Numerical analysis of the multilevel Milstein discretisation

Mike Giles

Kristian Debrabant (Univ. of Southern Denmark)

Andreas Rößler (TU Darmstadt)

mike.giles@maths.ox.ac.uk

Oxford-Man Institute of Quantitative Finance

Mathematical Institute, University of Oxford

(updated, expanded version of a talk given at MCQMC'12)

Multilevel Monte Carlo

Given a scalar SDE driven by a Brownian diffusion

$$\mathrm{d}S(t) = a(S,t)\,\mathrm{d}t + b(S,t)\,\mathrm{d}W(t), \qquad 0 < t < T$$

to estimate $\mathbb{E}[P]$ where the path-dependent payoff P 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 $\widehat{P}_{\ell-1}$,

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

MLMC Theorem

Theorem: *Let* P *be a functional of the solution of an SDE, and* $\widehat{P}_\ell$ *the discrete approximation using a timestep* $h_\ell = 2^{-\ell}\, T$.

If *there exist independent estimators* $\widehat{Y}_\ell$ *based on* N_ℓ *Monte Carlo samples, with computational complexity (cost)* C_ℓ*, and positive constants* $\alpha \geq \frac{1}{2}, \beta, c_1, c_2, c_3$ *such that*

i) $\left| \mathbb{E}[\widehat{P}_\ell - P] \right| \leq c_1\, h_\ell^\alpha$

ii) $\mathbb{E}[\widehat{Y}_\ell] = \begin{cases} \mathbb{E}[\widehat{P}_0], & \ell = 0 \\ \mathbb{E}[\widehat{P}_\ell - \widehat{P}_{\ell-1}], & \ell > 0 \end{cases}$

iii) $\mathbb{V}[\widehat{Y}_\ell] \leq c_2\, N_\ell^{-1} h_\ell^\beta$

iv) $C_\ell \leq c_3\, N_\ell\, h_\ell^{-1}$

MLMC Theorem

then *there exists a positive constant* c_4 *such that for any* $\varepsilon < e^{-1}$ *there are values* L *and* N_ℓ *for which the multilevel estimator*

$$\widehat{Y} = \sum_{\ell=0}^{L} \widehat{Y}_\ell,$$

has Mean Square Error $MSE \equiv \mathbb{E}\left[\left(\widehat{Y} - \mathbb{E}[P]\right)^2\right] < \varepsilon^2$

with a computational complexity C *with bound*

$$C \leq \begin{cases} c_4\, \varepsilon^{-2}, & \beta > 1, \\ c_4\, \varepsilon^{-2} (\log \varepsilon)^2, & \beta = 1, \\ c_4\, \varepsilon^{-2-(1-\beta)/\alpha}, & 0 < \beta < 1. \end{cases}$$

Numerical Analysis

If P is a Lipschitz function of $S(T)$, value of underlying path simulation at a fixed time, the strong convergence property

$$\left(\mathbb{E}\left[(\widehat{S}_N - S(T))^2\right]\right)^{1/2} = O(h^\gamma)$$

implies that $\mathbb{V}[\widehat{P}_\ell - P] = O(h_\ell^{2\gamma})$ and hence

$$V_\ell \;\equiv\; \mathbb{V}[\widehat{P}_\ell - \widehat{P}_{\ell-1}] = O(h_\ell^{2\gamma}).$$

Therefore $\beta = 1$ for Euler-Maruyama discretisation, and $\beta = 2$ for the Milstein discretisation.

However, in general, good strong convergence is neither necessary nor sufficient for good convergence for V_ℓ.

Numerics and Analysis

	Euler		Milstein	
option	numerics	analysis	numerics	analysis
Lipschitz	$O(h)$	$O(h)$	$O(h^2)$	$O(h^2)$
Asian	$O(h)$	$O(h)$	$O(h^2)$	$O(h^2)$
lookback	$O(h)$	$O(h)$	$O(h^2)$	$O(h^2 \lvert\log h\rvert^2)$
barrier	$O(h^{1/2})$	$o(h^{1/2-\delta})$	$O(h^{3/2})$	$o(h^{3/2-\delta})$
digital	$O(h^{1/2})$	$O(h^{1/2}\log h)$	$O(h^{3/2})$	$o(h^{3/2-\delta})$

Table: V_ℓ convergence observed numerically (for GBM) and proved analytically (for more general SDEs) for both the Euler and Milstein discretisations. δ can be any strictly positive constant.

