Stochastic Numerical Analysis

Prof. Mike Giles

mike.giles@maths.ox.ac.uk

Oxford University Mathematical Institute

Euler-Maruyama method

The simplest approximation for the scalar SDE

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

is the forward Euler scheme, which is known as the Euler-Maruyama approximation when applied to SDEs:

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

Here h is the timestep, $\widehat{S}_n$ is the approximation to $S(nh)$ and the ΔW_n are i.i.d. $N(0,h)$ Brownian increments.

Euler-Maruyama method

For ODEs, the forward Euler method has $O(h)$ accuracy, and other more accurate methods would usually be preferred.

However, SDEs are very much harder to approximate so the Euler-Maruyama method is used widely in practice.

Numerical analysis is also very difficult and even the definition of “accuracy” is tricky.

Weak convergence

In most applications, we are mostly concerned with **weak** errors, the error in the expected value of some output quantity, due to using a finite timestep h.

For an output which is a function of $S(T)$, the weak error is

$$\mathbb{E}[f(S(T))] - \mathbb{E}[f(\widehat{S}_{T/h})]$$

For an output which depends on the whole path, the weak error is

$$\mathbb{E}[f(S)] - \mathbb{E}[\widehat{f}(\widehat{S})]$$

where $f(S)$ is a function of the entire path $S(t)$, and $\widehat{f}(\widehat{S})$ is a corresponding approximation using the whole discrete path.

Weak convergence

Key theoretical result (Bally and Talay, 1995):

If $p(S)$ is the p.d.f. for $S(T)$ and $\widehat{p}(S)$ is the p.d.f. for $\widehat{S}_{T/h}$ computed using the Euler-Maruyama approximation, then under certain conditions on $a(S,t)$ and $b(S,t)$ (in particular that they are C^∞ with bounded derivatives)

$$p(S) - \widehat{p}(S) = O(h)$$

and hence

$$\mathbb{E}[f(S(T))] - \mathbb{E}[f(\widehat{S}_{T/h})] = O(h)$$

We will not go through the analysis – will instead focus on alternative strong convergence.

Weak convergence

Numerical demonstration: Geometric Brownian Motion

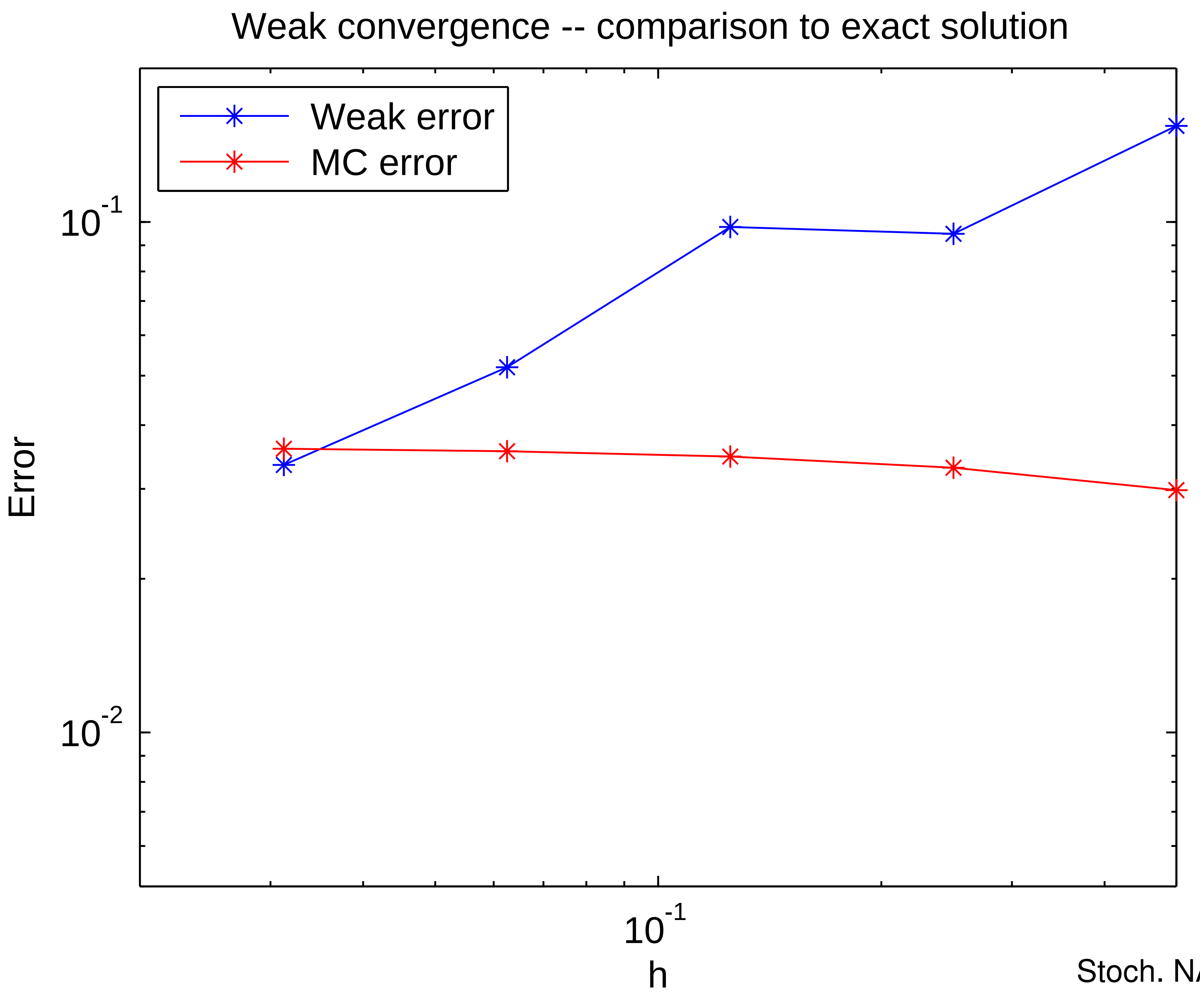
$$\mathrm{d}S = r\,S\,\mathrm{d}t + \sigma\,S\,\mathrm{d}W$$

