Introduction to stochastic numerics and computing ergodic limits

M.V. Tretyakov
School of Mathematical Sciences, University of Nottingham, UK

Woudschoten Conference, 5th October 2016

Plan of the talk

- What's for?
- Stochastics
  - SDE
  - Link to PDEs
- Numerics
  - Types of convergence
  - Mean-square approximations
  - Weak approximations
- Computing ergodic limits
  - Ensemble and time averaging and computational errors
  - Gradient system and 'non-Markovian' scheme
- Conclusions

What's for?

- Molecular dynamics (MD) - canonical ensemble (NVT) – see Part II on Friday
- Financial Engineering (pricing and hedging derivatives; management of couterparty risk, etc.)
- Bayesian statistics
- SDEs $\Longleftrightarrow$ linear parabolic and elliptic PDEs
- Other (models in biology, chemistry, economics, etc.)

Stochastics: SDEs

Assume that deterministic functions $b(t,x)$ and $b_r(t,x)$, $r = 1,\ldots,q$, have some good analytical properties. The solution of the (Ito) **stochastic differential equation**

$$dX = b(t,X)dt + \sum_{r=1}^{q} \sigma_r(t,X)dw_r(t), \quad X(t_0) = x, \tag{1}$$

is a stochastic process $X(t)$ such that

$$X(t) = x + \int_{t_0}^{t} b(s,X(s))ds + \sum_{r=1}^{q} \int_{t_0}^{t} \sigma_r(s,X(s))dw_r(s) \tag{2}$$

Stochastics: SDEs

Assume that deterministic functions $b(t,x)$ and $b_r(t,x)$, $r = 1,\ldots,q$, have some good analytical properties. The solution of the (Ito) **stochastic differential equation**

$$dX = b(t,X)dt + \sum_{r=1}^{q} \sigma_r(t,X)dw_r(t), \quad X(t_0) = x, \tag{1}$$

is a stochastic process $X(t)$ such that

$$X(t) = x + \int_{t_0}^{t} b(s,X(s))ds + \sum_{r=1}^{q} \int_{t_0}^{t} \sigma_r(s,X(s))dw_r(s) \tag{2}$$

Ito, Stratonovich?

Stochastics: SDEs

Assume that deterministic functions $b(t,x)$ and $b_r(t,x)$, $r=1,\ldots,q$, have some good analytical properties. The solution of the (Ito) **stochastic differential equation**

$$dX = b(t,X)dt + \sum_{r=1}^{q} \sigma_r(t,X)dw_r(t), \quad X(t_0) = x, \tag{1}$$

is a stochastic process $X(t)$ such that

$$X(t) = x + \int_{t_0}^{t} b(s,X(s))ds + \sum_{r=1}^{q} \int_{t_0}^{t} \sigma_r(s,X(s))dw_r(s) \tag{2}$$

Ito, Stratonovich?

$$\int_0^T \sigma_r(t,X(t))dw_r(t) = \underset{\delta\to 0}{\text{l.i.m.}} \sum_{i=0}^{N-1} \sigma_r(t_i,X(t_i)) \times (w_r(t_{i+1}) - w_r(t_i)),$$

$$\int_0^T \sigma_r(t,X(t))\circ dw_r(t) = \underset{\delta\to 0}{\text{l.i.m.}} \sum_{i=0}^{N-1} \sigma_r\left(t_i, \frac{X(t_i)+X(t_{i+1})}{2}\right)(w_r(t_{i+1})-w_r(t_i))$$

Feynman-Kac formula

$$Lu(t,x) \quad := \quad = \frac{\partial}{\partial t}u(t,x) + \frac{1}{2}\sum_{r=1}^{q}\sum_{i,j=1}^{d}\sigma_r^i\sigma_r^j\frac{\partial^2}{\partial x^i \partial x^j}u(t,x) \qquad (3)$$

$$+\sum_{i=1}^{d} b^i \frac{\partial}{\partial x^i}u(t,x)$$

The Cauchy problem for linear parabolic PDE:

$$Lu = 0, \quad t < T,\ x \in R^d, \qquad (4)$$

$$u(T,x) = f(x),\ x \in R^d. \qquad (5)$$

Feynman-Kac formula

$$Lu(t,x) \quad := \quad = \frac{\partial}{\partial t}u(t,x) + \frac{1}{2}\sum_{r=1}^{q}\sum_{i,j=1}^{d}\sigma_r^i\sigma_r^j\frac{\partial^2}{\partial x^i \partial x^j}u(t,x) \tag{3}$$

$$+\sum_{i=1}^{d} b^i \frac{\partial}{\partial x^i}u(t,x)$$

The Cauchy problem for linear parabolic PDE:

$$Lu = 0, \quad t < T,\ x \in R^d, \tag{4}$$

$$u(T,x) = f(x),\ x \in R^d. \tag{5}$$

Then

$$u(t_0,x) = Ef(X_{t_0,x}(T)), \tag{6}$$

where $X_{t_0,x}(t)$, $t \geq t_0$, is the solution of the Ito SDEs

$$dX = b(t,X)dt + \sum_{r=1}^{q}\sigma_r(t,X)dw_r(t), \quad X(t_0) = x. \tag{7}$$

Numerics

$$dX = b(t, X)dt + \sum_{r=1}^{q} \sigma_r(t, X)dw_r(t), \ X(0) = x, \tag{8}$$

Consider a numerical method for (8) based on the one-step approximation:

$$X_{t,x}(t+h) \simeq \bar{X}_{t,x}(t+h) = x + A(t, x, h; \xi), \ 0 \le t < t+h \le T, \tag{9}$$

where ξ is a random vector with moments of a sufficiently high order and A is a d-dimensional vector function. Introduce (for simplicity) the equidistant partition of the time interval $[0, T]$ into N parts with the step $h = T/N$: $0 = t_0 < t_1 < \cdots < t_N = T, \ \ t_{k+1} - t_k = h.$

$$dX = b(t, X)dt + \sum_{r=1}^{q} \sigma_r(t, X)dw_r(t), \ X(0) = x, \tag{8}$$

Consider a numerical method for (8) based on the one-step approximation:

$$X_{t,x}(t+h) \simeq \bar{X}_{t,x}(t+h) = x + A(t, x, h; \xi), \ 0 \le t < t+h \le T, \tag{9}$$

where ξ is a random vector with moments of a sufficiently high order and A is a d-dimensional vector function. Introduce (for simplicity) the equidistant partition of the time interval $[0, T]$ into N parts with the step $h = T/N$: $0 = t_0 < t_1 < \cdots < t_N = T$, $t_{k+1} - t_k = h$.
According to (9), we construct the sequence

$$X_0 = x, \ X_{k+1} = X_k + A(t_k, X_k, h; \xi_{k+1}), \ k = 0, \ldots, N-1, \tag{10}$$

where ξ_1 is independent of X_0 and ξ_{k+1} for $k > 0$ is independent of $X_0, \ldots, X_k$, $\xi_1, \ldots, \xi_k$.

