Numerical integration of stochastic differential equations

Luis M. Abia, Óscar Angulo, Juan Carlos López-Marcos
Miguel Ángel López-Marcos

Universidad de Valladolid
lmabia@mac.cie.uva.es, oscar@mat.uva.es, lopezmar@mac.uva.es,
malm@mac.uva.es

Lipari, 2009

Outline

Introduction: Probability space

Definition (σ-Algebra)

If Ω is a given set, then a σ-algebra $\mathcal{F}$ on Ω is a family $\mathcal{F}$ of subsets of Ω with the following properties:

1. $\emptyset \in \mathcal{F}$,
2. $F \in \mathcal{F} \Rightarrow F^C \in \mathcal{F}$, where $F^C = \Omega - F$ is the complement of F in Ω,
3. $A_1, A_2, \ldots \in \mathcal{F} \Rightarrow A := \bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$.

Definition (Measurable Space)

The pair $(\Omega, \mathcal{F})$, where Ω is a given set and $\mathcal{F}$ a σ-algebra on Ω is called a measurable set.

Introduction: Probability space

Definition (Probability measure)

A probability measure P on a measurable space $(\Omega, \mathcal{F})$ is a function $P : \mathcal{F} \longrightarrow [0,1]$ such that

1. $P(\emptyset) = 0,\ P(\Omega) = 1,$
2. If $A_1, A_2, \ldots \in \mathcal{F}$ and $\{A_i\}_{i=1}^{\infty}$ is disjoint (i.e. $A_i \cap A_j = \emptyset$ if $i \neq j$) then

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i).$$

Definition (Probability Space)

A probability space is a triple $(\Omega, \mathcal{F}, P)$ consisting of a set Ω (called the sample space), a σ-algebra $\mathcal{F}$ of subsets of Ω (these subsets are called events), and a probability measure P on $(\Omega, \mathcal{F})$.

Introduction: Random variable

Definition (Measurable function)

If $(\Omega, \mathcal{F}, P)$ is a given probability space, a function $Y : \Omega \longrightarrow \mathbb{R}^n$ is called $\mathcal{F}$-measurable if

$$Y^{-1}(B) := \{\omega \in \Omega; Y(\omega) \in B\} \in \mathcal{F}$$

for all Borel sets $B \in \mathbb{R}^n$.

Definition (Random Variable)

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A random variable X is a $\mathcal{F}$-measurable function $X : \Omega \to \mathbb{R}^n$.

Definition (Probability Distribution)

Let $(\Omega, \mathcal{F}, P)$ be a probability space. Given a random variable $X : \Omega \to \mathbb{R}^n$ a probability distribution of X on $\mathbb{R}^n$ is a probability measure $\mu_X : \mathbb{R}^n \to [0, 1]$, defined by

$$\mu_X(B) = P(X^{-1}(B)).$$

Introduction: Expected value

Definition (Expectation or Expected Value)

Let $(\Omega, \mathcal{F}, P)$ be a probability space and let $X : \Omega \to \mathbb{R}^n$ be a random variable defined on it, then the expected value of X, denoted by $\mathbb{E}(X)$, is defined as

$$\mathbb{E}(X) := \int_\Omega X(\omega)\, dP(\omega) = \int_{\mathbb{R}^n} x\, d\mu_X(x).$$

if $\int_\Omega |X(\omega)|\, dP(\omega) < \infty$.

If $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ is Borel measurable and if $\int_\Omega |f(X(\omega))|\, dP(\omega) < \infty$, then

$$\mathbb{E}(f(X)) := \int_\Omega f(X(\omega))\, dP(\omega) = \int_{\mathbb{R}^n} f(x)\, d\mu_X(x).$$

Introduction: Stochastic process

Definition (Stochastic process)

A stochastic process is a parametrized collection of random variables $\{X_t\}_{t \in T}$ defined on a probability space $(\Omega, \mathcal{F}, P)$ and assuming values in $\mathbb{R}^n$.

- For each $t \in T$ we have a random variable $\omega \longrightarrow X_t(\omega)$, $\omega \in \Omega$.
- Given $\omega \in \Omega$ we can consider the function $t \longrightarrow X_t(\omega)$, $t \in T$, which is call a *path* of X_t.

Introduction: Normal distribution

Definition (Normal distribution)

Let $(\Omega, \mathcal{F}, P)$ be a probability space. A random variable $X : \Omega \to \mathbb{R}$, is normal if the distribution of X has a density of the form

$$p_X(x) = \frac{1}{\sqrt{2\,\pi\,\sigma^2}} \exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right),$$

where σ and μ are constants. In other words:

$$P(X \in G) = \int_G p_X(x)\,dx, \quad \text{for all Borel sets } G \subset \mathbb{R}.$$

We denote as $X \sim N(\mu, \sigma^2)$

Then X has the next properties

1. $\mathbb{E}(X) = \mu$;
2. $\mathrm{var}(X) = \mathbb{E}((X-\mu)^2) = \sigma^2$.

The Standard Normal Distribution, has expected value zero and variance one $N(0,1)$.

Introduction: Brownian Motion or Wiener process

Definition (Brownian Motion or Wiener process)

A scalar standard Brownian motion, or standard Wiener process, over $[0, T]$ is a random variable $W(t)$ that depends continuously on $t \in [0, T]$ and satisfies the following three conditions.

1. $W(0) = 0$, (with probability 1).
2. For $0 \leq s < t \leq T$ the random variable given by the increment $W(t) - W(s)$ is normally distributed with mean zero and variance $t - s$; equivalently, $W(t) - W(s) \sim \sqrt{t-s}\, N(0,1)$, where $N(0,1)$ denotes a normally distributed random variable with zero mean and unit variance.
3. For $0 \leq s < t < u < v \leq T$ the increments $W(t) - W(s)$ and $W(v) - W(u)$ are independent.