$S_0 = 100, r = 0.05,\ \sigma = 0.5,\ T = 1$

Financial call option: $\mathbb{E}[\exp(-rT)\max(0, S(T)-K)]$
with $K = 110$ – there is a known analytic value for this.

Plot shows weak error versus analytic expectation when using 10^8 paths, and also Monte Carlo error (3 standard deviations)

Weak convergence



Weak convergence

Previous plot showed difference between exact expectation and numerical approximation.

What if the exact solution is unknown? Compare approximations with timesteps h and $2h$.

If

$$\mathbb{E}[f(S(T))] - \mathbb{E}[f(\widehat{S}^h_{T/h})] \approx a\, h$$

then

$$\mathbb{E}[f(S(T))] - \mathbb{E}[f(\widehat{S}^{2h}_{T/2h})] \approx 2\, a\, h$$

and so

$$\mathbb{E}[f(\widehat{S}^h_{T/h})] - \mathbb{E}[f(\widehat{S}^{2h}_{T/2h})] \approx a\, h$$

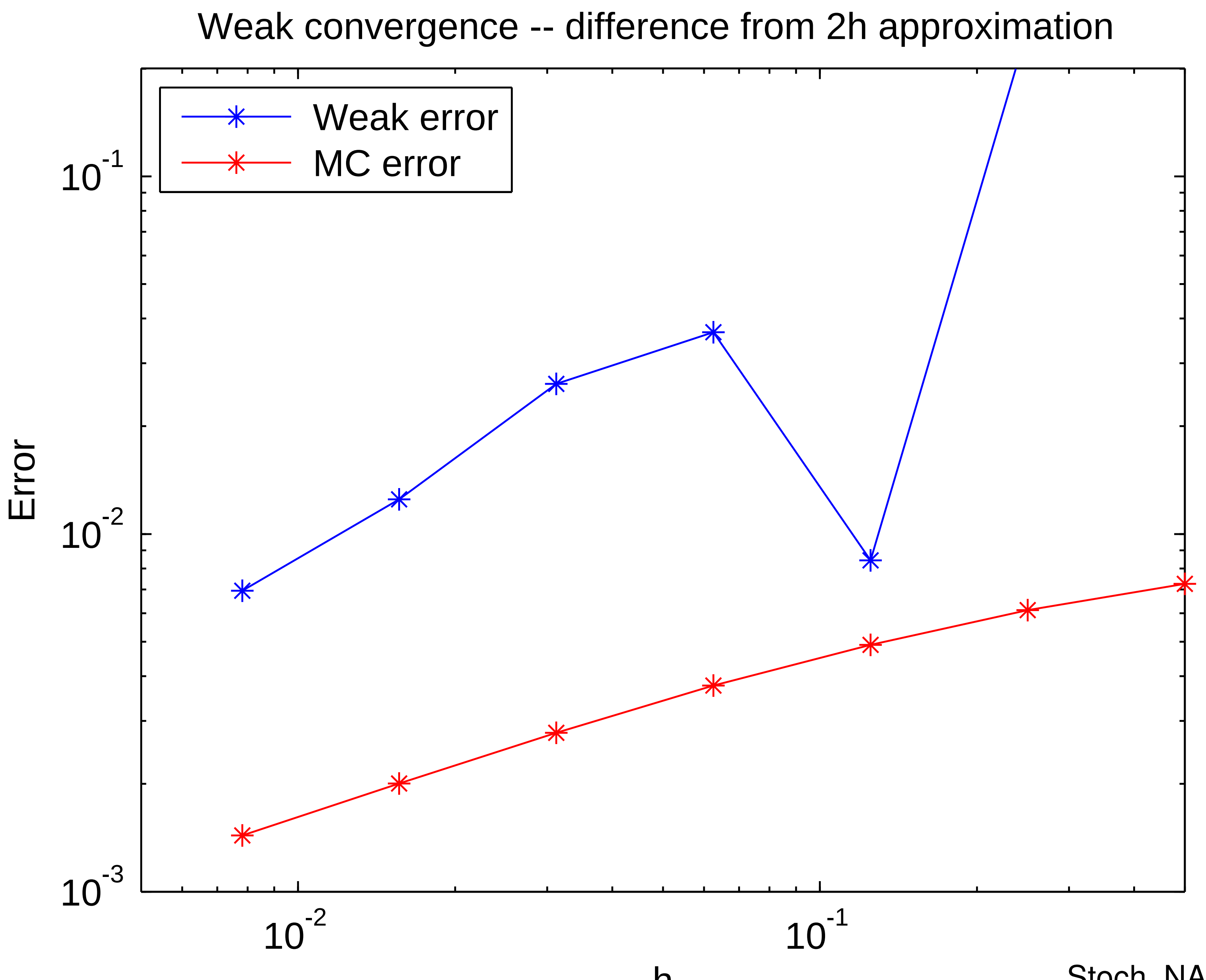
Weak convergence

To minimise the number of paths that need to be simulated, best to use **same** driving Brownian path when doing $2h$ and h approximations – i.e. take Brownian increments for h simulation and sum in pairs to get Brownian increments for $2h$ simulation.

The variance is lower because the h and $2h$ paths are close to each other (**strong** convergence).

