Advanced Monte Carlo Methods

P. Del Moral

INRIA Bordeaux

Nelder Fellow Lecture Series, March 2024

Part III - Continuous time models/tools

Outline

Continuous time processes

(Linear) Langevin diffusions

Nonlinear/interacting particle samplers/SMC

Performance/Convergence analysis

Continuous time processes
- Brownian motion - diffusion - jump processes
- Diffusion processes
- Jump-diffusion processes
- Nonlinear processes/generators
- Mean field/Interacting particle samplers

(Linear) Langevin diffusions

Nonlinear/interacting particle samplers/SMC

Performance/Convergence analysis

Brownian motion B_t or W_t

Discrete time version : "dt" time steps $\oplus$ fair coin tossing

$$W_t := W_{t-dt} + \begin{cases} +\sqrt{dt} & \text{if } \textbf{Heads} \\ -\sqrt{dt} & \text{if } \textbf{Tails} \end{cases}$$

or

$$W_t := W_{t-dt} + \sqrt{dt} \times \mathrm{N}(0,1)$$

$$\Downarrow$$

$$dt = 10^{-10000000}??$$

$$\downarrow$$

Brownian motion B_t or W_t

Discrete time version : "dt" time steps $\oplus$ fair coin tossing

$$W_t := W_{t-dt} + \begin{cases} +\sqrt{dt} & \text{if} \quad \textbf{Heads} \\ -\sqrt{dt} & \text{if} \quad \textbf{Tails} \end{cases}$$

or

$$W_t := W_{t-dt} + \sqrt{dt} \times \mathrm{N}(0,1)$$

$$\Downarrow$$

$$dt \;=\; 10^{-10000000}??$$

$$\downarrow$$

$\simeq_{dt\sim 0}$ **Continuous time model $\oplus$ stochastic calculus**

Brownian motion B_t or W_t

Brownian motion B_t or W_t

$\rightsquigarrow$ **Only "3 simple ingredients":**

$$\begin{aligned} dW_t \times dW_t &= \pm\sqrt{dt} \times \pm\sqrt{dt} = dt \\ dt \times dt &= 0 \\ dt \times dW_t &= dt \ \times \pm\sqrt{dt} = 0 \end{aligned}$$

1d - diffusion processes - Ito formula

$$dX_t = \underbrace{\mathbf{b_t(X_t)\ dt}}_{\text{drift term}} + \underbrace{\boldsymbol{\sigma}_\mathbf{t}\mathbf{(X_t)dW_t}}_{\text{diffusion term}} \iff X_{t+dt} = X_t + b_t(X_t)\ dt + \sigma_t(X_t)dW_t$$

1d - diffusion processes - Ito formula

$$dX_t = \underbrace{\mathbf{b_t(X_t)\ dt}}_{\text{drift term}} + \underbrace{\boldsymbol{\sigma}_\mathbf{t}\mathbf{(X_t)dW_t}}_{\text{diffusion term}} \iff X_{t+dt} = X_t + b_t(X_t)\ dt + \sigma_t(X_t)dW_t$$

$$\Downarrow$$

$$\begin{aligned} df(t, X_t) &= f(t + dt, X_t + dX_t) - f(t, X_t) \\ &= \frac{\partial f}{\partial t}(t, X_t)\ dt + \frac{\partial f}{\partial x}(t, X_t)\ dX_t + \frac{1}{2}\ \frac{\partial^2 f}{\partial x^2}(t, X_t)\ dX_t\ dX_t \end{aligned}$$

1d - diffusion processes - Ito formula

$$dX_t = \underbrace{\mathbf{b_t(X_t)\ dt}}_{\text{drift term}} + \underbrace{\boldsymbol{\sigma_t(X_t) dW_t}}_{\text{diffusion term}} \iff X_{t+dt} = X_t + b_t(X_t)\ dt + \sigma_t(X_t) dW_t$$

$$\Downarrow$$

$$\begin{aligned}
df(t, X_t) &= f(t + dt, X_t + dX_t) - f(t, X_t) \\
&= \frac{\partial f}{\partial t}(t, X_t)\ dt + \frac{\partial f}{\partial x}(t, X_t)\ dX_t + \frac{1}{2}\ \frac{\partial^2 f}{\partial x^2}(t, X_t)\ dX_t\ dX_t \\
&= \underbrace{\left[\frac{\partial f}{\partial t}(t, X_t) + \mathbf{b_t(X_t)}\ \frac{\partial \mathbf{f}}{\partial \mathbf{x}}(\mathbf{t}, \mathbf{X_t}) + \frac{\mathbf{1}}{\mathbf{2}}\ \boldsymbol{\sigma}_\mathbf{t}^\mathbf{2}(\mathbf{X_t})\ \frac{\partial^2 \mathbf{f}}{\partial \mathbf{x}^2}(\mathbf{t}, \mathbf{X_t})\right]}\ dt + \underbrace{\frac{\partial f}{\partial x}(t, X_t)\ \boldsymbol{\sigma}_\mathbf{t}(\mathbf{X_t})\ \mathbf{dW_t}} \\
&= \left(\frac{\partial}{\partial t} + L_t\right)(f)(t, X_t)\quad dt + \underbrace{dM_t(f)}_{\textbf{martingale}}
\end{aligned}$$

Jump processes ($d = 1$)

- **Between jump times T_n**

$$dX_t = b_t(X_t)\ dt \ + \ \sigma_t(X_t)\ dW_t \qquad T_n \leq t < T_{n+1}$$

Jump processes ($d = 1$)

- **Between jump times T_n**

$$dX_t = b_t(X_t)\ dt\ +\ \sigma_t(X_t)\ dW_t \qquad T_n \leq t < T_{n+1}$$

- **Jump times at rate/with intensity $\lambda_t(x) \geq 0$ i.e.**

$$T_{n+1} = \inf\left\{ T_n \leq t\ :\ \int_{T_n}^{t} \lambda_s(X_s)\ ds\ \geq E_{n+1} \right\}$$

E_n are i.i.d. exponential r.v. with unit parameter.

Jump processes ($d = 1$)

- **Between jump times T_n**

$$dX_t = b_t(X_t)\ dt \ + \ \sigma_t(X_t)\ dW_t \qquad T_n \leq t < T_{n+1}$$

- **Jump times at rate/with intensity $\lambda_t(x) \geq 0$ i.e.**

$$T_{n+1} = \inf \left\{ T_n \leq t \ : \ \int_{T_n}^{t} \lambda_s(X_s)\ ds \ \geq E_{n+1} \right\}$$

E_n are i.i.d. exponential r.v. with unit parameter.

