Strong convergence of path sensitivities

Mike Giles

Mathematical Institute, University of Oxford

Stochastic Numerics and Inverse Problems

June 19th, 2024

Outline

- "usual conditions" for analysis of SDE discretisations
- unusual features of SDE path sensitivities
- new analysis of strong convergence

This work is motivated by the use of Multilevel Monte Carlo (MLMC) methods to calculate sensitivities ("Greeks") in Mathematical Finance.

It seems to fill a gap in the existing literature, unless anyone knows otherwise?

Usual analysis of SDEs

When considering, for simplicity, the autonomous SDE

$$\mathrm{d}S_t = a(S_t)\,\mathrm{d}t + b(S_t)\,\mathrm{d}W_t$$

the "usual conditions" assume that $a(S)$ and $b(S)$ are globally Lipschitz, i.e. there exists L such that

$$\|a(v) - a(u)\| + \|b(v) - b(u)\| < L\,\|v-u\|, \quad \forall u, v.$$

Under these conditions, the SDE has a unique solution given initial S_0, and for any finite time interval $[0, T]$ and $p > 0$ there exist constants $c_p^{(1)}$, $c_p^{(2)}$ such that

$$\mathbb{E}\left[\sup_{0<t<T} \|S_t\|^p\right] \;\leq\; c_p^{(1)},$$

$$\mathbb{E}\left[\,\|S_t - S_{t_0}\|^p\right] \;\leq\; c_p^{(2)}\,(t-t_0)^{p/2}, \quad \text{for } 0 < t_0 < t < T.$$

Usual analysis of SDE discretisations

Furthermore, for the Euler-Maruyama discretisation

$$\widehat{S}_{(n+1)h} = \widehat{S}_{nh} + a(\widehat{S}_{nh})\, h + b(\widehat{S}_{nh})\, \Delta W_n,$$

with a uniform timestep of h, we have $O(h^{1/2})$ strong convergence so that for any $p>0$ there exists $c_p^{(3)}$ such that

$$\mathbb{E}\left[\sup_{0<t<T} \|\widehat{S}_t - S_t\|^p\right] \;\leq\; c_p^{(3)}\, h^{p/2}.$$

This strong convergence is important for the effectiveness and analysis of MLMC algorithms.

Pathwise sensitivities

Suppose now that S_t is scalar, and $a(\theta; S)$ and $b(\theta; S)$ depend smoothly on a scalar parameter θ as well as S

$$\mathrm{d}S_t = a(\theta; S_t)\,\mathrm{d}t + b(\theta; S_t)\,\mathrm{d}W_t$$

and we are interested in the expected value of a "payoff" function $P(S_T)$,

$$f(\theta) = \mathbb{E}\left[\, P(S_T(\theta; \{W_t\}_{0\leq t\leq T}))\,\right]$$

and want to compute its derivative

$$\dot{f} \equiv \frac{\mathrm{d}f}{\mathrm{d}\theta}$$

Pathwise sensitivities

If P is globally Lipschitz and piecewise smooth, then

$$\dot{f} \;\equiv\; \frac{\mathrm{d}}{\mathrm{d}\theta}\, \mathbb{E}[\, P(S_T)\,] \;=\; \mathbb{E}[\, \dot{P}(S_T)\,]$$

where

$$\dot{P} = \frac{\mathrm{d}P}{\mathrm{d}S}\; \dot{S}_T$$

and $\dot{S}_t \equiv \dfrac{\mathrm{d}S_t}{\mathrm{d}\theta}$ satisfies the SDE

$$\mathrm{d}\dot{S}_t = \big(\dot{a}(\theta;S_t) + a'(\theta;S_t)\,\dot{S}_t\big)\,\mathrm{d}t + \big(\dot{b}(\theta;S_t) + b'(\theta;S_t)\,\dot{S}_t\big)\,\mathrm{d}W_t$$

subject to initial $\dot{S}_0$, with $\dot{a} \equiv \dfrac{\partial a}{\partial \theta}$, $a' \equiv \dfrac{\partial a}{\partial S}$, and $\dot{b}, b'$ defined similarly.