Numerical Analysis

Analysis for Euler discretisations:

- lookback and barrier options: Giles, Higham & Mao (*Finance & Stochastics, 2009*)
  - lookback analysis follows from strong convergence
  - barrier analysis shows dominant contribution comes from paths which are near the barrier
  - similar analysis for digital options gives $O(h^{1/2-\delta})$ bound instead of $O(h^{1/2}\log h)$
- digital options: Avikainen (*Finance & Stochastics, 2009*)
  - method of analysis is rather different

Numerical Analysis

Analysis for Milstein discretisations:

- builds on approach in paper with Higham and Mao
- key idea for digital and barrier options is to
  - use boundedness of all moments to bound the contribution from "extreme" paths (e.g. with $\max_n |\Delta W_n| > h^{1/2-\delta}$ for some $\delta > 0$)
  - uses asymptotic analysis to bound the contribution from paths which are not "extreme"

Some preliminaries

These results come from Extreme Value Theory, which looks at the asymptotic distribution of the maximum of a large set of i.i.d. random variables.

Lemma 0.1 *If* $U^{(n)}, n=1,\ldots,N$ *are independent samples from a uniform distribution on the unit interval* $[0,1]$*, then for any positive integer* m

$$\mathbb{E}\left[\max_n |\log U^{(n)}|^m\right] = O((\log N)^m), \text{ as } N \to \infty.$$

Some preliminaries

Lemma 0.2 *If* $Z^{(n)}, n=1,\ldots,N$ *are independent samples from a standard Normal distribution, then for any positive integer* m

$$\mathbb{E}\left[\max_n |Z^{(n)}|^m\right] = O((\log N)^{m/2}), \ \textit{as}\ N\to\infty.$$

Corollary 0.3 *If* $W^{(n)}(t), n=1,\ldots,N$ *are independent Brownian paths on* $[0,1]$*, conditional on* $W_n(0)=W_n(1)=0$*, then for any positive integer* m

$$\mathbb{E}\left[\max_n \sup_{[0,1]} |W^{(n)}(t)|^m\right] = O((\log N)^{m/2}), \ \textit{as}\ N\to\infty.$$

Milstein Scheme

MLMC Theorem allows different approximations on the coarse and fine levels:

$$\widehat{Y}_\ell = N_\ell^{-1} \sum_{n=1}^{N_\ell} \left(\widehat{P}_\ell^f(\omega^{(n)}) - \widehat{P}_{\ell-1}^c(\omega^{(n)}) \right)$$

The telescoping sum still works provided

$$\mathbb{E}\left[\widehat{P}_\ell^f\right] \;=\; \mathbb{E}\left[\widehat{P}_\ell^c\right].$$

The key is to exploit this freedom to reduce the variance

$$\mathbb{V}\left[\widehat{P}_\ell^f - \widehat{P}_{\ell-1}^c\right].$$

Milstein Scheme

Fine path Brownian interpolation: within each timestep, model the behaviour as simple Brownian motion (constant drift and volatility) conditional on two end-points

$$\begin{aligned} \widehat{S}^f(t) &= \widehat{S}^f_n + \lambda(t)(\widehat{S}^f_{n+1} - \widehat{S}^f_n) \\ &\quad + b_n \Big(W(t) - W_n - \lambda(t)(W_{n+1} - W_n) \Big), \end{aligned}$$

where $\lambda(t) = \dfrac{t - t_n}{t_{n+1} - t_n}$.

There then exist analytic results for the distribution of the min/max/average over each timestep, and probability of crossing a barrier.

Brownian Bridge results

Conditional on $\widehat{S}_{n+1}, \widehat{S}_n$, a sample of the minimum of $\widehat{S}$ in the time interval $[t_n, t_{n+1}]$ is given by

$$\widehat{S}_{min} = \tfrac{1}{2}\left(\widehat{S}_{n+1} + \widehat{S}_n - \sqrt{(\widehat{S}_{n+1} - \widehat{S}_n)^2 - 2\, b_n^2\, h\, \log U_n} \right)$$

where U_n is a uniform $(0,1)$ r.v.

Similarly, a sample of the maximum is given by

$$\widehat{S}_{max} = \tfrac{1}{2}\left(\widehat{S}_{n+1} + \widehat{S}_n + \sqrt{(\widehat{S}_{n+1} - \widehat{S}_n)^2 - 2\, b_n^2\, h\, \log V_n} \right)$$

where V_n is a different uniform $(0,1)$ r.v.

