BIT Numer Math (2014) 54:801–821
DOI 10.1007/s10543-014-0482-4

BIT

A weak second-order split-step method for numerical simulations of stochastic differential equations

C. Perret · W. P. Petersen

Received: 17 March 2013 / Accepted: 19 February 2014 / Published online: 29 March 2014

Abstract In an analogy from symmetric ordinary differential equation numerical integrators, we derive a three-stage, weak 2nd-order procedure for Monte-Carlo simulations of Itô stochastic differential equations. Our composite procedure splits each time step into three parts: an $h/2$-stage of trapezoidal rule, an h-stage martingale, followed by another $h/2$-stage of trapezoidal rule. In n time steps, an $h/2$-stage deterministic step follows another $n-1$ times. Each of these adjacent pairs may be combined into a single h-stage, effectively producing a two-stage method with partial overlap between successive time steps.

Keywords Monte-Carlo · Stochastic differential equations · Stability

Mathematics Subject Classification (2010) 34F05, 65C05, 65C30

1 Introduction

It is often the case that a time evolving system of differential equations permits a splitting

$$\dot{x} = f(x) + g(x), \tag{1}$$

where each piece can be chosen for some computational advantage. For example, in Hamiltonian systems where $x = (q, p)$ and $\dot{x} = (\dot{q}, \dot{p}) = (p, g(q))$, the natural splitting contains two pieces: one containing only p variables on the right-hand side

Communicated by Raul Tempone.

C. Perret, W. P. Petersen (✉)
Seminar for Applied Mathematics, ETH Zürich, Zürich, Switzerland
e-mail: wpp@math.ethz.ch

of (1) and the other containing only q [6,17]. Similarly, when using a finite difference scheme for a 2-D partial differential equation, there are significant advantages to splitting the right-hand side of (1) into pieces which contain only differencing in orthogonal directions, say x_1, respectively x_2 [5,15].

In both cases, a composition of approximate flows for the $f(x)$ problem, then an approximation for the $g(x)$ problem yields methods with highly desirable properties. The idea is to write the separate approximate flows for f and g and the initial values (x_0, y_0) at the beginning of the step:

$$\begin{aligned} x(t) &= \phi_t(x_0), \quad \text{where } \dot{x} = f(x), \quad \text{and} \\ y(t) &= \psi_t(y_0), \quad \text{where } \dot{y} = g(y). \end{aligned}$$

One simple composition, called a symmetric method [6] or Yoshida splitting [17] is as follows:

$$x_h = \phi_{h/2} \circ \psi_h \circ \phi_{h/2}(x_0), \tag{2}$$

where x_0 is the value of the process at the beginning of the h-step. A similar analogy can be made in the stochastic case when

$$\begin{aligned} dx &= b(x(t))\,dt + \sigma(x(t))\,d\omega(t), \quad \text{orinintegralform}, \\ x_{t_0+h} &= x_0 + \underbrace{\int_{t_0}^{t_0+h} b(x(t))\,dt}_{\Delta\mathcal{A}} + \underbrace{\int_{t_0}^{t_0+h} \sigma(x(t))\,d\omega(t)}_{\Delta\mathcal{M}}. \end{aligned} \tag{3}$$

The drift term $\Delta\mathcal{A}(h) = \int_{t_0}^{t_0+h} b(x(t))dt$ has bounded variation, while $\Delta\mathcal{M}(h) = \int_{t_0}^{t_0+h} \sigma(x(t))d\omega(t)$ is a martingale [9]. The distinctly different diffusion term $\Delta\mathcal{M}$ characterizes the splitting:

$\boldsymbol{\phi}$: ϕ_h approximates one step of the drift, $\mathcal{A}(h)$, advancing one h-step, while
$\boldsymbol{\psi}$: ψ_h approximates the advancement of x one step of the martingale term $\mathcal{M}(h)$.

The paper is organized as follows. In Sect. 2, we give a simple justification for using trapezoidal rule to improve stability. Numerical results shown in Sect. 6.5 confirm this perspective. Our main result (10) of Sect. 4 is actually given by (2) in detailed terms specific to the semi-martingale problem (3). Section 3 frames the general multi-dimensional problem, and our main result is proven in Sect. 5. Our complex linear test problem with oscillating mean is given in Sect. 6 and its numerical simulation results are shown in Figs. 1 through 13.

2 Some remarks on stability

Our test examples have a complex oscillating mean: $\boldsymbol{E}[X(t)] = (X_0 - i)e^{it} + i$. Textbook treatments on numerical methods for ordinary differential equation stability usually examine the decreasing real process, $\dot{x} = -\lambda x$, with $\lambda > 0$ [4]. There, forward Euler stepping, $x_{nh} = (1 - h\lambda)^n x_0$, does not decrease the absolute size $|x_{nh}|$ as n increases if h is so large that $|1 - h\lambda| > 1$. Conversely, backward Euler stepping,

$x_{nh} = (1 + h\lambda)^{-n} x_0$, always decreases $|x_{nh}|$ for $\lambda > 0$. Canonically, the backward Euler method is thus considered more *stable* for transients with time scale $1/\lambda$.

Here we worry about a complex oscillator, $\dot{x} = i\lambda x$, where again we assume λ is real. In this case, forward Euler stepping, $x_{nh} = (1 + ih\lambda)^n x_0$, systematically increases the amplitude because $|1 + ih\lambda| = (1 + h^2\lambda^2)^{1/2} > 1$. Alternatively, backward Euler stepping, $x_{nh} = (1 - ih\lambda)^{-n} x_0$, systematically decreases the amplitude because $|(1 - ih\lambda)^{-1}| = (1 + h^2\lambda^2)^{-1/2} < 1$. One of our principal points in this paper is that choosing ϕ_h to be trapezoidal rule yields a more *neutral* stability. Namely, even if h is too large, the integrator will not diverge to infinity (forward Euler), nor to zero (backward Euler). Trapezoidal rule for the simple complex linear oscillator, $\dot{x} = i\lambda x$, is

$$x_{nh} = \left(\left(1 - i\frac{h}{2}\lambda \right)^{-1} \left(1 + i\frac{h}{2}\lambda \right) \right)^n x_0,$$

but now its one-step *fundamental matrix* [3] has norm

$$\left| \left(1 - i\frac{h}{2}\lambda \right)^{-1} \left(1 + i\frac{h}{2}\lambda \right) \right| = 1.$$

In the textbook case, $\dot{x} = -\lambda x$, trapezoidal rule, $x_{nh} = \left((1 + \frac{h}{2}\lambda)^{-1} (1 - \frac{h}{2}\lambda) \right)^n x_0$, graciously does not blow up if h is too big. This stability only helps somewhat if the size of the distribution grows, as we will see in Sect. 6.3.

3 The general problem in multi-dimensions

We first express (3) in component form the elements of the real n-dimensional process x indexed by $\alpha = 1, \ldots, n$, drift b is also n-dimensional and real, while σ is an $n \times n$ real array, possibly degenerate:

$$dx^\alpha = b^\alpha(x)dt + \sigma^{\alpha\beta}(x)\, d\omega^\beta, \tag{4}$$

driven by a vector $\omega \in \mathbb{R}^n$ of independent Brownian motions. Our notation in (4) uses the *summation convention*, wherein **repeated indices**, in this case the $\beta = 1, \ldots, n$, **are always summed over**. Additionally, the following shorthand notation for partial derivatives will be found convenient: $\partial_\alpha g \equiv \partial g / \partial x^\alpha$ (for some function $g(x)$).