(Note: analysis can be extended to P depending explicitly on θ)

Pathwise sensitivities

The Euler-Maruyama discretisation of the pathwise sensitivity SDE is

$$\begin{aligned}\widehat{\dot{S}}_{(n+1)h} \;&=\; \widehat{\dot{S}}_{nh} + \left(\dot{a}(\theta;\widehat{S}_{nh}) + a'(\theta;\widehat{S}_{nh})\,\widehat{\dot{S}}_{nh}\right) h \\ &\quad + \left(\dot{b}(\theta;\widehat{S}_{nh}) + b'(\theta;\widehat{S}_{nh})\,\widehat{\dot{S}}_{nh}\right) \Delta W_n\end{aligned}$$

This is also the equation one gets by differentiating the E-M discretisation of the original SDE.

Question: what is the order of strong convergence of $\widehat{\dot{S}}$ to $\dot{S}$?

Previous MLMC work has assumed the same strong convergence

$$\mathbb{E}\left[\sup_{0<t<T} \|\widehat{\dot{S}}_t - \dot{S}_t\|^p\right] = O(h^{p/2})$$

but I have not found a reference for this.

Pathwise sensitivities

The pathwise sensitivity SDE can be appended to the original SDE to form a vector SDE with $\mathbf{S}_t \equiv (S_t, \dot{S}_t)^T$

$$\mathrm{d}\mathbf{S}_t = \mathbf{a}(\theta; \mathbf{S}_t)\,\mathrm{d}t + \mathbf{b}(\theta; \mathbf{S}_t)\,\mathrm{d}W_t.$$

I think past work assumed this vector SDE satisfies the "usual conditions" and hence leads to $1/2$-order strong convergence for both $\widehat{S}$ and $\widehat{\dot{S}}$.

However, this is not true in general.

Pathwise sensitivities

Looking at the pathwise sensitivity SDE

$$\mathrm{d}\dot{S}_t = (\dot{a}(\theta; S_t) + a'(\theta; S_t)\, \dot{S}_t)\, \mathrm{d}t + (\dot{b}(\theta; S_t) + b'(\theta; S_t)\, \dot{S}_t)\, \mathrm{d}W_t$$

even if we assume all derivatives of $a(\theta; S)$ and $b(\theta; S)$ are bounded, then

$$\begin{aligned} a'(\theta; v_1)\, v_2 - a'(\theta; u_1)\, u_2 &= (a'(\theta; v_1) - a'(\theta; u_1))\, v_2 + a'(\theta; u_1)\, (v_2 - u_2) \\ &= a''(\theta; w)\, v_2\, (v_1 - u_1) + a'(\theta; u_1)\, (v_2 - u_2) \end{aligned}$$

for some $u_1 < w < v_1$.

The problem is that $|a''(\theta; w)\, v_2| \to \infty$ as $v_2 \to \infty$ unless $a''(\theta; w) = 0$, and something similar applies for $b'(\theta; S)\, \dot{S}$.

Pathwise sensitivities

If we use the shorthand $a_t \equiv a(\theta; S_t)$, $\dot{a}_t \equiv \dot{a}(\theta; S_t)$, $a'_t \equiv a'(\theta; S_t)$, and similarly for $b_t, \dot{b}_t, b'_t$ and higher derivatives, then the first order pathwise sensitivity SDE is

$$\mathrm{d}\dot{S}_t = (\dot{a}_t + a'_t\, \dot{S}_t)\, \mathrm{d}t + (\dot{b}_t + b'_t\, \dot{S}_t)\, \mathrm{d}W_t$$

The second order pathwise sensitivity SDE is then

$$\mathrm{d}\ddot{S}_t = (\ddot{a}_t + 2\dot{a}'_t\dot{S}_t + a''_t(\dot{S}_t)^2 + a'_t\ddot{S}_t)\,\mathrm{d}t + (\ddot{b}_t + 2\dot{b}'_t\dot{S}_t + b''_t(\dot{S}_t)^2 + b'_t\ddot{S}_t)\,\mathrm{d}W_t$$

and the $(\dot{S}_t)^2$ terms makes it even clearer that the "usual conditions" are not satisfied.

However, notice that $\dot{S}_t$ in the first equation, and $\ddot{S}_t$ in the second, are multiplied by a'_t and b'_t which are bounded

Pathwise sensitivities

There is a large literature on the approximation of SDEs which do not satisfy the usual conditions.

These use modified numerical approximations (e.g. tamed schemes, or adaptive timesteps) for which stability and strong convergence can be proved.

