High Weak Order Methods for Stochastic Differential Equations Based on Modified Equations

Abdulle, Assyr; Cohen, David; Vilmart, Gilles; Zygalakis, Konstantinos C.

How to cite

ABDULLE, Assyr et al. High Weak Order Methods for Stochastic Differential Equations Based on Modified Equations. In: SIAM journal on scientific computing, 2012, vol. 34, n° 3, p. A1800–A1823. doi: 10.1137/110846609

This publication URL: https://archive-ouverte.unige.ch/unige:41947
Publication DOI: 10.1137/110846609

HIGH WEAK ORDER METHODS FOR STOCHASTIC DIFFERENTIAL EQUATIONS BASED ON MODIFIED EQUATIONS

ASSYR ABDULLE*, DAVID COHEN†, GILLES VILMART‡, AND KONSTANTINOS C. ZYGALAKIS§

Abstract. Inspired by recent advances in the theory of modified differential equations, we propose a new methodology for constructing numerical integrators with high weak order for the time integration of stochastic differential equations. This approach is illustrated with the constructions of new methods of weak order two, in particular, semi-implicit integrators well suited for stiff (mean-square stable) stochastic problems, and implicit integrators that exactly conserve all quadratic first integrals of a stochastic dynamical system. Numerical examples confirm the theoretical results and show the versatility of our methodology.

Key words. weak convergence, modified equations, backward error analysis, stiff integrator, invariant preserving integrator

AMS subject classifications. 65C30, 60H35, 65L20

1. Introduction. The problem of computing the expectation of some functional of a random process appears in many practical situations, for example: in finance [37], in random mechanics [42], in nonlinear filtering [11] or bio-chemical processes [15], to mention a few examples. Here, we are interested in the situation where the random process is the solution of an Itô stochastic system differential equations (SDEs)

$$dX = f(X)dt + g(X)dW(t), \quad X(0) = X_0, \tag{1.1}$$

where $X(t), X_0$ are random variables with values in $\mathbb{R}^d$, $f : \mathbb{R}^d \to \mathbb{R}^d$ is the drift term, $g : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ is the diffusion term with $d \times m$ matrix values, and the components $W_{[j]}(t), j = 1, \dots, m$ of $W(t) = (W_{[1]}(t), \dots, W_{[m]}(t))^T$ are independent Wiener processes. We assume that the drift and diffusion terms are smooth enough, Lipschitz continuous and satisfy a growth bound, to ensure a unique (mean-square bounded) solution of (1.1) [5, 22]. Analytic solutions of SDEs are rarely known and their practical computation are usually done numerically. A one-step numerical method for the approximation of (1.1) is given by

$$X_{n+1} = \Psi(f, g, X_n, h, \xi_n), \tag{1.2}$$

where $\Psi(f, g, \cdot, h, \xi_n) : \mathbb{R}^d \to \mathbb{R}^d$, $X_n \in \mathbb{R}^d$ for $n \geq 0$, h denotes the timestep size, and ξ_n denotes a random vector. Of interest in this paper is the approximation of $\mathbb{E}(\phi(X(\tau)))$, where $\phi : \mathbb{R}^d \to \mathbb{R}$ is a smooth function, by $\mathbb{E}(\phi(X_N))$, $N = \tau/h$. For a practical computation, $\mathbb{E}(\phi(X_N))$ is further approximated by a Monte-Carlo method [22]. The efficiency of this later approximation, which is not addressed in the present

*ANMC, Mathematics Section, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland, Assyr.Abdulle@epfl.ch,

†Institut für Angewandte und Numerische Mathematik, Karlsruher Institut für Technologie, DE-76128 Karlsruhe, Germany, david.cohen@kit.edu ,

‡ANMC, Mathematics Section, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland. new address: École Normale Supérieure de Cachan, Antenne de Bretagne, av. Robert Schuman, F-35170 Bruz, France, Gilles.Vilmart@bretagne.ens-cachan.fr,

§ANMC Mathematics Section, École Polytechnique Fédérale de Lausanne, CH-1015 Lausanne, Switzerland. School of Mathematics, University of Southampton, Southampton SO17 1BJ, UK, k.zygalakis@soton.ac.uk

paper, is very important in practice and is still an active research topic. In particular, the methods developed in this paper could be combined with the recently proposed Multilevel Monte-Carlo method [16].

The accuracy of the approximation can be measured by the weak order of convergence of the numerical method. We recall that a numerical approximation (1.2), starting from the exact initial condition X_0 of (1.1) is said to have weak order p if for $\tau > 0$, we have

$$|\mathbb{E}(\phi(X_N)) - \mathbb{E}(\phi(X(t_N)))| \leq Ch^p, \tag{1.3}$$

for any fixed $t_N = Nh \in [0, \tau]$, for all h sufficiently small, and all functions $\phi : \mathbb{R}^d \to \mathbb{R} \in C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R})$, with a constant C independent of h. Here and in what follows, $C_P^\ell(\mathbb{R}^d, \mathbb{R})$ denotes the space of ℓ times continuously differentiable functions $\mathbb{R}^d \to \mathbb{R}$ with all partial derivatives with polynomial growth.

Remark 1.1. *A well-known theorem of Milstein [30] allows to infer the weak order from the error after one step. Assuming that f, g are Lipschitz continuous and satisfy $f \in C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R}^d)$, $g \in C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R}^{d\times m})$ that the moments of the exact solution of the SDE* (1.1) *exist and are bounded (up to a sufficiently high order) and that $\phi \in C_P^{2(p+1)}(\mathbb{R}^d, \mathbb{R})$, then, the local error bound*

$$|\mathbb{E}(\phi(X_1)) - \mathbb{E}(\phi(X(t_1)))| \leq Ch^{p+1} \tag{1.4}$$

for all initial values $X(0) = X_0$ and for all h sufficiently small, implies the global error bound (1.3)*. Here the constant C is again independent of h.*

The simplest method to approximate solutions to (1.1) is the so-called Euler-Maruyama method [28], which has weak order one. In many applications, it is of interest to approximate the moments of the solution of an SDE (or expectations of functionals of it) with a better accuracy. The construction of higher order schemes has been pursued by many authors. Classical approaches for getting high weak order numerical schemes for stochastic differential equations are based on weak Taylor approximation or Runge-Kutta type methods [8, 12]. For example, weak second order methods were proposed by Milstein [29, 30], Platen [36], Mackevicius [27], Talay [43] (see also [22, 32]) and Tocino and Vigo-Aguiar [45]. We mention also the extrapolation methods of Talay and Tubaro [44] and of [23] that combines methods with different stepsizes to achieve higher weak order convergence.

In this paper we propose yet another approach inspired by the construction of high order numerical integrators for deterministic problems proposed in [10] and the newly developed theory of modified equations for stochastic differential equations [14, 46, 41]. The basic idea of our new approach can be summarized as follows. Instead of applying the numerical method (1.2) to the SDE (1.1), we apply it to a suitably modified differential equation (a perturbation of (1.1)) so that the resulting numerical scheme yields a higher order approximation of the original SDE. This permits to fulfill automatically the order conditions, which can be very numerous for SDEs (for instance, 59 weak order two conditions have been listed for a class of stochastic Runge-Kutta type methods in [[39], Thm. 5.1]). We present a criterion (see Theorem 2.1) to construct weak methods of arbitrary order. Classical methods (Milstein or Talay methods) can be derived in a new way with our methodology. New methods will also be derived.

As an example, we propose a weak second order mean-square stable method suitable for the integration of so-called stiff problems. By stiff stochastic problems, we refer to mean-square stable problems with multiple scales for which classical explicit methods face a severe step size restriction. We also show how the methodology can be used to construct high weak order methods for random mechanical problems. In particular, we derive new weak second order methods preserving exactly all quadratic first integrals of the underlying SDE. As an illustration, we study the stochastic rigid body problem.

The paper is organized as follows. In Section 2, we present our new methodology and give a criterion for the construction of high weak order methods. In Section 3, we give explicit constructions of weak second order methods with emphasis on the numerical integration of stiff problems and random mechanical problems. Numerical examples illustrate the behavior of our new methods and corroborate the claimed weak orders of convergence.

