_{\xi(X)-z}\,dX$$

Effective dynamics using conditional expectations - 2

Chain rule:

$$d\left[\xi(X_t)\right] = \left(-\nabla V \cdot \nabla \xi + \beta^{-1} \Delta \xi\right)(X_t)\, dt + \sqrt{2\beta^{-1}}\ \ |\nabla \xi|(X_t)\ \ dB_t$$

Average of the drift and diffusion terms:

$$\begin{aligned} b(z) &:= \int \left(-\nabla V \cdot \nabla \xi + \beta^{-1} \Delta \xi\right)(X)\ \psi_{\rm Gibbs}(X)\, \delta_{\xi(X)-z}\, dX \\ \sigma^2(z) &:= \int |\nabla \xi(X)|^2\ \ \psi_{\rm Gibbs}(X)\, \delta_{\xi(X)-z}\, dX \end{aligned}$$

Eff. dyn. we propose: $\boxed{dz_t = b(z_t)\, dt + \sqrt{2\beta^{-1}}\, \sigma(z_t)\, dB_t}$

Effective dynamics using conditional expectations - 2

Chain rule:

$$d\left[\xi(X_t)\right] = \left(-\nabla V \cdot \nabla \xi + \beta^{-1}\Delta\xi\right)(X_t)\,dt + \sqrt{2\beta^{-1}}\;\;|\nabla\xi|(X_t)\;\;dB_t$$

Average of the drift and diffusion terms:

$$b(z) \;\;:=\;\; \int \left(-\nabla V \cdot \nabla \xi + \beta^{-1}\Delta\xi\right)(X)\;\psi_{\rm Gibbs}(X)\,\delta_{\xi(X)-z}\,dX$$

$$\sigma^2(z) \;\;:=\;\; \int |\nabla\xi(X)|^2\;\;\psi_{\rm Gibbs}(X)\,\delta_{\xi(X)-z}\,dX$$

Eff. dyn. we propose: $\boxed{dz_t = b(z_t)\,dt + \sqrt{2\beta^{-1}}\,\sigma(z_t)\,dB_t}$

The approximation makes sense if, in the manifold

$$\Sigma_z = \{X \in \mathbb{R}^n, \quad \xi(X) = z\},$$

X_t quickly reaches equilibrium: no metastability in Σ_z.

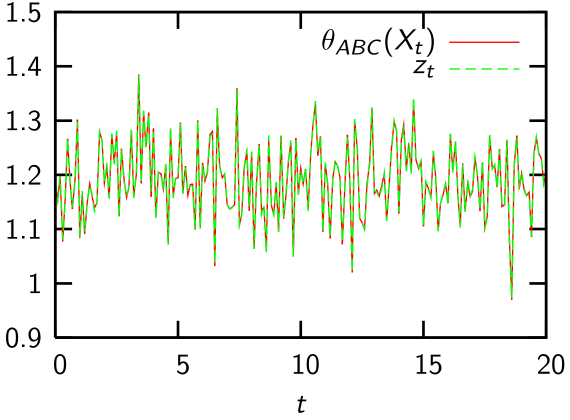
Accuracy of the effective dynamics: a pathwise approach

FL, T. Lelièvre, S. Olla, Stoch. Processes and their Applications 2017

Numerical result

Consider a three-atom molecule A-B-C in 2D, with bond length potentials much stiffer than the bond angle potential.

$$\xi(X) = \theta_{ABC}(X)$$



Assumptions

We focus on the case $\xi(X) = X^1$, and thus aim at controlling

$$\mathbb{E}\left[\sup_{0\leq t\leq T} \left|\xi(X_t) - z_t\right|^2\right] = \mathbb{E}\left[\sup_{0\leq t\leq T} \left|X_t^1 - z_t\right|^2\right]$$

Assumptions

We focus on the case $\xi(X) = X^1$, and thus aim at controlling

$$\mathbb{E}\left[\sup_{0\leq t\leq T} \left|\xi(X_t) - z_t\right|^2\right] = \mathbb{E}\left[\sup_{0\leq t\leq T} \left|X_t^1 - z_t\right|^2\right]$$