- **Jump selection (Markov) transition $J_t(x, dy)(\overset{\text{ex.}}{\propto} e^{-1/2(x-y)^2} dy)$:**

$$X_{T_{n+1}-} \ \rightsquigarrow \ X_{T_{n+1}} \text{ r.v. with distribution } J_{T_{n+1}}\left(X_{T_{n+1}-}, dx\right)$$

Bernoulli model on the time mesh "dt"

Given X_t, *description of the increment* $\Delta X_t = X_{t+dt} - X_t$

$$\mathbf{Y_t} = X_t + b_t(X_t)\ dt\ +\ \sigma_t(X_t)\ (W_{t+dt} - W_t)$$

$$\mathbb{P}\left(X_{t+dt} \in dx \mid \mathbf{Y_t}\right) = e^{-\lambda_t(\mathbf{Y_t})dt}\ \delta_{\mathbf{Y_t}}(dx) + \left(1 - e^{-\lambda_t(\mathbf{Y_t})dt}\right)\ J_t(\mathbf{Y_t}, dx)$$

Bernoulli model on the time mesh "dt"

Given X_t, *description of the increment* $\Delta X_t = X_{t+dt} - X_t$

$$\mathbf{Y_t} = X_t + b_t(X_t)\ dt\ +\ \sigma_t(X_t)\ (W_{t+dt} - W_t)$$

$$\mathbb{P}\left(X_{t+dt} \in dx \mid \mathbf{Y_t}\right) = e^{-\lambda_t(\mathbf{Y_t})dt}\ \delta_{\mathbf{Y_t}}(dx) + \left(1 - e^{-\lambda_t(\mathbf{Y_t})dt}\right)\ J_t(\mathbf{Y_t}, dx)$$

$$\Updownarrow$$

Bernoulli process

$$X_{t+dt} = (1 - \epsilon_t)\ \mathbf{Y_t}\ +\ \epsilon_t\ Z_t \quad \text{with} \quad Z_t\ \sim\ J_t(\mathbf{Y_t}, dx)$$

and the $\{0, 1\}$*-valued r.v.* ϵ_t *with jump probability*

$$\mathbb{P}\left(\epsilon_t = 1 \mid \mathbf{Y_t}\right) = 1 - e^{-\lambda_t(\mathbf{Y_t})dt} \simeq \lambda_t(\mathbf{Y_t})dt$$

Related Stoch. models

- $\sigma_t = 0 \rightsquigarrow$ Piecewise deterministic Markov processes (PDMP).

Related Stoch. models

- $\sigma_t = 0 \rightsquigarrow$ Piecewise deterministic Markov processes (PDMP).
- $b_t = 0 \oplus \sigma_t = 0 \rightsquigarrow$ Markov chain continuous time embedding

$$\mathbb{P}\left(\mathcal{X}_n \in dx \mid \mathcal{X}_{n-1}\right) = \mathcal{K}_n(\mathcal{X}_{n-1}, dx)$$

At jump times (exponential inter-times with unit parameter):

$$X_0 = \mathcal{X}_0 \quad \rightsquigarrow \quad X_{T_1} = \mathcal{X}_1 \quad \rightsquigarrow \quad X_{T_2} = \mathcal{X}_2 \ldots \quad \rightsquigarrow \quad X_{T_n} = \mathcal{X}_n$$

Also called *Pure jump processes*.

Related Stoch. models

- $\sigma_t = 0 \rightsquigarrow$ Piecewise deterministic Markov processes (PDMP).
- $b_t = 0 \oplus \sigma_t = 0 \rightsquigarrow$ Markov chain continuous time embedding

$$\mathbb{P}\left(\mathcal{X}_n \in dx \mid \mathcal{X}_{n-1}\right) = \mathcal{K}_n(\mathcal{X}_{n-1}, dx)$$

  At jump times (exponential inter-times with unit parameter):

$$X_0 = \mathcal{X}_0 \quad \rightsquigarrow \quad X_{T_1} = \mathcal{X}_1 \quad \rightsquigarrow \quad X_{T_2} = \mathcal{X}_2 \ldots \quad \rightsquigarrow \quad X_{T_n} = \mathcal{X}_n$$

  Also called *Pure jump processes*.
- $\lambda_t = 0 \rightsquigarrow$ Stochastic Diffusion equation (SDE).

Related Stoch. models

- $\sigma_t = 0 \rightsquigarrow$ Piecewise deterministic Markov processes (PDMP).
- $b_t = 0 \oplus \sigma_t = 0 \rightsquigarrow$ Markov chain continuous time embedding

$$\mathbb{P}\left(\mathcal{X}_n \in dx \mid \mathcal{X}_{n-1}\right) = \mathcal{K}_n(\mathcal{X}_{n-1}, dx)$$

At jump times (exponential inter-times with unit parameter):

$$X_0 = \mathcal{X}_0 \quad \rightsquigarrow \quad X_{T_1} = \mathcal{X}_1 \quad \rightsquigarrow \quad X_{T_2} = \mathcal{X}_2 \ldots \quad \rightsquigarrow \quad X_{T_n} = \mathcal{X}_n$$

Also called *Pure jump processes*.

- $\lambda_t = 0 \rightsquigarrow$ Stochastic Diffusion equation (SDE).
- $b_t = 0 \oplus \sigma_t = 0$
  - $\lambda_t(x) = \lambda \rightsquigarrow$ Poisson process $X_t = N_t$ with intensity λ.
  - $\lambda_t(x) = \lambda_t \rightsquigarrow$ Non-homogeneous Poisson process $\sim \lambda_t$.

Related Stoch. models

- $\sigma_t = 0 \rightsquigarrow$ Piecewise deterministic Markov processes (PDMP).
- $b_t = 0 \oplus \sigma_t = 0 \rightsquigarrow$ Markov chain continuous time embedding

$$\mathbb{P}\left(\mathcal{X}_n \in dx \mid \mathcal{X}_{n-1}\right) = \mathcal{K}_n(\mathcal{X}_{n-1}, dx)$$

At jump times (exponential inter-times with unit parameter):

$$X_0 = \mathcal{X}_0 \quad \rightsquigarrow \quad X_{T_1} = \mathcal{X}_1 \quad \rightsquigarrow \quad X_{T_2} = \mathcal{X}_2 \ldots \quad \rightsquigarrow \quad X_{T_n} = \mathcal{X}_n$$

Also called *Pure jump processes*.

- $\lambda_t = 0 \rightsquigarrow$ Stochastic Diffusion equation (SDE).
- $b_t = 0 \oplus \sigma_t = 0$
  - $\lambda_t(x) = \lambda \rightsquigarrow$ Poisson process $X_t = N_t$ with intensity λ.
  - $\lambda_t(x) = \lambda_t \rightsquigarrow$ Non-homogeneous Poisson process $\sim \lambda_t$.
- *General case* $\rightsquigarrow$ (Marked)-Jump-diffusion models.