Brownian Bridge results

Conditional on $\widehat{S}_{n+1}, \widehat{S}_n > B$, the probability that $\widehat{S}$ drops below B in the time interval $[t_n, t_{n+1}]$ is

$$\mathbb{P}(\widehat{S}_{min} < B \mid \widehat{S}_n, \widehat{S}_{n+1} > B) = \exp\left(- \frac{2\,(\widehat{S}_{n+1} - B)\,(\widehat{S}_n - B)}{b_n^2\ h} \right)$$

Similarly,

$$\mathbb{P}(\widehat{S}_{max} > B \mid \widehat{S}_n, \widehat{S}_{n+1} < B) = \exp\left(- \frac{2\,(B - \widehat{S}_{n+1})\,(B - \widehat{S}_n)}{b_n^2\ h} \right)$$

Milstein Scheme

Coarse path Brownian interpolation: exactly the same, but with double the timestep, so for even n

$$\begin{aligned}\widehat{S}^c(t) &= \widehat{S}^c_n + \lambda(t)(\widehat{S}^c_{n+2} - \widehat{S}^c_n) \\ &\quad + b_n \left(W(t) - W_n - \lambda(t)(W_{n+2} - W_n) \right),\end{aligned}$$

where $\lambda(t) = \dfrac{t - t_n}{t_{n+2} - t_n}$. Hence, in particular,

$$\begin{aligned}\widehat{S}^c_{n+1} \equiv \widehat{S}^c(t_{n+1}) &= \tfrac{1}{2}(\widehat{S}^c_n + \widehat{S}^c_{n+2}) \\ &\quad + b_n \left(W_{n+1} - \tfrac{1}{2}(W_n + W_{n+2}) \right),\end{aligned}$$

Milstein Scheme

Theorem: Under standard conditions,

- $$\mathbb{E}\left[\sup_{[0,T]} \left| \widehat{S}(t) - S(t) \right|^m \right] = O((h \log h)^m),$$

- $$\sup_{[0,T]} \mathbb{E}\left[\left| \widehat{S}(t) - S(t) \right|^m \right] = O(h^m),$$

- $$\mathbb{E}\left[\left(\int_0^T \widehat{S}(t) - S(t) \, \mathrm{d}t \right)^2 \right] = O(h^2).$$

Milstein Scheme

The first result comes from considering

$$\widehat{S}(t) - S(t) = (\widehat{S}(t) - \widehat{S}_{KP}(t)) + (\widehat{S}_{KP}(t) - S(t))$$

where $\widehat{S}_{KP}(t)$ is the Kloeden-Platen interpolant for which they prove

$$\mathbb{E}\left[\sup_{[0,T]} \left|\widehat{S}_{KP}(t) - S(t)\right|^m\right] = O(h^m).$$

It is easily shown that

$$\widehat{S}(t) - \widehat{S}_{KP}(t) \;=\; \tfrac{1}{2}\, b_n' b_n\, Y(t),$$

where . . .

Milstein Scheme

$$\begin{aligned} Y(t) &= \lambda\,(W_{n+1}-W_n)^2 - (W(t)-W_n)^2 \\ &= \lambda\,(1-\lambda)\,(W_{n+1}-W_n)^2 - \left(W(t) - W_n - \lambda\,(W_{n+1}-W_n)\right)^2 \\ &\quad - 2\,\lambda\,(W_{n+1}-W_n)\,(W(t) - W_n - \lambda\,(W_{n+1}-W_n)) \end{aligned}$$

and then the result comes from Hölder's inequality

$$\mathbb{E}\left[\sup_{[0,T]}\left|\widehat{S}(t)-\widehat{S}_{KP}(t)\right|^m\right] \le \sqrt{\mathbb{E}\left[\max_n |b'_n b_n|^{2m}\right]\;\mathbb{E}\left[\sup_{[0,T]}|Y(t)|^{2m}\right]}$$

and bounds on $\mathbb{E}\left[\sup_{[0,T]}|Y(t)|^{2m}\right]$ coming from

$$\begin{aligned} &\mathbb{E}[\sup_{[0,T]}|W(t) - W_n - \lambda\,(W_{n+1}-W_n)|^{2m}] \\ &\qquad = h^m\;\mathbb{E}[\max_n \sup_{[0,1]}|W^{(n)}(t) - t\,W_1^{(n)}|^m] \end{aligned}$$