For any z, the conditional probability measures $\psi^z_{\rm Gibbs}(x_2^n)$ satisfy a Poincaré inequality for a constant ρ independent of z: for any function v,

$$\int_{\mathbb{R}^{n-1}} \left(v - \int_{\mathbb{R}^{n-1}} v\, \psi^z_{\rm Gibbs}\right)^2 \psi^z_{\rm Gibbs} \leq \frac{1}{\rho} \int_{\mathbb{R}^{n-1}} \left|\widehat{\nabla} v\right|^2 \psi^z_{\rm Gibbs}$$

where $\widehat{\nabla} v = (\partial_2 v, \dots, \partial_n v)$.

The cross derivative $\widehat{\nabla}\partial_1 V$ is in $L^2(\psi_{\rm Gibbs})$:

$$\kappa^2 := \int_{\mathbb{R}^n} \left|\widehat{\nabla}\partial_1 V(x)\right|^2 \psi_{\rm Gibbs}(x) dx < \infty.$$

Remarks

- a Poincaré inequality holds on a probability measure $\exp(-V(x))\,dx$ under relatively mild assumptions on V.
- the probability measure $\exp(-V(x))\,dx$ satisfies a Poincaré inequality for a constant ρ if and only if, for the SDE

  $$dX_t = -\nabla V(X_t)\,dt + \sqrt{2}\,dW_t$$

  the probability distribution function $\psi(t,x)$ converges to $\exp(-V(x))\,dx$ (for any IC) at the rate $\exp(-\rho\,t)$
- the larger ρ is, the "easier" $\exp(-V(x))\,dx$ is to sample

Error estimate

Consider $(X_t)_{0 \le t \le T}$ solution to the reference dynamics and $(z_t)_{0 \le t \le T}$ solution to the effective dynamics over a bounded time interval $[0, T]$.

Then, there exists a constant C, which only depends on T, such that

$$\mathbb{E}\left[\sup_{0 \le t \le T} \left|X_t^1 - z_t\right|^2\right] \le C(T)\,\frac{\kappa^2}{\rho^2}.$$

Large time-scale separation $\sim$ large ρ $\sim$ z_t is a good approx of X_t^1

Extension to some non-reversible cases

$$dX_t = \mathcal{F}(X_t)\,dt + \sqrt{2}\,\sigma(X_t)\,dW_t$$

FL, T. Lelièvre, U. Sharma, Nonlinearity 2019

For the sake of simplicity, we focus here on

$$dX_t = \mathcal{F}(X_t)\,dt + \sqrt{2}\,dW_t, \qquad \mathcal{F} \text{ is not a gradient}, \qquad \xi(X) = X^1$$

Several options to coarse-grain

$$dX_t = \mathcal{F}(X_t)\,dt + \sqrt{2}\,dW_t \quad \text{in 2D}, \qquad \xi(X) = X^1$$

1. Let $\mu(x_1, x_2)$ be the invariant measure, and consider the conditional expectation as effective drift:

$$b(z) = \mathbb{E}_\mu\left[\,\mathcal{F}_1(X) \mid X^1 = z\right]$$

Several options to coarse-grain

$$dX_t = \mathcal{F}(X_t)\,dt + \sqrt{2}\,dW_t \quad \text{in 2D}, \qquad \xi(X) = X^1$$

1. Let $\mu(x_1, x_2)$ be the invariant measure, and consider the conditional expectation as effective drift:

$$b(z) = \mathbb{E}_\mu\left[\,\mathcal{F}_1(X) \mid X^1 = z\right]$$

2. Consider the dynamics in the fast variable for a frozen slow variable:

$$dX_t^2 = \mathcal{F}_2(z, X_t^2)\,dt + \sqrt{2}\,dW_t^2$$

Let $\theta_z(x_2)$ be the invariant measure, and consider as effective drift

$$b(z) = \mathbb{E}_{\theta_z}\left[\,\mathcal{F}_1(z, X^2)\right]$$

These choices generally lead to different formulas for b.