Generators/Ito calculus

$$dX_t = dX_t^c + \Delta X_t$$

Generators/Ito calculus

$$dX_t = dX_t^c + \Delta X_t = dX_t^c \quad \textbf{or} \quad \Delta X_t$$

Generators/Ito calculus

$$dX_t = dX_t^c + \Delta X_t = dX_t^c \quad \textbf{or} \quad \Delta X_t$$

$$\Downarrow$$

$$df(t, X_t)$$

$$= f(t + dt, X_t + dX_t) - f(t, X_t)$$

$$= \underbrace{\frac{\partial f}{\partial t}(t, X_t)dt \; + \; f'(X_t)\; dX_t^c + \frac{1}{2}\; f''(X_t)\; dX_t^c dX_t^c}_{=\left[\frac{\partial}{\partial t} + L_t^c\right](f)(t, X_t)\; dt \; + \; dM_t^c(f)} + \Delta f(t, X_t)$$

with the infinitesimal generator

$$L_t^c = b_t \; \frac{\partial}{\partial x} + \frac{1}{2}\; \sigma_t^2 \; \frac{\partial^2}{\partial x^2}$$

and the martingale increment

$$dM_t^c(f) = f'(X_t)\; \sigma_t(X_t)\; dW_t$$

Generators/Ito calculus

$$dX_t = dX_t^c + \Delta X_t = dX_t^c \ \textbf{or} \ \Delta X_t$$

$$\Downarrow$$

$$df(t, X_t)$$

$$= f(t + dt, X_t + dX_t) - f(t, X_t)$$

$$= \underbrace{\frac{\partial f}{\partial t}(t, X_t)dt \ + \ f'(X_t)\ dX_t^c + \frac{1}{2}\ f''(X_t)\ dX_t^c dX_t^c}_{=\left[\frac{\partial}{\partial t} + L_t^c\right](f)(t, X_t)\ dt \ + \ dM_t^c(f)} + \Delta f(t, X_t)$$

with the infinitesimal generator

$$L_t^c = b_t\ \frac{\partial}{\partial x} + \frac{1}{2}\ \sigma_t^2\ \frac{\partial^2}{\partial x^2}$$

and the martingale increment

$$dM_t^c(f) = f'(X_t)\ \sigma_t(X_t)\ dW_t \qquad \textbf{AND} \qquad \boldsymbol{\Delta f(t, X_t)} \quad \textbf{???}$$

The jump generator

$$\mathbb{P}\left(T = t + dt\ , X_{t+dt} \in dx \mid X_t\right) = \lambda_t(X_t)\ dt\ S_t(X_t, dx)$$

$$\Downarrow$$

$$\begin{aligned}\mathbb{E}\left(\Delta f(t, X_t) \mid X_t = x\right) &= \lambda_t(x)\ dt \int \left(f(t,y) - f(t,x)\right)\ J_t(x, dy) \\ &:= L_t^d(f)(x)\ dt\end{aligned}$$

$$\Downarrow$$

Predictable and martingale parts

$$\begin{aligned}\Delta f(t, X_t) &= \mathbb{E}\left(\Delta f(t, X_t) \mid \mathcal{F}_t\right) + \Delta f(t, X_t) - \mathbb{E}\left(\Delta f(t, X_t) \mid \mathcal{F}_t\right) \\ &= L_t^d(f)(X_t)\ dt + dM_t^d(f)\end{aligned}$$

The angle brackets

$$dM_t^d(f) = \Delta f(t, X_t) - \overbrace{\mathbb{E}\left(\Delta f(t, X_t) \mid \mathcal{F}_t\right)}^{=\ldots dt}$$

$$\Downarrow$$

$$\begin{aligned}
\mathbb{E}\left(\left(dM_t^d(f)\right)^2 \mid \mathcal{F}_t\right) &= \mathbb{E}\left(\left(\Delta f(t, X_t)\right)^2 \mid \mathcal{F}_t\right) \\
&= \lambda_t(X_t)\ dt \int \left(f(t, y) - f(t, X_t)\right)^2\ J_t(X_t, dy) \\
&= L_t^d[(f - f(x))^2](x)_{|\ x = X_t}\ dt
\end{aligned}$$

The angle brackets

$$dM_t^d(f) = \Delta f(t, X_t) - \overbrace{\mathbb{E}\left(\Delta f(t, X_t) \mid \mathcal{F}_t\right)}^{=\ldots dt}$$

$$\Downarrow$$

$$\begin{aligned}
\mathbb{E}\left(\left(dM_t^d(f)\right)^2 \mid \mathcal{F}_t\right) &= \mathbb{E}\left(\left(\Delta f(t, X_t)\right)^2 \mid \mathcal{F}_t\right) \\
&= \lambda_t(X_t)\ dt \int \ (f(t, y) - f(t, X_t))^2 \ J_t(X_t, dy) \\
&= L_t^d[(f - f(x))^2](x)_{|\ x=X_t}\ dt = \Gamma_{L_t^d}(f, f)(X_t)\ dt
\end{aligned}$$

The angle brackets

$$dM_t^d(f) = \Delta f(t, X_t) - \overbrace{\mathbb{E}\left(\Delta f(t, X_t) \mid \mathcal{F}_t\right)}^{=\ldots dt}$$

$$\Downarrow$$

$$\begin{aligned}
\mathbb{E}\left((dM_t^d(f))^2 \mid \mathcal{F}_t\right) &= \mathbb{E}\left((\Delta f(t, X_t))^2 \mid \mathcal{F}_t\right) \\
&= \lambda_t(X_t)\ dt \int (f(t, y) - f(t, X_t))^2\ J_t(X_t, dy) \\
&= L_t^d[(f - f(x))^2](x)_{|\ x = X_t}\ dt = \Gamma_{L_t^d}(f, f)(X_t)\ dt
\end{aligned}$$

$$\Downarrow$$

The angle bracket of the martingale $M_t^d(f)$

$$\langle M^d(f) \rangle_t = \int_0^t \Gamma_{L_s^d}(f, f)(X_s)\ ds$$

Finally... the general rule

$$
\begin{aligned}
df(t, X_t) &= \left[\frac{\partial}{\partial t} + L_t^c\right](f)(t, X_t)\ dt\ +\ dM_t^c(f)\ +\ L_t^c(f)(t, X_t)\ dt + dM^d(f) \\
&= \left[\frac{\partial}{\partial t} + L_t\right](f)(t, X_t)\ dt\ +\ dM_t(f)
\end{aligned}
$$

with

$$
L_t = L_t^c + L_t^d \quad \text{and} \quad M_t(f) = M_t^c(f) + M_t^d(f)
$$

Finally... the general rule

$$\begin{aligned} df(t, X_t) &= \left[\frac{\partial}{\partial t} + L_t^c\right](f)(t, X_t)\ dt\ +\ dM_t^c(f)\ +\ L_t^c(f)(t, X_t)\ dt + dM^d(f) \\ &= \left[\frac{\partial}{\partial t} + L_t\right](f)(t, X_t)\ dt\ +\ dM_t(f) \end{aligned}$$

with

$$L_t = L_t^c + L_t^d \quad \text{and} \quad M_t(f) = M_t^c(f) + M_t^d(f)$$

and a martingale $M_t(f)$ with angle bracket

$$\frac{d}{dt}\ \langle M(f)\rangle_t\ = \Gamma_{L_t^c}(f, f)(X_t)\ +\ \Gamma_{L_t^d}(f, f)(X_t) = \Gamma_{L_t}(f, f)(X_t)$$