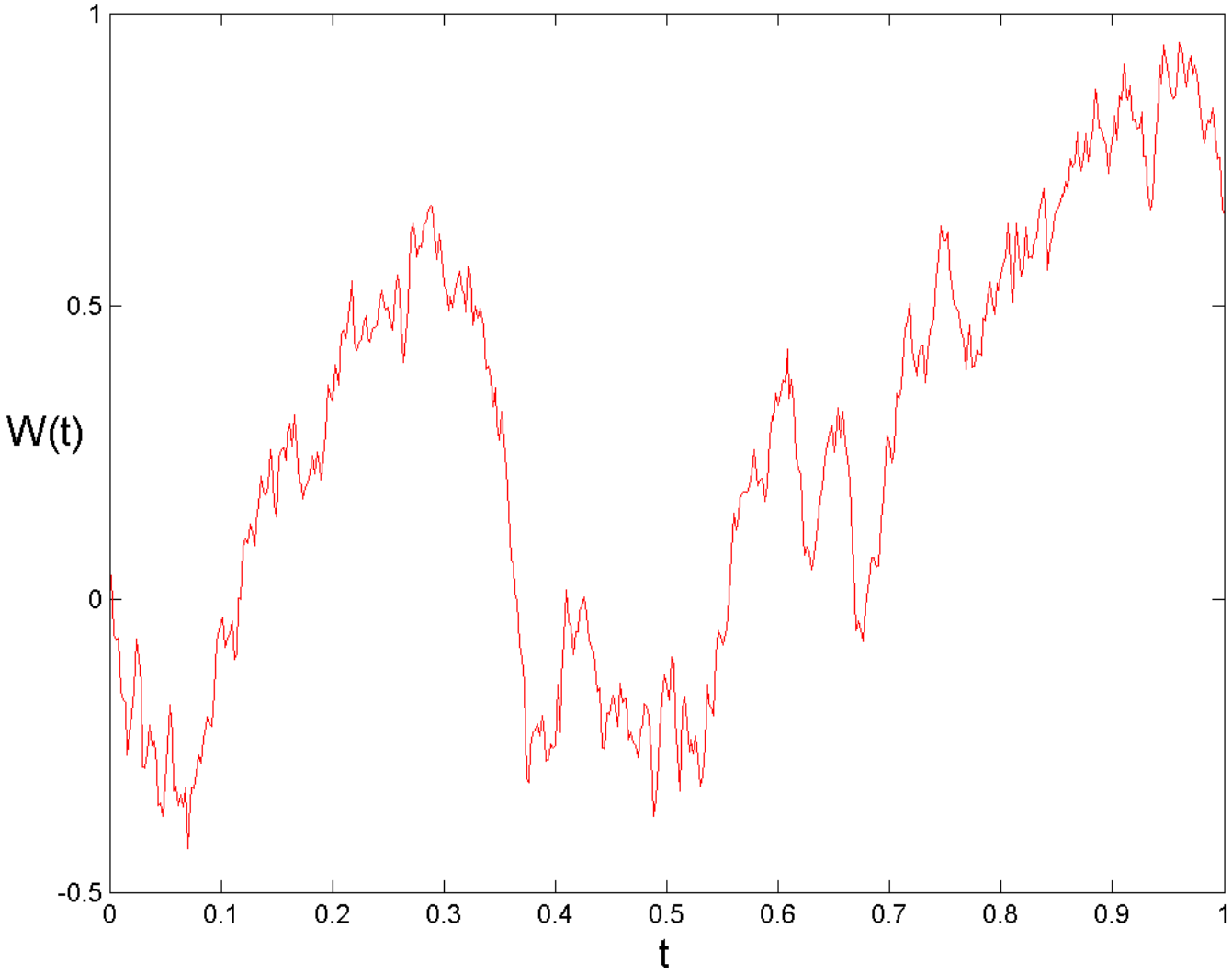
Discretized Brownian motion: $W_j = W(t_j)$, $t_j = j\,\delta t$, $\delta\, t = T/N$, $N \in \mathbb{N}$.

Condition 1: $W_0 = 0$ with probability 1, and conditions 2 and 3 means

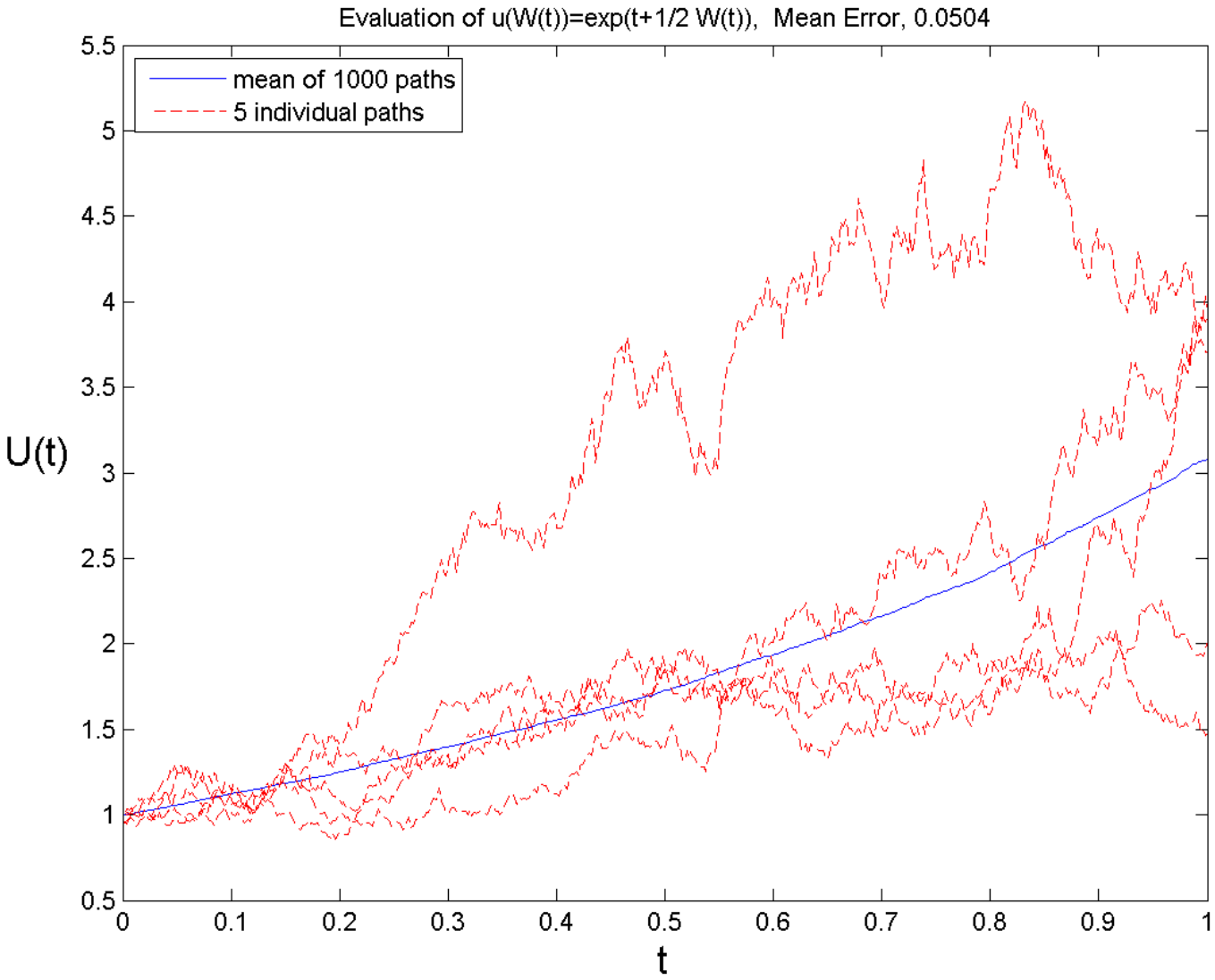
$$W_j = W_{j-1} + dWj, j = 1, 2, \ldots, N, \tag{1}$$

where each dW_j is an independent random variable $\sqrt{\delta\, t}\, N(0,1)$.

