- A theoretical result in higher dimension ($d,\ q \geqslant 1$):

Theorem (Multi-dimensional Milstein scheme)

Assume that b and σ are $\mathcal{C}^1$ on $\mathbb{R}^d$ with bounded α'_b and α'_σ-Hölder continuous existing partial derivatives respectively. Let $\alpha = \min(\alpha'_b, \alpha'_\sigma)$. Then, for every $p \in (0, +\infty)$, there exists a real constant $C_{p,b,\sigma,T} > 0$ such that for every $X_0 \in L^p(\mathbb{P})$, independent of the q-dimensional Brownian motion W, the error bound established when $d = q = 1$ remains valid i.e.

$$\Big\|\max_{0\leqslant k\leqslant n}\big|X_{t_k^n} - \bar{X}^{mil,n}_{t_k^n}\big|\Big\|_p \leqslant \Big\|\sup_{t\in[0,T]}\big|X_t - \bar{X}^{mil,n}_t\big|\Big\|_p \leqslant C_{b,\sigma,T,p}\Big(\frac{T}{n}\Big)^{\frac{1+\alpha}{2}}\big(1 + \|X_0\|_p\big)$$

where $\alpha = \min(\alpha_{b'}, \alpha_{\sigma'})$.

- When $\alpha_{b'} = \alpha_{\sigma'} = 1$ (Lipschitz continuous partial derivative) then order 1.
- A major drawback: In general the continuous Milstein scheme cannot be simulated.
- Mostly useful at time T.

Main result on Milstein scheme?

- Obvious but important corollary: if $\sigma(x) = \Sigma_0 \in \mathcal{M}(d, q, \mathbb{R})$ then the commutation condition holds (!) and Euler = Milstein (discrete time and genuine) so that...
- Main result on Milstein scheme?

Corollary (Euler scheme with constant diffusion coefficient)

If the drift $b \in \mathcal{C}^2(\mathbb{R}^d, \mathbb{R}^d)$ with bounded existing partial derivatives and if $\sigma(x) = \Sigma_0 \in \mathcal{M}(d, q, \mathbb{R})$ is constant, then the (discrete time) Euler and Milstein schemes coincide (in particular) at discretization times t_k^n.
As a consequence, for every $p \in (0, +\infty)$, there exists a real constant $C_{p,b,\sigma,T} > 0$

$$\Big\| \max_{0 \leqslant k \leqslant n} |X_{t_k^n} - \bar{X}^n_{t_k^n}| \Big\|_p \leqslant C_{p,b,\sigma,T} \frac{T}{n} \big(1 + \|X_0\|_p\big).$$

Stage report: What are we interested in ?

- We want to approximate $\mathbb{E}\, f(X_T)$ (vanilla payoffs) or $\mathbb{E}\, F\big((X_t)_{t\in[0,T]}\big)$ (Path-dependent/exotic payoffs), where $X = (X_t)_{t\in[0,T]}$ is a Brownian diffusion.
- X cannot be simulated in an exact way (except in some special cases...).
- We introduce time discretization schemes $(\widetilde{X}_{t_k^n})_{k=0,\ldots,n}$, say Euler, that can be simulated.
- It results, e.g. in the vanila case, a **bias**: $\mathbb{E}\, f(X_T) - \mathbb{E}\, f(\bar{X}_T^n)$.
- When f is Lipschitz continuous, elementary considerations yield

$$(**) \quad \Big|\mathbb{E}\, f(X_T) - \mathbb{E}\, f(\bar{X}_T^n)\Big| \leqslant [f]_{\rm Lip}\mathbb{E}\big|X_T - \bar{X}_T^n\big| = [f]_{\rm Lip}\, O\Big(\frac{1}{\sqrt{n}}\Big).$$

- With the Milstein scheme, when $\bar{X}_T^{n,mil}$ is simulable, we get $= O\Big(\dfrac{1}{n}\Big)$.
- $(**)$ is a brutal switch from weak to strong error.