Finally... the general rule

$$df(t,X_t) = \left[\frac{\partial}{\partial t} + L_t^c\right](f)(t,X_t)\;dt \;+\; dM_t^c(f) \;+\; L_t^c(f)(t,X_t)\;dt + dM^d(f)$$
$$= \left[\frac{\partial}{\partial t} + L_t\right](f)(t,X_t)\;dt \;+\; dM_t(f)$$

with

$$L_t = L_t^c + L_t^d \quad \text{and} \quad M_t(f) = M_t^c(f) + M_t^d(f)$$

and a martingale $M_t(f)$ with angle bracket

$$\frac{d}{dt}\,\langle M(f)\rangle_t \;= \Gamma_{L_t^c}(f,f)(X_t) \;+\; \Gamma_{L_t^d}(f,f)(X_t) = \Gamma_{L_t}(f,f)(X_t)$$

Linear evolution equation

$$\eta_t = \mathrm{Law}(X_t) \rightsquigarrow \partial_t\eta_t = \boldsymbol{\Lambda}_t(\eta_t) := \eta_t L_t$$

Nonlinear processes/generators

$\overline{X}_t$ = E-valued process with generator depending on $\eta_t = \mathrm{Law}(X_t)$

$$L_{t,\eta_t}(f)(x) = L^c_{t,\eta_t}(f)(x) + \lambda_{t,\eta_t}(x) \int \left(f(t,y) - f(t,x)\right) \; J_{t,\eta_t}(x,dy)$$

Ex. $E = \mathbb{R}^d$ and between jumps L^c_{t,η_t} generator of a diffusion

$$d\overline{X}_t = b_t\left(\eta_t, \overline{X}_t\right)\; dt + \sigma_t\left(\eta_t, \overline{X}_t\right)\; dW_t$$

$$\Downarrow$$

Nonlinear evolution equation

$$\eta_t = \mathrm{Law}(\overline{X}_t) \rightsquigarrow \partial_t \eta_t = \boldsymbol{\Lambda}_t(\eta_t) := \eta_t L_{t,\eta_t}$$

N-interacting Markov processes

Nonlinear evolution equation

$$\eta_t = \mathrm{Law}(\overline{X}_t) \rightsquigarrow \partial_t \eta_t = \eta_t L_{t,\eta_t}$$

$\rightsquigarrow$ **Mean field simulation = N-interacting Markov processes $\xi_t = (\xi_t^i)_{1\leq i\leq N}$ evolving on E^N with generator**

$$\mathcal{L}_t(F)(x_1,\ldots,x_N) = \sum_{1\leq i\leq N} L_{t,m(x)}(F_{x_{-i}})(x_i)$$

$$F_{x_{-i}}(\cdot) = F(x_1,\ldots,x_{i-1},\cdot,x_{i+1},\ldots,x_N) \quad \text{and} \quad m(x) := \frac{1}{N}\sum_{1\leq i\leq N}\delta_{x_i}$$

Stochastic perturbation analysis

Itô formula

$$dF(\xi_t) = \mathcal{L}_t(F)(\xi_t) + d\mathcal{M}_t(F) \quad \text{with} \quad \partial_t \langle \mathcal{M}(F) \rangle_t = \Gamma_{\mathcal{L}_t}(F, F)(\xi_t)$$

Stochastic perturbation analysis

Itô formula

$$dF(\xi_t) = \mathcal{L}_t(F)(\xi_t) + d\mathcal{M}_t(F) \quad \text{with} \quad \partial_t \langle \mathcal{M}(F) \rangle_t = \Gamma_{\mathcal{L}_t}(F,F)(\xi_t)$$

Note: for any $x = (x_1, \ldots, x_N)$ and $F(x) = m(x)(f)$ we have

$$\mathcal{L}_t(F)(x) = m(x) L_{t,m(x)}(f) \quad \text{and} \quad \Gamma_{\mathcal{L}_t}(F,F)(x) = \frac{1}{N}\, m(x) \Gamma_{L_{t,m(x)}}(f,f)$$

Stochastic perturbation analysis

Itô formula

$$dF(\xi_t) = \mathcal{L}_t(F)(\xi_t) + d\mathcal{M}_t(F) \quad \text{with} \quad \partial_t \langle \mathcal{M}(F) \rangle_t = \Gamma_{\mathcal{L}_t}(F,F)(\xi_t)$$

Note: for any $x = (x_1, \ldots, x_N)$ and $F(x) = m(x)(f)$ we have

$$\mathcal{L}_t(F)(x) = m(x) L_{t,m(x)}(f) \quad \text{and} \quad \Gamma_{\mathcal{L}_t}(F,F)(x) = \frac{1}{N}\, m(x) \Gamma_{L_{t,m(x)}}(f,f)$$

Stochastic perturbation formulation:

$$\begin{cases} d\eta_t &= \Lambda_t(\eta_t)\, dt \\ d\eta_t^N &= \Lambda_t(\eta_t^N)\, dt + \frac{1}{\sqrt{N}}\, dM_t \quad \text{with} \quad \partial_t \langle M(f) \rangle_t = \eta_t^N \Gamma_{L_{t,\eta_t^N}}(f,f) \end{cases}$$

Stochastic perturbation analysis

Itô formula

$$dF(\xi_t) = \mathcal{L}_t(F)(\xi_t) + d\mathcal{M}_t(F) \quad \text{with} \quad \partial_t \langle \mathcal{M}(F) \rangle_t = \Gamma_{\mathcal{L}_t}(F,F)(\xi_t)$$

Note: for any $x = (x_1, \ldots, x_N)$ and $F(x) = m(x)(f)$ we have

$$\mathcal{L}_t(F)(x) = m(x) L_{t,m(x)}(f) \quad \text{and} \quad \Gamma_{\mathcal{L}_t}(F,F)(x) = \frac{1}{N}\, m(x) \Gamma_{L_{t,m(x)}}(f,f)$$

Stochastic perturbation formulation:

$$\begin{cases} d\eta_t & = & \Lambda_t(\eta_t)\, dt \\ d\eta_t^N & = & \Lambda_t(\eta_t^N)\, dt + \frac{1}{\sqrt{N}}\, dM_t \quad \text{with} \quad \partial_t \langle M(f) \rangle_t = \eta_t^N \Gamma_{L_{t,\eta_t^N}}(f,f) \end{cases}$$

$\rightsquigarrow$ **Need to study the stability of nonlinear sg.** $\eta_t := \Phi_{s,t}(\eta_s)$

Stochastic perturbation analysis

Itô formula

$$dF(\xi_t) = \mathcal{L}_t(F)(\xi_t) + d\mathcal{M}_t(F) \quad \text{with} \quad \partial_t \langle \mathcal{M}(F) \rangle_t = \Gamma_{\mathcal{L}_t}(F,F)(\xi_t)$$

Note: for any $x = (x_1, \ldots, x_N)$ and $F(x) = m(x)(f)$ we have

$$\mathcal{L}_t(F)(x) = m(x) L_{t,m(x)}(f) \quad \text{and} \quad \Gamma_{\mathcal{L}_t}(F,F)(x) = \frac{1}{N}\, m(x) \Gamma_{L_{t,m(x)}}(f,f)$$

