

Discretization of SDEs: Euler Methods and Beyond

Christian Bayer

09-26-2006 / PRisMa 2006 Workshop

Outline

- 1 Introduction
 - Motivation
 - Stochastic Differential Equations
- 2 The Euler-Maruyama Scheme
 - Time Discretization of SDEs
 - Monte-Carlo Simulation
- 3 Higher Order Methods
 - Stochastic Taylor Schemes
 - The Milstein Scheme
 - The Milstein Scheme with Approximate Heat Kernels
- 4 Summary

Applications of SDEs

- In mathematical finance, financial markets are often modelled by the solution X_t of a stochastic differential equation (SDE). The price (at time 0) of a European option with maturity T and payoff $f(X_T)$ is given by an expression like

$$C(X_0) = E\left(e^{-rT} f(X_T)\right).$$

- More generally, given a differential operator L and let u denote the solution of the corresponding Cauchy problem

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) = Lu(t, x), & t \in]0, T], x \in \mathbb{R}^n \\ u(0, x) = f(x), & x \in \mathbb{R}^n \end{cases}.$$

A *stochastic representation* is the solution S_t of an SDE s. t.

$$u(t, x) = E(f(X_T) | X_t = x), \quad t \in]0, T], x \in \mathbb{R}^n.$$

Definition

Given a d -dimensional Brownian motion B defined on the probability space (Ω, \mathcal{F}, P) , $T > 0$, functions $a : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and $\sigma : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times d}$ being Lipschitz with at most linear growth.

Definition

By a *stochastic differential equation* we understand an equation

$$X_t^x = x + \int_0^t a(X_s^x) ds + \int_0^t \sigma(X_s^x) dB_s, \quad t \in [0, T], \quad x \in \mathbb{R}^n. \quad (1)$$

“ σdB_t ” should be interpreted as a matrix-vector multiplication.

We will often use the short-hand notation

$$\begin{cases} dX_t^x = a(X_t^x)dt + \sigma(X_t^x)dB_t, & t \in]0, T[\\ X_0^x = x \end{cases} . \quad (2)$$

Feynman-Kac Formula

Let L denote the second-order differential operator on \mathbb{R}^n given by

$$Lf(x) = \sum_{i=1}^n a^i(x) \frac{\partial f}{\partial x^i}(x) + \frac{1}{2} \sum_{i,j=1}^n b^{ij}(x) \frac{\partial^2 f}{\partial x^i \partial x^j}(x),$$

where $b = (b^{ij})$ is the $n \times n$ -matrix defined by $b = \sigma \sigma^T$.

Theorem (Feynman-Kac Formula)

Let $f \in C_c^2(\mathbb{R}^n)$ and $q \in C(\mathbb{R}^n)$ bounded from below and define

$$u(t, x) = E\left(\exp\left(\int_t^T q(X_s) ds\right) f(X_T) \mid X_t = x\right), \quad t \in [0, T], \quad x \in \mathbb{R}^n.$$

Then u solves the parabolic Cauchy problem

$$u_t(t, x) + Lu(t, x) + q(x)u(t, x) = 0, \quad u(T, x) = f(x).$$

Euler Scheme for SDEs

We present an approximation for the solution X_T^x of the SDE (2).

Definition

Fix $x \in \mathbb{R}^n$ and let $\Delta^{(N)}$ be the equidistant mesh of size N , i. e. $t_i = \frac{i}{N}T$, $i = 0, \dots, N$. Set $\Delta t_i = \frac{T}{N}$ and $\Delta B_i = B_{t_{i+1}} - B_{t_i}$, $i = 0, \dots, N-1$. Define the *uniform Euler-Maruyama approximation* to X_T^x by $\bar{X}_0^{(N)} = x$ and

$$\bar{X}_{i+1}^{(N)} = \bar{X}_i^{(N)} + a(\bar{X}_i^{(N)})\Delta t_i + \sigma(\bar{X}_i^{(N)})\Delta B_i, \quad i = 0, \dots, N-1. \quad (3)$$

Remark

Note that the discretization $\bar{X}_i^{(N)}$ – in this form – is only “discrete” in the time variable but not as a random variable.