- Same phenomenon with path-dependent payoffs functionals if $F : \underbrace{\mathbb{D}([0,T],\mathbb{R})}_{\text{space of càdlàg fcts}} \to \mathbb{R}$, supposed $\|\cdot\|_{\sup}$-Lipschitz i.e.

$$\forall\,(x,y)\in\mathbb{D}([0,T],\mathbb{R}), x=(x(t))_{t\in[0,T]},\; y=(y(t))_{t\in[0,T]}$$
$$\big|F(x)-F(y))\big|\leqslant [F]_{\rm Lip}\|x-y\|_{\sup}.$$

- Then

$$\begin{aligned}\Big|\mathbb{E}\,F\big((X_t)_{t\in[0,T]}\big)-\mathbb{E}\,F\big((\widetilde X_t)_{t\in[0,T]}\big)\Big| &\leqslant [F]_{\rm Lip}\mathbb{E}\Big[\sup_{t\in[0,T]}\big|X_t-\widetilde X^n_t\big|\Big]\\ &=[F]_{\rm Lip}\Big\|\sup_{t\in[0,T]}\big|X_t-\widetilde X^n_t\big|\Big\|_1\\ &=O\Big(\sqrt{\frac{1+\log n}{n}}\Big)\end{aligned}$$

- with either Euler or Milstein scheme.

Did we miss something?

We are more interested in the so-called weak error

$$\left|\mathbb{E}\, f(X_T) - \mathbb{E}\, f(\bar{X}^n_T)\right| \quad \text{or} \quad \left|\mathbb{E}\, F\big((X_t)_{t\in[0,T]}\big) - \mathbb{E}\, F\big((\widetilde{X}_t)_{t\in[0,T]}\big)\right|$$

than in the pathwise convergence itself of the scheme in L^p for the sup-norm.

Weak error expansions (Brownian diffusions and vanilla payoffs)

The following theorems have been established for the Euler scheme.

Theorem (I Smooth function f (Talay-Tubaro's Theorem))

Assume b and σ are $\mathcal{C}^\infty$ with bounded partial derivatives. Assume $f:\mathbb{R}^d\to\mathbb{R}$ is $\mathcal{C}^\infty$ with partial derivative having polynomial growth. Then, for every order $R\in\mathbb{N}$, the expansion

$$(\mathcal{E}_{R+1}) \;\equiv\; \mathbb{E}\,f(\bar{X}^{x,n}_T) = \mathbb{E}\,f(X^x_T) + \sum_{k=1}^{R}\frac{c_k}{n^k} + O(n^{-(R+1)}) \quad as \quad n\to+\infty$$

where the real valued coefficients $c_k = c_k(f, T, b, \sigma, x)$ depend on f, T, b, σ and x.

Remark. To obtain only $(\mathcal{E}_{R+1})$ for a fixed order R, replace $\mathcal{C}^\infty$ by $\mathcal{C}^{R+5}$.

Sketch of proof, Step I: Feynman-Kac formula

- Set $d = q = 1$, $b(t,x) = b(x)$ and $\sigma(t,x) = \sigma(x)$ for simplicity.
- Define the infinitesimal generator $\mathcal{L}$ of the diffusion on every $g \in C^2(\mathbb{R},\mathbb{R})$ by

$$\mathcal{L}g = bg' + \frac{1}{2}\sigma^2 g''$$

[in higher dimension $\mathcal{L}g = (b\,|\,\nabla g) + \frac{1}{2}\mathrm{Tr}(\sigma D^2 g \sigma^*)$].

- Then Itô's formula can be rewritten for $t \in [0,T]$,

$$g(X_t) = g(X_0) + \int_0^t Lg(X_s)ds + \int_0^t g'(X_s)\sigma(X_s)dW_s.$$

- Assume b, σ are smooth enough with bounded existing (partial) derivatives and and $f : \mathbb{R} \to \mathbb{R}$ is smooth enough having existing (partial) derivatives with polynomial growth.
- Then the *PDE*

$$\begin{cases} \frac{\partial u}{\partial t}(t,x) + \mathcal{L}u(t,x) = 0, & (t,x) \in [0,T) \times \mathbb{R}^d \\ u(T,\cdot) = f \end{cases}$$

has a (unique classical) solution $u \in C^{1,2}\big([0,T] \times \mathbb{R},\mathbb{R}\big)$ with space derivatives having polynomial growth.

- By Itô's formula:

$$
\begin{aligned}
f(X_T^x) &= u(T, X_T^x) \\
&= u(0,x) + \int_0^T \underbrace{\frac{\partial u}{\partial t} + \mathcal{L}u(s, X_s^x)}_{=0!}\, ds + \int_0^t \partial_x u(s, X_s^x)\sigma(X_s^x) dW_s \\
&= u(0,x) + \underbrace{\int_0^T \partial_x u(s, X_s^x)\sigma(X_s^x) dW_s}_{\text{true martingale !}}
\end{aligned}
$$

since $\sup_{s\in[0,T]} |X_s^x| \in L^p(\mathbb{P})$, $\forall\, p > 0$ and $\partial_x u(s,x)$ has polynomial growth.