Stochastic perturbation formulation:

$$\begin{cases} d\eta_t &= \Lambda_t(\eta_t)\, dt \\ d\eta_t^N &= \Lambda_t(\eta_t^N)\, dt + \dfrac{1}{\sqrt{N}}\, dM_t \quad \text{with} \quad \partial_t \langle M(f) \rangle_t = \eta_t^N \Gamma_{L_{t,\eta_t^N}}(f,f) \end{cases}$$

$\rightsquigarrow$ **Need to study the stability of nonlinear sg.** $\eta_t := \Phi_{s,t}(\eta_s)$

Ex.: EnKF $\rightsquigarrow \eta_t =$ **Law**$(\overline{X}_t) = \mathcal{N}(m_t, P_t)$**,...** $\rightsquigarrow$ **perturbation Kalman-filter**

Continuous time processes

(Linear) Langevin diffusions

- Stochastic gradient
- The MALA algorithm

Nonlinear/interacting particle samplers/SMC

Performance/Convergence analysis

MCMC/Langevin (gradient) diffusions

$X_t = (X_t^i)_{i \in I}$ diffusion $\in \mathbb{R}^I$

$$dX_t = -\beta \, \nabla U(X_t) \, dt + \sigma \, dB_t$$

with $\sigma = \sqrt{2}$ reversible w.r.t.

$$\pi_\beta(dx) \propto e^{-\beta \; U(x)} \underbrace{\nu(dx)}_{\text{Lebesgue measure on } \mathbb{R}^I}$$

Generator of X_t:

$$L(f) = \frac{1}{2} \sigma^2 \, e^{\frac{2\beta}{\sigma^2} \; U} \sum_{1 \leq k \leq I} \partial_{x_k} \left(e^{\frac{-2\beta}{\sigma^2} \; U} \, \partial_{x_k}(f) \right)$$

Reversibility property

$$\pi_\beta(f \, L(g)) = \pi_\beta(L(f) g)$$

MCMC/Langevin (gradient) diffusions

$X_t = (X_t^i)_{i\in I}$ diffusion $\in \mathbb{R}^I$

$$dX_t = -\beta\ \nabla U(X_t)\ dt + \ \sigma\ dB_t$$

with $\sigma = \sqrt{2}$ reversible w.r.t.

$$\pi_\beta(dx) \propto \quad e^{-\beta\ \ U(x)}\ \underbrace{\nu(dx)}_{\text{Lebesgue measure on } \mathbb{R}^I}$$

Generator of X_t:

$$L(f) = \frac{1}{2}\ \sigma^2\ e^{\frac{2\beta}{\sigma^2}\ \ U} \sum_{1\leq k\leq I} \partial_{x_k} \left(e^{\frac{-2\beta}{\sigma^2}\ \ U}\ \partial_{x_k}(f)\right)$$

Reversibility property

$$\pi_\beta(f\ L(g)) = \pi_\beta(L(f)g) \Longrightarrow \quad \pi_\beta L = 0$$

Nice and simple but in practice?

MALA

MALA = Metropolis Adjusted Langevin Algorithm

$$dX_t = -\beta \, \nabla U(X_t) + \sqrt{2} \, dW_t \quad \text{reversible w.r.t.} \quad \pi(x) \propto e^{-\beta U(x)}$$

Time discretisation $\Delta t = 1/m \rightsquigarrow$ **Discrete time Markov chain**

$$\mathcal{X}_{n+1} = \mathcal{X}_n - \beta \, \nabla U(\mathcal{X}_n)/m + \sqrt{2/m} \quad \underbrace{\mathcal{W}_{n+1}}_{\text{iid} \sim \mathcal{N}(0,\, I)}$$

MALA

MALA = Metropolis Adjusted Langevin Algorithm

$$dX_t = -\beta \nabla U(X_t) + \sqrt{2}\, dW_t \quad \text{reversible w.r.t.} \quad \pi(x) \propto e^{-\beta U(x)}$$

Time discretisation $\Delta t = 1/m \rightsquigarrow$ Discrete time Markov chain

$$\mathcal{X}_{n+1} = \mathcal{X}_n - \beta \nabla U(\mathcal{X}_n)/m + \sqrt{2/m} \underbrace{\mathcal{W}_{n+1}}_{\text{iid} \sim \mathcal{N}(0, I)}$$

NOT REVERSIBLE ANYMORE!! $\rightsquigarrow$ Solution??

MALA

MALA = Metropolis Adjusted Langevin Algorithm

$$dX_t = -\beta\, \nabla U(X_t) + \sqrt{2}\, dW_t \quad \text{reversible w.r.t.} \quad \pi(x) \propto e^{-\beta U(x)}$$

Time discretisation $\Delta t = 1/m \rightsquigarrow$ Discrete time Markov chain

$$\mathcal{X}_{n+1} = \mathcal{X}_n - \beta\, \nabla U(\mathcal{X}_n)/m + \sqrt{2/m}\, \underbrace{\mathcal{W}_{n+1}}_{\text{iid} \sim \mathcal{N}(0,I)}$$

NOT REVERSIBLE ANYMORE!! $\rightsquigarrow$ Solution??

Transition density

$$p_m(x,y) = \frac{1}{(4\pi/m)^{d/2}} \exp\left(-\frac{m}{4}\left\|y - x + \beta\, \nabla U(x)/m\right\|^2\right)$$

of the proposal

$$\mathcal{X}_n = x \rightsquigarrow \mathcal{Y}_{n+1} := \mathcal{X}_n - \beta\, \nabla U(\mathcal{X}_n)/m + \sqrt{2/m}\, \mathcal{W}_{n+1}$$

$\oplus$ **MH-acceptance/rejection with target $\pi_\beta(x)dx$!**

Continuous time processes

(Linear) Langevin diffusions

Nonlinear/interacting particle samplers/SMC
- Interpolating Gibbs measures
- Feynman-Kac formulation

Performance/Convergence analysis

Interpolating Gibbs/potential s.t. $\dot{U}_t := \partial_t U_t \geq 0$

$$\pi_{U_t}(dx) := \frac{1}{\nu(e^{-U_t})} \; e^{-U_t(x)} \; \nu(dx)$$

Interpolating Gibbs/potential s.t. $\dot{U}_t := \partial_t U_t \geq 0$

$$\pi_{U_t}(dx) := \frac{1}{\nu(e^{-U_t})}\; e^{-U_t(x)}\; \nu(dx)$$

Choose a π_{U_t}-shaker with generator L_t

$$\pi_{U_t} L_t = 0$$

Interpolating Gibbs/potential s.t. $\dot{U}_t := \partial_t U_t \geq 0$

$$\pi_{U_t}(dx) := \frac{1}{\nu(e^{-U_t})}\ e^{-U_t(x)}\ \nu(dx)$$

Choose a π_{U_t}-shaker with generator L_t

$$\pi_{U_t} L_t = 0$$

and note that

$$\partial_t \pi_{U_t}(f) \quad = \quad \frac{\nu(\partial_t U_t\ e^{-U_t})}{\nu(e^{-U_t})^2}\ \nu(fe^{-U_t}) - \frac{\nu(f\partial_t U_t e^{-U_t})}{\nu(e^{-U_t})}$$