2. Integrators based on modified equations. The general idea of constructing high order integrators based on modifying equation for SDEs can be summarized as follows. Consider a numerical method (1.2) for problem (1.1) and assume that its weak order of convergence (1.3) is $p \geq 1$. We then consider (1.1) with suitably modified drift and diffusion functions

$$d\widetilde{X} = f_h(\widetilde{X})dt + g_h(\widetilde{X})dW(t), \quad \widetilde{X}(0) = X_0, \tag{2.1}$$

where [1]

$$f_h(x) = f(x) + hf_1(x) + h^2 f_2(x) + \dots, \tag{2.2}$$
$$g_h(x) = g(x) + hg_1(x) + h^2 g_2(x) + \dots, \tag{2.3}$$

and apply the numerical method (1.2) to (2.1), i.e.,

$$\widetilde{X}_{n+1} = \Psi(f_h, g_h, \widetilde{X}_n, h, \xi_n).$$

The goal is to choose f_h, g_h in such a way that $(\widetilde{X}_n)_{n\geq 0}$ is a better weak approximation to the solution of the original SDE (1.1), i.e.,

$$|\mathbb{E}(\phi(\widetilde{X}_N)) - \mathbb{E}(\phi(X(t_N)))| \leq Ch^{p+r},$$

with $r > 0$.

Remark 2.1. *The above procedure should not be confused with a procedure called backward error analysis for SDEs [14, 46, 41] developed to study the long time behavior of numerical methods for SDEs. There, one tries to find a modified equation*

$$d\hat{X} = a_h(\hat{X})dt + b_h(\hat{X})dW(t), \quad \hat{X}(0) = X_0, \tag{2.4}$$

such that its exact solution is closer to the numerical solution (1.2)*, i.e.,*

$$|\mathbb{E}(\phi(X_N)) - \mathbb{E}(\phi(\hat{X}(t_N)))| \leq Ch^{p+q},$$

with $q > 0$. In general, the modified SDEs (2.4) *and* (2.1) *are different (see Remark 2.3 below).*

A natural way of looking at expectations of functionals of diffusion processes is by using the backward Kolmogorov equation associated to (1.1), which is the (deterministic) partial differential equation

$$\frac{\partial u}{\partial t} = \mathcal{L}u, \quad u(x,0) = \phi(x), \tag{2.5}$$

[1]Here, h is the timestep size of the numerical method (1.2).

where $\phi : \mathbb{R}^d \to \mathbb{R}$ is a smooth function, and the differential operator $\mathcal{L}$, called the generator of the SDE (1.1), is given by

$$\mathcal{L} := f \cdot \nabla_x + \frac{1}{2}(gg^T) : \nabla_x^2. \tag{2.6}$$

In (2.6), ∇_x and ∇_x^2 denote respectively the gradient and the Hessian matrix operator [2] with respect to x. In the case $m = d = 1$, the generator reduces to

$$\mathcal{L} = f\frac{\partial}{\partial x} + \frac{1}{2}g^2\frac{\partial^2}{\partial x^2}.$$

The probabilistic interpretation (see for example [33, 34, 38]) of the solution $u = u^{f,g}(\phi, x, t)$ to (2.5) is that

$$u^{f,g}(\phi, x, t) = \mathbb{E}\left(\phi(X(t))|X(0) = x\right),$$

where $X(t)$ solves (1.1). For the rest of the paper we assume for simplicity that the initial condition $X(0)$ is deterministic. We emphasize here that the results are still valid with random initial conditions, provided obvious notational changes. Using (2.5) one can easily derive the following formal Taylor expansion [14, 46]

$$u^{f,g}(\phi, x, h) - \phi(x) = \sum_{j=1}^{\infty}\frac{h^j}{j!}\mathcal{L}^j\phi(x).$$

Under appropriate smoothness assumptions on f, g and ϕ one can prove that

$$u^{f,g}(\phi, x, h) - \phi(x) = \sum_{j=1}^{k}\frac{h^j}{j!}\mathcal{L}^j\phi(x) + \mathcal{O}(h^{k+1}), \tag{2.7}$$

for all integer k. By defining

$$U^{f,g}(\phi, x, h) = \mathbb{E}\left(\phi\left(\Psi(f, g, X_0, h, \xi_0)|X_0 = x\right)\right), \tag{2.8}$$

for the numerical integrator (1.2), we see that the local weak error of the numerical integrator applied to (1.1) after one step is given by

$$\mathbb{E}(\phi(X_1)) - \mathbb{E}(\phi(X(t_1))) = U^{f,g}(\phi, x, h) - u^{f,g}(\phi, x, h). \tag{2.9}$$

Notice that (2.9) is the reformulation of the left-hand side of the local error bound (1.4) in terms of the solution of the backward Kolmogorov equation (2.5) associated to (1.1). Motivated by an expansion of (2.8) in Taylor series, we assume

Assumption 2.1. *The numerical solution* (2.8) *has the following expansion*

$$U^{f,g}(\phi, x, h) = \phi(x) + hA_0(f, g)\phi(x) + h^2A_1(f, g)\phi(x) + \ldots, \tag{2.10}$$

where $A_i(f, g),\ i = 0, 1, 2, \ldots$ are linear differential operators depending on the drift and diffusion functions of the SDE to which the numerical integrator is applied to. We further assume that these differential operators $A_i(f, g),\ i = 0, 1, 2, \ldots$ satisfy for all $f, \hat{f}, g, \hat{g}$ and $\varepsilon \to 0$,

$$A_i(f + \varepsilon\hat{f}, g + \varepsilon\hat{g}) = A_i(f, g) + \varepsilon\hat{A}_i(f, \hat{f}, g, \hat{g}) + \mathcal{O}(\varepsilon^2),$$

[2] Here, we consider the usual scalar product on matrices defined by $A : B = \mathrm{trace}(A^TB)$.

where $\hat{A}_i(f,\hat{f},g,\hat{g}),\ i=0,1,2,\dots$ *are again differential operators.*

The above smoothness assumption is usually satisfied by numerical integrators. For the expansion in integer powers of the stepsize h, special care has to be taken, as explained in the following remark.

Remark 2.2. *The assumption that the expansion* (2.10) *holds with integer powers of the timestep* h *is essential to avoid non integer powers of* h *in the modified equation* (2.1). *For instance, let us consider the scalar* θ*-Milstein method*

$$X_{n+1}=X_n+(1-\theta)hf(X_n)+\theta hf(X_{n+1})+g(X_n)\Delta W_n+\frac{1}{2}g'(X_n)g(X_n)((\Delta W_n)^2-h),$$

where ΔW_n *are independent* $\mathcal{N}(0,h)$ *distributed random variables and* θ *is a fixed parameter. More details on this scheme will be given in Section 3. The assumption* (2.10) *is not satisfied if one uses the Platen [22] approximation*

$$\frac{1}{2}g'(X_n)g(X_n)\approx\frac{1}{2\sqrt{h}}\Big(g(X_n+\sqrt{h}g(X_n))-g(X_n)\Big),$$

for approaching the derivative of the diffusion function in the θ*-Milstein method, because* (2.10) *would contain a term of size* $\mathcal{O}(h^{5/2})$. *However, if one considers instead the approximation used by Rößler [39] in which the noise part is evaluated as*

$$\frac{1}{2}g'(X_n)g(X_n)\approx\frac{1}{4\sqrt{h}}\Big(g(X_n+\sqrt{h}g(X_n))-g(X_n-\sqrt{h}g(X_n))\Big),$$

then, the assumption (2.10) *is satisfied. This can be checked by observing that the substitution* $\sqrt{h}\leftrightarrow-\sqrt{h}$ *leaves the definition of the method unchanged.*

Construction of modified equations. For a numerical method (1.2) with an expansion (2.10) satisfying (see (2.9))

$$U^{f,g}(\phi,x,h)-u^{f,g}(\phi,x,h)=\mathcal{O}(h^{p+1}),$$

i.e., of weak order p in view of Remark 1.1, the task now is to find a modified SDE (2.1) such that

$$U^{f_h,g_h}(\phi,x,h)-u^{f,g}(\phi,x,h)=\mathcal{O}(h^{p+r+1}),\tag{2.11}$$

i.e., a numerical method $(\widetilde{X}_n)_{n\geq 0}$ of weak order $p+r$ with $r>0$ for the original problem (1.1). A second assumption that we make on the numerical integrator is that it is consistent, i.e. of weak order at least one. This assumption implies $A_0(f,g)\phi=\mathcal{L}\phi$ and $A_0(f_h,g_h)\phi=\widetilde{\mathcal{L}}\phi$, where

$$\widetilde{\mathcal{L}}\phi:=f_h\cdot\nabla_x\phi+\frac{1}{2}(g_hg_h^T):\nabla_x^2\phi,\tag{2.12}$$

for all function ϕ. Substituting f_h,g_h given by (2.2),(2.3), respectively, in (2.12) yields the following expansion for $\widetilde{\mathcal{L}}$

$$\widetilde{\mathcal{L}}=\mathcal{L}+h\mathcal{L}_1+h^2\mathcal{L}_2+\dots,$$

where for $j=0,1,2,\dots$, $\mathcal{L}_j$ is given by

$$\mathcal{L}_j=f_j\cdot\nabla_x+\frac{1}{2}\sum_{k=0}^{j}(g_kg_{j-k}^T):\nabla_x^2,\tag{2.13}$$

where we used the notations $f_0 := f$ and $g_0 := g$. We will also sometimes write $\mathcal{L}_j = \mathcal{L}_j(f_j, g, g_1, \ldots, g_j)$ to emphasize the dependence of those operators on the functions f_j,g,g_1,..., g_j.