For a weak 2nd-order approximation, what is necessary is for any C^4 function $f(x)$ of vector $x = (x^1, \ldots, x^n)^\top$, and Δx is the change in x in a time step h, that we have

$$\begin{aligned} \boldsymbol{E}_{x_t}\left[f(x(t+h)) \right] = f(x(t)) + \frac{\partial f(x(t))}{\partial x^\alpha} \boldsymbol{E}_{x_t}\left[\Delta x^\alpha \right] \\ + \frac{1}{2} \frac{\partial^2 f(x(t))}{\partial x^\alpha \partial x^\beta} \boldsymbol{E}_{x_t}\left[\Delta x^\alpha \Delta x^\beta \right] \end{aligned}$$

$$+\frac{1}{3!}\frac{\partial^3 f(x(t))}{\partial x^\alpha \partial x^\beta \partial x^\gamma}\boldsymbol{E}_{x_t}\left[\Delta x^\alpha \Delta x^\beta \Delta x^\gamma\right]$$
$$+\frac{1}{4!}\frac{\partial^4 f(x(t))}{\partial x^\alpha \partial x^\beta \partial x^\gamma \partial x^\delta}\boldsymbol{E}_{x_t}\left[\Delta x^\alpha \Delta x^\beta \Delta x^\gamma \Delta x^\delta\right] + o(h^2), \quad (5)$$

where $\boldsymbol{E}_{x_t}[Y]$ is the conditional expectation of some process Y given $x(t) = x_t$ [13]. Increments of the n independent Brownian motions $\omega^\beta(t)$ are independent Gaussians and satisfy the following relations (for $\alpha, \beta = 1, \ldots, n$). The Kronecker delta, $\delta^{\alpha\beta} = 1$, if $\alpha = \beta$, zero otherwise; while $\delta(t_1 - t_2)$ is Dirac's [10]. We have:

$$\omega^\alpha(0) = 0, \quad (6a)$$

$$\boldsymbol{E}[d\omega^\alpha(t)] = 0, \quad (6b)$$

$$\boldsymbol{E}[d\omega^\alpha(t_1)\, d\omega^\beta(t_2)] = \delta^{\alpha\beta}\, \delta(t_1 - t_2)\, dt_1 dt_2. \quad (6c)$$

Because the infinitesimal ω-increments are Gaussian, all higher moments are determined by the relations (6). Equation (5) is satisfied by finite increments Δx satisfying a stochastic Taylor series. Fortunately, we do not need all the terms.

Simplified Taylor series

In [13], a stripped down Taylor series of weak order 2 is derived. Similar results are given in Kloeden and Platen [10], and Milstein [11]. Because only some terms of the increments $\{\Delta x^\alpha\}$ contribute to the expectations of each of their four monomials in (5), some simplifications are possible. This requires a little explanation. The model $\Xi^{\epsilon\gamma} \approx \int_{t_0}^{t_0+h} \omega^\epsilon d\omega^\gamma$ is well known (e.g. see Talay [16]) and is written out explicitly in Eq. (9). This approximation (denoted $\approx$) means all $\boldsymbol{E}_{x_t}\left[\Delta x^\alpha \cdots \Delta x^\delta\right]$ expectations in (5) are $O(h^2)$ accurate. In particular, $\boldsymbol{E}[\Xi^{\epsilon\gamma}\Delta\omega^\eta] = 0$ and $\boldsymbol{E}[\Xi^{\epsilon\gamma}\Xi^{\eta\iota}] = \frac{h^2}{2}\delta^{\epsilon\eta}\delta^{\gamma\iota}$. Two other facts must be laid out. Since we want the martingale increment $(\Delta\mathcal{M})$ to have vanishing expectation to all orders in stepsize h, any term of $O(h^2)$ in $\Delta\mathcal{M}(h)$ may be ignored. It will have zero expectation and thus cannot contribute to any moment in (5) to $O(h^2)$. Furthermore, the following substitution is permissible

$$h\Delta\omega^\epsilon \Delta\omega^\gamma \approx h^2\delta^{\epsilon\gamma}, \quad (7)$$

because the approximation on the right-hand side of (7) is the only contribution the left-hand side can make to any moments in (5) to $O(h^2)$. Likewise, any $O(h^{3/2})$ term, say F, in $\Delta\mathcal{M}$ which also has the property $\boldsymbol{E}[F\Delta\omega_1^\beta] = 0$ for all β, can also be ignored since it will never contribute to (5) to $O(h^2)$ either. Otherwise, the truncated Taylor series is a tedious but straightforward exercise in repeated substitution and its details are given in [13]. In every occurrence, the Brownian increments are approximated by

$\xi_k \approx \sqrt{h} z_k$, $k = 0, 1$, where z_0, z_1 are Gaussian random variables with mean zero and variance one. The simplified Taylor series of weak order 2 reads

$$
\begin{aligned}
x^\alpha_{t_0+h} = x^\alpha_0 & \\
& + b^\alpha_0 h + (\partial_\beta b^\alpha_0)\, \sigma^{\beta\gamma}_0 \frac{h}{2} \xi^\gamma_1 && \text{drift} \Delta\mathcal{A}^\alpha(h) \\
& + \frac{1}{2} \left\{ (\partial_\beta b^\alpha_0)\, b^\beta_0 h + \frac{1}{2} (\partial_\beta \partial_\gamma b^\alpha_0) \sigma^{\beta\epsilon}_0 \sigma^{\gamma\iota}_0 \xi^\epsilon_1 \xi^\iota_1 \right\} h && \\
& + \sigma^{\alpha\beta}_0 \xi^\beta_1 + (\partial_\beta \sigma^{\alpha\gamma}_0)\, \sigma^{\beta\epsilon}_0\, \Xi^{\epsilon\gamma} && \text{diffusion} \Delta\mathcal{M}^\alpha(h) \\
& + \frac{1}{2} \left\{ (\partial_\beta \sigma^{\alpha\gamma}_0)\, b^\beta_0 h + \frac{1}{2} (\partial_\beta \partial_\kappa \sigma^{\alpha\gamma}_0)\, \sigma^{\beta\mu}_0 \sigma^{\kappa\epsilon}_0 \xi^\mu_0 \xi^\epsilon_0 \right\} \xi^\gamma_1 . && (8)
\end{aligned}
$$

In Eq. (8), the model $\Xi^{\epsilon\gamma}$ for the stochastic integral $\int \omega^\epsilon d\omega^\gamma$ over one time step is usually called the Milstein term [11], as noted above. This is explicitly given by the model [13,16]

$$
\int_{t_0}^{t_0+h} \omega^\epsilon(s)\, d\omega^\gamma(s) \approx \Xi^{\epsilon\gamma} = \begin{cases} \frac{h}{2} (z^\epsilon_1 z^\gamma_1 - \tilde{z}^{\epsilon\gamma}), & \text{if } \epsilon > \gamma, \\ \frac{h}{2} (z^\epsilon_1 z^\gamma_1 + \tilde{z}^{\gamma\epsilon}), & \text{if } \epsilon < \gamma, \\ \frac{h}{2} ((z^\epsilon_1)^2 - 1), & \text{when } \epsilon = \gamma. \end{cases} \tag{9}
$$

In this model, the $\{\tilde{z}^{\gamma\epsilon}\}$ are $n(n-1)/2$ independent Gaussian random variables, each of mean zero, unit variance, and independent of the sets of random variables $\{z^\epsilon_0\}$ and $\{z^\epsilon_1\}$.