Setting (non-reversible case)

- Reference dynamics:

$$dX_t = \mathcal{F}(X_t)\,dt + \sqrt{2}\,dW_t, \qquad \text{unique stat. measure } \mu$$

- Non-closed dynamics:

$$dX_t^1 = \mathcal{F}_1(X_t)\,dt + \sqrt{2}\,dW_t^1$$

Setting (non-reversible case)

- Reference dynamics:

$$dX_t = \mathcal{F}(X_t)\,dt + \sqrt{2}\,dW_t, \qquad \text{unique stat. measure } \mu$$

- Non-closed dynamics:

$$dX_t^1 = \mathcal{F}_1(X_t)\,dt + \sqrt{2}\,dW_t^1$$

- Effective dynamics:

$$dz_t = b(z_t)\,dt + \sqrt{2}\,dW_t^1$$

with

$$b(z) = \mathbb{E}_\mu\left(\mathcal{F}_1(X) \,\middle|\, X^1 = z\right) = \int_{\mathbb{R}^{n-1}} \mathcal{F}_1(z, x_2^n)\ \mu^z(x_2^n)\,dx_2^n$$

where

$$\mu^z(x_2^n) = \text{condit. measure} = \frac{\mu(z, x_2^n)}{\displaystyle\int_{\mathbb{R}^{n-1}} \mu(z, x_2^n)\,dx_2^n}, \quad x_2^n = (x^2, \dots, x^n)$$

Assumptions (non-reversible case)

Similar assumptions as in the reversible case:

- For any z, the conditional probability measures $\mu^z(x_2^n)$ satisfy a Poincaré inequality for a constant ρ independent of z.
- The cross derivative $\widehat{\nabla}\mathcal{F}_1$ is in $L^2(\mu)$:

$$\kappa^2 := \int_{\mathbb{R}^n} \left| \widehat{\nabla}\mathcal{F}_1 \right|^2 \mu < \infty.$$

Error estimate (non-reversible case)

Consider $(X_t)_{0\leq t\leq T}$ solution to the reference dynamics and $(z_t)_{0\leq t\leq T}$ solution to the effective dynamics over a bounded time interval $[0, T]$.

Then, there exists a constant C, which only depends on T, such that

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\left|X_t^1 - z_t\right|^2\right] \leq C(T)\ \frac{\kappa^2}{\rho^2}.$$

Generalizations and open questions

Our result can be extended to the more general case

$$dX_t = \mathcal{F}(X_t)\,dt + \sqrt{2}\,\sigma(X_t)\,dW_t$$

for a non-degenerate diffusion matrix σ.

Open question: SDEs with degenerate noise, such as the Langevin-like equation:

$$dX_t = P_t\,dt, \qquad dP_t = \mathcal{F}(X_t)\,dt - P_t\,dt + \sqrt{2}\,\sigma(X_t)\,dW_t$$

Conclusions so far

- We proposed a "natural" way to obtain a closed equation on $\xi(X_t)$
- Rigorous, non-asymptotic error bounds
- Nice numerical results
- Our approach can be applied to the standard problem

$$\begin{cases} dX_t^{\varepsilon,1} = -\partial_1 V(X_t^\varepsilon)\, dt + \sqrt{2\beta^{-1}}\, dW_t^1 \\ dX_t^{\varepsilon,i} = -\varepsilon^{-1}\, \partial_i V(X_t^\varepsilon)\, dt + \sqrt{2\beta^{-1}\varepsilon^{-1}}\, dW_t^i \qquad \text{for } i = 2, \dots, n \end{cases}$$

without Lipschitz assumptions on ∇V.

FL, T. Lelièvre, S. Olla, Stoch. Processes and their Applications 2017
FL, T. Lelièvre, U. Sharma, Nonlinearity 2019

Beyond the effective dynamics

Assume now that the effective dynamics is not accurate enough, e.g. because the time scale separation is not large enough.

Concurrently use the approximate model and the reference model to design a more efficient algorithm for the integration of the reference model?

The parareal algorithm ...