We may now state in this section the main result of this paper. We show that under suitable assumptions, the weak order p of the numerical integrator (1.2) can be increased to $p + r$ with $r \geq 1$ by applying it to a suitably modified SDE (2.1), with modified drift and diffusion of the form

$$f_{h,s}(x) = f(x) + hf_1(x) + \ldots + h^s f_s(x), \tag{2.14}$$

$$g_{h,s}(x) = g(x) + hg_1(x) + \ldots + h^s g_s(x), \tag{2.15}$$

where $s = p + r - 1$. The integrator with improved weak order r can be written as

$$\widetilde{X}_{n+1} = \Psi(f_{h,p+r-1}, g_{h,p+r-1}, \widetilde{X}_n, h, \xi_n). \tag{2.16}$$

THEOREM 2.1. *Assume that the numerical scheme* (1.2) *has order* $p \geq 1$ *and that Assumption 2.1 holds. Let* $r \geq 1$ *and assume that the functions* f_j *and* g_j *for* $j = 1, \ldots, p{+}r{-}2$ *have been constructed such that* $\widetilde{X}_{n+1} = \Psi(f_{h,p+r-2}, g_{h,p+r-2}, \widetilde{X}_n, h, \xi_n)$ *has weak order* $p + r - 1$*. Consider the differential operator defined as*

$$\mathcal{L}_{p+r-1}\phi := \lim_{h\to 0} \frac{u^{f,g}(\phi, x, h) - U^{f_{h,p+r-1}, g_{h,p+r-1}}(\phi, x, h)}{h^{p+r}}, \tag{2.17}$$

where $u^{f,g}(\phi, x, h)$ *is expanded in* (2.7) *and* $U^{f,g}(\phi, x, h)$ *is defined in* (2.8)*. If there exist functions* $f_{p+r-1} : \mathbb{R}^d \to \mathbb{R}^d$ *and* $g_{p+r-1} : \mathbb{R}^d \to \mathbb{R}^{d\times m}$ *such that the differential operator* (2.17) *can be written in the form* (2.13)*, then the numerical integrator* (2.16) *applied to the SDE with the modified drift and diffusion* (2.14),(2.15) *has weak order of accuracy* $p + r$ *for the original system of SDEs* (1.1) *provided* $f_{h,p+r-1} \in C_P^{2(p+r+1)}(\mathbb{R}^d, \mathbb{R}^d)$ $g_{h,p+r-1} \in C_P^{2(p+r+1)}(\mathbb{R}^d, \mathbb{R}^{d\times m})$*. The error bound*

$$|\mathbb{E}(\phi(\widetilde{X}_N)) - \mathbb{E}(\phi(X(t_N)))| \leq Ch^{p+r},$$

holds for any fixed $t_N = Nh \in [0, \tau]$ *with* h *sufficiently small and for all functions* $\phi \in C_P^{2(p+r+1)}(\mathbb{R}^d, \mathbb{R})$.

Proof. By induction hypothesis, $\widetilde{X}_{n+1} = \Psi(f_{h,p+r-2}, g_{h,p+r-2}, \widetilde{X}_n, h, \xi_n)$ is an integrator of weak order $p + r - 1$. Thus, it has a weak expansion of the form

$$\begin{aligned}
U^{f_{h,p+r-2}, g_{h,p+r-2}}(\phi, x, h) &= \phi(x) + hA_0(f_{h,p+r-2}, g_{h,p+r-2})\phi(x) + \ldots \\
&\quad + h^{p+r} A_{p+r-1}(f_{h,p+r-2}, g_{h,p+r-2})\phi(x) + \mathcal{O}(h^{p+r+1}) \\
&= \phi(x) + h\mathcal{L}(f, g)\phi(x) + \ldots + \frac{h^{p+r-1}}{(p + r - 1)!}\mathcal{L}^{p+r-1}(f, g)\phi(x) \\
&\quad + h^{p+r} B_{p+r}(f, g)\phi(x) + \mathcal{O}(h^{p+r+1}),
\end{aligned}$$

where $B_{p+r}(f, g)$ is a certain differential operator. Using Assumption 2.1, we have the relation $A_i(f_{h,p+r-1}, g_{h,p+r-1}) = A_i(f_{h,p+r-2}, g_{h,p+r-2}) + \mathcal{O}(h^{p+r-1})$. The weak expansion of the modified integrator (2.16) can then be written as

$$\begin{aligned}
U^{f_{h,p+r-1}, g_{h,p+r-1}}(\phi, x, h) &= \phi(x) + h\mathcal{L}(f, g)\phi(x) + \ldots + \frac{h^{p+r-1}}{(p + r - 1)!}\mathcal{L}^{p+r-1}(f, g)\phi(x) \\
&\quad + h^{p+r}\big(\mathcal{L}_{p+r}(f_{p+r-1}, g, g_1, \ldots, g_{p+r-1}) + B_{p+r}(f, g)\big)\phi(x) \\
&\quad + \mathcal{O}(h^{p+r+1}),
\end{aligned}$$

where $\mathcal{L}_{p+r-1}$ is defined in (2.13). If f_{p+r-1} and g_{p+r-1} are such that

$$\mathcal{L}_{p+r-1}(f_{p+r-1}, g, g_1, \ldots, g_{p+r-1}) = \frac{\mathcal{L}^{p+r}}{(p+r)!} - B_{p+r}(f,g),$$

then (2.11) holds for $U^{f_{h,p+r-1}, g_{h,p+r-1}}(\phi, x, h)$. Now observing that the right-hand side of the above equality is equal to the right-hand side of (2.17) together with Remark 1.1 proves the theorem. ▯

Relation with backward error analysis.. We close this section by relating the previous construction of modified integrators with the backward error analysis for SDEs [14, 46, 41] mentioned in Remark 2.1. Applying the numerical integrator (1.2) to the original SDE (1.1), we search for a modified differential equation (2.4) such that

$$U^{f,g}(\phi, x, h) - u^{\hat{f}_h, \hat{g}_h}(\phi, x, h) = \mathcal{O}(h^{p+q+1}) \tag{2.18}$$

with $q > 0$. The aim in such a procedure is to better understand the behavior of the numerical method (1.2) (applied to (1.1)) by studying the modified SDE (2.4). The modified SDE (2.4), with $\hat{f}_h, \hat{g}_h$ given by an expansion

$$\hat{f}_h(x) = f(x) + h\hat{f}_1(x) + h^2\hat{f}_2(x) + \ldots,$$
$$\hat{g}_h(x) = g(x) + h\hat{g}_1(x) + h^2\hat{g}_2(x) + \ldots,$$

has an associated backward Kolmogorov equation (the formula (2.5) with $\mathcal{L}$ replaced by $\hat{\mathcal{L}}$) in (2.6), where

$$\hat{\mathcal{L}} = \hat{\mathcal{L}}_0 + h\hat{\mathcal{L}}_1 + h^2\hat{\mathcal{L}}_2 + \ldots, \tag{2.19}$$

where $\hat{\mathcal{L}}_0 = \mathcal{L}$ and for $j = 1, 2, \ldots$ $\hat{\mathcal{L}}_j$ is given by

$$\hat{\mathcal{L}}_j = \hat{f}_j \cdot \nabla_x + \frac{1}{2}\sum_{k=0}^{j} (\hat{g}_k \hat{g}_{j-k}^T) : \nabla_x^2,$$

where $\hat{g}_0 = g$. The Taylor expansion (2.7) becomes

$$u^{\hat{f}_h, \hat{g}_h}(\phi, x, h) - \phi(x) = \sum_{j=1}^{k} \frac{h^j}{j!} \hat{\mathcal{L}}^j \phi(x) + \mathcal{O}(h^{k+1}),$$

which gives in terms of the expansion (2.19) (see [46])

$$u^{\hat{f}_h, \hat{g}_h}(\phi, x, h) - \phi(x) = \sum_{j=1}^{k} h^j \sum_{i_1+i_2+\ldots+i_l+l=j} \frac{1}{l!} (\hat{\mathcal{L}}_{i_1} \cdots \hat{\mathcal{L}}_{i_l} \phi)(x) + \mathcal{O}(h^{k+1}).$$