In a later lecture, this forms the basis for the **Multilevel Monte Carlo** method (Giles, 2006)

Weak convergence



Strong convergence

Strong convergence looks instead at the average error in each individual path, either at a final time:

$$\mathbb{E}\left[\,\left|S(T)-\widehat{S}_{T/h}\right|\,\right] \quad \text{or} \quad \left(\mathbb{E}\left[\left(S(T)-\widehat{S}_{T/h}\right)^2\right]\right)^{1/2}$$

or a maximum over the path:

$$\mathbb{E}\left[\max_n \left|S(t_n)-\widehat{S}_n\right|\,\right] \quad \text{or} \quad \left(\mathbb{E}\left[\max_n \left(S(t_n)-\widehat{S}_n\right)^2\right]\right)^{1/2}$$

The main theoretical result (Kloeden & Platen 1992) is that for the Euler-Maruyama method under certain conditions on $a(S,t)$ and $b(S,t)$ these are both $O(\sqrt{h})$.

We will do the full analysis in lecture 6.

Strong convergence

Thus, each approximate path deviates by $O(\sqrt{h})$ from its true path.

How can the weak error be $O(h)$? Because the error

$$S(T) - \widehat{S}_{T/h}$$

has mean $O(h)$ even though the r.m.s. is $O(\sqrt{h})$.

(In fact to leading order it is normally distributed with zero mean and variance $O(h)$.)