Numerics - Example

Euler's method:

$$X_{k+1} = X_k + b_k h + \sum_{r=1}^{q} \sigma_{rk} \Delta_k w_r(h), \tag{11}$$

where $\Delta_k w_r(h) = w_r(t_{k+1}) - w_r(t_k)$, and the index k at σ_r and b indicates that these functions are evaluated at the point (t_k, X_k).

Numerics - Example

Euler's method:

$$X_{k+1} = X_k + b_k h + \sum_{r=1}^{q} \sigma_{rk} \Delta_k w_r(h), \tag{11}$$

where $\Delta_k w_r(h) = w_r(t_{k+1}) - w_r(t_k)$, and the index k at σ_r and b indicates that these functions are evaluated at the point (t_k, X_k).
G. Marujama (1955) showed the mean-square convergence of this method, while I.I. Gichman and A.V. Skorokhod (1968) proved that the order of accuracy of Euler's method is $1/2$, i.e.,

$$(E(X(t_k) - X_k)^2)^{1/2} \leq C h^{1/2}, \tag{12}$$

where C is a constant not depending on k and h.

Numerics - mean-square methods

We usually distinguish two types of convergence of numerical methods for SDEs: mean-square (also called strong) and weak. Mean-square methods are used for direct simulation of SDEs' trajectories which, e.g., can give information on general behavior of a stochastic model.

Numerics - mean-square methods

We usually distinguish two types of convergence of numerical methods for SDEs: mean-square (also called strong) and weak. Mean-square methods are used for direct simulation of SDEs' trajectories which, e.g., can give information on general behavior of a stochastic model.

Definition

If for some method we would have

$$(E(X(t_k) - X_k)^2)^{1/2} \leq Ch^p \,, \tag{13}$$

then we say that **the mean-square order of accuracy** *of the method is* p.

[e.g. Milstein, T.; Springer, 2004]

Numerics - weak convergence

Weak methods are sufficient for evaluation of mean values and are simpler than mean-square ones.

Numerics - weak convergence

Weak methods are sufficient for evaluation of mean values and are simpler than mean-square ones.

Definition

If an approximation $\bar{X}$ *is such that*

$$|Ef(\bar{X}(T)) - Ef(X(T))| \leq Kh^p \tag{14}$$

for f *from a sufficiently large class of functions, then we say that the* **weak order of accuracy** *of the approximation* $\bar{X}$ *(the method* $\bar{X}$*) is* p*.*

Numerics - weak convergence

Weak methods are sufficient for evaluation of mean values and are simpler than mean-square ones.

Definition

If an approximation $\bar{X}$ *is such that*

$$|Ef(\bar{X}(T)) - Ef(X(T))| \leq Kh^p \tag{14}$$

for f *from a sufficiently large class of functions, then we say that the* **weak order of accuracy** *of the approximation* $\bar{X}$ (*the method* $\bar{X}$) *is* p.

If a method converges with an order p in the mean-square sense, it also converges in the weak sense with order equal to or larger than p. The opposite is not true.

Numerics - weak convergence

Weak methods are sufficient for evaluation of mean values and are simpler than mean-square ones.

Definition

If an approximation $\bar{X}$ *is such that*

$$|Ef(\bar{X}(T)) - Ef(X(T))| \leq Kh^p \tag{14}$$

for f *from a sufficiently large class of functions, then we say that the* **weak order of accuracy** *of the approximation* $\bar{X}$ (*the method* $\bar{X}$) *is* p.

If a method converges with an order p in the mean-square sense, it also converges in the weak sense with order equal to or larger than p. The opposite is not true.
The weak order of accuracy of Euler's method (11) is 1.

[e.g. Milstein, T.; Springer, 2004]

Numerics - weak convergence

A crude method like

$$X_{k+1} = X_k + b_k h + \sqrt{h} \sum_{r=1}^{q} \sigma_{rk} \eta_{rk} \,, \tag{15}$$

where $\eta_{rk},\ r = 1, \ldots, q,\ k = 0, \ldots, N-1$, are independent random variables taking the values $+1$ and -1 with probabilities $1/2$, also has first order of accuracy in the sense of weak approximation. We usually call (15) as the *weak Euler scheme* (Milstein (1978); see also Talay (1984), Milstein (1985) and Milstein&T, Springer, 2004).

Numerics - weak convergence

A crude method like

$$X_{k+1} = X_k + b_k h + \sqrt{h} \sum_{r=1}^{q} \sigma_{rk} \eta_{rk} \,, \tag{15}$$

where $\eta_{rk},\ r = 1, \ldots, q,\ k = 0, \ldots, N-1$, are independent random variables taking the values $+1$ and -1 with probabilities $1/2$, also has first order of accuracy in the sense of weak approximation. We usually call (15) as the *weak Euler scheme* (Milstein (1978); see also Talay (1984), Milstein (1985) and Milstein&T, Springer, 2004).

The main interest in weak approximations lies in the hope to obtain simpler methods and, in particular, methods not requiring modeling of complicated random variables.

Numerics - Monte Carlo

The Monte Carlo technique:

$$u \equiv Ef(X(T)) \simeq \bar{u} \equiv Ef(X_N) \simeq \hat{u} \equiv \frac{1}{M} \sum_{m=1}^{M} f(X_N^{(m)}) \ , \qquad (16)$$

where $X_N^{(m)}$, $m = 1, \ldots, M$, are independent realizations of the random variable X_N.

Numerics - Monte Carlo

The Monte Carlo technique:

$$u \equiv Ef(X(T)) \simeq \bar{u} \equiv Ef(X_N) \simeq \hat{u} \equiv \frac{1}{M} \sum_{m=1}^{M} f(X_N^{(m)}) \ , \qquad (16)$$

where $X_N^{(m)}$, $m = 1, \ldots, M$, are independent realizations of the random variable X_N.
The error of the Monte Carlo method in (16) is evaluated by

$$\bar{\sigma} = c \, \frac{\left[Var \left\{ f(X_N) \right\} \right]^{1/2}}{M^{1/2}}$$

In practice:

$$\begin{aligned} \bar{u} &\in (\hat{u} - \frac{c}{\sqrt{M}} \sqrt{\hat{v}}, \hat{u} + \frac{c}{\sqrt{M}} \sqrt{\hat{v}}) \ , \qquad (17) \\ \hat{v} &\equiv \frac{1}{M} \sum_{m=1}^{M} \left[f({}_m X_N) \right]^2 - \hat{u}^2 \ , \end{aligned}$$

with probability 0.68 for $c = 1$, 0.95 for $c = 2$, and 0.997 for $c = 3$.
[e.g. Milstein, T.; Springer, 2004]

Ergodic limits

For many applications (in particular, molecular dynamics and Bayesian statistics), it is of interest to compute the mean of a given function with respect to the invariant law of the diffusion, i.e. the ergodic limit.

