Analysis of multilevel Milstein scheme without Lévy areas

Mike Giles

Lukas Szpruch

mike.giles,lukas.szpruch@maths.ox.ac.uk

Oxford University Mathematical Institute

Oxford-Man Institute of Quantitative Finance

FoCM, Budapest, July 8-10, 2011

Outline

- Milstein discretisation and multilevel method
- Clark & Cameron model problem
- antithetic treatment and analysis
- generalisation

Milstein discretisation

The Milstein discretisation of the SDE

$$\mathrm{d}S_i(t) = a_i(S)\,\mathrm{d}t + \sum_j b_{ij}(S)\,\mathrm{d}W_j(t), \quad 0<t<T$$

is

$$\begin{aligned}\widehat{S}_{i,n+1} &= \widehat{S}_{i,n} + a_i\,h + \sum_j b_{ij}\,\Delta W_{j,n} \\ &+ \tfrac{1}{2}\sum_{j,k,l} \frac{\partial b_{ij}}{\partial S_l}\, b_{lk} \left(\Delta W_{j,n}\,\Delta W_{k,n} - \Omega_{jk}\,h - A_{jk,n}\right)\end{aligned}$$

where Ω_{jk} is the correlation matrix, and the Lévy areas are

$$A_{jk,n} = \int_{t_n}^{t_{n+1}} (W_j(t){-}W_j(t_n))\;\mathrm{d}W_k - (W_k(t){-}W_k(t_n))\;\mathrm{d}W_j$$

Standard Multilevel approach

To estimate $\mathbb{E}[P]$, where the payoff $P = f(S_T)$ can be approximated by $\widehat{P}_\ell$ using 2^ℓ uniform timesteps, we use

$$\mathbb{E}[\widehat{P}_L] = \mathbb{E}[\widehat{P}_0] + \sum_{\ell=1}^{L} \mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}].$$

$\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$ is estimated using N_ℓ simulations with same $W(t)$ for both $\widehat{P}_\ell$ and $\widehat{P}_{\ell-1}$,

$$\widehat{Y}_\ell = N_\ell^{-1} \sum_{i=1}^{N_\ell} \left(\widehat{P}_\ell^{(i)} - \widehat{P}_{\ell-1}^{(i)} \right)$$

Because of strong convergence, on finer levels $\mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$ is small and so few paths are required.

Modified Multilevel approach

Sometimes better to use a different approximation for $\widehat{P}_\ell$ in $\mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}]$ and $\mathbb{E}[\widehat{P}_{\ell+1} - \widehat{P}_\ell]$. The decomposition

$$\mathbb{E}[\widehat{P}_L^f] = \mathbb{E}[\widehat{P}_0^f] + \sum_{\ell=1}^{L} \mathbb{E}[\widehat{P}_\ell^f - \widehat{P}_{\ell-1}^c]$$

is still a valid telescoping sum provided $\mathbb{E}[\widehat{P}_\ell^f] = \mathbb{E}[\widehat{P}_\ell^c]$.

In this work, we use $\widehat{P}_\ell^c = f(\widehat{S}_\ell^c)$ and

$$\widehat{P}_\ell^f = \tfrac{1}{2}\left(f(\widehat{S}_\ell^{f1}) + f(\widehat{S}_\ell^{f2}) \right)$$

where $f1$ is the fine path, and $f2$ is an “antithetic twin”.

Clark & Cameron problem

In their 1980 paper, Clark & Cameron considered the model problem:

$$\begin{aligned} \mathrm{d}X &= \mathrm{d}W_1 \\ \mathrm{d}Y &= X\, \mathrm{d}W_2 \end{aligned}$$

for independent Brownian paths W_1, W_2 and $X(0)\!=\!Y(0)\!=\!0$.