Strong convergence

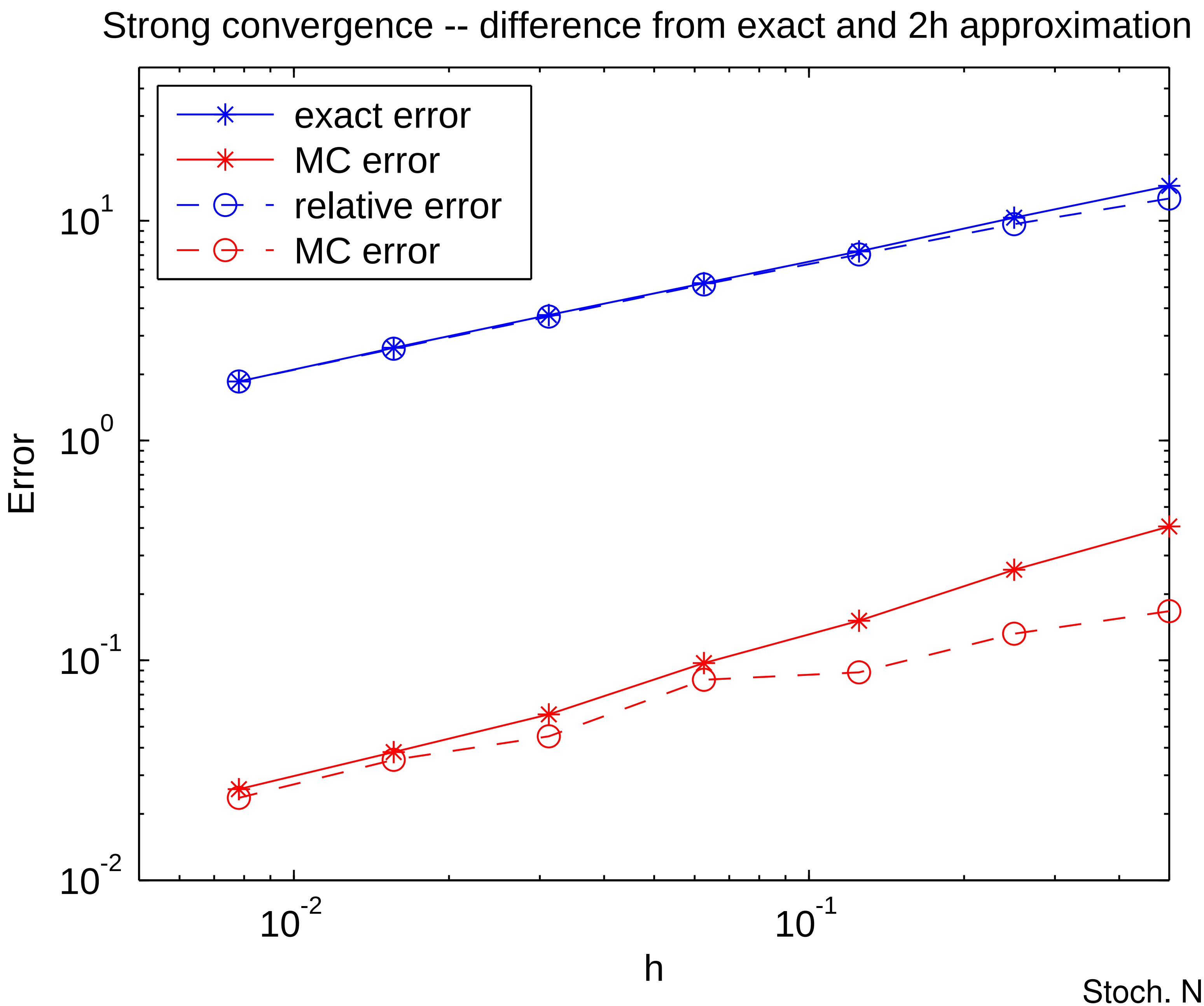
Numerical demonstration based on same Geometric Brownian Motion.