Discretized Brownian path



Average example



Introduction: Stochastic Integrals

Itô Integral

$$\int_0^T h(t)\,dt \approx \sum_{j=0}^{N-1} h(t_j)\,(t_{j+1} - t_j),\, t_j = j\,\delta\,t$$

$$\sum_0^{N-1} h(t_j)\,(W(t_{j+1}) - W(t_j)) \approx \int_0^T h(t)\,dW(t).$$

Stratonovich Integral

$$\sum_{j=0}^{N-1} h\left(\frac{t_j + t_{j+1}}{2}\right)(t_{j+1} - t_j) \approx \int_0^T h(t)\,dt$$

$$\sum_{j=0}^{N-1} h\left(\frac{t_j + t_{j+1}}{2}\right)(W(t_{j+1}) - W(t_j)) \approx \int_0^T h(t)\,dW(t).$$

The Euler-Maruyama Method

A scalar, autonomous SDE

$$X(t) = X_0 + \int_0^t f(X(s))\,ds + \int_0^t g(X(s))\,dW(s), 0 \leq t \leq T.$$

f and g are scalar functions and the initial condition X_0 is a random variable.
The second integral on the right-hand side is to be taken with respect to Brownian motion
The solution $X(t)$ is a random variable for each t.

SDE differential equation form

$$dX(t) = f(X(t))\,dt + g(X(t))\,dW(t), X(0) = X_0, 0 \leq t \leq T.$$

Deterministic SDE: $g \equiv 0$ and X_0 constant

$$dX(t)/dt = f(X(t)), \quad X(0) = X0.$$

The Euler-Maruyama Method

Numerical discretization

- $\Delta t = T/L$, $L \in \mathbb{N}$, $\tau_j = j\,\Delta t$.
- $X_j \approx X(\tau_j)$.

$$X(\tau_j) = X(\tau_{j-1}) + \int_{\tau_{j-1}}^{\tau_j} f(X(s))\,ds + \int_{\tau_{j-1}}^{\tau_j} g(X(s))\,dW(s).$$

The Euler-Maruyama (EM)

$$X_j = X_{j-1} + f(X_{j-1})\,\Delta t + g(X_{j-1})\,(W(\tau_j) - W(\tau_{j-1})), j = 1, 2, \ldots, L.$$

The Euler-Maruyama Method

Linear SDE

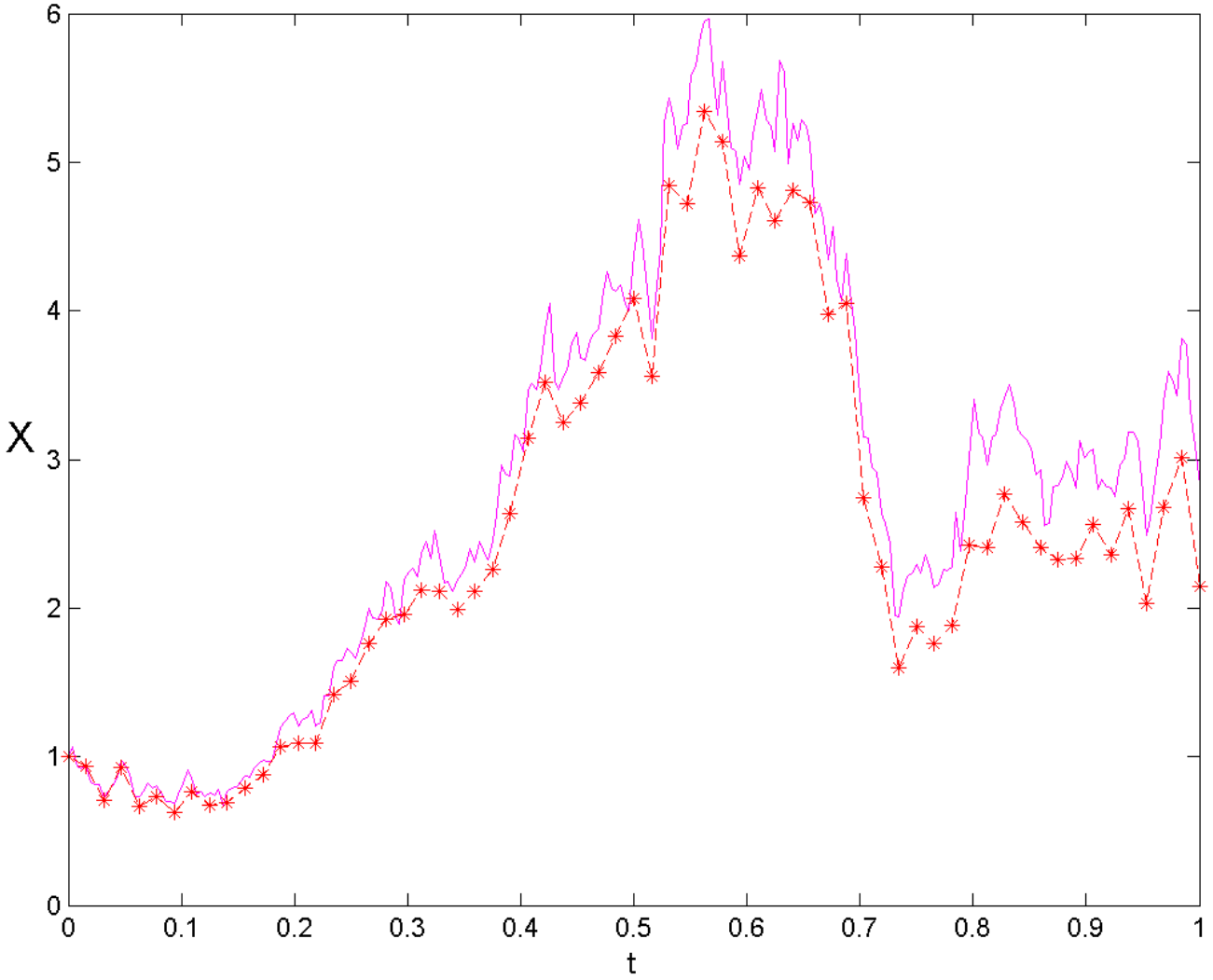
$$dX(t) = \lambda\, X(t)dt + \mu\, X(t)\, dW(t), \quad X(0) = X0,$$

$\lambda, \mu \in \mathbb{R}$.