Milstein Scheme

The second one comes from setting

$$\begin{aligned} W(t) - W_n &= \sqrt{\lambda\, h}\; Z_1 \\ W_{n+1} - W(t) &= \sqrt{(1-\lambda)\, h}\; Z_2 \end{aligned}$$

with Z_1, Z_2 independent standard Normal random variables. Then $|Y| \le h\, \max(Z_1^2, Z_2^2)$ leads to

$$|Y|^m \;\le\; h^m\, \max(Z_1^{2m}, Z_2^{2m}) \;\le\; h^m\, (Z_1^{2m} + Z_2^{2m})$$

and hence the assertion follows from

$$\mathbb{E}\left[\, \left| \widehat{S}(t) - \widehat{S}_{KP}(t) \right|^m \right] = 2^{-m} \mathbb{E}[\, |b_n' b_n|^m]\; \mathbb{E}[\, |Y|^m],$$

and standard bounds for moments of Normal random variables.

Milstein Scheme

For the third one, setting $X_n := \int_{t_n}^{t_{n+1}} Y(t)\, \mathrm{d}t$ gives

$$\mathbb{E}\left[\left(\int_0^T (\widehat{S}(t) - \widehat{S}_{KP}(t))\, \mathrm{d}t\right)^2\right] = \tfrac{1}{4}\mathbb{E}\left[\left(\sum_{n=0}^{N-1} b'_n b_n X_n\right)^2\right].$$

For $n\!>\!m$, $\mathbb{E}[b'_m b_m X_m b'_n b_n X_n] = 0$ since X_n is independent of $b'_m b_m X_m b'_n b_n$ and $\mathbb{E}[X_n] = 0$. In addition, the X_n are iid random variables, and therefore

$$\mathbb{E}\left[\left(\int_0^T (\widehat{S}(t) - \widehat{S}_{KP}(t))\, \mathrm{d}t\right)^2\right] = \tfrac{1}{4}\mathbb{E}[X_0^2] \sum_{n=0}^{N-1} \mathbb{E}[(b'_n b_n)^2].$$

The proof is completed by noting that $\mathbb{E}[X_0^2] = O(h^4)$ due to standard moment bounds for Brownian increments.

Lookback Option

Lookback options are a Lipschitz function of the minimum over the whole simulation path.

For the fine path, the minimum over one timestep is

$$\widehat{S}^f_{n,min} = \tfrac{1}{2}\left(\widehat{S}^f_n + \widehat{S}^f_{n+1} - \sqrt{\left(\widehat{S}^f_{n+1} - \widehat{S}^f_n\right)^2 - 2\left(b^f_n\right)^2 h_\ell \, \log U_n}\, \right)$$

where U_m is a $(0,1]$ uniform random variable.

For the coarse path, define $\widehat{S}^c_n$ for odd n using conditional Brownian interpolation, then use the same expression for the minimum with same U_n – this doesn't change the distribution of the computed minimum over the coarse timestep, so the telescoping sum is OK.

Lookback Option

Defining $\widehat{D}_n^{f/c} = \frac{1}{2}\sqrt{\left(\widehat{S}_{n+1}^{f/c} - \widehat{S}_n^{f/c}\right)^2 - 2\,(b_n^{f/c})^2\, h \log U_n}$, gives

$$\begin{aligned} \left|\widehat{S}_{min}^f - \widehat{S}_{min}^c\right| &\leq \max_n \left|\widehat{S}_{n,min}^f - \widehat{S}_{n,min}^c\right| \\ &\leq \max_n \left|\widehat{S}_n^f - \widehat{S}_n^c\right| + \max_n \left|\widehat{D}_n^f - \widehat{D}_n^c\right| \end{aligned}$$

Tedious analysis leads to

$$\left|\widehat{D}_n^f - \widehat{D}_n^c\right| \leq \left(\left|\widehat{S}_{n+1}^f - \widehat{S}_{n+1}^c\right| + \left|\widehat{S}_n^f - \widehat{S}_n^c\right|\right) + |b_n^f - b_n^c|\, \sqrt{h\,|\log U_n|}$$

and hence

$$\mathbb{E}\left[\left|\widehat{S}_{min}^f - \widehat{S}_{min}^c\right|^2\right] = O((h \log h)^2)$$

Barrier and Digital Options

The barrier option is based on the Brownian Bridge construction and the probability of crossing the barrier within each timestep.

The digital option is based on a Brownian extrapolation from one timestep before the end – the analysis is similar.