FL, T. Lelièvre, K. Myerscough, G. Samaey, arXiv 1912.09240

Setting

Reference model:

$$dX_t = F(X_t, Y_t)\,dt + \sqrt{2D(X_t, Y_t)}\,dU_t, \qquad X_t \in \mathbb{R},$$

$$dY_t = \frac{1}{\varepsilon} G(X_t, Y_t)\,dt + \sqrt{\frac{2E(X_t, Y_t)}{\varepsilon}}\,dV_t, \qquad Y_t \in \mathbb{R}^d$$

Effective dynamics in the limit $\varepsilon \ll 1$:

$$dz_t = b(z_t)\,dt + \sqrt{2\sigma(z_t)}\,dW_t, \qquad z_t \in \mathbb{R}$$

Quantity of interest:

- probability distribution function $\rho(t, x)$ of X (simpler than X_t)
- we thus consider the associated Fokker-Planck equations

Parallel in time algorithm for ODEs

$$\frac{dy}{dt} = f(y), \qquad y \in \mathbb{R}^d$$

The pareal algorithm (Lions, Maday and Turinici, 2001) is based upon two integrators to propagate the system over a time ΔT:

- a fine, accurate integrator $\mathcal{F}_{\Delta T}$
- a cheap coarse integrator $\mathcal{G}_{\Delta T}$

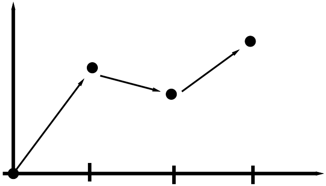
For instance,

$$\mathcal{F}_{\Delta T} = (\Phi_{\delta t_F})^{\Delta T/\delta t_F} \quad \text{and} \quad \mathcal{G}_{\Delta T} = (\Phi_{\delta t_G})^{\Delta T/\delta t_G} \quad \text{with} \quad \delta t_F \ll \delta t_G$$

The parareal iterative procedure

- Initialization: coarse propagation that yields $\left\{ y_n^{k=0} \right\}_n$:

$$\forall n, \quad y_{n+1}^{k=0} = \mathcal{G}_{\Delta T}(y_n^{k=0})$$

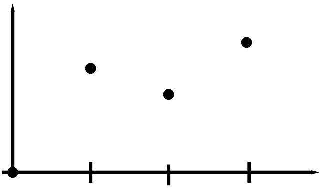


The parareal iterative procedure

- Initialization: coarse propagation that yields $\left\{ y_n^{k=0} \right\}_n$:

$$\forall n, \quad y_{n+1}^{k=0} = \mathcal{G}_{\Delta T}(y_n^{k=0})$$

- Iterate over $k \geq 0$:



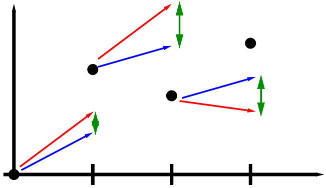
The parareal iterative procedure

- Initialization: coarse propagation that yields $\left\{y_n^{k=0}\right\}_n$:

$$\forall n, \quad y_{n+1}^{k=0} = \mathcal{G}_{\Delta T}(y_n^{k=0})$$

- Iterate over $k \geq 0$:
  - compute jumps (in parallel):

$$J_n^k = \mathcal{F}_{\Delta T}(y_n^k) - \mathcal{G}_{\Delta T}(y_n^k)$$



The parareal iterative procedure

- Initialization: coarse propagation that yields $\left\{y_n^{k=0}\right\}_n$:

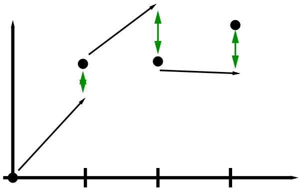
$$\forall n, \quad y_{n+1}^{k=0} = \mathcal{G}_{\Delta T}(y_n^{k=0})$$

- Iterate over $k \geq 0$:
  - compute jumps (in parallel):

$$J_n^k = \mathcal{F}_{\Delta T}(y_n^k) - \mathcal{G}_{\Delta T}(y_n^k)$$

  - sequential update to obtain $\left\{y_n^{k+1}\right\}_n$:

$$\forall n, \quad y_{n+1}^{k+1} = \mathcal{G}_{\Delta T}(y_n^{k+1}) + J_n^k$$



The parareal iterative procedure

- Initialization: coarse propagation that yields $\left\{y_n^{k=0}\right\}_n$:

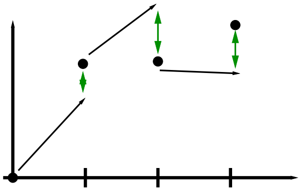
$$\forall n, \quad y_{n+1}^{k=0} = \mathcal{G}_{\Delta T}(y_n^{k=0})$$

- Iterate over $k \geq 0$:
  - compute jumps (in parallel):

$$J_n^k = \mathcal{F}_{\Delta T}(y_n^k) - \mathcal{G}_{\Delta T}(y_n^k)$$

  - sequential update to obtain $\left\{y_n^{k+1}\right\}_n$:

$$\forall n, \quad y_{n+1}^{k+1} = \mathcal{G}_{\Delta T}(y_n^{k+1}) + J_n^k$$



The fine solver is called only in the parallel part of the algorithm.

Specific difficulties

$$dX_t = F(X_t, Y_t)\,dt + \sqrt{2D(X_t, Y_t)}\,dU_t, \qquad X_t \in \mathbb{R},$$
$$dY_t = \frac{1}{\varepsilon} G(X_t, Y_t)\,dt + \sqrt{\frac{2E(X_t, Y_t)}{\varepsilon}}\,dV_t, \qquad Y_t \in \mathbb{R}^d$$

- The reference model and the effective model are written in different variables:

  $$(X, Y) \in \mathbb{R} \times \mathbb{R}^d \qquad \text{vs} \qquad z \in \mathbb{R}$$

  Easy to go from (X, Y) to z, but many ways to go from z to (X, Y)

  See FL, T. Lelièvre, G. Samaey, SISC 2013 for a similar problem

Specific difficulties

$$dX_t = F(X_t, Y_t)\, dt + \sqrt{2D(X_t, Y_t)}\, dU_t, \qquad X_t \in \mathbb{R},$$
$$dY_t = \frac{1}{\varepsilon} G(X_t, Y_t)\, dt + \sqrt{\frac{2E(X_t, Y_t)}{\varepsilon}}\, dV_t, \qquad Y_t \in \mathbb{R}^d$$

- The reference model and the effective model are written in different variables:

  $$(X, Y) \in \mathbb{R} \times \mathbb{R}^d \qquad \text{vs} \qquad z \in \mathbb{R}$$

  Easy to go from (X, Y) to z, but many ways to go from z to (X, Y)

  See FL, T. Lelièvre, G. Samaey, SISC 2013 for a similar problem

- The update formula

  $$y_{n+1}^{k+1} = \mathcal{G}_{\Delta T}(y_n^{k+1}) + \mathcal{F}_{\Delta T}(y_n^k) - \mathcal{G}_{\Delta T}(y_n^k)$$

  is not necessarily well-adapted to probability density functions (that should stay positive and of unit mass)

Four ways to iterate - 1

- Standard additive formula:

$$\rho_{n+1}^{k+1} = \mathcal{G}_{\Delta T}(\rho_n^{k+1}) + \mathcal{F}_{\Delta T}(\rho_n^k) - \mathcal{G}_{\Delta T}(\rho_n^k)$$

for which neither positivity nor unit mass is guaranteed

- Multiplicative formula:

$$\rho_{n+1}^{k+1} = \mathcal{G}_{\Delta T}(\rho_n^{k+1}) \, \frac{\mathcal{F}_{\Delta T}(\rho_n^k)}{\mathcal{G}_{\Delta T}(\rho_n^k)}$$

for which positivity (but not unit mass) is guaranteed

Four ways to iterate - 2

- Rotation formula (inspired by Maday and Turinici 2002):
  - embedd Probability Density Functions in $L^2(\mathbb{R})$ by considering $\sqrt{\rho}$
  - $\sqrt{\mathcal{G}_{\Delta T}(\rho_n^k)}$ and $\sqrt{\mathcal{F}_{\Delta T}(\rho_n^k)}$ define a rotation $\mathcal{R}$
  - rotate similarly $\sqrt{\mathcal{G}_{\Delta T}(\rho_n^{k+1})}$ and define $\sqrt{\rho_{n+1}^{k+1}} = \mathcal{R}\sqrt{\mathcal{G}_{\Delta T}(\rho_n^{k+1})}$
  - Both positivity and unit mass are guaranteed