Strong Convergence of the Euler-Maruyama Scheme

Definition

Given a sequence $\bar{X}^{(N)}$ of time-discrete approximations of X_T^x along time partitions $\Delta^{(N)}$. $\bar{X}_N^{(N)}$ converges *strongly* to X_T^x if it converges in $L^1(\Omega)$, i. e. if $\lim_{N \rightarrow \infty} E(|X_T^x - \bar{X}_N^{(N)}|) = 0$. It converges *strongly with order* $\gamma > 0$ if there is a constant $C > 0$ such that – with $|\Delta^{(N)}| = \max(t_{i+1} - t_i)$ –

$$E(|X_T^x - \bar{X}_N^{(N)}|) \leq C |\Delta^{(N)}|^\gamma.$$

Theorem

The uniform Euler-Maruyama scheme converges strongly of order $\gamma = \frac{1}{2}$ to the true solution X_T^x .

Weak Convergence of the Euler-Maruyama Scheme

Let $C_p^k(\mathbb{R}^n)$ denote the space of all C^k -functions of polynomial growth (together with their derivatives).

Definition

$\bar{X}_N^{(N)}$ converges weakly to X_T^x if, for some $k \in \mathbb{N}$,

$$\lim_{N \rightarrow \infty} E(f(\bar{X}_N^{(N)})) = E(f(X_T^x)), \quad \forall f \in C_p^k(\mathbb{R}^n).$$

$\bar{X}_N^{(N)}$ converges to X_T^x with *weak order* $\gamma > 0$ if $\exists C > 0$ s. t.

$$|E(f(\bar{X}_N^{(N)})) - E(f(X_T^x))| \leq C |\Delta^{(N)}|^\gamma, \quad \forall f \in C_p^{2(\gamma+1)}$$

and for all N large enough.

Weak Convergence of the Euler-Maruyama Scheme – 2

Remark

- 1 *The notion of weak convergence seems to be more appropriate for our problem than the – stronger – notion of strong convergence, since we actually want to approximate quantities of the form $E(f(X_T^x))$.*
- 2 *In probability theory, the notion of weak convergence is usually defined with respect to the space $C_b(\mathbb{R}^n)$ of continuous, bounded functions instead of the space $C_p^k(\mathbb{R}^n)$.*

Theorem

Let the data a and σ be C^4 . Then the uniform Euler-Maruyama scheme converges weakly with order 1.

Monte-Carlo Simulation

Proposition

Given an \mathbb{R}^n -valued random variable X and a measurable function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ such that $f(X) \in L^2(\Omega)$ with variance $\sigma^2(f)$. Let X_m , $m \in \mathbb{N}$, denote a sequence of independent copies of X . Then

$$\lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M f(X_m) = E(f(X)) \text{ a. s.}$$

and, in the sense of convergence in distribution,

$$\sqrt{M} \frac{\frac{1}{M} \sum_{m=1}^M f(X_m) - E(f(X))}{\sigma(f)} \xrightarrow{M \rightarrow \infty} \mathcal{N}(0, 1).$$

Error of Monte-Carlo Simulation

Remark

- ① *Monte-Carlo simulation allows approximation of $E(f(X))$ with error control provided that it is feasible to generate random numbers from X : heuristically, if M is large enough,*

$$\frac{1}{M} \sum_{m=1}^M f(X_m) - E(f(X)) \sim \mathcal{N}\left(0, \frac{\sigma^2(f)}{M}\right).$$

- ② *It is a method of order $\frac{1}{2}$. Given $\epsilon > 0$ and $0 < \delta < 1$, to guarantee that $|\frac{1}{M} \sum_{m=1}^M f(X_m) - E(f(X))| < \epsilon$ with probability $1 - \delta$ we have to choose $M \geq \left(\frac{p\sigma(f)}{\epsilon}\right)^2$, where p is the $(1 - \frac{\delta}{2})$ -quantile of a standard Gaussian random variable, i. e. $P(|Z| \geq p) = \delta$ for $Z \sim \mathcal{N}(0, 1)$.*

Monte-Carlo Simulation in the Euler-Maruyama Scheme

- We want to approximate the unknown value $E(f(X_T^x))$ by the value $E(f(\bar{X}_N^{(N)}))$, but – in general – it is impossible to calculate the latter expected value.
- Contrary to X_T^x , it is possible to generate random numbers according to the distribution of $\bar{X}_N^{(N)}$, see (3).
- This allows approximation of $E(f(\bar{X}_n^{(N)}))$ using *Monte-Carlo simulation*: for fixed N, M let $(\bar{X}_N^{(N,m)})_{m \in \mathbb{N}}$ denote a sequence of independent copies of $\bar{X}_N^{(N)}$. Then we approximate