Exact solution

$$X(t) = X(0)\, \exp\left(\left(\lambda - \frac{1}{2}\,\mu^2\right)\, t + \mu\, W(t)\right).$$

$$dX(t) = 2\,X(t)dt + X(t)\,dW(t), \quad X(0) = 1,$$



$dX(t) = 2\,X(t)dt + X(t)\,dW(t), \quad X(0) = 1,$

Error at $t = T$, $\Delta\, t = R\,\delta\, t$, $\delta\, t = 2^{-8}$

R	$\lvert X(T) - X_T \rvert$
4	0.6907
2	0.1595
1	0.0821

Strong Convergence of the EM Method

Definition (Strong converence)

A method is said to have strong order of convergence equal to γ if there exists a constant C such that

$$\mathbb{E}|X_n - X(\tau)| \leq C\,\Delta\,t^{\gamma}, \tau = n\,\Delta\,t \in [0, T], \tag{2}$$

and $\Delta\,t$ sufficiently small.

If f and g satisfy appropriate conditions, it can be shown that EM has strong order of convergence $\gamma = \frac{1}{2}$.

Strong Convergence of the EM Method

Strong error at the end-point

$$e_{\Delta t}^{\text{strong}} := \mathbb{E}|X_L - X(T)|, \text{where} L\,\Delta t = T,$$

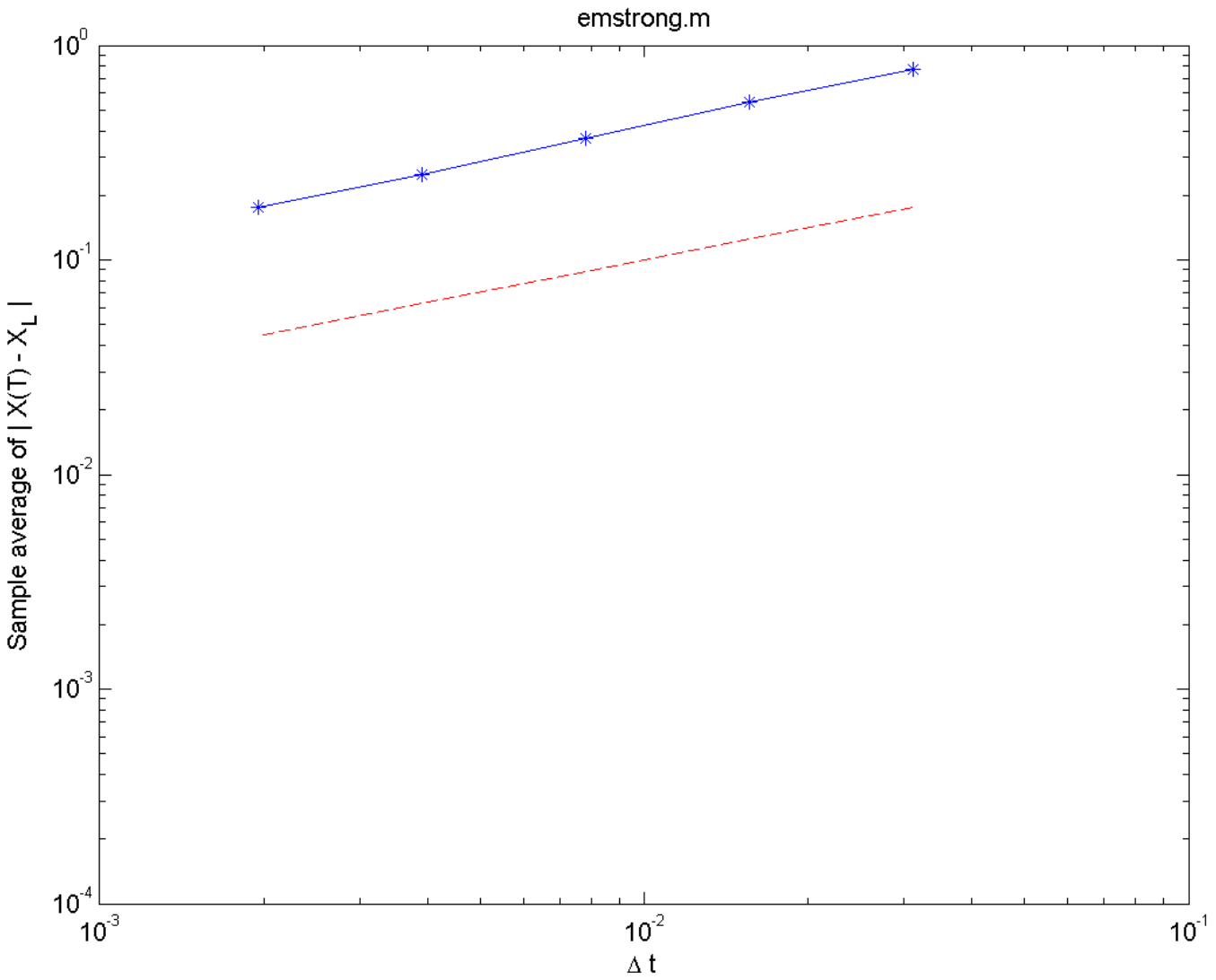
$$e_{\Delta t}^{\text{strong}} \leq C\,\Delta t^{\frac{1}{2}}$$

for sufficiently small Δt.

Test

- 1000 different discretized Brownian paths over $[0,1]$ with $\delta t = 2^{-9}$.
- 5 different stepsizes: $\Delta t = 2^{p-1}\,\delta t$ for $1 \leq p \leq 5$.
- The function mean is then used to average over all sample paths
- The pth element of the average is an approximation to $e_{\Delta t}^{\text{strong}}$, for $\Delta t = 2^{p-1}\,\delta t$.
- $\log e_{\Delta t}^{\text{strong}} \approx \log C + \frac{1}{2}\,\log \Delta t$.

$$dX(t) = 2\,X(t)dt + X(t)\,dW(t), \quad X(0) = 1,$$



Strong Convergence of the EM Method

Test

- 1000 different discretized Brownian paths over $[0, 1]$ with $\delta\, t = 2^{-9}$.
- 5 different stepsizes: $\Delta\, t = 2^{p-1}\, \delta\, t$ for $1 \leq p \leq 5$.
- The function mean is then used to average over all sample paths
- The pth element of the average is an approximation to $e^{\text{strong}}_{\Delta\, t}$, for $\Delta\, t = 2^{p-1}\, \delta\, t$.
- $\log e^{\text{strong}}_{\Delta\, t} \approx \log C + \frac{1}{2} \log \Delta\, t$.
- $e^{\text{strong}}_{\Delta\, t} = C\, \Delta\, t^q$ then
- $\log e^{\text{strong}}_{\Delta\, t} \approx \log C + q \log \Delta\, t$.
- $q = 0.5384$ with least squares residual of 0.0266.