Plot shows two curves, one showing the difference from the true solution

$$S(T) = S_0 \, \exp\left((r - \tfrac{1}{2}\sigma^2)T + \sigma\, W(T) \right)$$

and the other showing the difference from the $2h$ approximation

Strong convergence



Mean Square Error

If the true option value is $V = \mathbb{E}[f]$

and the discrete approximation is $\widehat{V} = \mathbb{E}[\widehat{f}]$

and the Monte Carlo estimate is $\widehat{Y} = \dfrac{1}{N}\displaystyle\sum_{n=1}^{N} \widehat{f}^{(n)}$

the Mean Square Error is

$$\begin{aligned}
\mathbb{E}\left[\left(\widehat{Y} - V\right)^2\right] &= \mathbb{E}\left[\left(\widehat{Y} - \mathbb{E}[\widehat{f}] \;+\; \mathbb{E}[\widehat{f}] - \mathbb{E}[f]\right)^2\right] \\
&= \mathbb{E}\left[(\widehat{Y} - \mathbb{E}[\widehat{f}])^2\right] + (\mathbb{E}[\widehat{f}] - \mathbb{E}[f])^2 \\
&= N^{-1}\mathbb{V}[\widehat{f}] + \left(\mathbb{E}[\widehat{f}] - \mathbb{E}[f]\right)^2
\end{aligned}$$

Mean Square Error

If there are M timesteps, the computational cost is proportional to $C = NM$ and the MSE is approximately

$$a\,N^{-1} + b\,M^{-2} = a\,N^{-1} + b\,C^{-2}N^{2}.$$

– can optimise N for a given accuracy.

To achieve a RMS error of ε requires $h = O(\varepsilon)$, and $N = O(\varepsilon^{-2})$ so the total cost is $O(\varepsilon^{-3})$.

Milstein Method

Starting from the integral equation:

$$S(t) = S(0) + \int_0^t a(S(s), s)\, \mathrm{d}s + \int_0^t b(S(s), s)\, \mathrm{d}W(s),$$

approximating this on interval $[0, h]$ using

$$a(S(t), t) \approx a(S(0), 0), \qquad b(S(t), t) \approx b(S(0), 0)$$

gives Euler-Maruyama method for first timestep

$$\widehat{S}_1 = \widehat{S}_0 + a_0\, h + b_0\, \Delta W_0.$$

Milstein Method

To leading order,

$$S(t) = S(0) + b(S(0),0)\, W(t) + O(h)$$

and hence

$$\begin{aligned} b(S(t),t)) &= b(S(0),0) + b'(S(0),0)\,(S(t) - S(0)) + O(h) \\ &= b(S(0),0) + b'(S(0),0)\; b(S(0),0)\; W(t) + O(h) \end{aligned}$$

This then leads to

$$S(h) = S(0) + a_0\, h + b_0\, W(h) + b'_0\, b_0 \int_0^h W(t)\, \mathrm{d}W(t) \;\; + O(h^{3/2})$$

where $a_0,\ b_0,\ b'_0$ are all evaluated at $(S(0),0)$.