4 Our main result

The following three-stage split-step procedure is proven to be 2nd-order weak accurate in Sect. 5. Applied to (3), the procedure begins at some time t_0 with process value x_0, and proceeds from $t_0 \mapsto t_0 + h$ via the following algorithm. The $\star$ and $\star\star$ indicate intermediate stages (see also Sect. 5).

$$
\begin{aligned}
& \phi_{h/2} : x_0 \mapsto x_\star \text{ is the mapping} \\
& \qquad x_\star = x_0 + \frac{h}{4}(b(x_\star) + b(x_0)), \quad \text{to be solved for } x_\star,
\end{aligned} \tag{10a}
$$

$$
\begin{aligned}
& \psi_h : x_\star \mapsto x_{\star\star} \text{ uses Eqs. 9 and 11 to approximate} \\
& \qquad x_{\star\star} = x_\star + \int_{t_0}^{t_0+h} \sigma(x(s)) d\omega(s), \quad \text{start at } x_\star, \text{ compute } x_{\star\star}, \text{ and}
\end{aligned} \tag{10b}
$$

$$
\begin{aligned}
& \phi_{h/2} : x_{\star\star} \mapsto x_{t_0+h} \text{ is the same as 10a, but starts at } x_{\star\star} \text{ instead of } x_0 : \\
& \qquad x_{t_0+h} = x_{\star\star} + \frac{h}{4}(b(x_{t_0+h}) + b(x_{\star\star})), \quad \text{to be solved for } x_{t_0+h}.
\end{aligned} \tag{10c}
$$

A similar approach was given in Higham et al. [8] and uses two stages. Their first stage is an implicit backward Euler h-stage for the drift, followed by an explicit h-stage for the martingale. Their approach is only 1st-order weak accurate, whereas (10) is 2nd-order weak accurate, as we prove below. Their idea is a good one because the implicit drift stage provides good stability and the stochastic h-stage is explicitly a discrete martingale. Ours shares the same properties and also ends up being a two stage method with only $h/2$-stages at the ends of the composition chain: (22) compared with (19) as we will show.

Our procedure (10) is the stochastic analog of Eq. (2) and has the following three qualities:

- The drift-stages are implicit for stability, but with enough flexibility that certain parts can be made explicit (for example, a Heun scheme [13]): see Sect. 5.1.
- Composite deterministic steps of pairs $\phi_{h/2} \circ \phi_{h/2}$ follow each ψ_h, but $\phi_h = \phi_{h/2} \circ \phi_{h/2}$ to the same $O(h^2)$ accuracy. These $\phi_{h/2} \circ \phi_{h/2}$ pairs from adjacent time steps may thus be compressed into ϕ_h steps in the n-chain as we will show in Sect. 5.1.
- The stochastic diffusion stage is an explicit martingale whose expectation value $\boldsymbol{E}[\Delta\mathcal{M}] = 0$ to all orders in the stepsize h.

5 Proof of order in the split-step procedure Eq. (10)

Given the simplified Taylor series (8), proven in [10,13], two things remain to be done. First, we need a general approximation to the martingale stage (10b), $\Delta\mathcal{M}$. For that, we use the following:

$$x^\alpha_{\star\star} = x^\alpha_\star + \frac{1}{2}\left\{\sigma^{\alpha\beta}\left(x_\star + \frac{1}{\sqrt{2}}\sigma_\star\xi_0\right) + \sigma^{\alpha\beta}\left(x_\star - \frac{1}{\sqrt{2}}\sigma_\star\xi_0\right)\right\}\xi_1^\beta + (\partial_\beta\sigma^{\alpha\delta}_\star)\,\sigma^{\beta\epsilon}_\star\,\Xi^{\epsilon\delta}. \tag{11}$$

Note that we used the shorthand notation $\sigma_\star = \sigma(x_\star)$, as in Sect. 4. Next, we need to show that an expansion of (10) gives the result (8). We will explore two special cases in Sect. 6.4 when x is a scalar: the additive noise case σ = constant, and a strictly linear multiplicative noise case, $\sigma(x) = \mu x$, where μ is also constant.

The proof that (10) is a weak 2nd-order approximation is a just a lengthy expansion. We start with the first stage (10a)

$$x^\alpha_\star = x^\alpha_0 + \frac{h}{4}\left(b^\alpha(x_\star) + b^\alpha(x_0)\right),$$

which when expanded to second order by repeated substitution yields

$$x^\alpha_\star = x^\alpha_0 + b^\alpha(x_0)\frac{h}{2} + (\partial_\gamma b^\alpha(x_0))b^\gamma(x_0)\frac{h^2}{8} + o(h^2). \tag{12}$$

Now, expanding (11) to order $O(h^{3/2})$ (we may ignore the $O(h^2)$ term which has vanishing expectation), gives

$$x^\alpha_{\star\star} = x^\alpha_\star + \underbrace{\sigma^{\alpha\beta}_\star \xi^\beta_1}_{H_1} + \underbrace{\frac{1}{4}(\partial_\gamma \partial_\eta \sigma^{\alpha\beta}_\star)\sigma^{\gamma\epsilon}_\star \sigma^{\eta\iota}_\star \xi^\epsilon_0 \xi^\iota_0 \xi^\beta_1}_{H_2} + \underbrace{(\partial_\beta \sigma^{\alpha\delta}_\star)\sigma^{\beta\epsilon}_\star \Xi^{\epsilon\delta}}_{H_3} + O(h^2), \tag{13}$$

where again the $O(h^2)$ term may be ignored because it has vanishing expectation. We examine these in reverse order, H_3, H_2, H_1, because there is little to do for H_3 and H_2 since $\sigma_\star = \sigma_0 + O(h)$, and because Ξ has vanishing expectation. In H_3, the $O(h) \cdot \Xi^{\epsilon\delta} = O(h^2)$ term may be ignored, thus

$$H_3 = (\partial_\beta \sigma^{\alpha\delta}_0)\sigma^{\beta\epsilon}_0 \Xi^{\epsilon\delta} + O(h^2). \tag{14}$$

The H_2 term is just as easy, because again using $\sigma_\star = \sigma_0 + O(h)$, we obtain

$$H_2 = \frac{1}{4}(\partial_\gamma \partial_\eta \sigma^{\alpha\beta}_0)\sigma^{\gamma\epsilon}_0 \sigma^{\eta\iota}_0 \xi^\epsilon_0 \xi^\iota_0 \xi^\beta_1 + O(h^{5/2}). \tag{15}$$

Thus, the only term in (13) which requires further expansion is H_1:

$$H_1 = \sigma^{\alpha\beta}_0 \xi^\beta_1 + (\partial_\gamma \sigma^{\alpha\beta}_0) b^\gamma_0 \frac{h}{2} \xi^\beta_1 + O(h^{5/2}). \tag{16}$$

Finally, the last stage (10c) gives the end value,

$$x^\alpha_h = x^\alpha_{\star\star} + \underbrace{b^\alpha_1 \frac{h}{2}}_{J_1} + \underbrace{(\partial_\gamma b^\alpha_1) b^\gamma_1 \frac{h^2}{8}}_{J_2}.$$

The last term J_2 is easy because it is already $O(h^2)$, thus

$$J_2 = (\partial_\gamma b^\alpha_0) b^\gamma_0 \frac{h^2}{8}. \tag{17}$$

From (13), we can get b^α_1 in $J_1 = b^\alpha_1 \frac{h}{2}$, with the result

$$b^\alpha_1 \frac{h}{2} = b^\alpha_0 \frac{h}{2} + (\partial_\kappa b^\alpha_0) \left\{ \sigma^{\kappa\lambda}_0 \xi^\lambda_1 + b^\kappa_0 \frac{h}{2} + (\partial_\phi \sigma^{\kappa\mu}_0)\sigma^{\phi\chi}_0 \Xi^{\chi\mu} \right\} + (\partial_\kappa \partial_\mu b^\alpha_0)\sigma^{\kappa\lambda}_0 \sigma^{\mu\phi}_0 \frac{h}{4} \xi^\lambda_1 \xi^\phi_1.$$