Interpolating Gibbs/potential s.t. $\dot{U}_t := \partial_t U_t \geq 0$

$$\pi_{U_t}(dx) := \frac{1}{\nu(e^{-U_t})} \; e^{-U_t(x)} \; \nu(dx)$$

Choose a π_{U_t}-shaker with generator L_t

$$\pi_{U_t} L_t = 0$$

and note that

$$\begin{aligned}
\partial_t \pi_{U_t}(f) &= \frac{\nu(\partial_t U_t \; e^{-U_t})}{\nu(e^{-U_t})^2} \; \nu(fe^{-U_t}) - \frac{\nu(f \partial_t U_t e^{-U_t})}{\nu(e^{-U_t})} \\
&= \pi_{U_t}(\dot{U}_t \; (\pi_{U_t}(f) - f)) = (\pi_{U_t} L^d_{t,\pi_{U_t}})(f)
\end{aligned}$$

Interpolating Gibbs/potential s.t. $\dot{U}_t := \partial_t U_t \geq 0$

$$\pi_{U_t}(dx) := \frac{1}{\nu(e^{-U_t})} \; e^{-U_t(x)} \; \nu(dx)$$

Choose a π_{U_t}-shaker with generator L_t

$$\pi_{U_t} L_t = 0$$

and note that

$$\begin{aligned}
\partial_t \pi_{U_t}(f) &= \frac{\nu(\partial_t U_t \; e^{-U_t})}{\nu(e^{-U_t})^2} \; \nu(fe^{-U_t}) - \frac{\nu(f \partial_t U_t e^{-U_t})}{\nu(e^{-U_t})} \\
&= \pi_{U_t}(\dot{U}_t \; (\pi_{U_t}(f) - f)) = (\pi_{U_t} L^d_{t,\pi_{U_t}})(f) = \pi_{U_t}(L^d_{t,\pi_{U_t}}(f))
\end{aligned}$$

with the jump generator

$$L^d_{t,\pi_{U_t}}(f)(x) = \dot{U}_t(x) \; \int (f(y) - f(x)) \; \pi_{U_t}(dy)$$

Interpolating Gibbs/potential s.t. $\dot{U}_t := \partial_t U_t \geq 0$

$$\pi_{U_t}(dx) := \frac{1}{\nu(e^{-U_t})}\; e^{-U_t(x)}\; \nu(dx)$$

Choose a π_{U_t}-shaker with generator L_t

$$\pi_{U_t} L_t = 0$$

and note that

$$\begin{array}{rcl}
\partial_t \pi_{U_t}(f) & = & \dfrac{\nu(\partial_t U_t\; e^{-U_t})}{\nu(e^{-U_t})^2}\; \nu(fe^{-U_t}) - \dfrac{\nu(f\partial_t U_t e^{-U_t})}{\nu(e^{-U_t})} \\
 & = & \pi_{U_t}(\dot{U}_t\; (\pi_{U_t}(f) - f)) = (\pi_{U_t} L^d_{t,\pi_{U_t}})(f) = \pi_{U_t}(L^d_{t,\pi_{U_t}}(f))
\end{array}$$

with the jump generator

$$L^d_{t,\pi_{U_t}}(f)(x) = \dot{U}_t(x)\; \int (f(y) - f(x))\; \pi_{U_t}(dy)$$

Conclusion:

$$\eta_t := \pi_{U_t} \quad \text{and} \quad L_{t,\eta_t} := L_t + L^d_{t,\eta_t} \Longrightarrow \partial_t \eta_t = \eta_t L_{t,\eta_t}$$

Feynman-Kac formulation - any $(\mathcal{U}_t, L_t) = (\dot{U}_t, L_t)$

Nonlinear updating/Markov transport:

$$\partial_t \eta_t = \eta_t L_{t,\eta_t}$$

with

$$L_{t,\eta_t}(f)(x) = L_t(f)(x) + \mathcal{U}_t(x) \int (f(y) - f(x))\ \eta_t(dy)$$

$$\Updownarrow \quad \text{with} \quad L_t = \text{generator of } X_t$$

Feynman-Kac formulation - any $(\mathcal{U}_t, L_t) = (\dot{U}_t, L_t)$

Nonlinear updating/Markov transport:

$$\partial_t \eta_t = \eta_t L_{t,\eta_t}$$

with

$$L_{t,\eta_t}(f)(x) = L_t(f)(x) + \mathcal{U}_t(x) \int (f(y) - f(x))\ \eta_t(dy)$$

$$\Updownarrow \quad \text{with} \quad L_t = \text{generator of } X_t$$

FK solution/semigroup/formulation:

$$\eta_t(f) = \gamma_t(f)/\gamma_t(1) \quad \text{with} \quad \gamma_t(f) = \mathbb{E}\left(f(X_t) \exp\left(-\int_0^t \mathcal{U}_s(X_s) ds\right)\right)$$

$\rightsquigarrow$ **Genetic algo. (mutation,killing/selection rate)** $\sim (L_t, \mathcal{U}_t)$

Continuous time processes

(Linear) Langevin diffusions

Nonlinear/interacting particle samplers/SMC

Performance/Convergence analysis
- Variational approaches
- V-norm contraction
- Stochastic perturbation

Stability/Variational approaches/Diffusion flows

Stochastic flow $X_{s,t}(x)$ **with** $t \in [s,\infty[$ **starting at** $X_{s,s}(x) = x \in \mathbb{R}^d$

$$dX_{s,t}(x) = b_t\left(X_{s,t}(x)\right)\ dt + \sigma_t\left(X_{s,t}(x)\right)\ dW_t$$

- $\nabla\sigma = 0 \rightsquigarrow$ **First Variational eq.:**

$$\partial_t \nabla X_{s,t}(x) = \nabla X_{s,t}(x)\ \nabla b_t(X_{s,t}(x)) \quad \text{with} \quad \nabla X_{s,s}(x) = I$$

Stability/Variational approaches/Diffusion flows

Stochastic flow $X_{s,t}(x)$ with $t \in [s, \infty[$ starting at $X_{s,s}(x) = x \in \mathbb{R}^d$

$$dX_{s,t}(x) = b_t\left(X_{s,t}(x)\right)\ dt + \sigma_t\left(X_{s,t}(x)\right)\ dW_t$$

- $\nabla\sigma = 0 \rightsquigarrow$ **First Variational eq.:**

$$\partial_t \nabla X_{s,t}(x) = \nabla X_{s,t}(x)\ \nabla b_t(X_{s,t}(x)) \quad \text{with} \quad \nabla X_{s,s}(x) = I$$

$$X_{s,t}(x) - X_{s,t}(y) = \int_0^1\ \nabla X_{s,t}(\epsilon x + (1-\epsilon)y)'(x-y)\ d\epsilon$$

$$\Longrightarrow \|X_{s,t}(x) - X_{s,t}(y)\|^2 \leq \left[\int_0^1\ \|\nabla X_{s,t}(\epsilon x + (1-\epsilon)y)\|_2^2\ d\epsilon\right]\ \|x-y\|^2$$

Log-norm $\rho(A) = \lambda_{max}(A_{sym})$ and $\|A\|_2 := \lambda_{max}(A'A)^{1/2}$