The task in this approach is to find $\hat{f}_h, \hat{g}_h$ such that for $U^{f,g}(\phi, x, h)$ given by (2.10) holds

$$A_j(f,g) = \sum_{i_1+i_2+\ldots+i_l+l=j} \frac{1}{l!} \hat{\mathcal{L}}_{i_1} \cdots \hat{\mathcal{L}}_{i_l},$$

for $j = p + q$. The above relation permits to define by induction the differential operators $\hat{\mathcal{L}}_j$ used to construct the modified equation for backward error analysis. We emphasize once more that the aim and the theory for integrators based on modified equations and backward error analysis are different. In the former approach, the modified SDE constitutes only a surrogate to obtain a better numerical approximation of the solution of the original SDE, in the latter approach, the modified SDE is a tool to study a numerical integrator applied to the original SDE.

Remark 2.3. *In the case $p = r = 1$, the above procedure yields for backward error analysis and for modified integrators the operators $\hat{\mathcal{L}}_1 = A_1 - \frac{1}{2}\mathcal{L}^2$ and $\mathcal{L}_1 = \frac{1}{2}\mathcal{L}^2 - A_1$, respectively. Thus, the perturbations $\hat{f}_1, \hat{g}_1$ in the modified equations for backward error analysis and f_1, g_1 for modified integrators are identical up to the multiplicative factor -1.*

3. High weak order methods with application to stiff problems and geometric integration. In this section we show two applications of the methodology developed in Section 2. We first derive a class of weak second order methods based on first order methods. Classical methods (Milstein or Talay methods) will be recovered, but new methods will also be derived. In particular, we derive a new weak second order method which is mean-square stable, suitable for the integration of so-called stiff problems. This method belongs to a general class of weak second order methods derived by Milstein [30], but seems not to have appeared explicitly in the literature. Secondly, we show how our methodology can be applied to structure preserving integrators and derive weak second order methods preserving quadratic invariants. As an example, we consider the stochastic rigid body problem.

3.1. Weak second order methods with application to stiff stochastic problems. To illustrate our methodology based on modifying equations, we derive here a family of weak second order methods. For that, we pick a weak first order method

$$X_1 = \Psi(f, g, X_0, h, \xi_0),$$

consider the modified equation

$$dX = [f(X) + hf_1(X)]\,dt + [g(X) + hg_1(X)]\,dW(t), \quad X(0) = X_0, \tag{3.1}$$

and apply Theorem 2.1. Accordingly, we have to find f_1, g_1 such that

$$\mathcal{L}_1 = \frac{\mathcal{L}^2}{2} - A_1(f, g), \tag{3.2}$$

where $\mathcal{L}_1 := f_1\nabla_x + \frac{1}{2}(gg_1^T + g_1g^T) : \nabla_x^2$, and where the differential operator A_1 depends on the choice of the weak first order method.

3.1.1. One-dimensional case. For the sake of simplicity let us first consider a one dimensional SDE with one dimensional noise. The simplest weak first order method is the Euler-Maruyama method. However, for reasons explained in Remark 3.2 below, this is not a suitable method to start with. A fairly general class of weak first order methods that can be used for our purpose is the θ-Milstein method [18] (that we denote θ-M method in what follows)

$$X_{n+1} = X_n + (1-\theta)hf(X_n) + \theta hf(X_{n+1}) + g(X_n)\Delta W_n + \frac{1}{2}g'(X_n)g(X_n)((\Delta W_n)^2 - h), \tag{3.3}$$

where ΔW_n are independent $\mathcal{N}(0, h)$ distributed random variables and $X_0 = x$. For $\theta = 0$, (3.3) reduces to the classical explicit Milstein method, while for $\theta > 0$ it yields semi-implicit methods (implicit in the drift function and explicit in the diffusion function). We expand in Taylor series the function ϕ up to the 4th order, $\phi(X_1) = \sum_{i=0}^{4} \frac{1}{i!}\phi^{(i)}(x)F^i + ...$, where

$$F = (1-\theta)hf(x) + \theta hf(X_1) + g(x)\Delta W_0 + \frac{1}{2}g'(x)g(x)((\Delta W_0)^2 - h),$$

and obtain [3]

$$U^{f,g}(\phi, x, h) = \mathbb{E}(\phi(X_1)|X_0 = x) = \phi(x) + h\mathcal{L}\phi(x) + h^2 A_1(f,g)\phi(x) + \mathcal{O}(h^3),$$

where

$$\begin{aligned} A_1(f,g)\phi(x) = {} & \theta\left[f'(x)f(x) + \frac{1}{2}f''(x)g^2(x)\right]\phi'(x) \\ & + \frac{1}{2}\left[f^2(x) + 2\theta f'(x)g^2(x) + \frac{1}{2}(g'(x)g(x))^2\right]\phi''(x) \\ & + \frac{1}{2}\left[g'(x)g^3(x) + g^2 f(x)\right]\phi'''(x) + \frac{h^2}{8}g^4(x)\phi^{(4)}(x). \end{aligned}$$

Applying the method (3.3) to the modified equation (3.1) we obtain $U^{f_h,g_h}(\phi, x, h)$ which is a second order approximation of $u^{f,g}(\phi, x, h)$ if we can find f_1, g_1 such that (3.2) holds. A simple computation reveals that

$$\begin{aligned} \left(\frac{1}{2}\mathcal{L}^2\phi - A_1(f,g)\phi\right)(x) = {} & \left(\frac{1}{2}-\theta\right)\left(f'(x)f(x) + \frac{1}{2}f''(x)g^2(x)\right)\phi'(x) \\ & + \left(\left(\frac{1}{2}-\theta\right)f'(x)g(x) + \frac{1}{2}g'(x)f(x) + \frac{1}{4}g^2(x)g''(x)\right)g(x)\phi''(x). \end{aligned}$$

We see from the above formula that we can define the appropriate operator $\mathcal{L}_1$ with

$$\begin{aligned} f_1(x) &= \left(\frac{1}{2}-\theta\right)f'(x)f(x) + \frac{1}{2}\left(\frac{1}{2}-\theta\right)f''(x)g^2(x), \\ g_1(x) &= \left(\frac{1}{2}-\theta\right)f'(x)g(x) + \frac{1}{2}g'(x)f(x) + \frac{1}{4}g^2(x)g''(x). \end{aligned}$$

Now setting $f_{h,1} = f + hf_1$ and $g_{h,1} = g + hg_1$, we obtain the following new integrator

$$\begin{aligned} X_{n+1} = {} & X_n + (1-\theta)hf_{h,1}(X_n) + \theta hf_{h,1}(X_{n+1}) + g_{h,1}(X_n)\Delta W_n \\ & + \frac{1}{2}g'(X_n)g(X_n)((\Delta W_n)^2 - h), \end{aligned} \tag{3.4}$$

which has weak order two for the SDE (1.1) in dimension one.