Weak Convergence of the EM Method

Definition (Weak convergence)

A method is said to have weak order of convergence equal to γ if there exists a constant C such that for all functions p in some class

$$|\mathbb{E}p(X_n) - \mathbb{E}p(X(T))| \leq C\,\Delta\,t^{\gamma}, \quad \tau = n\,\Delta\,t \in [0, T]$$

and $\Delta\,t$ sufficiently small.

If f and g satisfy appropriate conditions, it can be shown that EM has weak order of convergence $\gamma = 1$.

Weak Convergence of the EM Method

Weak error at the end-point

$$e_{\Delta t}^{\text{weak}} := |\mathbb{E}X_L - \mathbb{E}X(T)|, \text{where} L\,\Delta t = T,$$

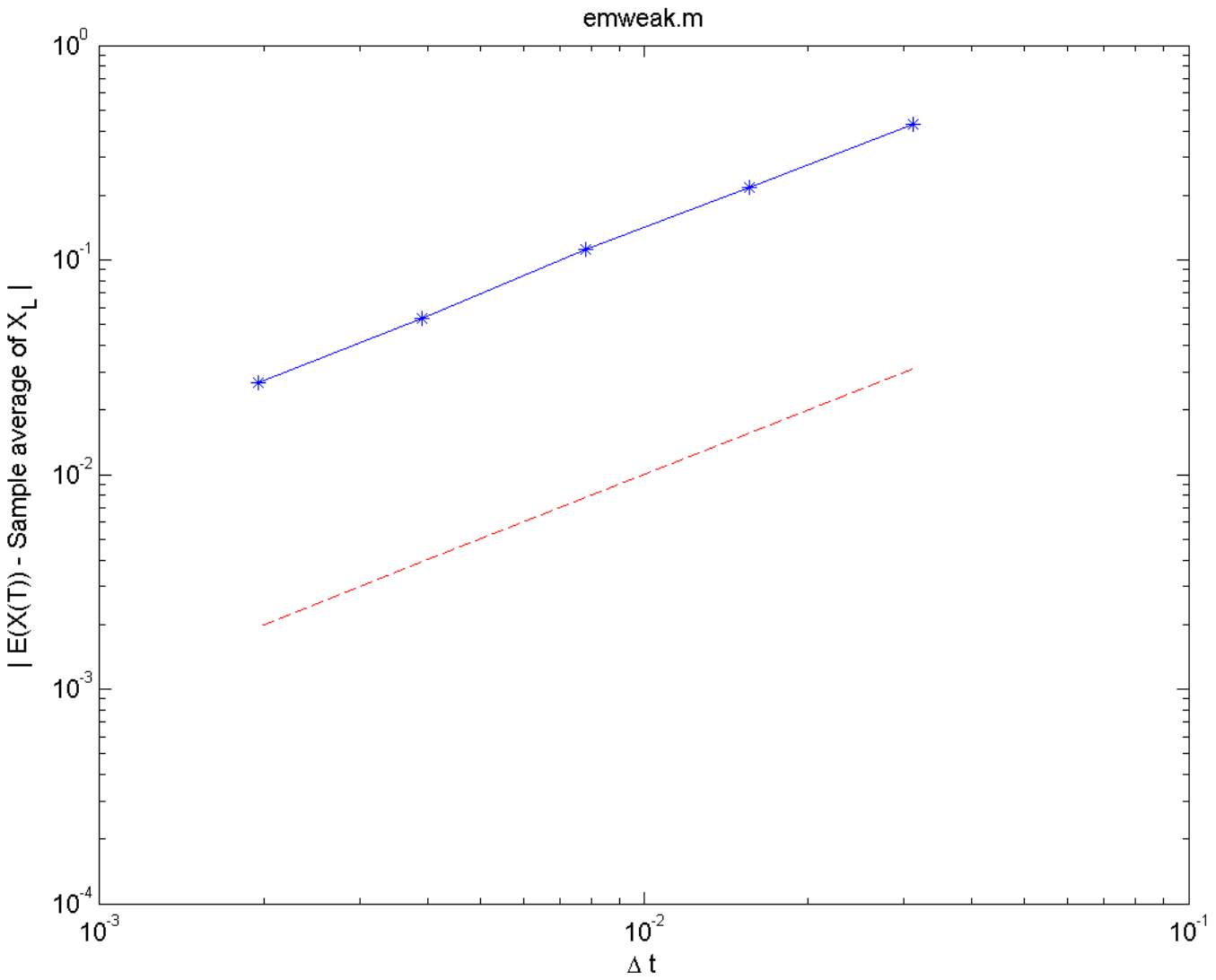
$$e_{\Delta t}^{\text{weak}} \leq C\,\Delta t$$

for sufficiently small Δt.

Test

- 50000 different discretized Brownian paths over $[0,1]$ $\delta = \Delta t$.
- 5 different stepsizes: $\Delta t = 2^{p-10}$ for $1 \leq p \leq 5$.
- The function mean is then used to average over all sample paths, $\mathbb{E}X(T) = e^{\lambda T}$.

$dX(t) = 2\,X(t)dt + X(t)\,dW(t), \quad X(0) = 1,$



Weak Convergence of the EM Method

Test

- 50000 different discretized Brownian paths over $[0,1]$ $\delta = \Delta\, t$.
- 5 different stepsizes: $\Delta\, t = 2^{p-10}$ for $1 \leq p \leq 5$.
- The function mean is then used to average over all sample paths, $\mathbb{E}X(T) = e^{\lambda\, T}$.
- $e^{\mathrm{weak}}_{\Delta\, t} = C\, \Delta\, t^q$ then
- $\log e^{\mathrm{weak}}_{\Delta\, t} \approx \log C + q\, \log \Delta\, t$.
- $q = 0.9858$ with least squares residual of 0.0508.