Log-norm stability condition:

$$\frac{1}{2}\ \left(\nabla b_t(x) + \nabla b_t'(x)\right) \leq -\alpha\ I \Longrightarrow \rho\left(\nabla b_u(x)\right) \leq -\alpha$$

$\mathbb{L}_2$-norm for time varying/possibly random linear systems ($\nabla\sigma = 0$):

$$\log \|\nabla X_{s,t}(x)\|_2 \leq \int_s^t\ \rho\left(\nabla b_u(X_{s,u}(x))\right)\ du$$

Proof: log-norm....

$$\dot{v_t} = A_t v_t \Longrightarrow \partial_t \log \|v_t\| = \frac{\langle v_t, A_t v_t\rangle}{\|v_t\|^2} = \ldots$$

Log-norm $\rho(A) = \lambda_{max}(A_{sym})$ and $\|A\|_2 := \lambda_{max}(A'A)^{1/2}$

Log-norm stability condition:

$$\frac{1}{2}\ \left(\nabla b_t(x) + \nabla b_t'(x)\right) \leq -\alpha\ I \Longrightarrow \rho\left(\nabla b_u(x)\right) \leq -\alpha$$

$\mathbb{L}_2$-norm for time varying/possibly random linear systems ($\nabla\sigma = 0$):

$$\log \|\nabla X_{s,t}(x)\|_2 \leq \int_s^t\ \rho\left(\nabla b_u(X_{s,u}(x))\right)\ du$$

Proof: log-norm....

$$\dot{v}_t = A_t v_t \Longrightarrow \partial_t \log \|v_t\| = \frac{\langle v_t, A_t v_t\rangle}{\|v_t\|^2} = \ldots$$

$$\Downarrow$$

Theo. [Lyapunov first/indirect method (A.M. Lyapunov, PhD 1892)]

$$\rho\left(\nabla b_u(x)\right) \leq -\lambda \Longrightarrow \|X_{s,t}(x) - X_{s,t}(y)\| \leq e^{-\lambda(t-s)}\ \|x - y\|$$

Ex.: MCMC/Langevin (gradient) diffusions ($X_{0,t} = X_t$)

$$dX_t = -\beta \, \nabla U(X_t) \, dt + \sigma \, dB_t$$

$$\Downarrow$$

Direct proof with Frobenius norms:

$$\nabla^2 U \geq \lambda \, I \quad \Longrightarrow \quad \partial_t \nabla X_t = -\beta \, \nabla X_t \, \nabla^2 \mathbf{U}(\mathbf{X_t})$$

$$\Longrightarrow \quad \partial_t \nabla X_t \nabla X_t' = -2\beta \, \nabla X_t \, \nabla^2 U(X_t) \nabla X_t' \leq -2\alpha \nabla X_t \nabla X_t'$$

$$\Longrightarrow \quad \|\nabla X_t(z)\|_{Frob} \leq \sqrt{d} \, e^{-\lambda t}$$

$$\Longrightarrow \quad \|X_t(x) - X_t(y)\| \leq c \, e^{-\lambda t} \, \|x - y\|$$

More general diffusion flows

First variational equation

$$d\,\nabla X_{s,t} = \nabla X_{s,t}\ dC_{s,t} \quad \text{with} \quad \nabla X_{s,s}(x) = I$$

with the stochastic matrix diffusion

$$dC_{s,t} := \nabla b_t\,(X_{s,t})\ \ dt + \sum_{1\leq k\leq r} \nabla \sigma_{t,k}\,(X_{s,t})\ \ dW_t^k$$

More general diffusion flows

First variational equation

$$d\,\nabla X_{s,t} = \nabla X_{s,t}\; dC_{s,t} \quad \text{with} \quad \nabla X_{s,s}(x) = I$$

with the stochastic matrix diffusion

$$dC_{s,t} := \nabla b_t\left(X_{s,t}\right)\; dt + \sum_{1\leq k\leq r} \nabla \sigma_{t,k}\left(X_{s,t}\right)\; dW_t^k$$

Theorem [Lyapunov method for the tangent process]

$$\nabla b_t + \nabla b_t' + \sum_{1\leq k\leq r} \nabla \sigma_{k,t} \nabla \sigma_{k,t}' \leq -2\lambda\; I$$

$$\Downarrow$$

$$\mathbb{E}\left(\left\|\nabla X_{s,t}(x)\right\|^2\right)^{1/2} \leq c\; e^{-\lambda(t-s)}$$

V-norms contractions (discrete time approach)

V-Dobrushin contraction coef.: $\exists \tau > 0$ s.t.

$$\beta_V(P_{s,s+\tau}) \leq 1 - \delta_\tau \quad \text{and} \quad \sup_{|t-s|\leq\tau} |||P_{s,t}|||_V < \infty$$

V-norms contractions (discrete time approach)

V-Dobrushin contraction coef.: $\exists \tau > 0$ s.t.

$$\beta_V(P_{s,s+\tau}) \leq 1 - \delta_\tau \quad \text{and} \quad \sup_{|t-s| \leq \tau} |||P_{s,t}|||_V < \infty$$

$$\Downarrow$$

V-contraction: $\exists \lambda > 0$ s.t.

$$|||(\mu - \eta)P_{s,t}|||_V \leq c \; e^{-\lambda(t-s)} \; |||\mu - \eta|||_V$$

Time homogeneous $(P_{s,t} = P_{t-s})$: $\exists! \eta_\infty = \eta_\infty P \in \mathcal{P}_V(E)$

$$|||\mu P_t - \eta_\infty|||_V \leq c \; e^{-\lambda t} \; |||\mu - \eta_\infty|||_V$$

Diffusion flow interpolation (1 dim to simplify)

A couple of diffusion flows $(s \leq t)$ starting at x at time s:

$$
\begin{aligned}
dX_{s,t}(x) &= b_t\left(X_{s,t}(x)\right)\, dt + \sigma_t\left(X_{s,t}(x)\right)\, dW_t \\
d\overline{X}_{s,t}(x) &= \overline{b}_t\left(\overline{X}_{s,t}(x)\right)\, dt + \overline{\sigma}_t\left(\overline{X}_{s,t}(x)\right)\, dW_t
\end{aligned}
$$

Diffusion flow interpolation (1 dim to simplify)

A couple of diffusion flows $(s \leq t)$ starting at x at time s:

$$\begin{aligned} dX_{s,t}(x) &= b_t\left(X_{s,t}(x)\right)\ dt + \sigma_t\left(X_{s,t}(x)\right)\ dW_t \\ d\overline{X}_{s,t}(x) &= \overline{b}_t\left(\overline{X}_{s,t}(x)\right)\ dt + \overline{\sigma}_t\left(\overline{X}_{s,t}(x)\right)\ dW_t \end{aligned}$$

Semigroup formula

$$\forall s \leq u \leq t \qquad X_{s,t} = X_{u,t} \circ X_{s,u} \quad \text{and} \quad X_{s,s} = Id$$

Diffusion flow interpolation (1 dim to simplify)

A couple of diffusion flows $(s \leq t)$ starting at x at time s:

$$
\begin{aligned}
dX_{s,t}(x) &= b_t\left(X_{s,t}(x)\right)\ dt + \sigma_t\left(X_{s,t}(x)\right)\ dW_t \\
d\overline{X}_{s,t}(x) &= \overline{b}_t\left(\overline{X}_{s,t}(x)\right)\ dt + \overline{\sigma}_t\left(\overline{X}_{s,t}(x)\right)\ dW_t
\end{aligned}
$$

Semigroup formula

$$\forall s \leq u \leq t \qquad X_{s,t} = X_{u,t} \circ X_{s,u} \quad \text{and} \quad X_{s,s} = Id$$

Estimate $(\overline{X}_{s,t} - X_{s,t})$?