Ergodic limits

For many applications (in particular, molecular dynamics and Bayesian statistics), it is of interest to compute the mean of a given function with respect to the invariant law of the diffusion, i.e. the ergodic limit.

Consider the system of Ito SDEs

$$dX = b(X)dt + \sum_{l=1}^{q} \sigma_r(X)dw_r(t),\ X(0) = x \tag{18}$$

We assume

(A1) The coefficients of (18) are sufficiently smooth functions in $\mathbf{R}^d$.

(A2) The solution of (18) is *regular*, i.e., it is defined for all $t \geq 0$.

(A3) The process $X(t)$ is *ergodic*, i.e., there exists a unique invariant measure μ of X and independently of $x \in \mathbf{R}^d$ there exists the limit

$$\lim_{t\to\infty} E\varphi(X(t;x)) = \int \varphi(x)\, d\mu(x) := \varphi^{erg} \tag{19}$$

for any function $\varphi(x)$ with polynomial growth at infinity.

(A4) The Markov transition function $P(t, x, dy)$ and the invariant measure $\mu(dx)$ have sufficiently smooth densities $p(t, x, y)$ and $\rho(x)$, respectively.

[Hasminskii 1980; 2012]

Ergodic limits

Ergodicity. We are interested here in systems which solutions satisfy a stronger condition than (A3):

(A3e) The process $X(t)$ is *exponentially ergodic*, i.e., for any $x \in \mathbf{R}^d$ and any function φ with a polynomial growth we have the following strengthening of (19):

$$|E\varphi(X(t;x)) - \varphi^{erg}| \le Ce^{-\lambda t}, \quad t \ge 0, \tag{20}$$

where $C > 0$ and $\lambda > 0$ are some constants.

It follows from (20) (and (19)) that for any $\varepsilon > 0$ there exists $T_0 > 0$ such that for all $T \ge T_0$

$$|E\varphi(X(T;x)) - \varphi^{erg}| \le \varepsilon. \tag{21}$$

Thus, the problem of computing the ergodic limit is reduced to evaluating the expectation $E\varphi(X(T))$ at a finite (though usually large) time T.

Example: Langevin equations

Example

The Langevin equations are very popular in, e.g., molecular simulation:

$$\begin{aligned} dP &= f(Q)dt - \gamma P dt + \sqrt{\frac{2\gamma}{\beta}} dw(t), \quad P(0) = P_0 = p, \qquad (22) \\ dQ &= M^{-1} P dt, \quad Q(0) = Q_0 = q, \end{aligned}$$

where P, Q, $f(q) = -\nabla U(q)$ are n-dimensional column-vectors, $\gamma > 0$ is a friction parameter, M is a non-singular, diagonal matrix, and $w(t) = (w_1(t), \ldots, w_n(t))^\top$ and $w_l(t)$ are independent standard Wiener processes.
Let $U(q) \geq 0$ for all $q \in \mathbf{R}^n$ and there exists an $\alpha_1 > 0$ and $0 < \alpha_2 < 1$ such that $\frac{1}{2}(\nabla U(q), q) \geq \alpha_2 U(q) + \gamma^2 \frac{\alpha_2(2-\alpha_2)}{8(1-\alpha_2)}|q|^2 - \alpha_1$. Then the solution of (22) is exponentially ergodic [Mattingly, Stuart, Higham (2002)] and its invaraint measure is Gibbsian with the density

$$\rho(p, q) \propto \exp\left(-\beta\left\{\frac{1}{2} p^\top M^{-1} p + U(q)\right\}\right) \qquad (23)$$

Ensemble averaging

$$|E\varphi(X(T;x)) - \varphi^{erg}| \leq \varepsilon.$$

Ensemble averaging

$$|E\varphi(X(T;x)) - \varphi^{erg}| \leq \varepsilon. \tag{24}$$

Monte Carlo estimate for the ergodic limit φ^{erg}:

$$\hat{\varphi}^{erg} = \frac{1}{M}\sum_{m=1}^{M} \varphi\left(\bar{X}^{(m)}(T;x)\right), \tag{25}$$

where M is the number of independent approximate realizations and $\bar{X}(T;x)$ is a weak approximation of $X(T;x)$ with order $p > 0$.

The total error

$$R_{\hat{\varphi}^{erg}} := \hat{\varphi}^{erg} - \varphi^{erg} \tag{26}$$

consists of three parts: the error ε of the ergodic limit approximation; the error of numerical integration Kh^p; and the Monte Carlo error; i.e.,

$$R_{\hat{\varphi}^{erg}} \sim Kh^p + \varepsilon + O\left(\frac{1}{\sqrt{M}}\right).$$

[Milstein, T.; Physica D (2007)]

Ensemble averaging

More precisely,

$$Bias(\hat{\varphi}^{erg}) = |E\hat{\varphi}^{erg} - \varphi^{erg}| \leq Kh^p + \varepsilon \tag{27}$$

and the estimator's variance is

$$Var(\hat{\varphi}^{erg}) = O(1/M). \tag{28}$$

Time averaging

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \varphi(X(s;x))ds = \varphi^{erg} \quad a.s., \tag{29}$$

where the limit does not depend on x.

Time averaging

$$\lim_{t\to\infty} \frac{1}{t} \int_0^t \varphi(X(s;x))ds = \varphi^{erg} \quad a.s., \tag{29}$$

where the limit does not depend on x.