Milstein Method

Already shown that

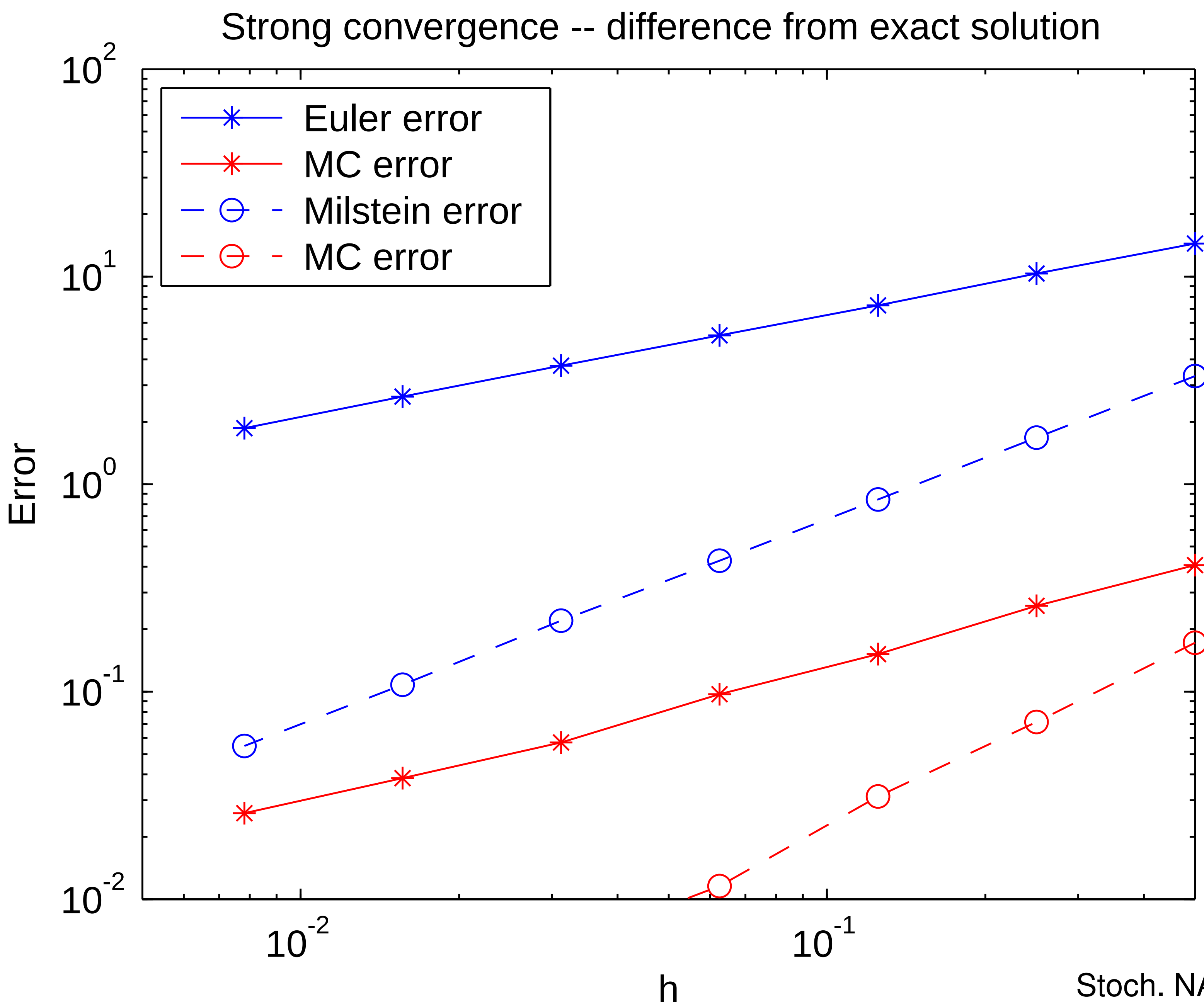
$$\int_0^h W(t)\,\mathrm{d}W(t) = \tfrac{1}{2}\left(W^2(h) - h\right)$$

which then gives us the Milstein scheme:

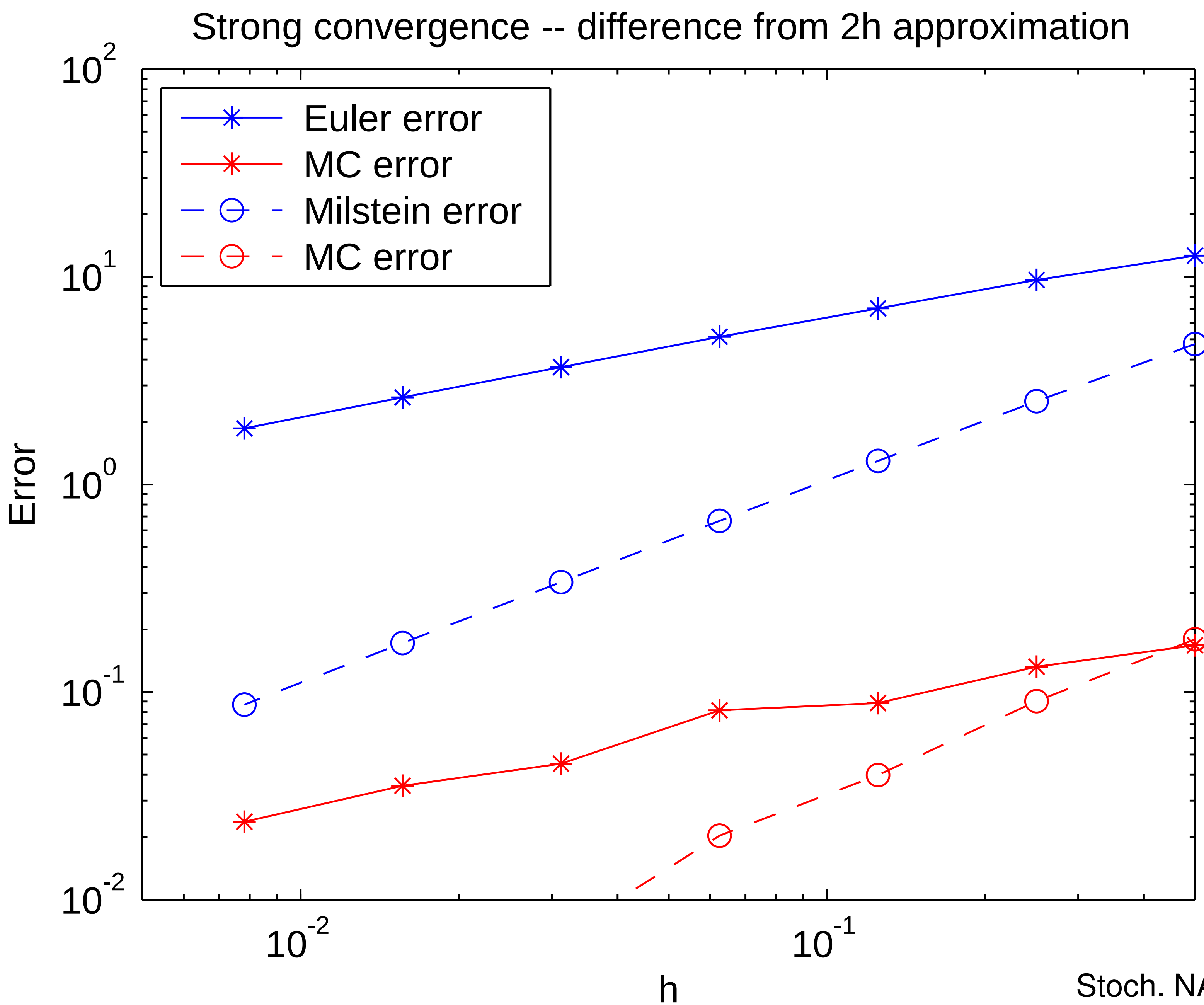
$$\begin{aligned}\widehat{S}_{n+1} &= \widehat{S}_n + a(\widehat{S}_n, t_n)\,h + b(\widehat{S}_n, t_n)\,\Delta W_n \\ &\quad + \tfrac{1}{2}\,b'(\widehat{S}_n, t_n)\,b(\widehat{S}_n, t_n)\left(\Delta W_n^2 - h\right)\end{aligned}$$

The weak error is still $O(h)$ for Lipschitz outputs but the strong error is now $O(h)$.

Strong convergence



Strong convergence



Predictor-Corrector Method

Predictor step:

$$\widehat{S}^{(p)}_{n+1} = \widehat{S}_n + a(\widehat{S}_n, t_n)\, h + b(\widehat{S}_n, t_n)\, \Delta W_n$$

Corrector step:

$$\widehat{S}_{n+1} = \widehat{S}_n + \tfrac{1}{2}\left(a(\widehat{S}_n, t_n) + a(\widehat{S}^{(p)}_{n+1}, t_{n+1}) \right) h + b(\widehat{S}_n, t_n)\, \Delta W_n$$

The weak error is $O(h)$ for Lipschitz outputs, and the strong error is $O(h^{1/2})$. Advantage of this approximation is that it is $O(h^2)$ when $b \equiv 0$, so good for applications when $b \ll a$.

Generalisations of this are mentioned in the book by Kloeden & Platen.

Implicit Euler Method

Nonlinear version:

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

Linearised version:

$$\widehat{S}_{n+1} = \widehat{S}_n + \left(a(\widehat{S}_n, t_n) + a'(\widehat{S}_n, t_n)(\widehat{S}_{n+1} - \widehat{S}_n) \right) h + b(\widehat{S}_n, t_n)\, \Delta W_n$$

Again, the weak error is $O(h)$, strong error is $O(h^{1/2})$.

Advantage of these is that they are stable for applications with rapid linear or nonlinear reversion:

$$\begin{aligned} \mathrm{d}S_t &= -\kappa\, S_t\, \mathrm{d}t + \sigma\, \mathrm{d}W_t \\ \mathrm{d}S_t &= -\kappa\, S_t^3\, \mathrm{d}t + \sigma\, S_t\, \mathrm{d}W_t \end{aligned}$$