Here, $h\xi_1^\lambda \xi_1^\phi \approx h^2 \delta^{\lambda\phi}$ to order $O(h^2)$, as in (7). Combining everything, Eq. (12) for $x_\star$, (16) for H_1, (15) for H_2, and (14) for H_3 to form $x_{\star\star}$ in Eq. (13), along with J_1 and J_2 (17) to form the final stage x_h, we get

$$\begin{aligned} x^\alpha_{t_0+h} = x^\alpha_0 &+ \sigma_0^{\alpha\beta}\xi_1^\beta + b_0^\alpha h + (\partial_\beta \sigma_0^{\alpha\gamma})\sigma_0^{\beta\epsilon}\,\Xi^{\epsilon\gamma} \\ &+ (\partial_\gamma \sigma_0^{\alpha\beta}) b_0^\gamma \frac{h}{2}\xi_1^\beta + (\partial_\kappa b_0^\alpha)\sigma_0^{\kappa\lambda}\frac{h}{2}\xi_1^\lambda + \frac{1}{4}(\partial_\gamma \partial_\eta \sigma_0^{\alpha\beta})\sigma_0^{\gamma\epsilon}\sigma_0^{\eta\iota}\xi_0^\epsilon \xi_0^\iota \xi_1^\beta \\ &+ (\partial_\gamma b_0^\alpha) b_0^\gamma \frac{h^2}{2} + (\partial_\kappa \partial_\mu b_0^\alpha)\sigma_0^{\kappa\lambda}\sigma_0^{\mu\lambda}\frac{h^2}{4} + o(h^2). \end{aligned} \tag{18}$$

Again, we used (7) in the last contributing term. Comparing (18) with (8), term by term, the proof that (10) is second-order weak accurate is now complete. □

5.1 Compression

An n-step computation of x_{t_0+nh} starting with x_0 is a composition rule with n triplets (2) of $\phi_{h/2} \circ \psi_h \circ \phi_{h/2}$:

$$x_{t_0+nh} = \underbrace{\phi_{h/2} \circ \psi_h \circ \phi_{h/2}}_{x_{nh} \leftarrow\!\shortmid x_{(n-1)h}} \circ \cdots \circ \overbrace{\phi_{h/2} \circ \phi_{h/2}}^{\approx \phi_h} \circ \psi_h \circ \phi_{h/2} \circ \underbrace{\phi_{h/2} \circ \psi_h \circ \phi_{h/2}}_{x_h \leftarrow\!\shortmid x_0}(x_0). \tag{19}$$

To the same order, $O(h^2)$, of accuracy, each triplet overlaps its next neighbor to the left, that is $\phi_{h/2} \circ \phi_{h/2}$, by essentially a full h-step trapezoidal rule ϕ_h, where

$$\phi_h : x_{\star\star} \mapsto x_\star, \text{ in which } x_\star = x_{\star\star} + \frac{h}{2}\left(b(x_\star) + b(x_{\star\star})\right), \tag{20}$$

is to be solved for the intermediate value $x_\star$, which is the input to the following martingale step ψ_h. Straightforward repeated back-substitutions in the composition $\phi_{h/2} \circ \phi_{h/2}(x_{\star\star})$ compared to that of one full h-step $\phi_h(x_{\star\star})$ from (20), show that

$$\phi_h(x_{\star\star}) - \phi_{h/2} \circ \phi_{h/2}(x_{\star\star}) = \frac{h^3}{16}\left((b_{\star\star})^2 b''_{\star\star} + (b'_{\star\star})^2 b_{\star\star}\right) + O(h^4), \tag{21}$$

where b, b', and b'' are evaluated at $x_{\star\star}$. By occasionally computing *both* ϕ_h and $\phi_{h/2} \circ \phi_{h/2}$, an estimate of the truncation error/step is possible, suggesting a step-size monitor [14] (sec. 7.2.5). This truncation error estimate is four times the $O(h^3)$ term of (21). The *compression* eliminates $n-1$ trapezoidal rule steps.

A compressed version of (19) is thus

$$x_{t_0+nh} = \phi_{h/2} \circ \psi_h \circ \phi_h \circ \cdots \circ \phi_h \circ \psi_h \circ \phi_h \circ \psi_h \circ \phi_{h/2}(x_0), \tag{22}$$

where only the first and last of the compositions have half-steps $\phi_{h/2}$. This compression reduces the n triplets ($3n$ operations of ϕ or ψ) down to $2n+2$ such operations: that

is, one ψ_h, $n-1$ ($\psi_h \circ \phi_h$) pairs, and the ends. The end corrections are familiar from trapezoidal formula integrations: see Sect. 18.9 of [7]. Also notice that (22) remains left–right symmetric, as is (19).

5.2 A semi-implicit variant

A further remark is in order. Our basic trapezoidal rule steps of the form (20) require solving non-linear algebraic equations for $x_\star$. In general this will be awkward, particularly in the vector case. An easier version is apparent when $b(x) = Ax + q(x)$ has a linear part. Here, $q(x)$ is a non-linear part which we may be hesitant to deal with in implicit form. In this situation, a semi-implicit procedure

$$x_\star = x_{\star\star} + \frac{h}{2}\left(Ax_\star + q(x_E) + b(x_{\star\star})\right),$$

where $x_E = x_{\star\star} + hb(x_{\star\star})$ is the Euler estimate, can be solved easily. Namely,

$$x_\star = \left(1 - \frac{h}{2}A\right)^{-1}\left(x_{\star\star} + \frac{h}{2}\left(q(x_E) + b(x_{\star\star})\right)\right)$$

only requires the solution of one linear system. The order of this semi-implicit method remains $O(h^2)$ accurate: see [13], or Section 2.3 of [4]. It works here because our $\phi - \psi$ splitting defines ϕ_h to be deterministic. Alternatively, replacing any $x_{\star\star}$ term in $b(x_{\star\star})$ with x_E would lead to another variant on this theme.

6 Test example

An SDE example having both linear drift and diffusion is given by

$$dX = (a_0 + a_1 X)dt + (c_0 + c_1 X)dw(t), \qquad \text{with initial value } X(0) = X_0. \quad (23)$$

Here X is a complex scalar process driven by a real one-dimensional Brownian motion $w(t)$ (see Kloeden and Platen [10], equation (4.4.9)). If $a_1, c_1 \neq 0$, one of the constants a_0 or c_0 may be set to zero without loss of generality. To see this, either set $X(t) = Z(t) - a_0/a_1$, which eliminates a_0 in the drift, or set $X(t) = Z(t) - c_0/c_1$ to eliminate c_0 in the diffusion, with suitable changes to $\{a_1, c_0, c_1\}$, or $\{a_0, a_1, c_1\}$, respectively, in the resulting SDE for $Z(t)$. It is not particularly difficult to get a formal solution to (23) using an integrating factor [9,10], as we now show.

6.1 A formal solution to the test example

Let $M(t) = \exp(-c_1 w(t) - (a_1 + c_1^2/2)t)$, which has the Itô SDE [9]

$$dM = -a_1 M dt - c_1 M dw(t). \quad (24)$$

Furthermore, for all $t < \infty$, $M(t)$ has the inverse

$$M^{-1}(t) = e^{c_1 w(t) + (a_1 + c_1^2/2)t}.$$

Thus, using (23) and (24),

$$\begin{aligned} d(MX) &= dM\, X + M\, dX + dM\, dX \\ &= (a_0 - c_0 c_1) M dt + c_0 M dw(t) - c_1^2 M X dt. \end{aligned}$$

By using the integrating factor $e^{c_1^2 t}$, we get

$$d(e^{c_1^2 t} MX) = e^{c_1^2 t}((a_0 - c_0 c_1) M dt + c_0 M dw(t)),$$

and thus

$$e^{c_1^2 t}(MX)(t) - X(0) = (a_0 - c_0 c_1) \int_0^t e^{c_1^2 s} M(s) ds + c_0 \int_0^t e^{c_1^2 s} M(s) dw(s).$$

Now let

$$P(t) = e^{-c_1^2 t} M^{-1}(t) = e^{c_1 w(t) + (a_1 - c_1^2/2)t},$$

which satisfies the SDE [compare with (24)]

$$dP = a_1 P dt + c_1 P dw(t). \tag{25}$$

This yields the formal solution [10]

$$X(t) = X(0)P(t) + (a_0 - c_0 c_1) P(t) \int_0^t P^{-1}(s) ds + c_0 P(t) \int_0^t P^{-1}(s) dw(s), \tag{26}$$

where, of course,

$$P^{-1}(t) = e^{-c_1 w(t) - (a_1 - c_1^2/2)t}.$$

6.2 $\boldsymbol{E}[X(t)]$ test statistic

This solution (26) is only formal in the sense that neither integral in (26) can be evaluated as a function of $(t, \omega(t))$. However, Eq. (26) can be used to compute the expected value $\boldsymbol{E}[X(t)]$, which we will use as a test statistic for our integrator (10).