Then by approximating a single trajectory, one gets for a sufficiently large $\tilde{T}$:

$$\frac{1}{\tilde{T}} \int_0^{\tilde{T}} \varphi(X(s;x))ds \approx \check{\varphi}^{erg} = \check{\varphi}_L^{erg} := \frac{1}{L} \sum_{l=1}^{L} \varphi(\bar{X}(t_l;x)), \tag{30}$$

where $Lh = \tilde{T}$.
[Talay (1990); Talay, Tubaro (1990); Milstein, T (2007); Mattingly, Stuart, T. (2010)]

Time averaging

Under some conditions on the SDE coefficients:

$$\mathrm{Bias}(\check{\varphi}_L^{erg}) \leq K_1 h^p + \frac{K_2}{\tilde{T}} \ ; \tag{31}$$

$$\mathrm{Var}(\check{\varphi}_L^{erg}) \leq \frac{K}{\tilde{T}}; \tag{32}$$

there exists a deterministic constant K so that for h sufficiently small, positive $\varepsilon > 0$, and $\tilde{T}$ sufficiently large one has:

$$|\check{\varphi}_L^{erg} - \varphi^{erg}| \leq K h^p + \frac{C(\omega)}{\tilde{T}^{1/2-\varepsilon}} \quad \text{a.s.}, \tag{33}$$

where $C(\omega) > 0$ is an a.s. bounded random variable depending on ε and the particular φ.

There are three types of errors:

1. numerical integration error (estimated by Kh^p);
2. the error due to the distance from the stationary distribution (i.e., the error due to the finite time of integration $\tilde{T}$ estimated by $K/\tilde{T}$);

Time averaging

Under some conditions on the SDE coefficients:

$$\mathrm{Bias}(\check{\varphi}_L^{erg}) \leq K_1 h^p + \frac{K_2}{\tilde{T}} \ ; \tag{31}$$

$$\mathrm{Var}(\check{\varphi}_L^{erg}) \leq \frac{K}{\tilde{T}}; \tag{32}$$

there exists a deterministic constant K so that for h sufficiently small, positive $\varepsilon > 0$, and $\tilde{T}$ sufficiently large one has:

$$|\check{\varphi}_L^{erg} - \varphi^{erg}| \leq K h^p + \frac{C(\omega)}{\tilde{T}^{1/2-\varepsilon}} \quad \text{a.s.}, \tag{33}$$

where $C(\omega) > 0$ is an a.s. bounded random variable depending on ε and the particular φ.

There are three types of errors:

1. numerical integration error (estimated by Kh^p);
2. the error due to the distance from the stationary distribution (i.e., the error due to the finite time of integration $\tilde{T}$ estimated by $K/\tilde{T}$);
3. the statistical error ($\sim 1/\sqrt{\tilde{T}}$).

[Talay (1990); Mattingly, Stuart, T. (2010)]

Time averaging: stats error

Run a long trajectory $M\tilde{T}$ split into M blocks of a large length $\tilde{T} = hL$ each. We evaluate the estimators ${}_m\check{\varphi}_L^{erg}$, $m = 1, \ldots, M$, for each block. Since $\tilde{T}$ is big and a time decay of correlations is usually fast, ${}_m\check{\varphi}_L^{erg}$ can be considered as almost uncorrelated. We compute the sampled variance

$$\hat{D} = \frac{1}{M-1} \sum_{m=1}^{M} \left({}_m\check{\varphi}_L^{erg}\right)^2 - \left(\frac{1}{M} \sum_{m=1m}^{M} \check{\varphi}_L^{erg}\right)^2.$$

For a sufficiently large $\tilde{T}$ and M, $\mathbf{E}\,\check{\varphi}_L^{erg}$ belongs to the confidence interval

$$\mathbf{E}\,\check{\varphi}_L^{erg} \in \left(\check{\varphi}_{LM}^{erg} - c\frac{\sqrt{\hat{D}}}{\sqrt{M}}, \check{\varphi}_{LM}^{erg} + c\frac{\sqrt{\hat{D}}}{\sqrt{M}}\right) ,$$

with probability, for example 0.95 for $c = 2$ and 0.997 for $c = 3$. Note that $\mathbf{E}\,\check{\varphi}_L^{erg}$ contains the two errors forming the bias.

Stochastic gradient system

[Leimkuhler, Matthews, T.; Proc. R. Soc. A (2014)]

$$\mathrm{dX} = a(\mathrm{X})\mathrm{d}t + \sigma \mathrm{dw}, \ \mathrm{X}(0) = \mathrm{X}_0, \tag{34}$$

$$a(x) := -\nabla U(x), \quad x \in \mathbf{R}^d, \tag{35}$$

$\sigma = \sqrt{2/\beta}$, $d = 3n$, and $\mathrm{w}(t)$ is a standard d-dimensional Wiener process.

Stochastic gradient system

[Leimkuhler, Matthews, T.; Proc. R. Soc. A (2014)]

$$\mathrm{dX} = a(\mathrm{X})\mathrm{d}t + \sigma \mathrm{dw}, \ \mathrm{X}(0) = \mathrm{X}_0, \tag{34}$$

$$a(x) := -\nabla U(x), \quad x \in \mathbf{R}^d, \tag{35}$$

$\sigma = \sqrt{2/\beta}$, $d = 3n$, and $\mathrm{w}(t)$ is a standard d-dimensional Wiener process.
We use the following notation for the solution of (34): $\mathrm{X}(t) = \mathrm{X}_{t_0,x}(t)$ when $\mathrm{X}(t_0) = x$, $t \geq t_0$, and also we will write $\mathrm{X}_x(t)$ when $t_0 = 0$.

Stochastic gradient system

[Leimkuhler, Matthews, T.; Proc. R. Soc. A (2014)]

$$\mathrm{dX} = a(\mathrm{X})\mathrm{d}t + \sigma \mathrm{dw}, \ \mathrm{X}(0) = \mathrm{X}_0, \tag{34}$$

$$a(x) := -\nabla U(x), \quad x \in \mathbf{R}^d, \tag{35}$$

$\sigma = \sqrt{2/\beta}$, $d = 3n$, and $\mathrm{w}(t)$ is a standard d-dimensional Wiener process.
We use the following notation for the solution of (34): $\mathrm{X}(t) = \mathrm{X}_{t_0,x}(t)$ when $\mathrm{X}(t_0) = x$, $t \geq t_0$, and also we will write $\mathrm{X}_x(t)$ when $t_0 = 0$.

Stochastic gradient system

[Leimkuhler, Matthews, T.; Proc. R. Soc. A (2014)]

$$\mathrm{dX} = a(\mathrm{X})\mathrm{d}t + \sigma \mathrm{dw}, \ \mathrm{X}(0) = \mathrm{X}_0, \tag{34}$$

$$a(x) := -\nabla U(x), \quad x \in \mathbf{R}^d, \tag{35}$$

$\sigma = \sqrt{2/\beta}$, $d = 3n$, and $\mathrm{w}(t)$ is a standard d-dimensional Wiener process.
We use the following notation for the solution of (34): $\mathrm{X}(t) = \mathrm{X}_{t_0,x}(t)$ when $\mathrm{X}(t_0) = x$, $t \geq t_0$, and also we will write $\mathrm{X}_x(t)$ when $t_0 = 0$.