This can be integrated to give $X(t) = W_1(t)$ and

$$\begin{aligned} Y(t) &= \int_0^t W_1(s)\, \mathrm{d}W_2(s) \\ &= \tfrac{1}{2}\, W_1(t)\, W_2(t) + \tfrac{1}{2} \int_0^t W_1(s)\, \mathrm{d}W_2(s) - W_2(s)\, \mathrm{d}W_1(s) \end{aligned}$$

Clark & Cameron problem

If we consider a set of times $t_n = n\, h$, then we get

$$Y(t_{n+1}) = Y(t_n) + X(t_n)\, \Delta W_{2,n} + \tfrac{1}{2} \Delta W_{1,n}\, \Delta W_{2,n} + \tfrac{1}{2} A_n,$$

where $\Delta W_{j,n} \equiv W_j(t_{n+1}) - W_j(t_n)$ and

$$A_n = \int_{t_n}^{t_{n+1}} W_1(s)\, \mathrm{d}W_2(s) - W_2(s)\, \mathrm{d}W_1(s).$$

This matches exactly the Milstein discretisation – i.e. the Milstein discretisation is exact for this problem

Clark & Cameron problem

Summing over n gives

$$Y(T) = \sum_n \left(X(t_n)\, \Delta W_{2,n} + \tfrac{1}{2} \Delta W_{1,n}\, \Delta W_{2,n} + \tfrac{1}{2} A_n \right)$$

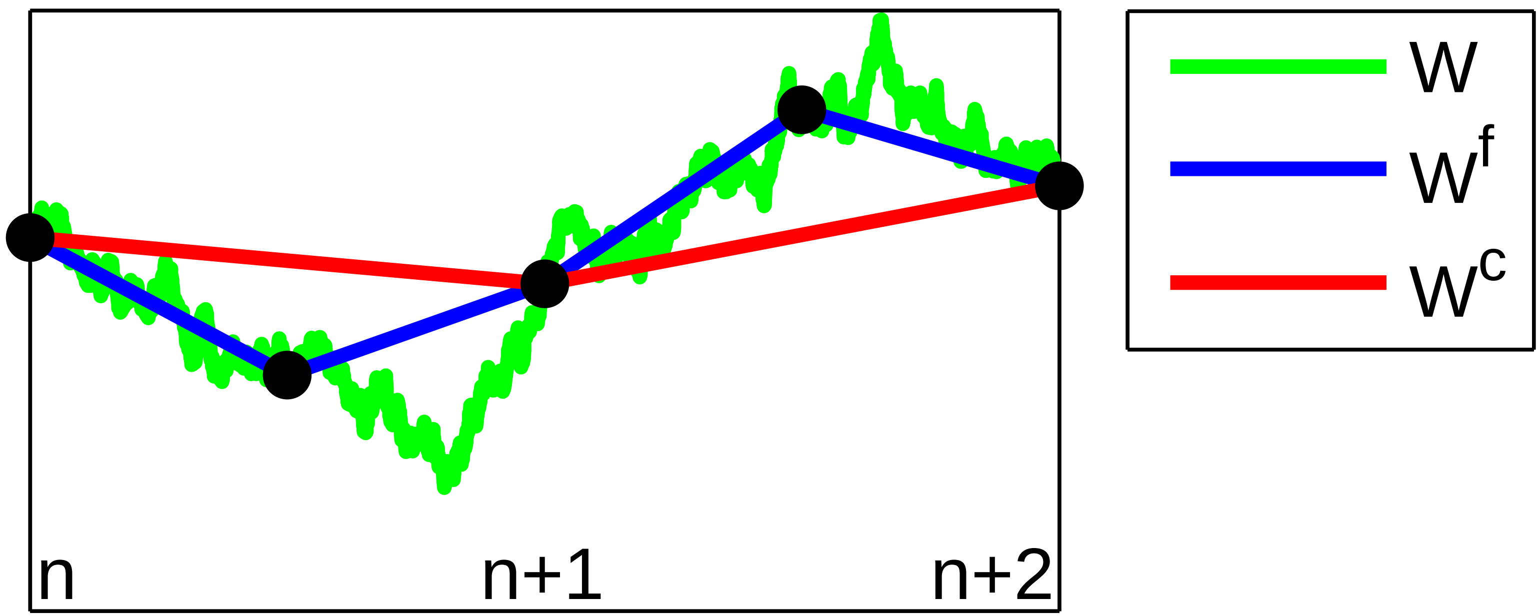
Key point of their paper: conditional on ΔW increments,

- $$\mathbb{E}\left[Y(T)\,|\,\Delta W \right] = \sum_n \left(X(t_n)\, \Delta W_{2,n} + \tfrac{1}{2} \Delta W_{1,n}\, \Delta W_{2,n} \right)$$
- $$\mathbb{V}\left[Y(T)\,|\,\Delta W \right] = \tfrac{1}{4} \sum_n \mathbb{V}[A_n] = O(h)$$

Hence, any numerical discretisation which uses only Brownian increments <u>cannot</u> in general achieve better than $O(\sqrt{h})$ strong convergence.