Remark 3.1. *In principle one should also replace g by $g_{h,1}$ in the last term of (3.4), but omitting hg_1 for this term does not affect the weak order two of accuracy. Indeed,*

$$g'_{h,1}(X_n)g_{h,1}(X_n) = g'(X_n)g(X_n) + C(X_n)h + \mathcal{O}(h^2),$$

where $C(x)$ is a smooth function. Using $\mathbb{E}(C(X_n)h((\Delta W_n)^2 - h)) = 0$, we deduce

$$\mathbb{E}\left(g'_{h,1}(X_n)g_{h,1}(X_n)((\Delta W_n)^2 - h)\right) = \mathbb{E}(g'(X_n)g(X_n)((\Delta W_n)^2 - h)) + \mathcal{O}(h^3),$$

[3] Recall that $\mathbb{E}(\Delta W_0) = \mathbb{E}(\Delta W_0^3) = 0$ and $\mathbb{E}(\Delta W_0^2) = h, \mathbb{E}(\Delta W_0^4) = 3h^2$.

and thus it does not influence the accuracy of the method because it induces a $\mathcal{O}(h^3)$ *perturbation of* $\mathbb{E}(\phi(X_{n+1}))$. Notice that the integrator (3.4) belongs to a sub-class of a general family of weak second order methods introduced by Milstein [30]. For $\theta = 0$ it has also been considered by Talay who proved its order of convergence [43]. For $\theta = 1/2$ the method was considered by Milstein who showed its good stability behavior for scalar SDEs with additive noise. For $\theta = 1$, the method does not seem to have appeared explicitly in the literature. We will show below that it has favorable stability properties for scalar SDEs with multiplicative noise (mean-square stability).

Remark 3.2. *Notice that* $\mathcal{L}^2$ *is a differential operator of order four in general. Thus, the difference* $\frac{1}{2}\mathcal{L}^2 - A_1(f,g)$ *is a differential operator of order two of the same form as* $\mathcal{L}$ *only if* $A_1(f,g)$ *contains the same third and fourth order derivatives of* ϕ *as* $\frac{1}{2}\mathcal{L}^2$. *As explained further, this is true for the Milstein method. However, this would not be the case for the Euler-Maruyama method where the term* $\frac{1}{2}g'(x)g(x)^3\phi'''(x)$ *involving the third derivative of* ϕ *in* $\frac{1}{2}\mathcal{L}^2\phi$ *is not cancelled in general (unless* $g' = 0$, *i.e. for additive noise). Therefore, as observed in [46], a modified SDE cannot be constructed for the Euler-Maruyama method.*

3.1.2. Multi-dimensional case. The formula derived for the one-dimensional case can easily be extended to the multi-dimensional case. Consider the multi-dimensional SDE (1.1), where f is a column vector of size d and g is a $d \times m$ matrix (below, we denote by $\cdot_{[i]}$ the ith component of a vector in $\mathbb{R}^d$ and by $\cdot_{[i,j]}$ the coefficients of a $d \times m$ matrix). For a fixed parameter θ, consider the θ-M method

$$X_{n+1} = X_n + (1-\theta)hf(X_n) + \theta hf(X_{n+1}) + g(X_n)\Delta W_n + M(X_n, W), \qquad (3.5)$$

where the Milstein term $M(X_n, W)$ is defined for $i = 1, \dots, d$ by

$$M_{[i]} = \Xi_i(X_n) : I = \sum_{j_1,j_2=1}^{m} \Xi_{i[j_1,j_2]} I_{[j_1,j_2]}.$$

The coefficients of the $m \times m$ matrix Ξ_i are defined for $i = 1, \dots, d$ by

$$\Xi_{i[j_1,j_2]} = \sum_{k=1}^{d} \frac{\partial g_{[i,j_2]}}{\partial x_k} g_{[k,j_1]},$$

and the coefficients of the $m \times m$ matrix I of multiple integrals are given by

$$I_{[j_1,j_2]} = \int_{t_n}^{t_{n+1}} \left(\int_{t_n}^{s} dW_{j_1}(t) \right) dW_{j_2}(s). \qquad (3.6)$$

Following the procedure for the one-dimensional case, we obtain the modified θ-M method of weak order two

$$X_{n+1} = X_n + (1-\theta)hf_{h,1}(X_n) + \theta hf_{h,1}(X_{n+1}) + g_{h,1}\Delta W_n + M(X_n, W), \qquad (3.7)$$

where $f_{h,1} = f + hf_1$, $g_{h,1} = g + hg_1$, and f_1, g_1 are given (componentwise) by

$$\begin{aligned} f_{1,[i]} &= \left(\frac{1}{2} - \theta\right)(f'f)_{[i]} + \frac{1}{2}\left(\frac{1}{2} - \theta\right) gg^T : f''_{[i]}, \\ g_{1,[i,j]} &= \left(\frac{1}{2} - \theta\right)(f'g)_{[i,j]} + \frac{1}{2}g'_{[i,j]}f + \frac{1}{4}gg^T : g''_{[i,j]}, \end{aligned} \qquad (3.8)$$

for all $i = 1, \ldots, d$ and $j = 1, \ldots, m$. We observe that the above method contains derivatives of the drift and diffusion functions. This is a general feature of the methods obtained using modified equations. In some cases, these derivatives are easy and cheap to compute (see for example the stochastic mechanical problem in Section 3.2.1). In general, these derivatives can be approximated. In particular, one can use formulas based on finite differences. Some care is however required for an efficient implementation (i.e., a low number of function evaluations in dependence on the number of Wiener processes [13]).

Remark 3.3. *The multiple integral matrix I in* (3.6) *is difficult to evaluate in general and needs to be approximated. One can use for instance the following weak approximation for the matrix I in the definition* (3.6) *(see for instance [22, eq. (5.12.9)])*

$$J = \frac{1}{2}\big(\Delta W_n \Delta W_n^T + E_n\big),$$

where E_n is a random skew-symmetric matrix whose coefficients $E_{n,[j_1,j_2]}$ are independent two-point distributed random variables,

$$\mathbb{P}(E_{n,[j_1,j_2]} = \pm h) = 1/2, \quad \text{for all} \quad 1 \le j_1 < j_2 \le d$$

and $E_{n,[j_1,j_2]} = -E_{n,[j_2,j_1]}$ for all $j_1, j_2 = 1, \ldots, m$. Using J instead of I does not alter the weak order two of accuracy of the method (3.7) *(it does however decrease the strong order of the method from* 1 *to* 1/2*). The independent Gaussian variables $\Delta W_{n,[j]}$ can also be replaced by independent three point random variables with*

$$\mathbb{P}(\Delta W_{n,[j]} = \pm\sqrt{3h}) = \frac{1}{6}, \quad \mathbb{P}(\Delta W_{n,[j]} = 0) = \frac{2}{3}, \tag{3.9}$$

without decreasing the weak order two of the method.

3.1.3. A mean-square stable modified θ-M method. In this section we show that we can construct a second order modified θ-M method with favorable mean-square stability. To study the stability in the mean-square sense of numerical integrators, a widely used test equation introduced in [40] for SDEs is the following scalar SDE with multiplicative noise

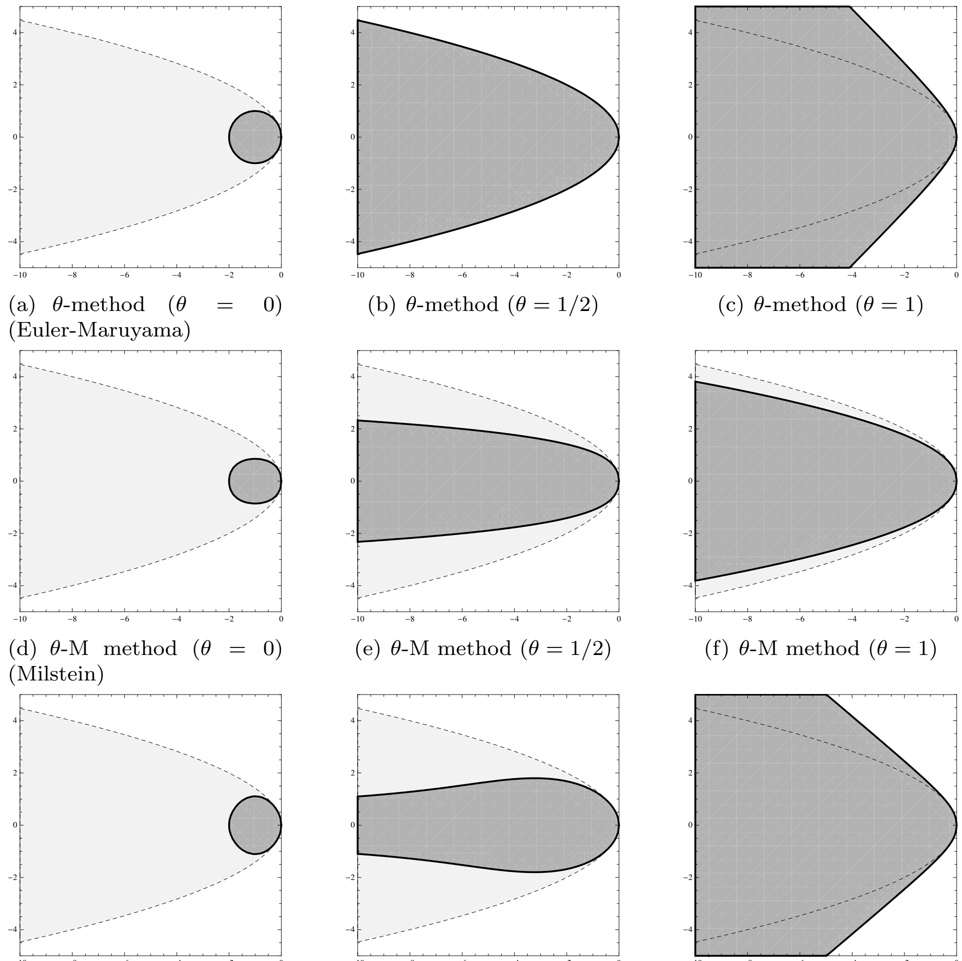
$$dX = \lambda X dt + \mu X dW(t), \tag{3.10}$$