The solution $\mathrm{X}(t)$ of (34) is exponentially ergodic and

$$\rho(x) \propto \exp\left(-\frac{2}{\sigma^2} U(x)\right)$$

if there exist $c_0 \in \mathbf{R}$ and $c_1 > 0$ such that

$$(x, a(x)) \leq c_0 - c_1 |x|^2. \tag{36}$$

[Hasminskii (1980), Mattingly, Stuart, Higham (2002)]

Stochastic gradient system

The Euler scheme:

$$\mathrm{X}_{k+1} = \mathrm{X}_k + h a(\mathrm{X}_k) + \sigma\sqrt{h}\xi_{k+1}, \tag{37}$$

where $\xi_k = (\xi_k^1, \ldots, \xi_k^d)^\top$ and ξ_k^i, $i = 1, \ldots, d$, $k = 1, \ldots,$ are i.i.d. random variables with the law $\mathcal{N}(0, 1)$.

Stochastic gradient system

The Euler scheme:

$$\mathrm{X}_{k+1} = \mathrm{X}_k + h a(\mathrm{X}_k) + \sigma\sqrt{h}\xi_{k+1}, \tag{37}$$

where $\xi_k = (\xi_k^1, \ldots, \xi_k^d)^\top$ and ξ_k^i, $i = 1, \ldots, d$, $k = 1, \ldots,$ are i.i.d. random variables with the law $\mathcal{N}(0, 1)$.

Heun's scheme:

$$\begin{aligned} \hat{\mathrm{X}}_{k+1} &= \mathrm{X}_k + h a(\mathrm{X}_k) + \sigma\sqrt{h}\xi_{k+1}, \\ \mathrm{X}_{k+1} &= \mathrm{X}_k + \frac{h}{2}\left[a(\hat{\mathrm{X}}_{k+1}) + a(\mathrm{X}_k)\right] + \sigma\sqrt{h}\xi_{k+1}. \end{aligned} \tag{38}$$

Stochastic gradient system

The Euler scheme:

$$\mathrm{X}_{k+1} = \mathrm{X}_k + h a(\mathrm{X}_k) + \sigma\sqrt{h}\xi_{k+1}, \tag{37}$$

where $\xi_k = (\xi_k^1, \ldots, \xi_k^d)^\top$ and ξ_k^i, $i = 1, \ldots, d$, $k = 1, \ldots,$ are i.i.d. random variables with the law $\mathcal{N}(0, 1)$.

Heun's scheme:

$$\begin{aligned} \hat{\mathrm{X}}_{k+1} &= \mathrm{X}_k + h a(\mathrm{X}_k) + \sigma\sqrt{h}\xi_{k+1}, \\ \mathrm{X}_{k+1} &= \mathrm{X}_k + \frac{h}{2}\left[a(\hat{\mathrm{X}}_{k+1}) + a(\mathrm{X}_k)\right] + \sigma\sqrt{h}\xi_{k+1}. \end{aligned} \tag{38}$$

Weak convergence:

$$|E\varphi(X(T)) - E\varphi(X_N)| \leq Kh^p \tag{39}$$

$N = T/h$; Euler – $p = 1$, Heun – $p = 2$ [e.g. Milstein, T.; Springer 2004]

Stochastic gradient system

The Euler scheme:

$$\mathrm{X}_{k+1} = \mathrm{X}_k + ha(\mathrm{X}_k) + \sigma\sqrt{h}\xi_{k+1}, \tag{37}$$

where $\xi_k = (\xi_k^1, \ldots, \xi_k^d)^\top$ and ξ_k^i, $i = 1, \ldots, d$, $k = 1, \ldots,$ are i.i.d. random variables with the law $\mathcal{N}(0, 1)$.

Heun's scheme:

$$\begin{aligned} \hat{\mathrm{X}}_{k+1} &= \mathrm{X}_k + ha(\mathrm{X}_k) + \sigma\sqrt{h}\xi_{k+1}, \\ \mathrm{X}_{k+1} &= \mathrm{X}_k + \frac{h}{2}\left[a(\hat{\mathrm{X}}_{k+1}) + a(\mathrm{X}_k)\right] + \sigma\sqrt{h}\xi_{k+1}. \end{aligned} \tag{38}$$

Weak convergence:

$$|E\varphi(X(T)) - E\varphi(X_N)| \leq Kh^p \tag{39}$$

$N = T/h$; Euler – $p = 1$, Heun – $p = 2$ [e.g. Milstein, T.; Springer 2004]

Ergodic limits [Talay (1990); Talay, Tubaro (1990); Mattingly, Stuart, Higham (2002); Milstein, T (2007); Mattingly, Stuart, T. (2010)]:

$$|\varphi^{erg} - E\varphi(X_N)| \leq Kh^p + Ce^{-\lambda T} \tag{40}$$

Non-Markovian scheme

$$\mathrm{X}_{k+1} = \mathrm{X}_k + ha(\mathrm{X}_k) + \sigma \frac{\sqrt{h}}{2}(\xi_k + \xi_{k+1}), \tag{41}$$

where $\xi_k = (\xi_k^1, \ldots, \xi_k^i)^\top$ defined on $(\Omega, P, \mathcal{F})$ and ξ_k^i, $i = 1, \ldots, d$, $k = 1, \ldots,$ are i.i.d. random variables with the law $\mathcal{N}(0, 1)$
[Leimkuhler, Matthews (2013)]

Example

Let $a(x) = -\alpha x$ with $\alpha > 0$, then $\mathrm{X}(t)$ from (34) is the Ornstein-Uhlenbeck process, which is Gaussian with $\mathbf{E}\mathrm{X}_x(t) = xe^{-\alpha t}$, $Cov(\mathrm{X}_x(s), \mathrm{X}_x(t)) = \dfrac{\sigma^2}{2\alpha}(e^{-\alpha(t-s)} - e^{-\alpha(t+s)})$ for $s \leq t$ and $Var(\mathrm{X}_x(t)) = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t})$.