The analysis for both uses the idea of “extreme” paths which are highly improbable – the variance comes mainly from non-extreme paths for which one can use asymptotic analysis.

Extreme Paths

Lemma: If X_ℓ is a random variable on level l, and $\mathbb{E}[\,|X_\ell|^m] \leq C_m$ is uniformly bounded, then, for any $\delta > 0$,

$$\mathbb{P}[\,|X_\ell| > h_\ell^{-\delta}] = o(h_\ell^p), \qquad \forall p > 0.$$

Proof: Markov inequality $\mathbb{P}[\,|X_\ell|^m > h_\ell^{-m\delta}] \;<\; h_\ell^{-m\delta}\, \mathbb{E}[\,|X_\ell|^m]$.

Lemma: If Y_ℓ is a random variable on level ℓ, $\mathbb{E}[\,Y_\ell^2]$ is uniformly bounded, and the indicator function $\mathbf{1}_{E_\ell}$ satisfies $\mathbb{E}[\mathbf{1}_{E_\ell}] = o(h_\ell^p),\ \forall p > 0$ then

$$\mathbb{E}[\,|Y_\ell|\,\mathbf{1}_{E_\ell}] = o(h_\ell^p), \qquad \forall p > 0.$$

Proof: Hölder inequality $\mathbb{E}[\,|Y_\ell|\,\mathbf{1}_{E_\ell}] \leq \sqrt{\mathbb{E}[\,Y_\ell^2]\ \mathbb{E}[\mathbf{1}_{E_\ell}]}$

Extreme Paths

Theorem: For any $\gamma > 0$, the probability that $W(t)$, its increments ΔW_n and the corresponding SDE solution $S(t)$ and approximations $\widehat{S}^f_n$ and $\widehat{S}^c_n$ satisfy any of the following "extreme" conditions

$$
\begin{aligned}
\max_n \left(\max(|S(nh)|,\ |\widehat{S}^f_n|,\ |\widehat{S}^c_n|) \right) &> h^{-\gamma} \\
\max_n \left(\max(|S(nh) - \widehat{S}^c_n|,\ |S(nh) - \widehat{S}^f_n|,\ |\widehat{S}^f_n - \widehat{S}^c_n|) \right) &> h^{1-\gamma} \\
\max_n |\Delta W_n| &> h^{1/2-\gamma}
\end{aligned}
$$

is $o(h^p)$ for all $p > 0$.

Non-extreme paths

Furthermore, there exist constants c_1, c_2, c_3, c_4 such that if none of these conditions is satisfied, and $\gamma < \frac{1}{2}$, then

$$
\begin{aligned}
\max_n |\widehat{S}^f_n - \widehat{S}^f_{n-1}| &\leq c_1\, h^{1/2-2\gamma} \\
\max_n |b^f_n - b^f_{n-1}| &\leq c_2\, h^{1/2-2\gamma} \\
\max_n \left(|b^f_n| + |b^c_n|\right) &\leq c_3\, h^{-\gamma} \\
\max_n |b^f_n - b^c_n| &\leq c_4\, h^{1/2-2\gamma}
\end{aligned}
$$

where b^c_n is defined to equal b^c_{n-1} if n is odd.

Barrier and Digital Options

For barrier options, split paths into 3 subsets:

- extreme paths
- paths with a minimum within $O(h^{1/2-\gamma})$ of the barrier
- rest

– dominant contribution comes from the second subset.

For digital options, again split paths into 3 subsets:

- extreme paths
- paths with final $S(T)$ within $O(h^{1/2-\gamma})$ of the strike
- rest

– dominant contribution again from the second subset.

Other numerical analysis

- multi-dimensional Milstein – next lecture
- multilevel scalar finite rate jump-diffusion, including path-dependent Poisson rate – Yuan Xia
  - algorithm and numerical results presented at MCQMC'10
  - numerical analysis completed in DPhil, 2014
- multilevel Greeks – Sylvestre Burgos
  - algorithm and numerical results presented at MCQMC'10
  - numerical analysis completed in DPhil, 2014

Conclusions

- numerical analysis of multilevel variance achieves bounds which match numerical experiments for Milstein discretisation of scalar SDEs and all common payoffs
- Brownian interpolation is key to obtaining rapid convergence of the multilevel variance for complex payoffs
- excluding the significance of “extreme” paths and using asymptotic analysis for the rest is a non-standard approach to numerical analysis, but seems quite flexible

Multilevel papers are available from:
`people.maths.ox.ac.uk/gilesm/mlmc.html`