- Quantile formula (see also Gear 2001):
  - instead of working with ρ, work with quantile function q:

$$c(x) = \int_{-\infty}^{x} \rho, \qquad q(p) = c^{-1}(p) \text{ for any } p \in [0,1]$$

  - update q by standard additive formula: this yields q_{n+1}^{k+1} and next ρ_{n+1}^{k+1}
  - Both positivity and unit mass are guaranteed (but restricted to 1D)

Parareal for Fokker-Planck equations

Focus on a toy problem: consider the scalar SDE

$$dX_t = F(X_t)\,dt + \sqrt{2D}\,dW_t$$

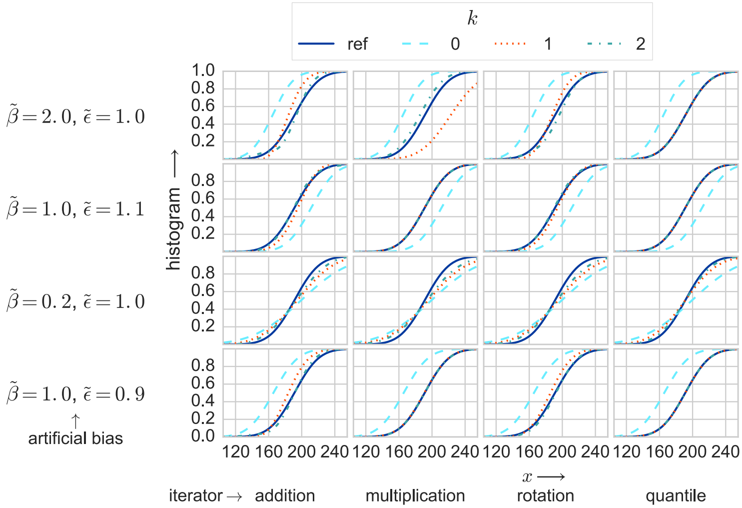
with

$$F(x) = -1/\varepsilon \qquad \text{and} \qquad D = 1/\beta$$

- Fine propagator: (exact) integration of the FP equation with the reference values $\varepsilon = \beta = 1$.
- Coarse propagator: (exact) integration of the FP equation with some approximate values $\widetilde{\varepsilon}$ and $\widetilde{\beta}$.

The fact that $\varepsilon \neq \widetilde{\varepsilon}$ and $\beta \neq \widetilde{\beta}$ stands for the fact that, in general, the coarse propagator is not accurate enough.

Numerical results (reference values: $\varepsilon = \beta = 1$)



With the quantile iterator, convergence to the exact result as soon as $k = 1$ (only one parareal correction is needed).

Next steps

Go back to the fast / slow Fokker-Planck equation associated to

$$dX_t = F(X_t, Y_t)\,dt + \sqrt{2D(X_t, Y_t)}\,dU_t, \qquad X_t \in \mathbb{R},$$
$$dY_t = \frac{1}{\varepsilon} G(X_t, Y_t)\,dt + \sqrt{\frac{2E(X_t, Y_t)}{\varepsilon}}\,dV_t, \qquad Y_t \in \mathbb{R}^d$$

and couple

- the reference, high-dimensional FP (discretized with e.g. a weighted Monte Carlo method) with solution $\psi(t, x, y)$ for $(x, y) \in \mathbb{R} \times \mathbb{R}^d$
- with the one-dimensional FP of the effective dynamics with solution $\phi(t, z)$ for $z \in \mathbb{R}$

The way we transfer information between (x, y) and z appears to be less sensitive than the choice of the parareal iterator.

FL, T. Lelièvre, K. Myerscough, G. Samaey, arXiv 1912.09240