$$E(f(\bar{X}_N^{(N)})) \approx \frac{1}{M} \sum_{m=1}^M f(\bar{X}_N^{(N,m)}).$$

A New Error Regime

- Monte-Carlo simulation introduces an additional, stochastic error to the Euler scheme.
- The total error naturally splits into two parts according to

$$\begin{aligned} \text{Error} &= \left| E(f(X_T^x)) - \frac{1}{M} \sum_{m=1}^M f(\bar{X}_N^{(N,m)}) \right| \\ &\leq \left| E(f(X_T^x)) - E(f(\bar{X}_N^{(N)})) \right| \\ &\quad + \left| E(f(\bar{X}_N^{(N)})) - \frac{1}{M} \sum_{m=1}^M f(\bar{X}_N^{(N,m)}) \right|. \end{aligned}$$

The first term is called *time-discretization error* and the second term is called *statistical error*.

A Crude Error Analysis

Remark

- *The slow convergence rates of the Euler-Maruyama scheme seem to indicate the need for higher order methods.*
- *Fix an error tolerance $\epsilon > 0$ and $0 < \lambda < 1$. Reserve $\lambda\epsilon$ for the discretization and $(1 - \lambda)\epsilon$ for the statistical error. Using a p -order discretization method, we need*

$$N(\lambda\epsilon) = C_1(\lambda\epsilon)^{-1/p}, \quad M((1 - \lambda)\epsilon) = C_2((1 - \lambda)\epsilon)^{-2}.$$

- *Consequently, the total work is proportional to*

$$N(\lambda\epsilon)M((1 - \lambda)\epsilon) = C_1 C_2 \lambda^{-1/p} (1 - \lambda)^{-2} \epsilon^{-(2 + \frac{1}{p})}.$$

The Stratonovich Stochastic Integral

Definition

For a continuous semimartingale Y , the *Stratonovich integral* of Y with respect to Brownian motion is defined by

$$\int_0^t Y_s \circ dB_s^i = \int_0^t Y_s dB_s^i + \frac{1}{2}[Y, B^i]_t.$$

Remark

We can switch between SDEs in Stratonovich and Itô form by

$$\int_0^t h(X_s) \circ dB_s = \int_0^t h(X_s) dB_s + \frac{1}{2} \sum_{j=1}^n \sum_{k=1}^d \sigma^{jk}(X_s) \frac{\partial h^k}{\partial x^j}(X_s) ds.$$

Stochastic Taylor Expansion

Theorem

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be C^∞ -bounded. Define differential operators $A^j = \sum_{k=1}^n \sigma^{kj} \frac{\partial}{\partial x^k}$ and $A^0 = \sum_{k=1}^n (a^k - \frac{1}{2} \sum_{j=1}^d A^j \sigma^{kj}) \frac{\partial}{\partial x^k}$, $j = 1, \dots, d$. For $\alpha = (i_1, \dots, i_k) \in \{0, \dots, d\}^k$ set $f_\alpha = A^{i_1} \dots A^{i_k} f$ and construct the iterated Stratonovich integrals by $J_t^\alpha = \int_{0 < t_1 < \dots < t_k < t} \circ dB_{t_1}^{i_1} \dots \circ dB_{t_k}^{i_k}$ with the convention that $\circ dB_s^0 = ds$. For $m \in \mathbb{N}$ we have

$$f(X_t^x) = \sum_{\substack{k \in \mathbb{N}, \alpha \in \{0, \dots, d\}^k, \\ k + \#\{i_j=0\} \leq m}} f_\alpha(x) J_t^\alpha + R_m(t, x, f), \quad (4)$$

where $\sup_{x \in \mathbb{R}^n} \sqrt{E(R_m^2)} = \mathcal{O}\left(t^{\frac{m+1+1}{2}}_{1, \infty} \llbracket t \rrbracket\right)$.

Taylor Schemes

Remark

- 1 *The stochastic Taylor expansion is the analogue of the Taylor expansion for deterministic functions. Note that the iterated Stratonovich integrals J^α play the rôle of the polynomials.*
- 2 *The intriguing summation boundary " $k + \#\{i_j = 0\}$ " comes from the different scaling of Brownian motion – corresponding to non-zero indices – and time – corresponding to the index 0 – which is sometimes described by " $dB_t \approx \sqrt{dt}$ ".
Consequently, time indices have to be counted twice.*

It seems plausible to generalize the Euler-Maruyama scheme by schemes where the increment (3) is obtained by truncating the stochastic Taylor expansion at some point. These schemes are called *Taylor schemes*.