Milstein's Higher Order Method

$$X(\tau_j) = X(\tau_{j-1}) + \int_{\tau_{j-1}}^{\tau_j} f(X(s))\, ds + \int_{\tau_{j-1}}^{\tau_j} g(X(s))\, dW(s).$$

$$\begin{aligned} X_j &= X_{j-1} + \Delta\, f(Xj-1) + g(Xj-1)\,(W(\tau_j) - W(\tau_{j-1})) \\ &\quad + \frac{1}{2}\, g(X_{j-1})\, g'(X_{j-1})\,(W(\tau_j) - W(\tau_{j-1}))^2 - \Delta\, t, \quad j = 1, 2, \ldots, L. \end{aligned}$$

Milstein's Higher Order Method

Population dynamics

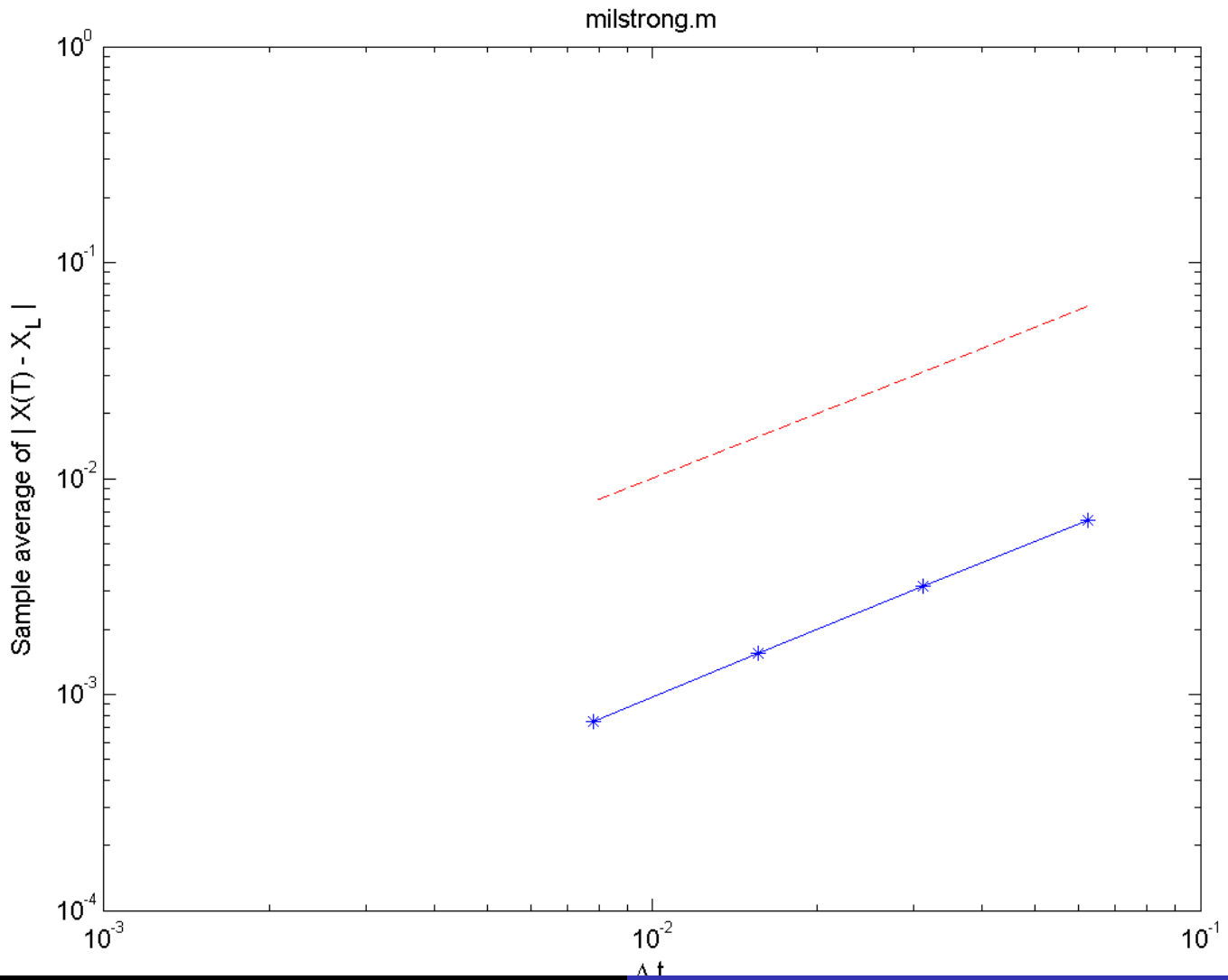
$$dX(t) = r\,X(t)\,(K - X(t))dt + \beta\,X(t)\,dW(t), \quad X(0) = X_0,$$

r, K, and β are constants.

Test

- $r = 2$, $K = 1$, $\beta = 0.25$ and $X_0 = 0.5$
- 500 different discretized Brownian paths over $[0, 1]$ with $\delta\,t = 2^{-11}$.
- Exact solution: $\Delta\,t = \delta\,t$
- 5 different stepsizes: $\Delta\,t = 2^p\,\delta\,t$ for $3 \leq p \leq 7$.
- The function mean is then used to average over all sample paths
- The pth element of the average is an approximation to $e^{\text{strong}}_{\Delta\,t}$, for $\Delta\,t = 2^p\,\delta\,t$.
- $\log e^{\text{strong}}_{\Delta\,t} \approx \log C + \log \Delta\,t$.

$$dX(t) = 2\,X(t)\,(1 - X(t))dt + 0.25\,X(t)\,dW(t), \quad X(0) = 0.5,$$



Milstein's Higher Order Method

Test

- $r = 2$, $K = 1$, $\beta = 0.25$ and $X_0 = 0.5$
- 500 different discretized Brownian paths over $[0, 1]$ with $\delta\, t = 2^{-11}$.
- Exact solution: $\Delta\, t = \delta\, t$
- 5 different stepsizes: $\Delta\, t = 2^p\, \delta\, t$ for $3 \leq p \leq 7$.
- The function mean is then used to average over all sample paths
- The pth element of the average is an approximation to $e^{\text{strong}}_{\Delta\, t}$, for $\Delta\, t = 2^p\, \delta\, t$.
- $\log e^{\text{strong}}_{\Delta\, t} \approx \log C + \log \Delta\, t$.
- $e^{\text{strong}}_{\Delta\, t} = C\, \Delta\, t^q$ then
- $\log e^{\text{strong}}_{\Delta\, t} \approx \log C + q\, \log \Delta\, t$.
- $q = 1.0184$ and $resid = 0.0350$.

Linear Stability

Ordinary differential equations

- $dX/dt = \lambda X$, where $\lambda \in \mathbb{C}$, is constant parameter.
- $X_0 \exp(\lambda t)$
- $\lim_{t \to \infty} X(t) = 0$ for any X_0,
- $\Re\{\lambda\} < 0$

Linear Stability

Stochastic differential equations

- $dX(t) = \lambda\, X(t)\, dt + \mu\, X(t)\, dW(t)$, $X(0) = X_0$, $\lambda, \mu \in \mathbb{C}$
- Mean-square stability

$$\lim_{t\to\infty} \mathbb{E}X(t)^2 = 0 \iff \Re\{\lambda\} + \frac{1}{2}\,|\mu|^2 < 0,$$

- Asymptotic stability

$$\lim_{t\to\infty} |X(t)| = 0, \text{with probability} 1 \iff \Re\{\lambda + \frac{1}{2}\,|\mu|^2\} < 0.$$

- Si $\mu = 0$, equivalent to ODE theory.