We have

$$\boldsymbol{E}[X(t)] = X(0)\underbrace{\boldsymbol{E}[P(t)]}_{E_0} + (a_0 - c_0c_1)\underbrace{\boldsymbol{E}\left[P(t)\int_0^t P^{-1}(s)ds\right]}_{E_1}$$
$$+c_0\underbrace{\boldsymbol{E}\left[P(t)\int_0^t P^{-1}(s)dw(s)\right]}_{E_2}.$$

Finding that $E_0 = e^{a_1 t}$ is easy enough. Likewise, evaluating E_1 is straightforward:

$$\begin{aligned} E_1 &= \int_0^t \boldsymbol{E}[P(t)P^{-1}(s)]ds \\ &= \int_0^t e^{(a_1 - \frac{c_1^2}{2})(t-s)}\boldsymbol{E}\left[e^{c_1(w(t)-w(s))}\right]ds \\ &= \int_0^t e^{(a_1 - \frac{c_1^2}{2})(t-s)}e^{\frac{c_1^2}{2}(t-s)}ds \\ &= e^{a_1 t}\int_0^t e^{-a_1 s}ds = \frac{e^{a_1 t}-1}{a_1}. \end{aligned}$$

Finally, although E_2 looks awkward, considering its differential dE_2 and remembering $dP(t) = a_1 P dt + c_1 P d\omega(t)$ from (25) are helpful. We have

$$\begin{aligned} dE_2 &= \boldsymbol{E}\left[(a_1 P(t)dt + c_1 P(t)d\omega t)\left(\int_0^t P^{-1}(s)d\omega(s)\right)\right] \\ &\quad + \boldsymbol{E}\left[P(t)P^{-1}(t)d\omega(t) + c_1 P(t)P^{-1}(t)dt\right] \\ &= a_1 E_2\, dt + c_1\, dt. \end{aligned} \tag{27}$$

To get (27), we used $\boldsymbol{E}[P(t)d\omega(t)\int_0^t P^{-1}(s)d\omega(s)] = 0$. To see why this vanishes, apply $\boldsymbol{E}$ applied to the Itô martingale: $\boldsymbol{E}\left[\int_0^t d\omega(u)(P(u)\int_0^u P^{-1}(s)d\omega(s))\right] = 0$. This must hold for all t and therefore $t+dt$. From (27), the integrating factor $\exp(-a_1 t)$ gives

$$E_2 = c_1\frac{e^{a_1 t}-1}{a_1}.$$

All together, we have

$$\boldsymbol{E}[X(t)] = \left(X(0) + \frac{a_0}{a_1}\right)e^{a_1 t} - \frac{a_0}{a_1}. \tag{28}$$

It is important to note that this expectation result is independent of c_0, c_1, although the $X(t)$ distributions are very different for each choice of these coefficients.

6.3 Numerical test examples

In Sect. 6.4, we test two complex oscillating cases with $a_1 = i$: a scalar additive noise problem (**N1**), and a multiplicative noise case (**N2**). For convenience, the shorthand notations $X_t = X(t)$ and $Z_t = Z(t)$ will be freely used.

N1: The choices $a_0 = 1, a_1 = i, c_0 = 1$ and $c_1 = 0$ give the SDE

$$dX = (1 + iX)dt + d\omega(t), \quad \text{whose mean} \tag{29}$$

$$\boldsymbol{E}[X(t)] = (X(0) - i)e^{it} + i, \tag{30}$$

is known from (28). For this problem, we can compute the absolute variance using the substitution $Z(t) = X(t) - i$: we get $\boldsymbol{E}[|X(t) - ((X_0 - i)e^{it} + i)|^2] = \boldsymbol{E}[|Z - Z_0 e^{it}|^2]$. Therefore,

$$\begin{aligned}
\boldsymbol{E}\big[|Z - Z_0 e^{it}|^2\big] &= \boldsymbol{E}\left[\left(e^{it}\int_0^t e^{-is}d\omega(s)\right)\left(e^{-it}\int_0^t e^{iu}d\omega(u)\right)\right] \\
&= \int_0^t e^{-is}\int_0^t e^{iu}\boldsymbol{E}[d\omega(s)d\omega(u)] \\
&= \int_0^t e^{-is}\left(\int_0^t e^{iu}\delta(u-s)du\right)ds \\
&= t.
\end{aligned} \tag{31}$$

The variance for **N1** thus grows linearly with t, and we might expect poor numerical results for $\boldsymbol{E}[X(t)]$ when $t \gg |\boldsymbol{E}Z(t)|^2 = |Z_0|^2 = 2$, that is when $X_0 = 1$, as in our examples.

N2: From the choices $a_0 = 1, a_1 = i, c_0 = 0$, and $c_1 = 1$ we have the SDE

$$dX = (1 + iX)dt + Xd\omega, \tag{32}$$

whose mean, $\boldsymbol{E}[X(t))]$ is also given by (28) and is the same as (30). For this problem, we can also compute the absolute variance by again using the substitution $Z(t) = X(t) - i$:

$$\begin{aligned}
\boldsymbol{E}\big[|Z_t - Z_0 e^{it}|^2\big] &= \boldsymbol{E}\left[\left(\int_0^t e^{-is}(Z_s + i)d\omega(s)\right)\left(\int_0^t e^{iu}(\overline{Z}_u - i)d\omega(u)\right)\right] \\
&= \boldsymbol{E}\left[\int_0^t Q(s)d\omega(s)\int_0^t \overline{Q}(u)d\omega(u)\right],
\end{aligned}$$

where we used the abbreviation $Q(t) = e^{-it}(Z_t + i)$ and $\overline{Z}$ is the complex conjugate of Z. By the same trick used to get (27), we have

$$\begin{aligned} d\boldsymbol{E}[|Z_t|^2] &= d\left(\boldsymbol{E}[|Z_t|^2] - |Z_0|^2\right) \\ &= \boldsymbol{E}[|Q(t)|^2]dt \\ &= \boldsymbol{E}[|Z_t + i|^2]dt \\ &= \left(\boldsymbol{E}[|Z_t|^2] + 2\,\mathrm{Im}\{\boldsymbol{E}[Z_t]\} + 1\right)dt \\ &= \left(\boldsymbol{E}[|Z_t|^2] + 2\,\mathrm{Im}\{Z_0 e^{it}\} + 1\right)dt. \end{aligned}$$

So, by using the integrating factor e^{-t}, we get an expression for the variance

$$\begin{aligned} var &= \boldsymbol{E}[|Z_t|^2] - |Z_0|^2 \\ &= (|Z_0|^2 + 1)(e^t - 1) + \mathrm{Im}\,\{Z_0(1+i)(e^{it} - e^t)\}. \end{aligned}$$