However, with these pathwise equations there is no problem using the standard Euler-Maruyama discretisation – all that is needed is a new numerical analysis to prove it has the observed $O(h^{1/2})$ strong convergence order.

We will perform the analysis for the first order pathwise sensitivities, but inductively it applies to higher orders too.

Numerical analysis

The numerical analysis is not difficult – essentially retraces the steps of the standard analysis.

Focussing on the first order sensitivity equation, the key is that in the drift and diffusion terms $\dot{S}_t$ is multiplied by the bounded a'_t and b'_t.

Arbitrary moments of all other terms are bounded due to standard results for S_t and $\widehat{S}_t$.

Beyond this, the methodology is standard: use Jensen, Hölder, and Burkholder-Davis-Gundy inequalities to set things up for finally using Grönwall's inequality.

Numerical analysis: step 1

Lemma

For a given time interval $[0,T]$, and any $p\geq 2$, there exists a constant $c_p^{(1)}$ such that

$$\sup_{0<t<T} \mathbb{E}\left[\,|\dot{S}_t|^p\right] \leq c_p^{(1)}.$$

Proof.

For even integer $p\geq 2$, if we define $P_t = \dot{S}_t^p$ then Ito's lemma gives us

$$\begin{aligned} \mathrm{d}P_t &= \left(p\,\dot{S}_t^{p-1}(\dot{a}_t + a'_t\dot{S}_t) + \tfrac{1}{2}\,p\,(p{-}1)\,\dot{S}_t^{p-2}(\dot{b}_t + b'_t\dot{S}_t)^2\right)\mathrm{d}t \\ & \quad + p\,\dot{S}_t^{p-1}(\dot{b}_t + b'_t\dot{S}_t)\,\mathrm{d}W_t, \end{aligned}$$

$$\Longrightarrow \quad \mathrm{d}\mathbb{E}[P_t] \;\leq\; (p\,L_a + p\,(p{-}1)\,L_b^2)\,(1+2\,\mathbb{E}[P_t])\,\mathrm{d}t$$

Then apply Grönwall's inequality, and use Jensen inequality for intermediate p. □

Numerical analysis: step 2

Theorem

For a given time interval $[0, T]$, and any $p \geq 2$, there exists a constant $c_p^{(1)}$ such that

$$\mathbb{E}\left[\sup_{0<t<T} |\dot{S}_t|^p\right] \leq c_p^{(1)}.$$

Proof.

Starting from

$$\dot{S}_t = \dot{S}_0 + \int_0^t (\dot{a}_s + a'_s \dot{S}_s)\,\mathrm{d}s + \int_0^t (\dot{b}_s + b'_s \dot{S}_s)\,\mathrm{d}W_s,$$

and defining $\dot{M}_t^{(p)} = \mathbb{E}\left[\sup_{0<s<t} |\dot{S}_s|^p\right]$, *then …*

Numerical analysis: step 2

Proof (continued).

Jensen's inequality gives

$$\dot{M}_t^{(p)} \;\leq\; 5^{p-1} \left(\, |\dot{S}_0|^p + \mathbb{E}\left[\sup_{0<s<t} \left| \int_0^s \dot{a}_u \, \mathrm{d}u \right|^p \right] + \mathbb{E}\left[\sup_{0<s<t} \left| \int_0^s a'_u \dot{S}_u \, \mathrm{d}u \right|^p \right] \right.$$
$$\left. + \; \mathbb{E}\left[\sup_{0<s<t} \left| \int_0^s \dot{b}_u \, \mathrm{d}W_u \right|^p \right] + \mathbb{E}\left[\sup_{0<s<t} \left| \int_0^s b'_u \dot{S}_u \, \mathrm{d}W_u \right|^p \right] \, \right).$$

Bounding each term, using BDG inequality for stochastic integrals, leads to an equation of the form

$$\dot{M}_t^{(p)} \leq c_1 + c_2 \int_0^t \dot{M}_u^{(p)} \, \mathrm{d}u$$

and then Grönwall's inequality gives the desired result. □

Numerical analysis: step 3

Lemma

For a given time interval $[0, T]$*, and any* $p \geq 2$*, there exists a constant* $c_p^{(2)}$ *such that*

$$\mathbb{E}\left[|\dot{S}_t - \dot{S}_{t_0}|^p \right] \leq c_p^{(2)} (t - t_0)^{p/2}$$

for any $0 \leq t_0 \leq t \leq T$*.*

Proof.