where the parameters $\lambda, \mu \in \mathbb{C}$. We notice that other test equations have been considered recently in [6, 7], to better account for the stability behavior of numerical integrators when applied to systems of SDEs or scalar equations with several multiplicative noise terms. The mean-square stability domain of (3.10) is given by

$$\mathcal{S} = \{(\lambda, \mu) \in \mathbb{C}^2 \,; \Re\lambda + \frac{1}{2}|\mu|^2 < 0\}. \tag{3.11}$$

The set of (λ, μ) that fulfill condition (3.11) can be visualized, for $\lambda, \mu \in \mathbb{R}$, as the shaded area with a boundary given by the dotted parabolas in Figure 3.1. The θ-M method applied to the linear test equation (3.10) yields

$$X_{n+1} = \frac{(1 + p(1-\theta) + qV_n + \frac{1}{2}q^2(V_n^2 - 1))}{1 - \theta p} X_n,$$



(a) θ-method ($\theta = 0$) (Euler-Maruyama) (b) θ-method ($\theta = 1/2$) (c) θ-method ($\theta = 1$)

(d) θ-M method ($\theta = 0$) (Milstein) (e) θ-M method ($\theta = 1/2$) (f) θ-M method ($\theta = 1$)

(g) modified θ-M ($\theta = 0$) (Milstein-Talay) (h) modified θ-M ($\theta = 1/2$) (Milstein) (i) modified θ-M ($\theta = 1$)

FIG. 3.1. *Comparison of mean-square stability domains (dark gray area) of the θ-methods, the θ-M methods, and the modified θ-M methods. The modified θ-M methods with $\theta = 0, 1/2$ have been proposed by Milstein and Talay. Horizontal axis: p, vertical axis: q.*

where V_n are independent Gaussian variables with a $\mathcal{N}(0,1)$ distribution and $p = \lambda h$ and $q = \mu\sqrt{h}$. Squaring the result and taking the expectation, we obtain the relation $\mathbb{E}(|X_{n+1}|^2) = R_{\theta,M}(p,q)\mathbb{E}(|X_n|^2)$, where

$$R_{\theta,M}(p,q) = \frac{|1 + p(1-\theta)|^2 + |q|^2 + |q|^4/2}{|1-\theta p|^2}. \tag{3.12}$$

We next define the set

$$\mathcal{S}_{\theta,M} = \{(p,q) \in \mathbb{C}^2 \; ; R_{\theta,M}(p,q) < 1\}.$$

The method is called mean-square (MS) stable if

$$R_{\theta,M}(p,q) \leq 1, \qquad \text{for all } (p,q) \in \mathcal{S},$$

or alternatively if $\mathcal{S} \subset \mathcal{S}_{\theta,M}$. It is readily seen that there does not exist a value of $\theta \in [0,1]$ such that the θ-M method is MS stable. Furthermore, MS stability is recovered for $\theta = 3/2$ [18].

Remark 3.4. *In contrast, the θ-methods (the methods* (3.5) *with $M \equiv 0$), whose stability function reads*

$$R_\theta(p,q) = \frac{|1+p(1-\theta)|^2 + |q|^2}{|1-\theta p|^2},$$

can be shown to be MS stable if and only if $\theta \geq 1/2$ as reported in [19]. We come now to study the stability properties of the modified θ-M methods (3.4) whose stability functions can be easily deduced from (3.12) and read

$$\widetilde{R}_{\theta,M}(p,q) = \frac{|1+\widetilde{p}(1-\theta)|^2 + |\widetilde{q}|^2 + |q|^4/2}{|1-\theta\widetilde{p}|^2},$$

where $\widetilde{p} = p + (\frac{1}{2} - \theta)p^2$, $\widetilde{q} = q + (1-\theta)pq$. A simple calculation shows that this method is MS stable if and only if $\theta = 1$. Thus, for $\theta = 1$ we have constructed the weak second order method (3.4) which is MS stable. This method is thus suitable for the numerical integration of stiff systems of SDEs as illustrated in the numerical example below.

In Figure 3.1 we plot the mean-square stability domain for the standard θ-methods, the standard θ-M methods and the modified θ-M methods (the light-dark region which lies inside the dotted parabola is the stability domain $\mathcal{S}$ of the exact solution of the test problem). For $\theta \in [0,1]$, it can be seen that the θ-M methods are never MS stable, and that only for $\theta = 1$ is the modified θ-M method MS stable.

Numerical experiments. We illustrate the numerical behavior of the modified θ-M methods previously constructed. We consider an economy model for asset prices proposed in [21], see also [20]. It is an Itô system of SDEs in dimension $d = 3$, with $m = 2$ non-commutative noises, given by

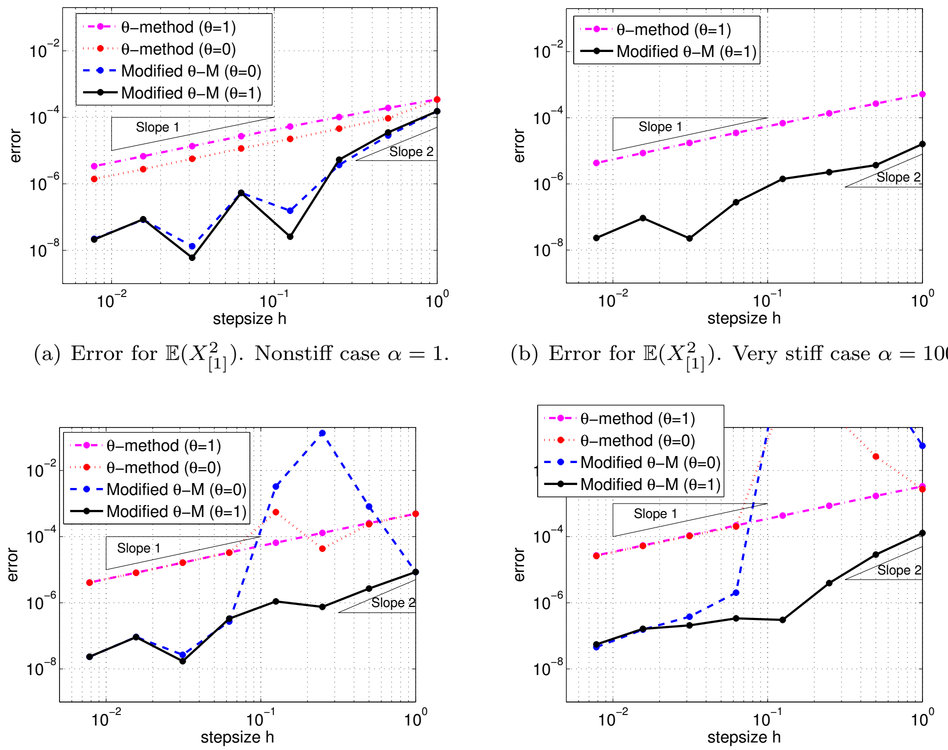
$$\begin{aligned} dX_{[1]} &= \beta_1 X_{[1]} X_{[2]} dW_{[1]}(t), \\ dX_{[2]} &= -(X_{[2]} - X_{[3]})dt + \beta_2 X_{[2]} dW_{[2]}(t), \\ dX_{[3]} &= \alpha(X_{[2]} - X_{[3]})dt, \end{aligned} \tag{3.13}$$

where $X_{[1]}(t), X_{[2]}(t)$ and $X_{[3]}(t)$ represent the asset price, the instantaneous volatility and the average volatility, respectively. We take the parameters $\beta_1 = 1, \beta_2 = 0.3$, the initial value $X(0) = (X_{[1]}(0), X_{[2]}(0), X_{[3]}(0))^T = (1, 0.1, 0.1)^T$, and consider the time interval $[0,1]$ as in [20]. The parameter $\alpha > 0$ corresponds to the volatility parameter, and is related to the strength of the past dependence of the average volatility. We refer to [21] for details in the context of economy modeling. For large values of α, the largest eigenvalue in modulus of the Jacobian of the drift function is $-|\alpha|$ and the SDE problem (3.13) becomes stiff.