- Then, and that's why *PDE* pricers exist

$$u(0,x) = \mathbb{E}\, f(X_T^x).$$

- More generally writing Itô's formula between t and T yields

$$\begin{aligned} f(X_T^x) &= u(T, X_T^x) \\ &= u\big(t, X_t^x\big) + \int_t^T \partial_x u(s, X_s^x)\sigma(X_s^x)dW_s \end{aligned}$$

so that

$$u(t, X_t^x) = \mathbb{E}\big(f(X_T^x)\,|\,\mathcal{F}_t\big).$$

- As $u(t, X_t^x)$ is a function of X_t^x which is $\mathcal{F}_t$-measurable, taking conditional expectation $\mathbb{E}(\cdot\,|\,X_t^x)$ then yields

$$u(t, X_t^x) = \mathbb{E}\big(f(X_T^x)\,|\,X_t^x\big).$$

- Which in turn implies

$$\forall\, t \in [0, T],\ \forall\, y \in \mathbb{R}, \quad u(t, y) = \mathbb{E}\big(f(X_T)\,|\,X_t = y\big) = \mathbb{E}\, f(X_{T-t}^y).$$

Step II: The *PDE* method.

- Domino decomposition (having in mind that $\bar X_0^{n,x} = X_0^x = x$)

$$\mathbb{E}\, f(\bar X_T^{n,x}) - \mathbb{E}\, f(X_T^x) = \mathbb{E}\big(u(T,\bar X_T^{n,x}) - u(0,x)\big)$$
$$= \sum_{k=1}^{n} \mathbb{E}\big(u(t_k,\bar X_{t_k}^{n,x}) - u(t_{k-1},\bar X_{t_{k-1}}^{n,x})\big)$$

- Apply Itô's formula to $u(t_k,\bar X_{t_k}^{n,x}) - u(t_{k-1},\bar X_{t_{k-1}}^{n,x})$, namely

$$u(t_k,\bar X_{t_k}^{n,x}) - u(t_{k-1},\bar X_{t_{k-1}}^{n,x}) = \int_{t_{k-1}}^{t_k} \frac{\partial u}{\partial t} + \bar{\mathcal{L}} u(t_{k-1},s,\bar X_s^n)ds + \int_{t_{k-1}}^{t_k} [\cdots]dW_s$$

with the local pseudo-infinitesimal generator reading on u at t,s,y:

$$\bar{\mathcal{L}} u(t,s,y) = b(\bar X_t^n)\partial_x u(s,y) + \tfrac{1}{2}\sigma^2(\bar X_t^n)\partial_{x^2}^2 u(s,y).$$

- Hence

$$\mathbb{E}\, f(\bar X_T^{n,x}) - \mathbb{E}\, f(X_T^x) = \sum_{k=1}^{n} \int_{t_{k-1}}^{t_k} \mathbb{E}\left[\frac{\partial u}{\partial t} + \bar{\mathcal{L}} u(t_{k-1},s,\bar X_s^n)\right] ds.$$

- Using that u solves the *PDE* i.e. $\frac{\partial u}{\partial t} = -Lu$ one derives

$$\mathbb{E}\, f(\bar{X}^{n,x}_T) - \mathbb{E}\, f(X^x_T) = \mathbb{E}\big(u(T,\bar{X}^{n,x}_T) - u(0,x)\big)$$
$$= \sum_{k=1}^{n} \mathbb{E}\Big[\int_{t_{k-1}}^{t_k} \big(\bar{\mathcal{L}} - \mathcal{L}\big)u(t_{k-1}, s, \bar{X}^n_s)ds\big)\Big]$$

- where

$$(\bar{\mathcal{L}} - \mathcal{L})u(t_{k-1}, s, \bar{X}^n_s) = \big(b(\bar{X}^n_{t_{k-1}}) - b(\bar{X}^n_s)\big)\partial_x u(s,\bar{X}^n_s)$$
$$+ \tfrac{1}{2}\big(\sigma^2(\bar{X}^n_{t_{k-1}}) - \sigma^2(\bar{X}^n_s)\big)\partial^2_{x^2} u(s,\bar{X}^n_s).$$

- Two more Itô's formulas applied to the increments in b and σ + Strong error (at last) show that

$$\left|\mathbb{E}\Big[\int_{t_{k-1}}^{t_k} \big(\bar{\mathcal{L}} - \mathcal{L}\big)u(t_{k-1}, s, \bar{X}^n_s)ds\right| \leqslant \frac{C}{n^2}.$$

- Which finally yields the $\dfrac{C}{n}$ first order weak error bound.