Almost identical to the previous proof, but starting from

$$\dot{S}_t - \dot{S}_{t_0} = \int_{t_0}^{t} (\dot{a}_s + a'_s \dot{S}_s)\, \mathrm{d}s + \int_{t_0}^{t} (\dot{b}_s + b'_s \dot{S}_s)\, \mathrm{d}W_s,$$

and defining

$$\dot{M}_t^{(p)} = \mathbb{E}\left[\sup_{t_0 < s < t} |\dot{S}_s - \dot{S}_{t_0}|^p \right].$$

□

Numerical analysis: step 4

Lemma

For a given time interval $[0, T]$, and any $p \geq 2$, there exists a constant $c_p^{(1)}$ such that

$$\mathbb{E}\left[\sup_{0<t<T} |\widehat{\dot{S}}_t|^p\right] \leq c_p^{(1)}.$$

Proof.

The proof follows the same approach used for $\mathbb{E}\left[\sup_{0<t<T} |\dot{S}_t|^p\right]$. □

Numerical analysis: step 5

Finally we come to the strong convergence theorem.

Theorem

Given the boundedness of all first and second derivatives, for a given time interval $[0,T]$*, and any* $p \geq 2$*, there exists a constant* $c_p^{(3)}$ *such that*

$$\mathbb{E}\left[\sup_{0<t<T} |\widehat{\dot{S}}_t - \dot{S}_t|^p\right] \leq c_p^{(3)} h^{p/2}.$$

Proof.

The continuous-time Euler-Maruyama discretisation can be written as

$$\widehat{\dot{S}}_t = \widehat{\dot{S}}_0 + \int_0^t (\widehat{\dot{a}}_{\underline{s}} + \widehat{a}'_{\underline{s}} \widehat{\dot{S}}_{\underline{s}})\, \mathrm{d}s + \int_0^t (\widehat{\dot{b}}_{\underline{s}} + \widehat{b}'_{\underline{s}} \widehat{\dot{S}}_{\underline{s}})\, \mathrm{d}W_s,$$

where $\underline{s}$ *denotes* s *rounded downwards to the nearest timestep, and* $\widehat{\dot{a}}_{\underline{s}}$ *denotes* $\dot{a}(\theta, \widehat{S}_{\underline{s}})$ *with similar meanings for* $\widehat{a}'_{\underline{s}}$, $\widehat{\dot{b}}_{\underline{s}}$ *and* $\widehat{b}'_{\underline{s}}$.

Numerical analysis: step 5

Proof (continued).

Defining $E_t = \widehat{\dot{S}}_t - \dot{S}_t$, *the difference between the two is*

$$
\begin{aligned}
E_t \;=\; & \int_0^t (\widehat{\dot{a}}_{\underline{s}} - \dot{a}_s) + (\widehat{a}'_{\underline{s}}\widehat{\dot{S}}_{\underline{s}} - a'_s \dot{S}_s)\,\mathrm{d}s + \int_0^t (\widehat{\dot{b}}_{\underline{s}} - \dot{b}_s) + (\widehat{b}'_{\underline{s}}\widehat{\dot{S}}_{\underline{s}} - b'_s \dot{S}_s)\,\mathrm{d}W_s \\
=\; & \int_0^t (\widehat{\dot{a}}_{\underline{s}} - \dot{a}_{\underline{s}}) + (\widehat{a}'_{\underline{s}}\widehat{\dot{S}}_{\underline{s}} - a'_{\underline{s}} \dot{S}_{\underline{s}}) + (\dot{a}_{\underline{s}} - \dot{a}_s) + (a'_{\underline{s}}\dot{S}_{\underline{s}} - a'_s \dot{S}_s)\,\mathrm{d}s \\
+\; & \int_0^t (\widehat{\dot{b}}_{\underline{s}} - \dot{b}_{\underline{s}}) + (\widehat{b}'_{\underline{s}}\widehat{\dot{S}}_{\underline{s}} - b'_{\underline{s}} \dot{S}_{\underline{s}}) + (\dot{b}_{\underline{s}} - \dot{b}_s) + (b'_{\underline{s}}\dot{S}_{\underline{s}} - b'_s \dot{S}_s)\,\mathrm{d}W_s \\
=\; & \int_0^t (\widehat{\dot{a}}_{\underline{s}} - \dot{a}_{\underline{s}}) + (\widehat{a}'_{\underline{s}} - a'_{\underline{s}})\widehat{\dot{S}}_{\underline{s}} + (\dot{a}_{\underline{s}} - \dot{a}_s) + (a'_{\underline{s}} - a'_s)\dot{S}_{\underline{s}} + a'_s(\dot{S}_{\underline{s}} - \dot{S}_s)\,\mathrm{d}s \\
+\; & \int_0^t (\widehat{\dot{b}}_{\underline{s}} - \dot{b}_{\underline{s}}) + (\widehat{b}'_{\underline{s}} - b'_{\underline{s}})\widehat{\dot{S}}_{\underline{s}} + (\dot{b}_{\underline{s}} - \dot{b}_s) + (b'_{\underline{s}} - b'_s)\dot{S}_{\underline{s}} + b'_s(\dot{S}_{\underline{s}} - \dot{S}_s)\,\mathrm{d}W_s \\
+\; & \int_0^t a'_{\underline{s}} E_{\underline{s}}\,\mathrm{d}s + \int_0^t b'_{\underline{s}} E_{\underline{s}}\,\mathrm{d}W_s.
\end{aligned}
$$