We shall consider various values of the volatility parameter α in the numerical experiments.

Since the drift vector field in (3.13) is linear, the modified θ-M methods (3.14) for (3.13) are linearly implicit and using the formulas (3.7) and (3.8), it can be written as

$$(Id - \theta hA)X_{n+1} = (Id + (1-\theta)hA)X_n + g_{h,1}(X_n)\Delta W_n + M(X_n, \Delta W_n), \tag{3.14}$$



(a) Error for $\mathbb{E}(X_{[1]}^2)$. Nonstiff case $\alpha = 1$.

(b) Error for $\mathbb{E}(X_{[1]}^2)$. Very stiff case $\alpha = 100$.

(c) Error for $\mathbb{E}(X_{[1]}^2)$. Stiff case $\alpha = 25$.

(d) Error for $\mathbb{E}(X_{[2]}^2)$. Stiff case $\alpha = 25$.

FIG. 3.2. *Finance model* (3.13). *Comparison of weak convergence rates for the modified θ-M with $\theta = 1$ (solid lines), the θ-method with $\theta = 1$ (dashed-dotted lines), the θ-method with $\theta = 0$ (Euler-Maruyama method, dotted line), and the modified θ-M with $\theta = 0$ (Milstein-Talay method, dashed line).*

where A denotes the matrix

$$A = \left[1 - h\left(\frac{1}{2} - \theta\right)(1+\alpha)\right]\begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & \alpha & -\alpha \end{pmatrix},$$

and where

$$g_{h,1}(X)\Delta W = \begin{pmatrix} \beta_1 X_{[1]}\big(X_{[2]} + \frac{h}{2}(X_{[3]} - X_{[2]})\big)\Delta W_{[1]} \\ \beta_2\Big(\big(1 - h(\frac{1}{2} - \theta)\big)X_{[2]} + \frac{h}{2}(X_{[3]} - X_{[2]})\Big)\Delta W_{[2]} \\ \alpha\beta_2 h\big(\frac{1}{2} - \theta\big)X_{[2]}\Delta W_{[2]} \end{pmatrix},$$

$$M(X, \Delta W) = \begin{pmatrix} \frac{1}{2}\beta_1^2 X_{[1]}X_{[2]}^2(\Delta W_{[1]}^2 - h) + \beta_1\beta_2 X_{[1]}X_{[2]}I_{[2,1]} \\ \frac{1}{2}\beta_2^2 X_{[2]}(\Delta W_{[2]}^2 - h) \\ 0 \end{pmatrix}.$$

We take for the random variables $\Delta W_n = (\Delta W_{n,[1]}, \Delta W_{n,[2]})^T$ independent Gaussian variables with mean zero and variance h. Notice that similar results have been obtained when considering instead discrete random variables satisfying (3.9). The above multiple integral $I_{[2,1]}$ in the Milstein term $M(X, \Delta W)$ is approximated by $I_{[2,1]} \approx (\Delta W_{[1]}\Delta W_{[2]} + \xi_n h)/2$, where ξ_n are independent random variables satisfying $\mathbb{P}(\xi_n = \pm 1) = 1/2$ as detailed in Remark 3.3.

To confirm the weak order two of convergence of the modified θ-M method (3.7), we compute the relative errors in the quantities $\mathbb{E}(X_{[1]}^2)$ in Figures 3.2(a)-(c) and $\mathbb{E}(X_{[2]}^2)$ in Figure 3.2(d) at the final time $t = 1$ for the stepsizes $h = 2^{-i}$, $i = 0, \ldots, 7$. The reference solutions are computed using the small timestep $h = 2^{-14}$. To check carefully the accuracy of the methods up to small time steps, we need to drastically reduce the Monte-Carlo error. We thus approximate the required moments of the numerical solutions by averages over 500 millions of trajectories computed in FORTRAN, using the random number generator [35]. For a fair comparison, notice that we use the same set of random numbers for each numerical integrator. We observe in Figure 3.2 the expected lines of slope two (solid lines) both in the nonstiff case ($\alpha = 1$) and the stiff cases ($\alpha = 25$ and $\alpha = 100$). Notice that for small timesteps ($h < 0.25$) the zigzag that we observe is due to the Monte-Carlo error, which could be further reduced by increasing the number of samples. For comparison, we also plot the results for the classical semi-implicit θ-method ($\theta = 1$) (weak order one), obtained from (3.14) by removing the Milstein term $M(X_n, \Delta W_n)$ and setting $h = 0$ in the definitions of A and $g_{h,1}(X)\Delta W$. We also compare with two classical explicit integrators, the Euler-Maruyama method (weak order one) and the Talay method (weak order two), obtained by taking $\theta = 0$ in (3.14). Notice these two explicit methods are not (unconditionally) mean-square stable. Indeed, since for large α, the largest eigenvalue in the drift function of (3.13) has size α, the mean-square stability constraint for these explicit methods has the form $\alpha h \leq C$, where C is a constant of moderate size, independent of h and α. We observe in Figures 3.2(c)-(d) that these methods are indeed unstable for $h > 2^{-4}$ for the (moderately) stiff case $\alpha = 25$. For the very stiff case $\alpha = 100$ (Figure 3.2(b)), these methods show too much instability to fit in the scales of the figure and are thus omitted.

This numerical experiment shows that the modified θ-M method with $\theta = 1$ has the same (unconditional) mean-square stability as the standard θ-method ($\theta \geq 1/2$), but with an improved accuracy by several orders of magnitude due to the improved weak order two of convergence.

3.2. High weak order integrators preserving quadratic invariants. In this section, we construct numerical integrators for Stratonovich SDEs of high weak order two which exactly conserve all quadratic first integrals (up to machine precision). We consider the SDE (1.1) in Stratonovich form with a one-dimensional noise

$$dX = f(X)dt + g(X) \circ dW(t), \quad X(0) = X_0, \tag{3.15}$$

where the notation $\circ dW(t)$ emphasizes that the Stratonovich stochastic integrals are considered for (3.15). As a basic numerical integrator to apply our methodology of modified equation, we choose the (fully) implicit midpoint rule, as first introduced in [31],

$$X_{n+1} = X_n + hf\left(\frac{X_n + X_{n+1}}{2}\right) + g\left(\frac{X_n + X_{n+1}}{2}\right)\Delta W_n, \tag{3.16}$$

where ΔW_n is a scalar random variable. It is shown in [31] that (3.16) has weak and strong orders one in the case of a one-dimensional or commutative multi-dimensional noise. Notice however that for general SDEs with multi-dimensional noise, the strong order is 1/2 and the weak order is 1.

Remark 3.5. *The method* (3.16) *is implicit with respect to both the drift and the noise terms. In the case where* ΔW_n *is a standard Gaussian variable, the unboundedness of* ΔW_n *for arbitrarily small* h *leads to non-uniqueness of solutions to the non-linear system* (3.16) *and the integrator is not well defined. One way to address this problem, is to replace* ΔW_n, *with a suitable chosen bounded random variable [31] (see also [32, Sect. 1.3]). Here we shall simply consider discrete random variable, e.g.* (3.9)*, as in Remark 3.3, which are obviously bounded.*

First integral conservation. A smooth quantity $C(x)$ is called a first integral of the system (3.15) if it is exactly conserved along time for all realizations of the Wiener process $W(t)$, i.e. $C(X(t)) = C(X_0)$ for all time t and all initial condition $X(0) = X_0$. Given a smooth function $C(x)$, the identity[4] $dC(X) = \nabla C(X) \cdot f(X)dt + \nabla C(X) \cdot g(X) \circ dW(t)$ shows that $C(X)$ is a first integral of (3.15) if and only if

$$\nabla C(x) \cdot f(x) = \nabla C(x) \cdot g(x) = 0 \quad \text{for all } x \in \mathbb{R}^d. \tag{3.17}$$

Proposition 3.1. *The implicit midpoint rule* (3.16) *exactly preserves all quadratic first integrals* $C(x)$ *of* (3.15)*, i.e.* $C(x_{n+1}) = C(x_n)$ *for all* h *and all realizations of* ΔW_n.