Theorem (II (Uniformly) elliptic diffusion (Bally-Talay's Theorem))

If b and σ are bounded, $\mathcal{C}^\infty$ with bounded partial derivatives and if σ is uniformly elliptic i.e.

$$\exists\, \varepsilon_0 > 0 \text{ such that } \forall\, x \in \mathbb{R}^d, \quad \sigma\sigma^*(x) \geqslant \varepsilon_0 I_d,$$

then the above expansions holds true $(\mathcal{E}_{R+1})$ for any bounded Borel function f.

- Uniform ellipticity can be relaxed into

$$\forall\, x \in \mathbb{R}^d, \quad \mathrm{Span}\big(\sigma_{.j}(x),\, j = 1 : q\big) = \mathbb{R}^d \qquad (\Rightarrow q \geqslant d).$$

  or even parabolic Hörmander assumption.
- The Milstein scheme satisfies the same weak error expansions under similar assumptions.
- Taking this into account, the Milstein scheme may seem at this stage of little interest since it is more complex.However it requires less regularity (when simulable . . .).

- Striking (and unexpected) result !
- Why? Because the weak error of Milstein scheme is also of order 1.
- The competitive advantage of Milstein scheme becomes less clear, even when simulable: (sometimes) less regularity on b and σ and f.

How to take advantage of higher order weak error to reduce/kill the bias?

Application to Richardson-Romberg extrapolation

- We consider a priori two diffusions satisfying (under strong uniqueness assumption) [weak uniqueness is ok !]

$$dX_t^{(i)} = b(X_t^{(i)})dt + \sigma(X_t^{(i)})dW_t^{(i)}, \quad X_0^{(i)} = x_0,\ i = 1,2$$

where $W^{(i)}$, $i = 1,2$ are two standard q-Brownian motions on $(\Omega, \mathcal{A}, \mathbb{P})$. The processes $X^{(i)}$ have the same distribution, say of $(X_t)_{t\in[0,T]}$ (solution to $dX_t = b(X_t)dt + \sigma(X_t)dW_t$, $X_0 = x_0$) i.e.

$$\mathcal{L}\big((X_t^{(1)})_{t\in[0,T]}\big) \stackrel{d}{=} \mathcal{L}\big((X_t^{(2)})_{t\in[0,T]}\big)$$

- We associate the two (genuine) Euler schemes with respective time steps $\frac{T}{n}$ for $X^{(1)}$ and $\frac{T}{2n}$ for $X^{(2)}$ (each with "its own $\underline{t}$")

$$d\bar{X}_t^{n,(1)} = b(\bar{X}_{\underline{t}_n}^{n,(1)})dt + \sigma(\bar{X}_{\underline{t}_n}^{n,(1)})dW_t^{(1)}, \quad X_0^{(1)} = x_0,$$

$$d\bar{X}_t^{2n,(2)} = b(\bar{X}_{\underline{t}_{2n}}^{2n,(2)})dt + \sigma(\bar{X}_{\underline{t}_{2n}}^{2n,(2)})dW_t^{(2)}, \quad X_0^{(2)} = x_0.$$

- Let $f : \mathbb{R}^d \to \mathbb{R}$. Under the former appropriate smoothness/ellipticity assumptions, assume one has $(\mathcal{E}_3)$

$$\mathbb{E}\, f\big(\bar{X}_T^{n,(1)}\big) = \mathbb{E}\, f\big(X_T\big) + \frac{c_1}{n} + \frac{c_2}{n^2} + O(n^{-3})$$

$$\mathbb{E}\, f\big(\bar{X}_T^{2n,(2)}\big) = \mathbb{E}\, f\big(X_T\big) + \frac{c_1}{2n} + \frac{c_2}{4n^2} + O((2n)^{-3})$$

- Then

$$2\,\mathbb{E}\, f\big(\bar{X}_T^{2n,(2)}\big) - \mathbb{E}\, f\big(\bar{X}_T^{n,(1)}\big) = (2-1)\mathbb{E}\, f\big(X_T\big) + c_1\Big(\frac{2}{2n} - \frac{1}{n}\Big) + \frac{c_2}{n^2}\Big(\frac{2}{4} - 1\Big) + O(n^{-3})$$

- i.e.

$$\mathbb{E}\Big[2\, f\big(\bar{X}_T^{2n,(2)}\big) - f\big(\bar{X}_T^{n,(1)}\big)\Big] = \mathbb{E}\, f\big(X_T\big) - \frac{c_2}{2n^2} + O(n^{-3}).$$

- This meta-scheme provides a one order gain for the weak error i.e. the bias is fades at order 1 with n.
- (First) Price to pay. The complexity is tripled (rather low price).