Clark & Cameron problem

If A_n is not known, best approximation sets it to zero,
– equivalent to a piecewise linear interpolation of the driving Brownian path.

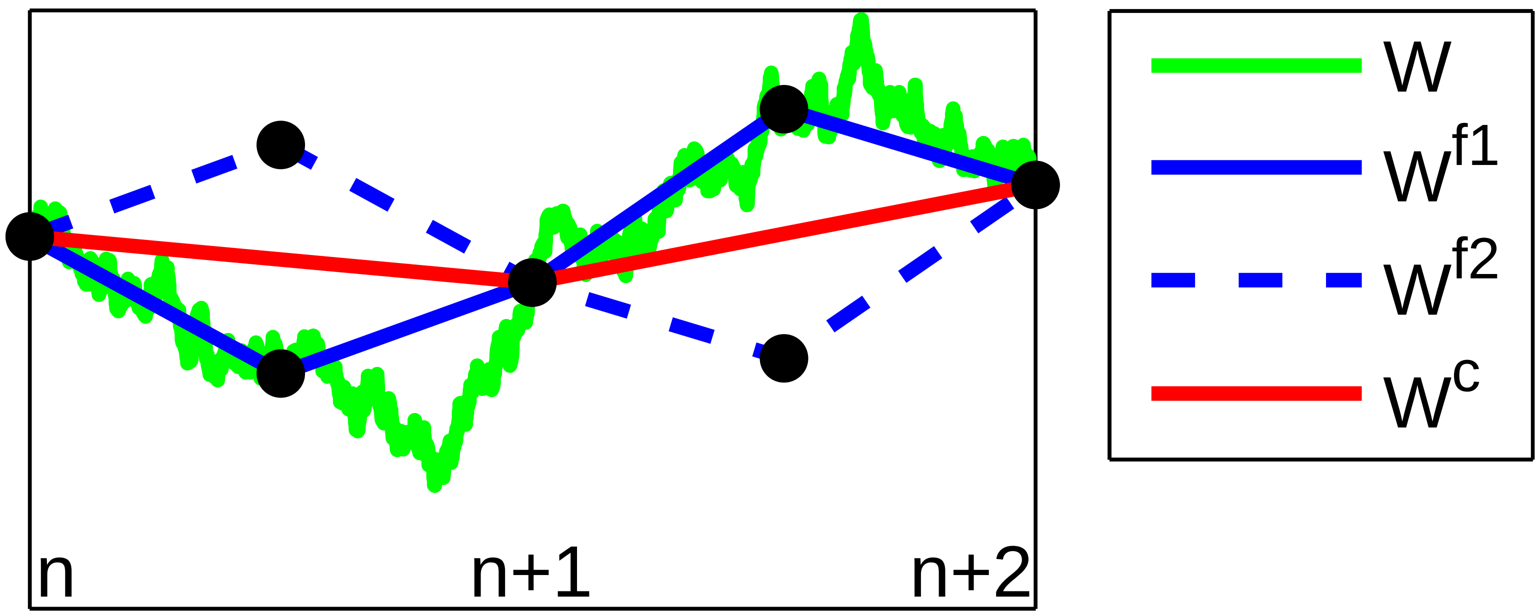


Coarse and fine paths use different interpolations

$$Y^f - Y^c = \sum_n A_n \Longrightarrow \mathbb{V}[Y^f - Y^c] = O(h)$$

Clark & Cameron problem

Fine path "antithetic twin" swaps Brownian increments for odd and even timesteps – average of two piecewise linear Brownian paths matches coarse one



$$A_n^{f2} = -A_n^{f1} \quad \Longrightarrow \quad (Y^{f2} - Y^c) = -(Y^{f1} - Y^c)$$