From the Taylor expansion $e^t + 2\cos(t) + e^{-t} - 4 = 4\sum_{k=1}^{\infty} \frac{t^{4k}}{(4k)!} \geq 0$, we get $2(e^t - 1)^2 \geq |e^t - e^{it}|^2$, which yields a useful sharp lower bound for var:

$$\begin{aligned} var &\geq (|Z_0|^2 + 1)(e^t - 1) - \sqrt{2}|Z_0||e^t - e^{it}| \\ &\geq (|Z_0|^2 + 1)(e^t - 1) - 2|Z_0||e^t - 1| \\ &= (|Z_0| - 1)^2(e^t - 1). \end{aligned} \tag{33}$$

Hence, the variance for **N2** grows exponentially as t gets large. We can estimate when the variance becomes much larger than $|\boldsymbol{E}[Z(t)]|^2 = |Z_0|^2 = |X_0 - i|^2$, namely if

$$(|Z_0| - 1)^2(e^t - 1) \gg |Z_0|^2.$$

For our test, $X_0 = 1$, so $|Z_0|^2 = 2$, and thus the variance is larger than this and growing rapidly when

$$t > \log\left\{1 + \frac{2}{(\sqrt{2} - 1)^2}\right\} \approx 2.54.$$

This is consistent with the results in Figs. 8, 10, and 12, where, by eyeball inspection, the simulated means move away from the circle when $t > 2.5$.

6.4 Numerical solutions

In turn, we can write down the algorithm (10) for the problems **N1** and **N2** of Sect. 6.3. In the last case, from (9) we have $\Xi^{11} = (\xi_1^2 - h)/2$, and in both cases $\xi_1 = \sqrt{h}z_1$ (see Sect. 3). All our **N1** and **N2** plots shown here were from the simple forms of **stage1**, **stage2**, and **stage3**, neither combined nor alternate.

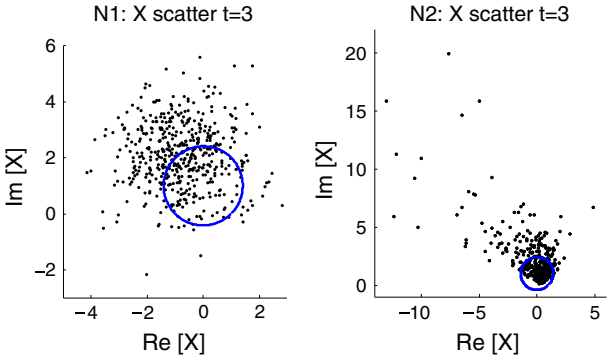

Fig. 1 *Left*: **N1** scatter-plot, $N = 500$, $t = 3$. *Right*: same for **N2**. Both radius $= \sqrt{2}$ orbits of $\boldsymbol{E}\,[X(t)]$ are shown as *circles*: **N2**'s looks very small. Method (10) was used with $h = 0.01$

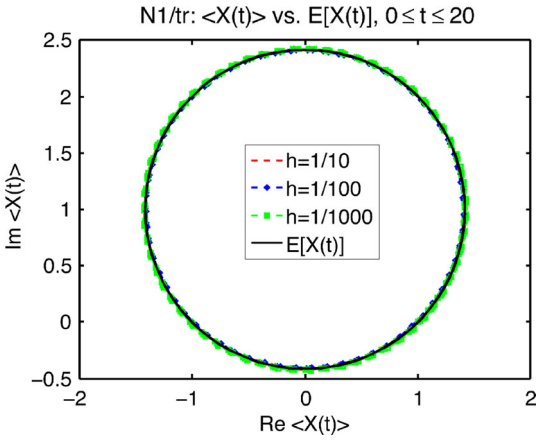

Fig. 2 Numerical results for problem **N1** using our 3-stage (10) method with $h = 1/10, 1/100, 1/1{,}000$ and all other parameters given in the notes on page 16

N1's numerical solution. In the additive noise SDE (29) for **N1**, we have $\sigma = 1$, a constant. The stages of (10) in this case are:

stage1

$$X_\star = \left(1 - \frac{ih}{4}\right)^{-1} \left(\left(1 + \frac{ih}{4}\right) X_0 + h/2\right) \tag{34}$$

stage2 is simply

$$X_{\star\star} = X_\star + \xi_1 \tag{35}$$

stage3

$$X_h = \left(1 - \frac{ih}{4}\right)^{-1} \left(\left(1 + \frac{ih}{4}\right) X_{\star\star} + h/2\right) \tag{36}$$

stage3 and the next **stage1** combined together have the analytic form

$$X_\star = \phi_h(X_{\star\star}) = e^{ih} X_{\star\star} + i(1 - e^{ih}). \tag{37}$$

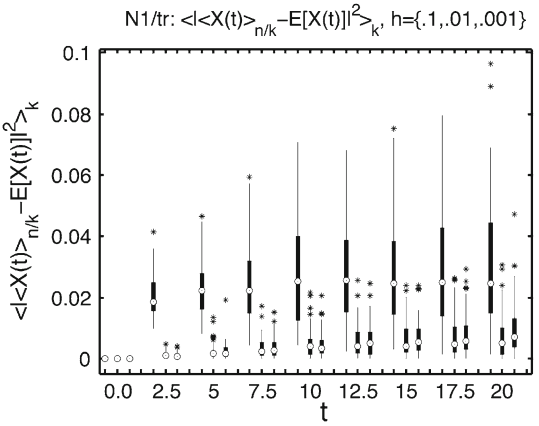

Fig. 3 Fiber-box plots of the distributions for **N1** of the variance around the analytic mean $\boldsymbol{E}[X(t)]$ using trapezoidal rule (10). *Each box* triplet ($h = 10^{-1}, 10^{-2}, 10^{-3}$) shares the same t-value indicated in the labeling. See the notes on page 16

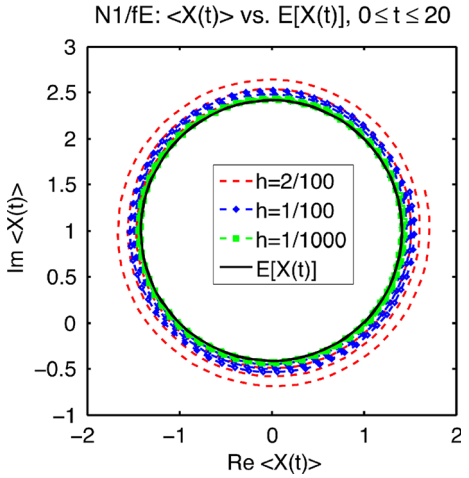

Fig. 4 Numerical results for problem **N1** using Euler's method, which is the same as Milstein's in this case. Note the outward spiraling for as step size h increases. See the notes on page 16

N2's numerical solution. From (32), the stages of (10) are:

stage1 is the same as (34).

stage2, where again Ξ^{11} is from eq. (9) and page 15

$$X_{\star\star} = \left(1 + \xi_1 + \Xi^{11}\right) X_{\star},$$

an alternate stage2 would be the explicit version

$$X_{\star\star} = e^{\xi_1 - h/2} X_{\star}. \tag{38}$$

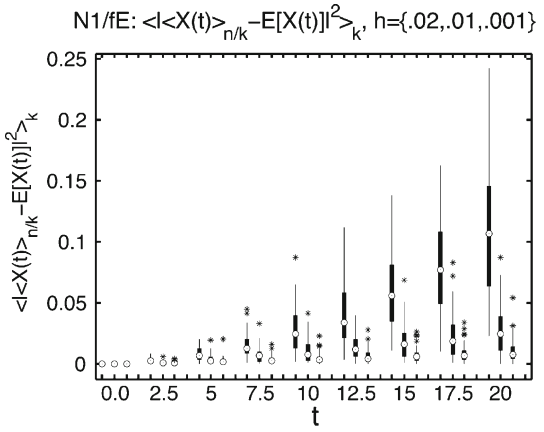

Fig. 5 Fiber-box plots of the distribution over sub-sample means for **N1** using the forward Euler method. *Each box* triplet ($h = 0.02, 0.01, 0.0001$) shares the same t-value indicated in the labeling. See the notes on page 16