Example

Let $a(x) = -\alpha x$ with $\alpha > 0$, then $\mathrm{X}(t)$ from (34) is the Ornstein-Uhlenbeck process, which is Gaussian with $\mathbf{E}\mathrm{X}_x(t) = xe^{-\alpha t}$, $Cov(\mathrm{X}_x(s), \mathrm{X}_x(t)) = \frac{\sigma^2}{2\alpha}(e^{-\alpha(t-s)} - e^{-\alpha(t+s)})$ for $s \le t$ and $Var(\mathrm{X}_x(t)) = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t})$.
For the Euler scheme (37):

$$\begin{aligned} \mathbf{E}\mathrm{X}_N &= x_0(1-\alpha h)^N = x_0 e^{-\alpha T}(1 + \mathcal{O}(h)), \\ Var(\mathrm{X}_N) &= \frac{\sigma^2}{2\alpha}\frac{1-(1-\alpha h)^{2N}}{1+\alpha h} \\ &= \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha T}) - \frac{\sigma^2}{2}h + e^{-2\alpha T}\mathcal{O}(h) + \mathcal{O}(h^2), \ \alpha h < 1, \end{aligned}$$

where $|\mathcal{O}(h^p)| \le Kh$ with $K > 0$ independent of T.

Example

Let $a(x) = -\alpha x$ with $\alpha > 0$, then $\mathrm{X}(t)$ from (34) is the Ornstein-Uhlenbeck process, which is Gaussian with $\mathbf{E}\mathrm{X}_x(t) = xe^{-\alpha t}$, $Cov(\mathrm{X}_x(s), \mathrm{X}_x(t)) = \frac{\sigma^2}{2\alpha}(e^{-\alpha(t-s)} - e^{-\alpha(t+s)})$ for $s \le t$ and $Var(\mathrm{X}_x(t)) = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha t})$.
For the Euler scheme (37):

$$\begin{aligned} \mathbf{E}\mathrm{X}_N &= x_0(1-\alpha h)^N = x_0 e^{-\alpha T}(1 + \mathcal{O}(h)), \\ Var(\mathrm{X}_N) &= \frac{\sigma^2}{2\alpha}\frac{1-(1-\alpha h)^{2N}}{1+\alpha h} \\ &= \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha T}) - \frac{\sigma^2}{2}h + e^{-2\alpha T}\mathcal{O}(h) + \mathcal{O}(h^2), \ \alpha h < 1, \end{aligned}$$

where $|\mathcal{O}(h^p)| \le Kh$ with $K > 0$ independent of T.
For the scheme (41):

$$\begin{aligned} \mathbf{E}\mathrm{X}_N &= x_0(1-\alpha h)^N = x_0 e^{-\alpha T}(1 + \mathcal{O}(h)), \\ Var(\mathrm{X}_N) &= \frac{\sigma^2}{2\alpha}\left[1 - \frac{(1-\alpha h)^{2N}}{1-\alpha h}\right] = \frac{\sigma^2}{2\alpha}(1 - e^{-2\alpha T}) + e^{-2\alpha T}\mathcal{O}(h). \end{aligned}$$

Assumptions

Assumption 1 *There exist* $c_0 \in \mathbf{R}$ *and* $c_1 > 0$ *such that*

$$(x, a(x)) \leq c_0 - c_1|x|^2.$$

Assumption 2 *The potential* $U(x) \in C^7(\mathbf{R}^d)$, *its first-order derivatives grow not faster than a linear function at infinity and higher derivatives are bounded. The function* $\varphi(x) \in C^6(\mathbf{R}^d)$ *and it and its derivatives grow not faster than a polynomial function at infinity.*

The most restrictive condition in Assumption 2 is the requirement for $a(x) = -\nabla U$ to be globally Lipschitz:

$$|a(x)|^2 \leq K(1 + |x|^2), \tag{42}$$

where $K > 0$ is independent of $x \in \mathbf{R}^d$, which can be relaxed via [Milstein, T. (2005) and (2007): the concept of rejecting exploding trajectories]

Introduce the operator L

$$L := \frac{\partial}{\partial t} + \mathcal{L},$$

where $\mathcal{L}$

$$\mathcal{L} := \sum_{i=1}^{d} a^i(x) \frac{\partial}{\partial x^i} + \frac{\sigma^2}{2} \sum_{i=1}^{d} \frac{\partial^2}{(\partial x^i)^2}. \tag{43}$$

The function

$$u(t,x) = \mathbf{E}\varphi(\mathrm{X}_{t,x}(T)) \tag{44}$$

satisfies the Cauchy problem for the backward Kolmogorov equation

$$\begin{aligned} Lu &= 0, \\ u(T,x) &= \varphi(x). \end{aligned} \tag{45}$$

Main theorem

Theorem (1; [Leimkuhler, Matthews, T.; Proc. R. Soc. A (2014)])

Let Assumptions 1-2 hold. Then the scheme (41) is first order weakly convergent and for all sufficiently small $h > 0$ its error has the form

$$\mathbf{E}\varphi(\mathrm{X}_x(T)) - \mathbf{E}\varphi(\mathrm{X}_N) = C_0(T,x)h + C(T,x)h^2, \tag{46}$$

$$C_0(T,x) = \mathbf{E}\int_0^T B_0(t,\mathrm{X}_x(t))\mathrm{d}t, \tag{47}$$

$$B_0(t,x) = \frac{1}{2}\left[\sum_{i,j=1}^{d} a^j(x)\frac{\partial a^i(x)}{\partial x^j}\frac{\partial u(t,x)}{\partial x^i}\right.$$

$$\left. + \frac{\sigma^2}{2}\sum_{i,j=1}^{d}\frac{\partial a^i(x)}{\partial x^j}\frac{\partial^2 u(t,x)}{\partial x^i \partial x^j} + \frac{\sigma^2}{2}\sum_{i,j=1}^{d}\frac{\partial^2 a^i(x)}{(\partial x^j)^2}\frac{\partial u(t,x)}{\partial x^i}\right],$$

$$|C(T,x)| \le K(1+|x|^{\varkappa}e^{-\lambda T}),$$

for some $K > 0$, $\varkappa \in \mathbf{N}$ and $\lambda > 0$ independent of h and T.

Corollary

Theorem 1:

$$\mathbf{E}\varphi(\mathrm{X}_x(T)) - \mathbf{E}\varphi(\mathrm{X}_N) = C_0(T,x)h + C(T,x)h^2,$$

$$C_0(T,x) = \mathbf{E}\int_0^T B_0(t,\mathrm{X}_x(t))\mathrm{d}t.$$

Corollary

Theorem 1:

$$\mathbf{E}\varphi(\mathrm{X}_x(T)) - \mathbf{E}\varphi(\mathrm{X}_N) = C_0(T,x)h + C(T,x)h^2,$$

$$C_0(T,x) = \mathbf{E}\int_0^T B_0(t,\mathrm{X}_x(t))\mathrm{d}t.$$

Theorem (2; [Leimkuhler, Matthews, T.; Proc. R. Soc. A (2014)])

Let Assumptions 1-2 hold. Then the coefficient $C_0(T,x)$ from (47) goes to zero as $T \to \infty$:

$$|C_0(T,x)| \leq K(1+|x|^{\varkappa})e^{-\lambda T} \tag{48}$$

for some constants $K > 0$, $\varkappa \in \mathbf{N}$ and $\lambda > 0$, i.e., over a long integration time the scheme (41) is of order two up to exponentially small correction.