Stability of Euler-Maruyama method

Suppose that the parameters λ and μ are chosen so that the SDE is stable in the mean-square or asymptotic sense

For what range of $\Delta\, t$ is the EM solution stable in an analogous sense?

Stability of Euler-Maruyama method

- Mean-square stability

$$\lim_{j\to\infty} \mathbb{E}X_j^2 = 0 \iff |1+\Delta\, t\,\lambda|^2 + \Delta\,|\mu|^2 < 1,$$

- Asymptotic stability

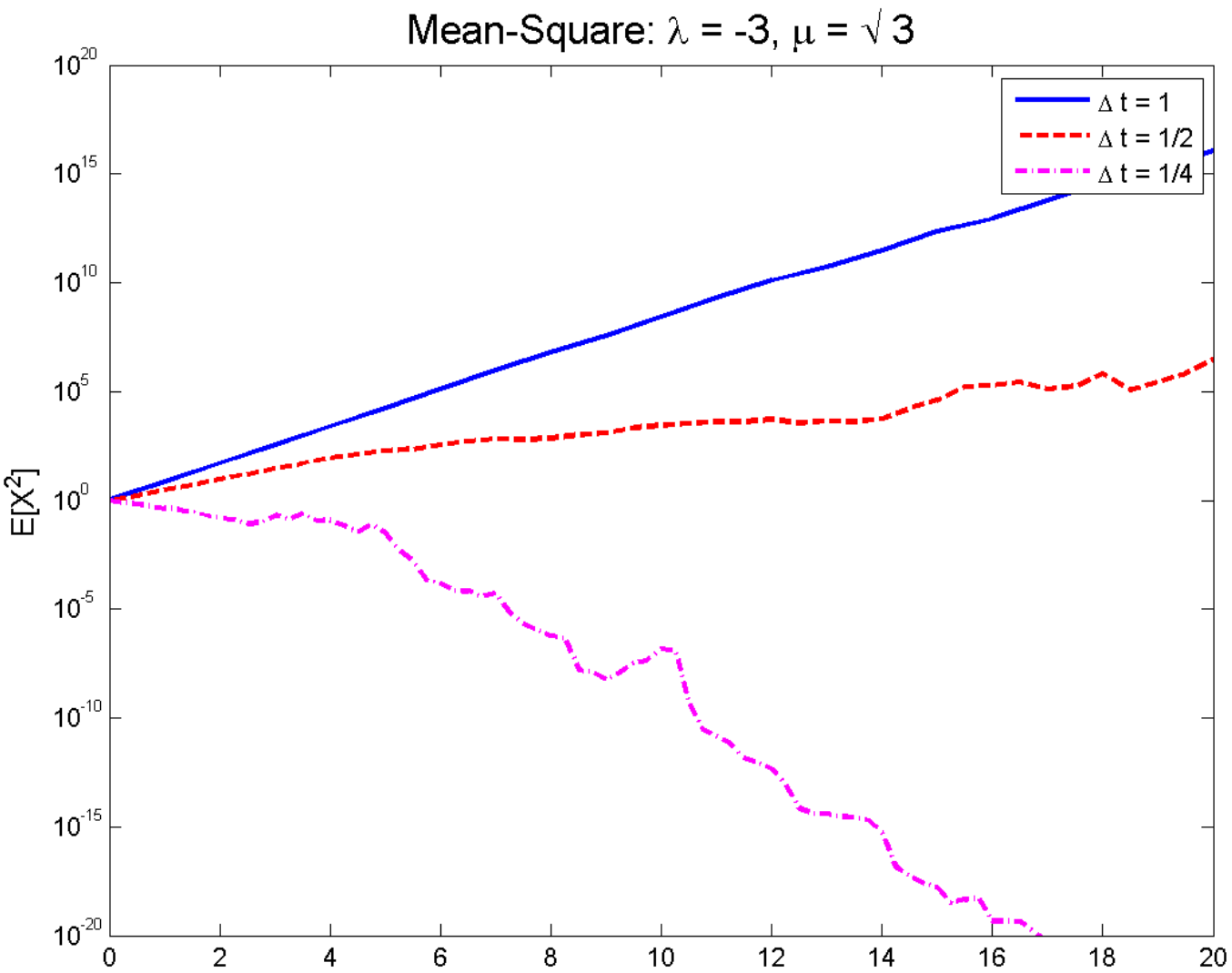
$$\lim_{j\to\infty} |X_j| = 0, \text{with probability} 1 \iff \mathbb{E}\log\left|1+\Delta\, t\,\lambda + \sqrt{\Delta\, t}\,\mu\, N(0,1)\right| < 0.$$

Stability of Euler-Maruyama method

Test: Mean-Square Stability

- $X_0 = 1$ constant, $[0, 20]$
- $\lambda = -3$ and $\mu = \sqrt{3}$ Mean-square stable
- 50000 different discretized Brownian paths. $\Delta\, t = 1, 1/2, 1/4$.