Numerical analysis: step 5

Proof (continued).

Defining

$$Z_t = \mathbb{E}\left[\sup_{0<s<t} |E_s|^p\right]$$

and bounding each of the terms in turn, using Hölder's inequality for products, such as

$$\mathbb{E}\left[\,|(a'_{\underline{s}}-a'_s)\dot{S}_{\underline{s}}|^p\right] \leq \mathbb{E}\left[\,|a'_{\underline{s}}-a'_s|^{2p}\right]^{1/2}\ \mathbb{E}\left[\,|\dot{S}_{\underline{s}}|^{2p}\right]^{1/2},$$

we end up with

$$Z_t \leq c_1\, h^{p/2} + c_2 \int_0^t Z_s\, \mathrm{d}s,$$

and Grönwall's inequality gives the desired result.

□

Numerical analysis: extensions

- higher derivatives – no problem based on this lemma

Lemma

If u_i, v_i $i=1,2,\ldots k$ are scalar random variables, and for any $p \geq 2$ there are finite constants C_p, D_p such that

$$\mathbb{E}[\,|u_i|^p] \leq C_p, \quad \mathbb{E}[\,|v_i|^p] \leq C_p, \quad \mathbb{E}[\,|u_i - v_i|^p] \leq D_p$$

for all i, then

$$\mathbb{E}\left[\,\left|\prod_{i=1}^{k} u_i - \prod_{i=1}^{k} v_i\right|^p\,\right] \leq k^p\, C_{pk}^{1-1/k}\, D_{pk}^{1/k}$$

- vector SDEs – no problem
- non-autonomous SDEs – no problem if a and b have bounded derivs in θ, S, t (probably OK if θ, S derivatives are $1/2$-Hölder in time)
- other discretisations – probably fine for Milstein discretisation

Conclusions

Pathwise sensitivity analysis has been used extensively for many years.

In the literature, the focus has been on conditions under which

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\,\mathbb{E}[\,P(S_T)\,] \;=\; \mathbb{E}\left[\,\frac{\mathrm{d}P}{\mathrm{d}S}\,\frac{\mathrm{d}S_T}{\mathrm{d}\theta}\,\right]$$

This work fills in an apparent gap in the literature concerning the strong convergence of the numerical approximations – this is essential for MLMC analysis.

It is also needed for new research on Multilevel Function Approximation, building on the original research of Stefan Heinrich in approximating

$$f(\theta) = \mathbb{E}[\,g(\theta;\omega)\,]$$

– subject of talks at MCQMC'24

References

F. De Angelis. PhD thesis, in preparation, Oxford University, 2024.

S. Burgos. 'The computation of Greeks with multilevel Monte Carlo'. PhD thesis, Oxford University, 2013.

P. L'Ecuyer. 'A unified View of the IPA, SF, and LR gradient estimation techniques'. Management Science, 36(11):1364-1383, 1990.

M.B. Giles. 'Multilevel Monte Carlo path simulation'. Operations Research, 56(3):607-617, 2008.

P. Glasserman. 'Monte Carlo Methods in Financial Engineering'. Springer, 2003.

P.E. Kloeden, E. Platen. 'Numerical Solution of Stochastic Differential Equations'. Springer, 1992.