Sketch of the proof

$$
\begin{aligned}
C_0(T,x) &= \int_0^T \mathbf{E} B_0(t, \mathrm{X}_x(t))\mathrm{d}t = \int_0^T \int_{\mathbf{R}^d} B_0(t,y) p(t,x,y)\mathrm{d}y\mathrm{d}t \quad (49)\\
&= \int_0^T \int_{\mathbf{R}^d} B_0(t,y)\rho(y)\mathrm{d}y\mathrm{d}t \\
&\quad + \int_0^T \int_{\mathbf{R}^d} B_0(t,y)[p(t,x,y) - \rho(y)]\mathrm{d}y\mathrm{d}t,
\end{aligned}
$$

where $p(t,x,y)$ is the transition density for (34) and $\rho(y)$ is the invariant density.

Sketch of the proof

$$
\begin{aligned}
C_0(T,x) &= \int_0^T \mathbf{E} B_0(t, \mathrm{X}_x(t))\mathrm{d}t = \int_0^T \int_{\mathbf{R}^d} B_0(t,y)p(t,x,y)\mathrm{d}y\mathrm{d}t \quad (49)\\
&= \int_0^T \int_{\mathbf{R}^d} B_0(t,y)\rho(y)\mathrm{d}y\mathrm{d}t \\
&\quad + \int_0^T \int_{\mathbf{R}^d} B_0(t,y)[p(t,x,y)-\rho(y)]\mathrm{d}y\mathrm{d}t,
\end{aligned}
$$

where $p(t,x,y)$ is the transition density for (34) and $\rho(y)$ is the invariant density.

$$
\int_{\mathbf{R}^d} B_0(t,y)\exp\left(-\frac{2}{\sigma^2}U(y)\right)\mathrm{d}y = 0. \quad (50)
$$

Discussion

1. We emphasize that the fact that the average of $B_0(t, x)$ with respect to the invariant measure is equal to zero is the reason why the scheme (41) is second order accurate in approximating ergodic limits.

Discussion

1. We emphasize that the fact that the average of $B_0(t,x)$ with respect to the invariant measure is equal to zero is the reason why the scheme (41) is second order accurate in approximating ergodic limits.
2. In the case of the Euler scheme (37) we get the same error expansion as (46) for the scheme (41) but with a different $B_0(t,x) = B_0^E(t,x)$ (see [Milstein, T.; Springer 2004]):

$$B_0^E(t,x) = \frac{1}{2}\left[\sum_{i,j=1}^{d} a^j \frac{\partial u}{\partial x^j} a^i \frac{\partial u}{\partial x^i}\right.$$
$$+\frac{\sigma^2}{2}\sum_{i,j}^{d}\frac{\partial^2 a^j}{(\partial x^i)^2}\frac{\partial u}{\partial x^j} + \frac{\sigma^2}{2}\sum_{i,j=1}^{d} a^i \frac{\partial^3 u}{\partial x^i (\partial x^j)^2}$$
$$\left.+\sigma^2\sum_{i,j=1}^{d}\frac{\partial a^j}{\partial x^i}\frac{\partial^2 u}{\partial x^j \partial x^i} + \frac{\sigma^4}{6}\sum_{i,j=1}^{d}\frac{\partial^4 u}{(\partial x^i)^2 (\partial x^j)^2}\right].$$

The average of $B_0^E(t,x)$ with respect to the invariant measure is not equal to zero and, consequently, the Euler scheme (37) approximates ergodic limits with order one – the same order as its weak convergence over a finite time interval.

Discussion

3. Let a one-step weak approximation $\bar{\mathrm{X}}_{t,x}(t+h)$ of the solution $\mathrm{X}_{t,x}(t+h)$ of (34) generate a method of order p. The global error of the method:

$$\begin{aligned} R \quad &: \quad = \mathbf{E}\varphi(\mathrm{X}_x(T)) - \mathbf{E}\varphi(\bar{\mathrm{X}}_x(T)) \\ &= \quad C_0(T,x)h^p + \cdots + C_n(T,x)h^{p+n} + O(h^{p+n+1}), \end{aligned} \tag{51}$$

where $n \in \mathbf{N}$ and the functions $C_0(T,x), \ldots, C_n(T,x)$ are independent of h which can be presented in the form

$$C_i(T,x) = \int_0^T \mathbf{E}B_i(s, \mathrm{X}_x(s))\mathrm{d}s.$$

One can deduce from the proof of Theorem 2 that if the averages of $B_i(s,x)$ $0 \le i \le n$, with respect to the invariant measure are equal to zero then in the limit of $T \to \infty$ the scheme has $p+n$ order of accuracy in h.

Discussion

3. Let a one-step weak approximation $\bar{\mathrm{X}}_{t,x}(t+h)$ of the solution $\mathrm{X}_{t,x}(t+h)$ of (34) generate a method of order p. The global error of the method:

$$\begin{aligned} R &:= \mathbf{E}\varphi(\mathrm{X}_x(T)) - \mathbf{E}\varphi(\bar{\mathrm{X}}_x(T)) \\ &= C_0(T,x)h^p + \cdots + C_n(T,x)h^{p+n} + O(h^{p+n+1}), \end{aligned} \tag{51}$$

where $n \in \mathbf{N}$ and the functions $C_0(T,x), \ldots, C_n(T,x)$ are independent of h which can be presented in the form

$$C_i(T,x) = \int_0^T \mathbf{E}B_i(s, \mathrm{X}_x(s))\mathrm{d}s.$$

One can deduce from the proof of Theorem 2 that if the averages of $B_i(s,x)$ $0 \le i \le n$, with respect to the invariant measure are equal to zero then in the limit of $T \to \infty$ the scheme has $p+n$ order of accuracy in h.