What about the variance ?

- Under appropriate assumptions (Lipschitz continuity on b and σ, continuity and polynomial growth for f)

$$\mathrm{Var}\big(2\,f\big(\bar X_T^{2n,(2)}\big) - f\big(\bar X_T^{n,(1)}\big)\big) \stackrel{n\to+\infty}{\longrightarrow} \mathrm{Var}\big(2\,f\big(X_T^{(2)}\big) - f\big(X_T^{(1)}\big)\big)$$

- We know that Standard deviation $\mathrm{Std}(U) = \sqrt{\mathrm{Var}(U)} = \|U - \mathbb{E}\,U\|_2$ satisfies a triangle inequality

$$\mathrm{Std}(U+V) \leqslant \mathrm{Std}(U) + \mathrm{Std}(V).$$

with equality iff $V = v \in \mathbb{R}$ *a.s.* or $U = \lambda(V+\mu)$ *a.s.* for some (deterministic) real numbers $\lambda > 0$ and μ).

- Using the resulting reverse triangle inequality with $U \stackrel{d}{=} V \in L^2(\mathbb{P})$,

$$\mathrm{Std}(2U - V) \geqslant \mathrm{Std}(2U) - \mathrm{Std}(V) = 2\,\mathrm{Std}(U) - \mathrm{Std}(V) = \mathrm{Std}(U)$$

with equality (Exercise !) iff $U = V$.

Optimal choice

- Richardson-Romberg extrapolation (R-R) increases the asymptotic variance.
- Except if one choose the same two underlying diffusions $X^{(1)} = X^{(2)}$ i.e. in practice
- the same standard q-Brownian motions

$$W^{(1)} = W^{(2)} = W$$

- ℵ (Toward) Practitioner's corner. Joint simulations of the two (discrete time) Euler schemes from the same underlying Brownian motion W means

Two consistent $\mathcal{N}(0, I_q)$ white noises.

ℵ Practitioner's corner

- **Example (Vanilla options).**

$$\bar{Y}^{(h)} = f(\bar{X}^n_T) \quad \text{and} \quad \bar{Y}^{(\frac{h}{2})} = f(\bar{X}^{2n}_T)$$

- How to simulate consistent Brownian increments?

  Set $t^n_k = k\frac{T}{n} = kh$ and $t^{2n}_k = k\frac{T}{2n} = k\frac{h}{2}$.

$$W_{t^n_{k+1}} - W_{t^n_k} = \big(W_{t^{2n}_{2(k+1)}} - W_{t^{2n}_{2k+1}}\big) + \big(W_{t^{2n}_{2k+1}} - W_{t^{2n}_{2k}}\big).$$

- We start by the fine scheme

$$W_{t^{2n}_{2k+1}} - W_{t^{2n}_{2k}} = \sqrt{\frac{T}{2n}}Z_{2k+1} \quad \text{and} \quad W_{t^{2n}_{2(k+1)}} - W_{t^{2n}_{2k+1}} = \sqrt{\frac{T}{2n}}Z_{2(k+1)}$$

and, for the coarse scheme

$$W_{t^n_{k+1}} - W_{t^n_k} = \sqrt{\frac{T}{2n}}Z_{2k+1} + \sqrt{\frac{T}{2n}}Z_{2(k+1)} = \sqrt{\frac{T}{n}}\frac{Z_{2k+1} + Z_{2(k+1)}}{\sqrt{2}}$$

- **Warning !!** A lazy approach (calling sequentially the instruction "Random") amounts to assume

$$W^{(1)} \perp\!\!\!\perp W^{(2)}$$

yields

$$\begin{aligned}
\mathrm{Var}\big(2\,f(\bar X^{2n,(2)}_T) - f(\bar X^{n,(1)}_T)\big) &\overset{n\to+\infty}{\longrightarrow} \mathrm{Var}\big(2\,f(X^{(2)}_T) - f(X^{(1)}_T)\big)\\
&= \mathrm{Var}\big(2\,f(X^{(2)}_T)\big) + \mathrm{Var}\big(f(X^{(1)}_T)\big)\\
&= (4+1)\mathrm{Var}\big(f(X_T)\big) = 5\,\mathrm{Var}\big(f(X_T)\big)!!
\end{aligned}$$