Definition of the Milstein Scheme

Definition

Given a (uniform) partition as before. The Milstein scheme is defined by $\bar{X}_0^{(N)} = x$ and, for $i = 0, \dots, N-1$,

$$\bar{X}_{i+1}^{(N)} = \bar{X}_i^{(N)} + a(\bar{X}_i^{(N)})\Delta t_i + \sigma(\bar{X}_i^{(N)})\Delta B_i + \sum_{j,k=1}^d A^j \sigma^k(\bar{X}_i^{(N)}) I_{\Delta t_i}^{(j,k)},$$

where $I_{\Delta t_i}^{(j,k)} = \int_{t_i}^{t_{i+1}} (B_t^j - B_{t_i}^j) dB_t^k$.

Remark

The Milstein scheme is obtained by using the second order stochastic Taylor expansion (after rewriting it in Itô form).

Properties of the Milstein Scheme

Theorem

The Milstein scheme converges strongly and weakly and is of order 1 in both senses.

Remark

The weak orders of the Milstein and the Euler-Maruyama schemes coincide. Since we are mostly interested in weak convergence, it might not seem sensible to use the Milstein scheme. Note, however, that one can construct Taylor schemes of any order and the Milstein scheme already has most of the properties of higher order Taylor schemes. Therefore, we may use it exemplorarily for general higher order schemes.

Sampling Iterated Itô Integrals

- If we want to apply the Milstein scheme we need to be able to sample random numbers according to the distribution of the $d(d+1)$ -dimensional random variable

$$(\Delta B_i^1, \dots, \Delta B_i^d, I_{\Delta t_i}^{(1,1)}, \dots, I_{\Delta t_i}^{(j,k)}, \dots, I_{\Delta t_i}^{(d,d)}). \quad (5)$$

- $I_{\Delta t_i}^{(j,j)}$, $1 \leq j \leq d$, can be expressed in terms of Brownian motion in an algebraic way by $I_{\Delta t_i}^{(j,j)} = \frac{1}{2}((\Delta B_i^j)^2 - \Delta t_i)$.
- The mixed terms $I_{\Delta t_i}^{(i,j)}$ depend on B in a non-trivial way. This makes sampling from the afore-mentioned random vector a costly task.

Embedding Higher Order Taylor Schemes in a Lie Group

- For the scheme constructed by using the stochastic Taylor expansion of order m , we need to be able to sample from the distribution of the random vector of all J_t^α with α running through all the multi-indices in $\{0, \dots, d\}$ such that $|\alpha| + \#\{i_j = 0\} \leq m$, c. f. (4).
- The analysis of this random vector is greatly simplified if we interpret the vector of iterated Stratonovich integrals of order up to m as a stochastic process taking values in some Lie group $G_{d,1}^m$, which is nilpotent of order m .

Definition of the Lie Group

- Let $\mathbb{A}_{d,1}^m$ be the (associative, *non-commutative*) algebra of all polynomials in e_0, \dots, e_d with degree less or equal to m , where all occurrences of e_0 are counted twice. We truncate all terms of higher degrees when multiplying two such polynomials.
- We encode the vector of iterated Stratonovich integrals as an element of the algebra by setting $Y_t^1 = 1 + \sum J_t^\alpha e_\alpha$, where the sum runs over $|\alpha| + \#\{j_j = 0\} \leq m$ and $e_{(i_1, \dots, i_k)} = e_{i_1} \cdots e_{i_k}$.

Definition

We define $G_{d,1}^m = \{Y_t^1(\omega) \mid \omega \in \Omega\}$ for any $t > 0$.

Remark

$G_{d,1}^m$ is a Lie group having a global chart. Note that we would only need all paths ω of bounded variation in the definition of $G_{d,1}^m$.

The Vector of Iterated Integrals as a Process in $G_{d,1}^m$

For $y \in G_{d,1}^m$ consider the stochastic differential equation in $G_{d,1}^m$

$$dY_t^y = Y_t^y e_0 dt + \sum_{i=1}^d Y_t^y e_i \circ dB_t^i. \quad (6)$$

Proposition (Teichmann 2006)

The vector Y_t^1 of iterated Stratonovich integrals coincides with the solution to the SDE (6) with initial value 1.

Remark

The infinitesimal generator to (6) is the sub-Laplacian on $G_{d,1}^m$. Therefore, we call the fundamental solution of the SDE – and density of the iterated integrals – the heat kernel on $G_{d,1}^m$.