[Abdulle, Vilmart, Zygalakis 2014-15]

Numerical experiments

Anharmonic scalar model: the one-dimensional potential energy $U(x) = \cos(x)$

$$L_2 \text{ error: } \sqrt{\sum_i (\hat{\rho}_i - \rho_i)^2}$$

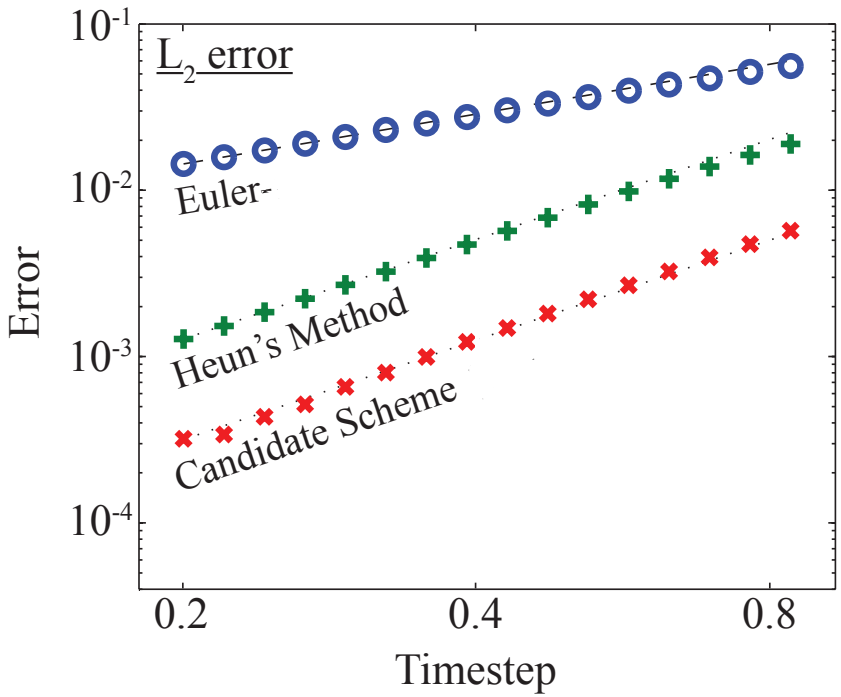

Figure: The error in computed distributions is plotted for each scheme.

[Leimkuhler, Matthews, T.; Proc. R. Soc. A (2014)]

Error in finite time

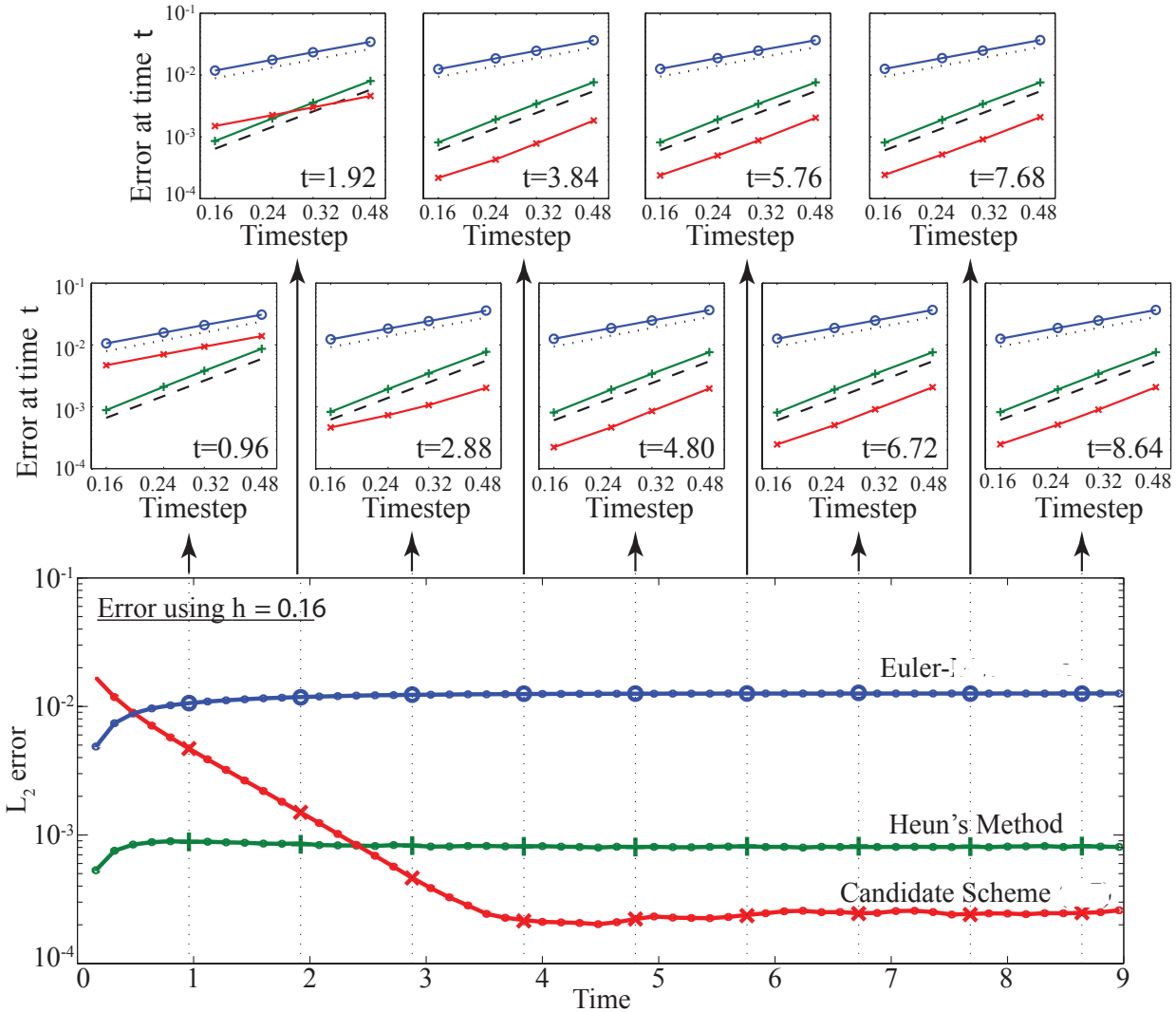

Figure: The lower plot shows the error in the distribution after time t, as computed using each scheme at $h = 0.16$. In the plots at the top, we compare the error growth with respect to stepsize h at multiples of $t = 0.96$.

Conclusions

We discussed

- Some basics of stochastic numerics
- How to compute ergodic limits
- A ‘non-Markovian’ scheme tailored towards computing ergodic limits

Conclusions

We discussed

- Some basics of stochastic numerics
- How to compute ergodic limits
- A 'non-Markovian' scheme tailored towards computing ergodic limits

Further numerical aspects for SDEs:

- Variance reduction/reduction of complexity
- Geometric integration and computing ergodic limits (see Part II on Friday)
- DEs with small noise, DEs with coloured noise
- Convergence/approximation of schemes in the case of nonglobally Lipschitz coefficients of SDEs
- Mean-square and weak approximation of SDEs in bounded domains (absorbtion, reflection)

The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA
EPSRC
Engineering and Physical Sciences Research Council
The Leverhulme Trust