Mean-Square stability of Euler-Maruyama method

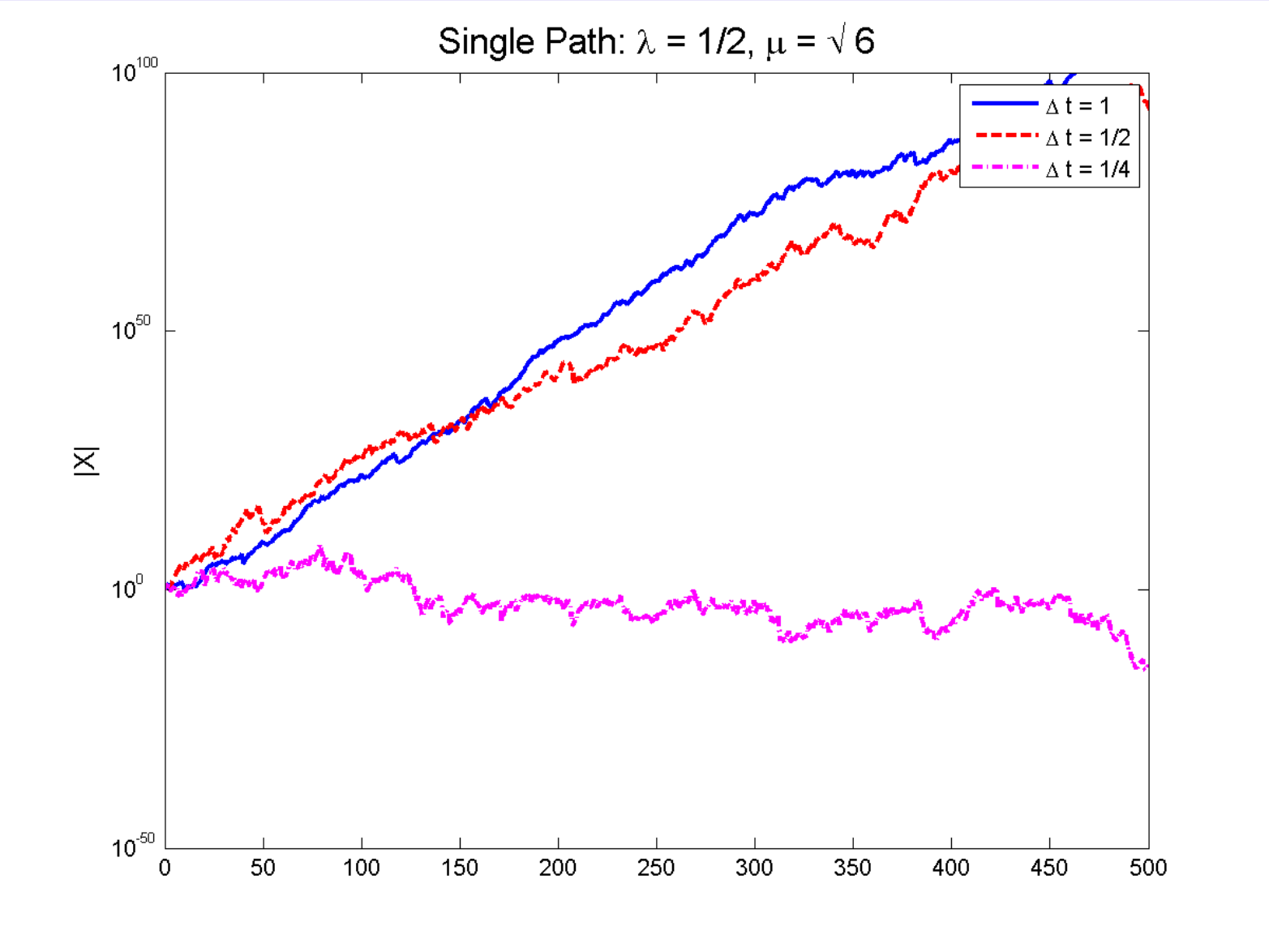
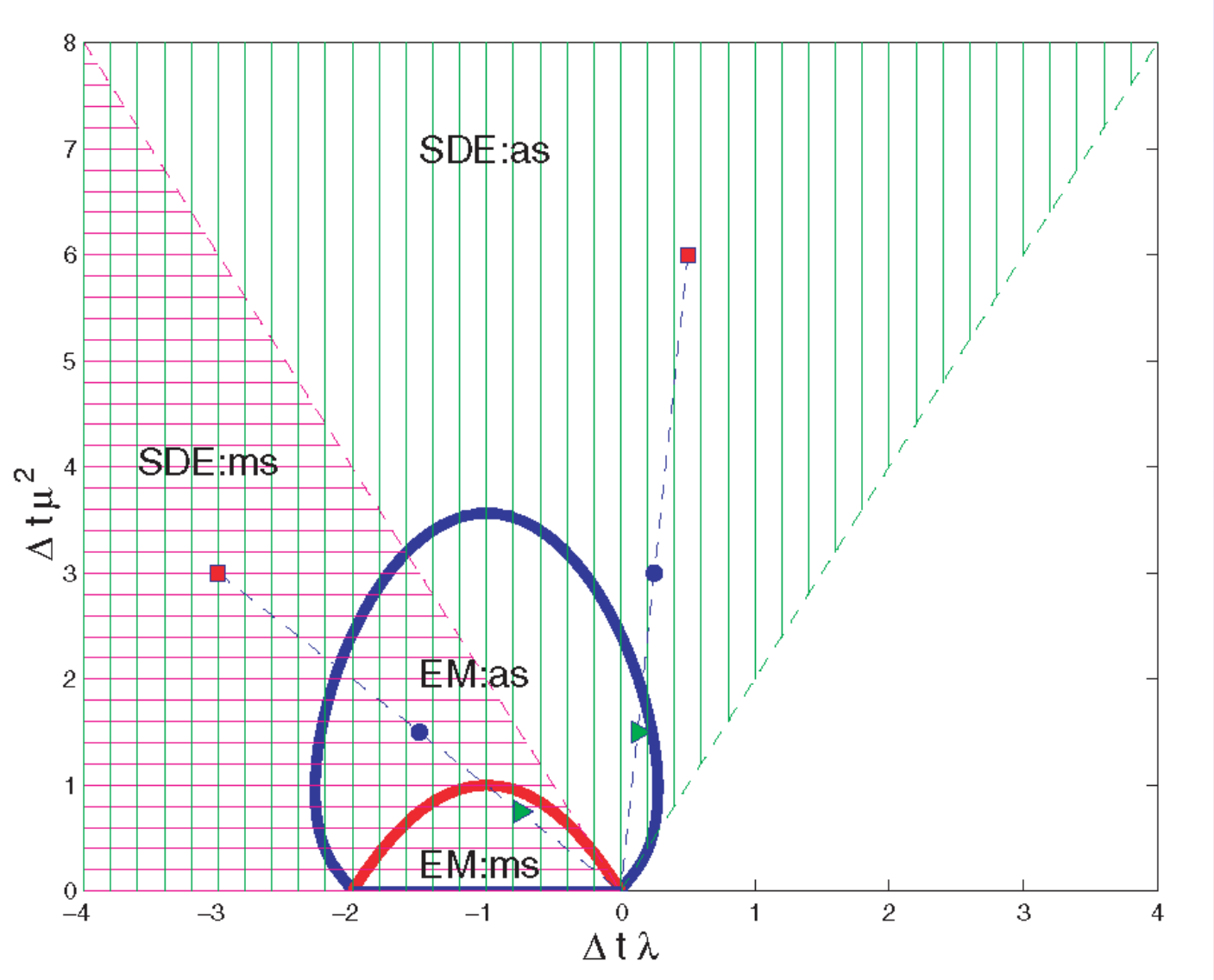


Stability of Euler-Maruyama method

Test: Asymptotic Stability

- $X_0 = 1$ constant, $[0, 500]$
- $\lambda = 1/2$ and $\mu = \sqrt{6}$ Asymptotically stable, but not mean-square stable.
- 1 Brownian path. $\Delta\, t = 1, 1/2, 1/4$.

Asymptotic stability of Euler-Maruyama method



Stability region of Euler-Maruyama method