Approximation of the Heat Kernel

- Note that the coefficient of Y_t^1 in e_0 is equal to t . Consequently, the heat kernel as density of Y_t^1 does not exist albeit in a generalized function sense.
- If we factorize $\mathbb{A}_{d,1}^m$ w. r. t. the subspace generated by e_0 , then one can show that the heat kernel exists as a function in the Schwartz space.
- This suggests approximating the heat kernel by Hermite functions. Interpret Y_t^1 as a process taking values in some \mathbb{R}^K . Let Σ_t be a suitable covariance matrix thereon and let $p_t(z)$ and $h_\beta(t, z)$ respectively denote the density of $\mathcal{N}(0, \Sigma_t)$ and the corresponding Hermite polynomials, $z \in \mathbb{R}^K$, $\beta \in \mathbb{N}^K$. There is an efficient procedure for calculating the coefficients a_t^β of the expansion

$$\frac{p_t(\cdot)}{r_t(\cdot)} = \sum a_t^\beta h_\beta(t, \cdot). \quad (7)$$

Milstein Scheme with Approximate Densities

- In the step from t_i to t_{i+1} , the increment of the Milstein scheme depends on (an independent copy of) $Y_{\Delta t_i}^1$, consequently $\bar{X}_N^{(N)}$ depends on Z_0, \dots, Z_{N-1} , where $Z_i \sim Y_{\Delta t_i}^1$ independent of each other. We write $\bar{X}_N^{(N)}(Z_0, \dots, Z_{N-1})$.
- Therefore, we immediately get

$$E(f(\bar{X}_N^{(N)})) = \int_{\mathbb{R}^{KN}} f(\bar{X}^{(N)}(z_0, \dots, z_{N-1})) p_{\Delta t_0}(z_0) \cdots p_{\Delta t_{N-1}}(z_{N-1}) dz_0 \cdots dz_{N-1}.$$

- Fix an approximation \tilde{p}_t of the heat kernel by truncating the Hermite expansion at some point and a probability Q_t on \mathbb{R}^K with density $q_t > 0$, e. g. $Q_t = \mathcal{N}(0, \Sigma_t)$.

Milstein Scheme with Approximate Densities – 2

Proposition

For $M \in \mathbb{N}$ let $Z_i^{(m)}$, $m = 1, \dots, M$, $i = 0, \dots, N-1$, be a sequence of independent, $Q_{\Delta t_i}$ -distributed random variables.

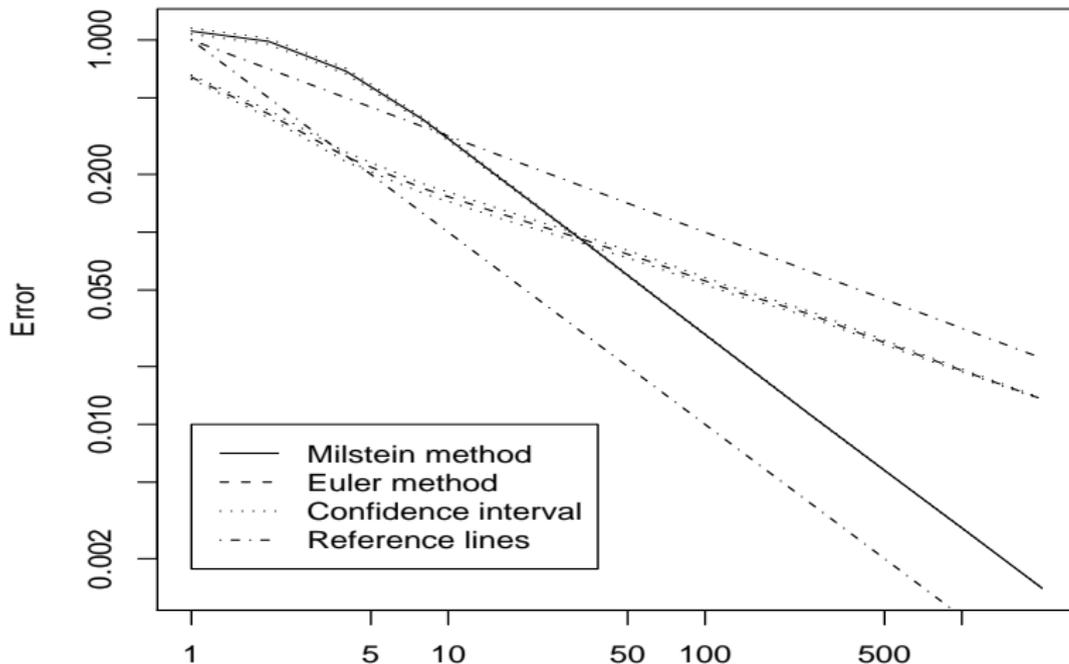
$$E(f(\bar{X}_N^{(N)})) = \lim_{M \rightarrow \infty} \frac{1}{M} \sum_{m=1}^M f(\bar{X}_N^{(N)}(Z_0^{(m)}, \dots, Z_{N-1}^{(m)}))$$
$$\frac{\tilde{p}_{\Delta t_0}(Z_0^{(m)})}{q_{\Delta t_0}(Z_0^{(m)})} \dots \frac{\tilde{p}_{\Delta t_{N-1}}(Z_{N-1}^{(m)})}{q_{\Delta t_{N-1}}(Z_{N-1}^{(m)})}.$$