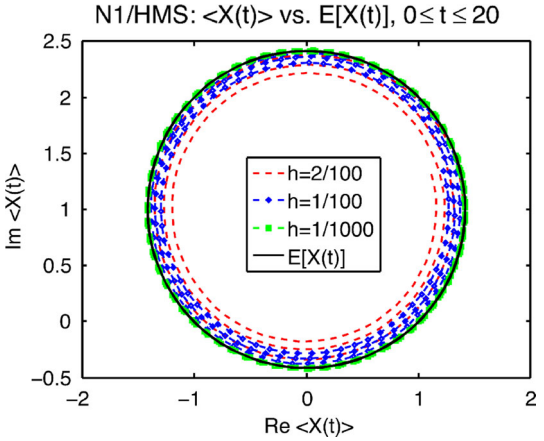

Fig. 6 Numerical orbits for **N1** using HMS showing inward spiral. See the notes on page 16

stage3 is the same as (36).
stage3 combined with **stage1** $\phi_h(X_{\star\star}) = \phi_{h/2} \circ \phi_{h/2}(X_{\star\star})$ is also eq. (37).

6.5 Numerical results

The simulation conditions were as follows. Programs providing data for all the figures were written in C, and compiled with `gcc 4.2.1 (-O3)`. The plotters were from MATLAB R2012B. The machine was a **Mac Mini**, OS 10.6.8, with an Intel Core 2 Duo chip. The sample sizes there were $N = 2^{17} = 131072$, with fixed time steps having sizes noted in the figures: for $h = .001$, runtime was ∼5 min. The random number generator was the `WELL512a` [12], with linear congruential (`gg1`) seeding [2]. To test

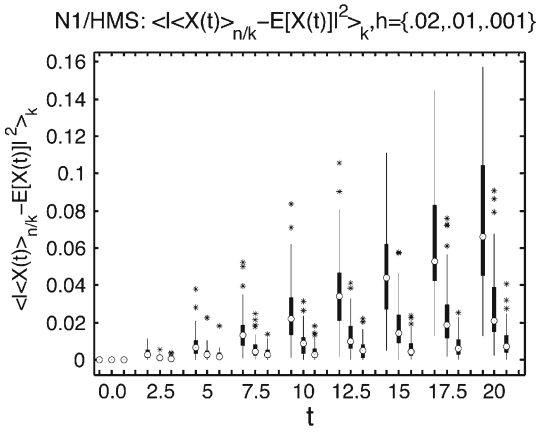

Fig. 7 Numerical results for problem **N1** using the HMS method [8]. *Each box* triplet ($h = 0.02, 0.01, 0.001$) shares the same t-value indicated in the labeling. See the notes on page 16

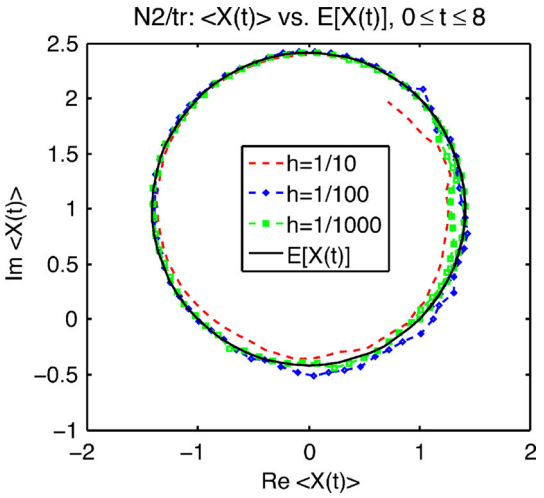

Fig. 8 Numerical mean orbits for problem **N2** using method (10). See the notes on page 16

consistency, k independent sub-samples were taken from total N, $k = (2N)^{1/3} = 64$, with $N/k = 2^{11} = 2{,}048$ independent paths per sub-sample.

Notes on the figures

- Monte-Carlo mean estimates for $\boldsymbol{E}[X(t)]$ are denoted by $\langle X\rangle = \frac{1}{N}\sum_{l=1}^{N} X_l$ for the full N sample, and
- $\langle X\rangle_{N/k}$ for a size N/k sub-sample.
- Functionals over all sub-samples are written $\langle f(X)\rangle_k$

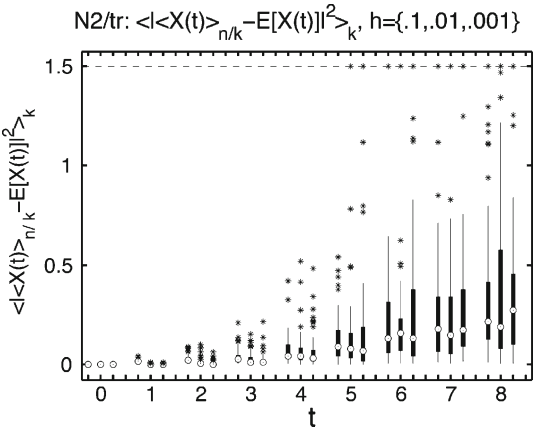

Fig. 9 Fiber-box plots of the mean distributions for problem **N2** using (10). *Each box* triplet ($h = 10^{-1}, 10^{-2}, 10^{-3}$) shares the same t-value in the labeling. See the notes on page 16

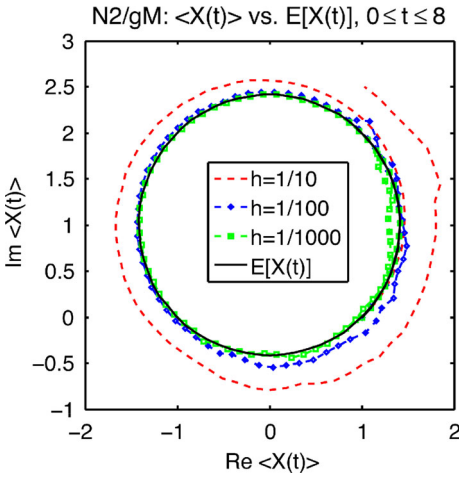

Fig. 10 Numerical results on the orbits for problem **N2** using Milstein's method. The forward Euler treatment of the drift causes outward spiraling. See the notes on page 16

In particular, for sub-sample m which begins at $X_{1+(m-1)\frac{N}{k}}$, its mean is

$$\langle X\rangle[m] = \frac{k}{N}\sum_{j=1}^{N/k} X_{j+(m-1)\frac{N}{k}}$$

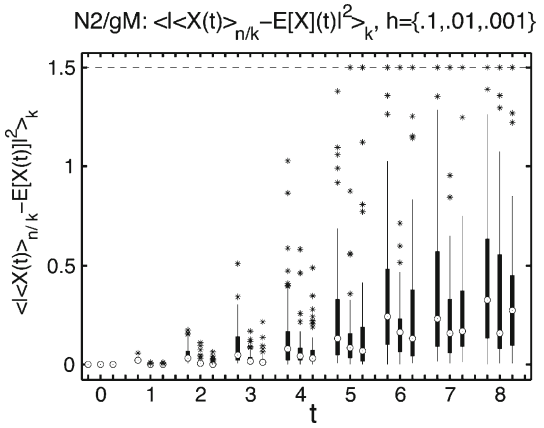

Fig. 11 Fiber *box* plots for the means of problem **N2** using Milstein's method. *Each box* triplet ($h = 10^{-1}, 10^{-2}, 10^{-3}$) shares the same t-value indicated in the labeling. See the notes on page 16