Hence $\frac{1}{2}(Y^{f1} + Y^{f2}) = Y^c$

Clark & Cameron problem

If the payoff function $f(X,Y)$ is twice-differentiable,

$$\begin{aligned} \tfrac{1}{2}\left(f(X,Y^{f1}) + f(X,Y^{f2})\right) - f(X,Y^c) &= \tfrac{1}{2}\,\frac{\partial^2 f}{\partial Y^2}\,(Y^{f1}\!-\!Y^c)^2 \\ &= O(h) \end{aligned}$$

Hence, $\mathbb{V}[\widehat{P}_\ell\!-\!\widehat{P}_{\ell-1}] = O(h^2)$ – much better than before.

If $f(X,Y)$ is Lipschitz and twice-differentiable except on K, and (X,Y^c) is within $O(\sqrt{h})$ of K with probability $O(\sqrt{h})$, then a local analysis gives $\mathbb{V}[\widehat{P}_\ell\!-\!\widehat{P}_{\ell-1}] = O(h^{3/2})$

Generalisation

For the general SDE

$$\mathrm{d}S_i(t) = a_i(S)\,\mathrm{d}t + \sum_j b_{ij}(S)\,\mathrm{d}W_j(t), \quad 0<t<T$$

we define the driving Brownian paths in the same way:

- fine path $W^{f1}(t)$ is piecewise linear interpolation with interval $h/2$
- fine path $W^{f2}(t)$ is “antithetic twin”, swapping odd and even increments
- coarse path $W^c(t)$ is piecewise linear interpolation with interval h, and also average of the two fine paths

Generalisation

If we define differences $\widehat{D}^f_n \equiv \widehat{S}^f_n - \widehat{S}^c_n$, then $\widehat{D}^{f1}_n$ and $\widehat{D}^{f2}_n$ are both $O(\sqrt{h})$, as before.

However, the average $\widehat{D}_n \equiv \frac{1}{2}\left(\widehat{D}^{f1}_n + \widehat{D}^{f2}_n\right)$ is no longer zero and instead satisfies a recurrence equation:

$$\widehat{D}_{i,n+1} = \widehat{D}_{i,n} + \sum_j \frac{\partial a_i}{\partial S_j}\, \widehat{D}_{j,n}\, h + \sum_{j,k} \frac{\partial b_{ij}}{\partial S_k}\, \widehat{D}_{k,n}\, \Delta W_{j,n} \; + \; R_n$$

The “remainder” R_n has properties

$$\mathbb{E}[R_n] = O[h^2], \quad \mathbb{V}[R_n] = O[h^3],$$

and hence we eventually obtain $\widehat{D}_n = O(h)$.

Generalisation

If the payoff function $f(S_T)$ is twice differentiable then

$$\begin{aligned}
\tfrac{1}{2}\left(f(\widehat{S}^{f1}) + f(\widehat{S}^{f2}) \right) - f(\widehat{S}^{c}) \\
&\approx \tfrac{1}{2}\left(\widehat{D}^{f1} + \widehat{D}^{f2} \right) \cdot \nabla f(\widehat{S}^{c}) \\
&\quad + \tfrac{1}{4}\left((\widehat{D}^{f1})^T H_f \widehat{D}^{f1} \ + \ (\widehat{D}^{f2})^T H_f \widehat{D}^{f2} \right) \\
&= O(h)
\end{aligned}$$

where $H_f \equiv \dfrac{\partial^2 f}{\partial S^2}$ is the Hessian matrix for f.

Hence, again get $\mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] = O(h^2)$, and it becomes $O(h^{3/2})$ when $f(S)$ is not twice-differentiable everywhere.

Conclusions

- MCQMC10 presentation gave numerical results showing effectiveness for Heston stochastic volatility model
- also gave an asymptotic analysis explanation
- new numerical analysis supports the observations and previous explanation
- further analysis treats case in which we approximate the Lévy areas by sub-sampling the Brownian path within each timestep – needed for discontinuous payoffs