Diffusion flow interpolation (1 dim to simplify)

A couple of diffusion flows $(s \leq t)$ starting at x at time s:

$$\begin{aligned} dX_{s,t}(x) &= b_t\left(X_{s,t}(x)\right)\ dt + \sigma_t\left(X_{s,t}(x)\right)\ dW_t \\ d\overline{X}_{s,t}(x) &= \overline{b}_t\left(\overline{X}_{s,t}(x)\right)\ dt + \overline{\sigma}_t\left(\overline{X}_{s,t}(x)\right)\ dW_t \end{aligned}$$

Semigroup formula

$$\forall s \leq u \leq t \qquad X_{s,t} = X_{u,t} \circ X_{s,u} \quad \text{and} \quad X_{s,s} = Id$$

Estimate $(\overline{X}_{s,t} - X_{s,t})? \rightsquigarrow$ **key idea** $=$ Backward-Forward **Interpolation**

$$\overline{X}_{s,t} - X_{s,t} = \int_s^t d_u(X_{u,t} \circ \overline{X}_{s,u})$$

Refs.: `Backward Ito-Ventzell and stochastic interpolation formulae` (Arxiv-19/SPA-22) & `(CRAS-20)`

Forward-Backward interpolation $(\overline{a}, a) := (\overline{\sigma}^2, \sigma^2)$

$$d_u(X_{u,t} \circ \overline{X}_{s,u})(x) = (d_u X_{u,t})(\overline{X}_{s,u}(x))$$

$$+ (\partial X_{u,t})\,(\overline{X}_{s,u}(x)) \underbrace{d_u \overline{X}_{s,u}(x)}_{\overline{b}_u(\overline{X}_{s,u}(x))du + \overline{\sigma}_u(\overline{X}_{s,u}(x))dW_u} + \frac{1}{2}\,(\partial^2 X_{u,t})\,(\overline{X}_{s,u}(x))\; \overline{a}_u(\overline{X}_{s,u}(x))\; du$$

Forward-Backward interpolation $(\overline{a}, a) := (\overline{\sigma}^2, \sigma^2)$

$$d_u(X_{u,t} \circ \overline{X}_{s,u})(x) = (d_u X_{u,t})(\overline{X}_{s,u}(x))$$

$$+ (\partial X_{u,t})\, (\overline{X}_{s,u}(x)) \quad \underbrace{d_u \overline{X}_{s,u}(x)}_{\overline{b}_u(\overline{X}_{s,u}(x))du + \overline{\sigma}_u(\overline{X}_{s,u}(x))dW_u} \quad + \frac{1}{2}\, (\partial^2 X_{u,t})\, (\overline{X}_{s,u}(x))\ \overline{a}_u(\overline{X}_{s,u}(x))\ du$$

with the backward term

$$(d_u X_{u,t})(\overline{X}_{s,u}(x))$$

$$= -\partial X_{u,t}(\overline{X}_{s,u}(x))' \ \left(b_u(\overline{X}_{s,u}(x))du + \sigma_u(\overline{X}_{s,u}(x))\ dW_u \right)$$

$$-\frac{1}{2}\ \partial^2 X_{u,t}(\overline{X}_{s,u}(x))\ a_u(\overline{X}_{s,u}(x))\ du$$

Forward-Backward interpolation formula

$$d_u(X_{u,t} \circ \overline{X}_{s,u}) = \partial X_{u,t}(\overline{X}_{s,u}) \; \big((\overline{b}_u - b_u)(\overline{X}_{s,u}) \; du + (\overline{\sigma}_u - \sigma_u)(\overline{X}_{s,u}) \; dW_u\big)$$

$$+\frac{1}{2} \; \partial^2 X_{u,t}(\overline{X}_{s,u}) \; (\overline{a}_u - a_u)(\overline{X}_{s,u}) \; du$$

Forward-Backward interpolation formula

$$d_u(X_{u,t} \circ \overline{X}_{s,u}) = \partial X_{u,t}(\overline{X}_{s,u}) \ \left((\overline{b}_u - b_u)(\overline{X}_{s,u})\ du + (\overline{\sigma}_u - \sigma_u)(\overline{X}_{s,u})\ dW_u\right)$$

$$+\frac{1}{2}\ \partial^2 X_{u,t}(\overline{X}_{s,u})\ (\overline{a}_u - a_u)(\overline{X}_{s,u})\ du$$

Stability terms $\subset$ Gradient and Hessian of the stochastic flow $X_{s,t}(x)$

Tangent process, first variational equations, spectral techniques,...

Case: $\sigma = 0$ & $\overline{b} = b$ & $\overline{\sigma} = \epsilon\, \widehat{\sigma}$ (small perturbation)

$$\overline{X}_{s,t} - X_{s,t}$$

$$= \epsilon \overbrace{\int_s^t \partial X_{u,t}(\overline{X}_{s,u})\ \widehat{\sigma}_u(\overline{X}_{s,u})\ dW_u}^{\text{1st order=fluctuations}}$$

$$+\epsilon^2 \underbrace{\frac{1}{2} \int_s^t \partial^2 X_{u,t}(\overline{X}_{s,u})\ \widehat{a}_u(\overline{X}_{s,u})\ du}_{\text{2nd order= bias}}$$

Case: $\sigma = 0$ & $\overline{b} = b$ & $\overline{\sigma} = \epsilon\, \widehat{\sigma}$ (small perturbation)

$$\overline{X}_{s,t} - X_{s,t}$$

$$= \epsilon \overbrace{\int_s^t \partial X_{u,t}(\overline{X}_{s,u})\, \widehat{\sigma}_u(\overline{X}_{s,u})\, dW_u}^{\text{1st order=fluctuations}} + \epsilon^2 \underbrace{\frac{1}{2} \int_s^t \partial^2 X_{u,t}(\overline{X}_{s,u})\, \widehat{a}_u(\overline{X}_{s,u})\, du}_{\text{2nd order= bias}}$$

Stochastic perturbation analysis for mean field particle models:

- Stochastic Riccati *diffusions in matrix spaces* ($\sim$ EnKF) (+Bishop (Arxiv-17)/(IHP-20),...).
- McKean-Vlasov diff. = diff. $\in$ Hilbert/Fréchet derivatives (+Arnaudon (Arxiv-2018)/(SAA-19), (Arxiv-19)/(AAP-20)); Interacting-jumps, + Arnaudon $\rightsquigarrow$ (Arxiv-18)/(EJP-00).