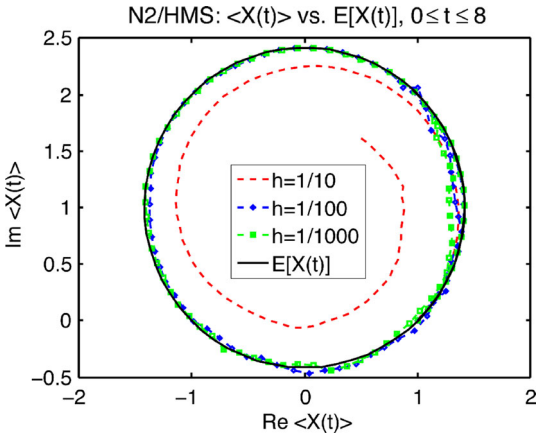

Fig. 12 Numerical orbits of the mean in problem **N2** using the HMS method. Backward Euler treatment of the drift causes inward spiraling. See the notes on page 16

and we compare these sub-sample means to the analytic expectation $\boldsymbol{E}[X(t)] = (1 - i)e^{it} + i$ by computing the variance

$$\langle |\langle X\rangle_{N/k} - \boldsymbol{E}[X(t)]|^2\rangle_k = \frac{1}{k}\sum_{m=1}^{k} |\langle X\rangle[m] - \boldsymbol{E}[X(t)]|^2,$$

which defines the labeling in the fiber-box plots. Consistent sampling means $\langle X\rangle[m]$ should give a good approximation to $\boldsymbol{E}[X(t)]$ for each m. Because the forward Euler method spirals toward infinity while backward Euler (HMS) goes toward zero for large h, we only plotted values $h = 0.02, 0.01, 0.001$ for these methods on **N1**. Other

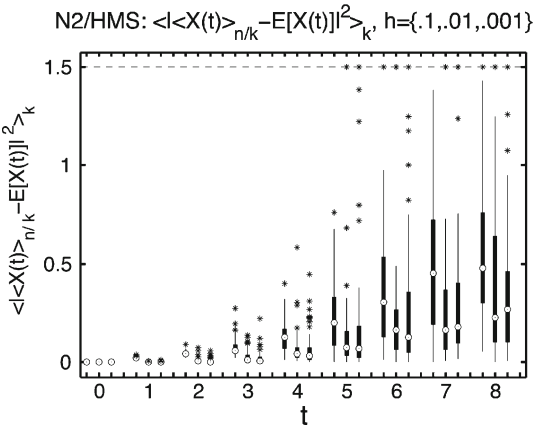

Fig. 13 Fiber *box* plots for the means of problem **N2** using the HMS method [8]. *Each box* triplet ($h = 10^{-1}, 10^{-2}, 10^{-3}$) shares the same t-value as labeled on this axis. See the notes on page 16

figures have $h = 10^{-1}, 10^{-2}, 10^{-3}$. Each box shows the median of the 64 sub-sample means $\langle X \rangle_{n/k}$ in *target* mode [1]. Box ends denote first (lower) and the fourth (upper) quartiles. Exponential growth in variance (33), required outliers for **N2** be cut off at value 1.5 for plot legibility. It is worthwhile to see what examples **N1**, Eq. (31), and **N2**, Eq. (33), really mean. Figure 1 shows scatter-plots of 500 path samples at time $t = 3$. In particular, the exponential growth (33) in the variance, i.e. scatter, in problem **N2** indicates why the $\boldsymbol{E}[X(t)]$ statistic becomes poorly determined as t gets large. Other notation: **tr** means trapezoidal rule, **gM** denotes G. Milstein's method [11], and **HMS** refers to the method in [8].

7 Conclusions

To summarize, we have devised a three-stage, 2nd-order weak accurate integration scheme for Monte-Carlo integration of stochastic differential equations driven by Brownian motion. The procedure is based on trapezoidal rule, but the stochastic stage is isolated, as in [8]. By combining adjacent $h/2$-sized drift steps, it becomes a two-stage method, as we showed in Sect. 5.1. For oscillating problems, like our **N1** and **N2** examples, the numerical results for the statistic $\boldsymbol{E}[X(t)]$ stay on the proper orbit—for long times in case **N1**, but for shorter times in case **N2**. Trapezoidal rule does a good job keeping the sample paths on circular orbits for reasons noted on page 3. It is well known (see Figs. 1.2 and 1.4 in [6]) for oscillating ordinary differential equations that forward Euler solutions tend to spiral out, while backward Euler solutions tend to spiral in. Our results for both problems **N1** and **N2** show these same properties and are illustrated in Figs. 2, 4, 6, 8, 10, and 12. Problem **N2** has an exponentially growing variance, so long time estimates of $\boldsymbol{E}[X(t)]$ become inaccurate. All three integrators—Milstein's method [11], Higham et al. [8], and ours (10)—simulate reasonably well what the process should do. Thus, integrations of problem **N2** behave as expected for all three of these integrators but the comportment of method (10) is better for larger

step sizes. Finally, we also ran the simulations with **alternate** and **combined** forms enumerated in Sect. 6.4 with results virtually indistinguishable from those shown in Figs. 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 and 13.

Acknowledgments We are grateful to our most critical reviewer. His/her hard work helped improve this paper significantly.

References

1. Documentation for boxplot. http://www.mathworks.com/help/stats/boxplot.html
2. http://www.roguewave.com/products/imsl-numerical-libraries.aspx Routine `ggl` is the same as `IMSL`'s `RNUM`. A single precision `C` version is available from this URL http://people.inf.ethz.ch/arbenz/book/Chapter2/dex1.c
3. Coddington, E.A., Levinson, N.: Theory of Ordinary Differential Equations. McGraw-Hill, New York (1955)
4. Gear, C.W.: Numerical Initial Value Problems in Ordinary Differential Equations. Prentice-Hall, Englewood Cliffs (1971)
5. Godunov, S.K., Ryabenkii, V.S.: Difference Schemes. Elsevier Publishers, Amsterdam (1987)
6. Hairer, E., Lubich, C., Wanner, E.: Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations. Springer, Berlin (2006)
7. Hamming, R.W.: Numerical Methods for Scientists and Engineers. Dover Books, New York (1980)
8. Higham, D.J., Mao, X., Stuart, A.M.: Strong convergence of Euler-type methods for nonlinear stochastic differential equations. SIAM J. Numer. Anal. **40**(3), 1041–1063 (2002)
9. Ikeda, N., Watanabe, S.: Stochastic Differential Equations and Diffusion Processes, pp. 218–220. Elsevier, North Holland (1981)
10. Kloeden, P.E., Platen, E.: Numerical Solution of Stochastic Differential Equations. Springer, Berlin (1992)
11. Milstein, G. N.: Numerical Solution of Stochastic Differential equations, (revised and translated version of 1988 Russian work published by Ural State University Press). Kluwer Academic (1995)
12. Panneton, F., L'Ecuyer, P., Matsumoto, M.: Improved long-period generators based on linear recurrences modulo 2. ACM Trans. Math. Softw. (TOMS) **32**(1), 1–16 (2006)
13. Petersen, W.P.: A general implicit splitting for stabilizing numerical simulations of Itô stochastic differential equations. SIAM J. Numer. Anal. **35**(4), 1439–1451 (1998). (electronic)
14. Stoer, J., Bulirsch, R.: Introduction to Numerical Analysis. Springer, New York (1980)
15. Strang, W.G.: On the construction and comparison of difference schemes. SIAM J. Numer. Anal. **5**, 506–517 (1968)
16. Talay, D.: Discrétisation d'une équation différentielle stochastique et calcul approché d'espérances de fonctionnelles de la solution. RAIRO Modél. Math. Anal. Numér. **20**(1), 141–179 (1986)
17. Yoshida, H.: Construction of higher order symplectic integrators. Phys. Lett. A **150**, 262–268 (1990)