Remark

- The choice of Σ_t and Q_t is crucial.
- Note that \tilde{p}_t is not a true density and takes negative values.

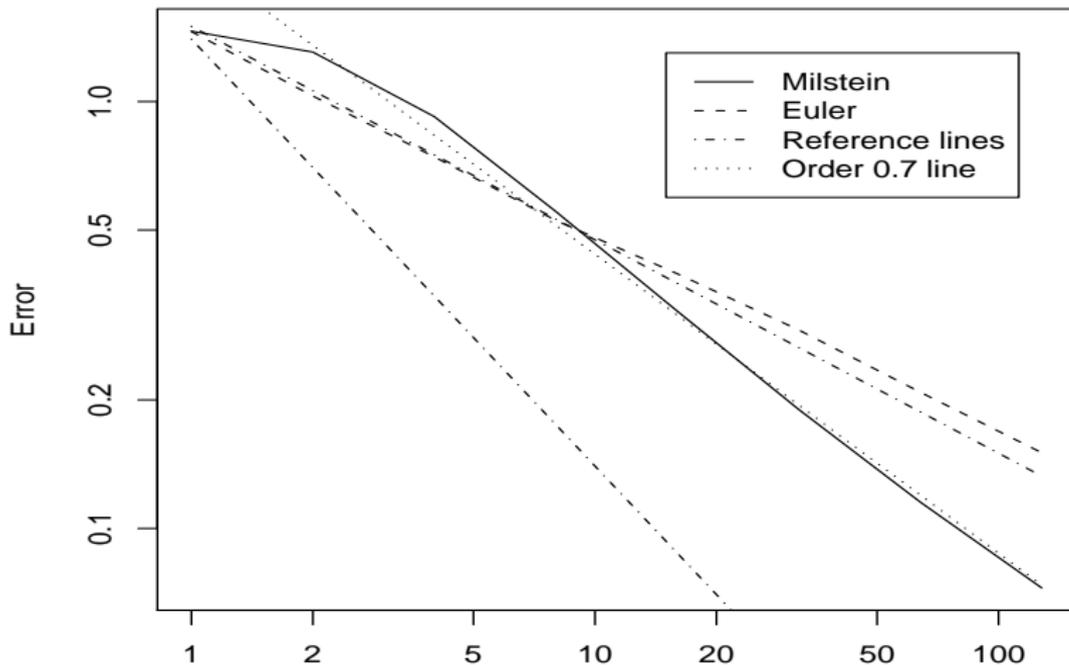
First Results for the Weak Error

Call on the sphere (non-hypo., non-comm.)



First Results for the Strong Error

Strong error sphere (non-comm. matrices)



Summary

- The Euler-Maruyama scheme for discretization of SDEs is simple to understand and implement, but suffers from a low order of convergence, especially in the strong sense.
- Typically, the cost of the Monte-Carlo simulation overshadows this defect and it seems to be more important to accelerate the simulation than to improve the order.
- In situations with dominating time discretization error, stochastic Taylor methods provide methods of any order, but the simulation is costly. One possible alternative might be the use of approximate heat kernels.
- One should also consider Richardson extrapolation – a generalization of Romberg extrapolation – as a mean to get higher order methods.

References

-  Peter E. Kloeden and Eckhard Platen, *Numerical solutions of stochastic differential equations*, Applications of Mathematics 23, 1992.
-  Denis Talay and Luciano Tubaro, *Expansion of the global error for numerical schemes solving stochastic differential equations*, Stochastic Anal. Appl. 8, No. 4, 483–509, 1990.