Proof. Since $C(x)$ is quadratic, we write $C(x) = x^T Sx$, where S is a constant symmetric matrix. Using (3.17), with $X = (X_{n+1} + X_n)/2$ and $R(X) = hf(X) + g(X)\Delta W_n$, a short computation yields

$$C(X_{n+1}) - C(X_n) = X_n^T SR(X) + R(X)^T SX_{n+1} = 2X^T SR(X) = \nabla C(X) \cdot R(X) = 0.$$

□

3.2.1. New invariant preserving integrators of high weak order. Using the framework of integrators based on modified equations, we introduce the following new numerical integrator

$$X_{n+1} = X_n + hf_{h,1}\left(\frac{X_n + X_{n+1}}{2}\right) + g_{h,1}\left(\frac{X_n + X_{n+1}}{2}\right)\Delta W_n, \tag{3.18}$$

where $f_{h,1} = f + hf_1$ and $g_{h,1} = g + hg_1$ and show below that for

$$f_1 = \frac{1}{4}\left(\frac{1}{2}f''(g,g) - g'f'g\right) \qquad g_1 = \frac{1}{4}\left(\frac{1}{2}g''(g,g) - g'g'g\right), \tag{3.19}$$

the numerical integrator is a weak second order method for the SDE (3.15) which preserves all quadratic first integrals. We notice that if we consider the modified Stratonovich SDE

$$dX = [f(X) + hf_1(X)]\,dt + [g(X) + hg_1(X)] \circ dW(t), \tag{3.20}$$

then (3.18) is equivalent to applying the original midpoint rule (3.16) to the modified Stratonovich SDE (3.20).

Theorem 3.2. *The integrator* (3.18) *for a system of Stratonovich SDEs* (3.15) *with* $m = 1$ *noise has weak order* 2*. It exactly conserves all quadratic first integrals of* (3.15).

[4] Notice that Stratonovich calculus is used here.

Proof. The Stratonovich SDE (3.15) is equivalent to the Itô SDE

$$dX = \Big(f(X) + \frac{1}{2}g'(X)g(X)\Big)\,dt + g(X)dW(t), \tag{3.21}$$

where compared to the Itô system of SDEs (1.1), the vector field f is replaced by $f + \frac{1}{2}g'g$. This permits to deduce an expansion analogue to (2.7) associated to the Itô SDE (3.21). The weak expansion of (3.18) (applied to (3.15), equivalent to the Itô SDE (3.21)) is computed as follows. First we have (for $X_0 = x$) $X_1 = x + F = X_0 + hf(x + F/2) + g(x + F/2)\Delta W_0$, where

$$F = hf + \frac{h}{2}f'F + \frac{h}{8}f''(F,F) + g\Delta W_0 + \frac{1}{2}g'F\Delta W_0 + \frac{1}{8}g''(F,F)\Delta W_0 + \frac{1}{48}g'''(F,F,F)\Delta W_0 + \mathcal{O}(h^{5/2}).$$

For the computation of $A_1(f,g)\phi$ we consider the expansion

$$\phi(X_1) = \phi(x+F) = \phi(X_0) + \sum_k F_{[k]}\partial_k\phi + \frac{1}{2}\sum_{kl} F_{[k]}F_{[l]}\partial_{kl}\phi + \frac{1}{6}\sum_{klm} F_{[k]}F_{[l]}F_{[m]}\partial_{klm}\phi + \frac{1}{24}\sum_{klmi} F_{[k]}F_{[l]}F_{[m]}F_{[i]}\partial_{klmi}\phi + \dots$$

We then compute $\mathbb{E}\big(\phi(X_1)|X_0 = x\big) = \mathbb{E}\big(\phi(x+F)\big)$, identify the differential operator multiplying the term h^2, and obtain after some tedious but straightforward computations,

$$\begin{aligned}\Big(\frac{1}{2}\mathcal{L}^2\phi - A_1(f,g)\Big)\phi &= \frac{1}{4}\Big(\frac{1}{2}f''(g,g) - g'f'g + \frac{1}{4}g'''(g,g,g) - \frac{1}{4}g'g''(g,g) - g'g'g'g\Big)\cdot\nabla_x\phi \\ &\quad + \frac{1}{8}\Big(g\Big(\frac{1}{2}g''(g,g) - g'g'g\Big)^T + \Big(\frac{1}{2}g''(g,g) - g'g'g\Big)g^T\Big) : \nabla_x^2\phi \\ &= \Big(f_1 + \frac{1}{2}(g_1'g + g'g_1)\Big)\cdot\nabla_x\phi + \frac{1}{2}\Big(gg_1^T + g_1g^T\Big) : \nabla_x^2\phi,\end{aligned}$$

where we define $f_1 = \frac{1}{4}\big(\frac{1}{2}f''(g,g) - g'f'g\big)$ and $g_1 = \frac{1}{4}\big(\frac{1}{2}g''(g,g) - g'g'g\big)$. The modified Itô SDE of Theorem 2.1 then reads

$$dX = \Big(f_{h,1} + \frac{1}{2}g'g + \frac{h}{2}(g_1'g + g'g_1)\Big)(X)dt + g_{h,1}(X)dW(t),$$

where $f_{h,1} = f + hf_1$ and $g_{h,1} = g + hg_1$. Using $g_{h,1}'g_{h,1} = g'g + h(g_1'g + g'g_1) + \mathcal{O}(h^2)$ and neglecting the $\mathcal{O}(h^2)$ terms, the above Itô SDE can be converted to the Stratonovich SDE (3.20). This proves that (3.18)-(3.19) is a weak second order method for the SDE (3.21). Finally, the conservation of quadratic first integrals by (3.18) is an immediate consequence of Proposition 3.1 and Lemma 3.3 below. □

LEMMA 3.3. *Any quadratic first integral $C(y)$ of (3.15) is a first integral of (3.20).*

Proof. Consider the original midpoint rule (3.16) applied to (3.15). Using Remark 2.3, we obtain that the modified SDE up to second order for backward error analysis associated to (3.16) is given by (3.20) with h replaced by $-h$,

$$d\hat{X} = \Big[f(\hat{X}) - hf_1(\hat{X})\Big]dt + \Big[g(\hat{X}) - hg_1(\hat{X})\Big]\circ dW(t), \quad \hat{X}(0) = X_0,$$

and we have from (2.18) with $p=1, q=0$, $\mathbb{E}(\phi(X_1)) - \mathbb{E}(\phi(\hat{X}(h))) = \mathcal{O}(h^2)$. Using Proposition 3.1, we have $C(X_1) = C(X_0)$. On the one hand, replacing $\phi(x)$ by $\phi(C(x))$, we obtain $\mathbb{E}(\phi(C(X_0))) - \mathbb{E}(\phi(C(\hat{X}(h)))) = \mathcal{O}(h^2)$. On the other hand, using (3.17), we have

$$d\phi(C(\hat{X})) = -h\phi'(C(x))(\nabla C(x)\cdot f_1(\hat{X})dt + \nabla C(x)\cdot g_1(\hat{X})\circ dW(t)),$$

where $d\phi(C(\hat{X}))$ has size $\mathcal{O}(h)$. This yields $\mathbb{E}(\phi(C(X_0))) - \mathbb{E}(\phi(C(\hat{X}(h)))) = 0$ for all test functions ϕ, and thus $C(\hat{X}(h)) = C(X_0)$. We obtain $\nabla C(x)\cdot f_1(x) = \nabla C(x)\cdot g_1(x) = 0$. □

We close this section by indicating that a further modification of the integrator (3.18) allows yet an even better accuracy. Consider

$$X_{n+1} = X_n + hf_{h,2}\left(\frac{X_n + X_{n+1}}{2}\right) + g_{h,1}\left(\frac{X_n + X_{n+1}}{2}\right)\Delta W_n, \tag{3.22}$$

where $f_{h,2} = f + hf_1 + h^2 f_2$ and $g = g + hg_1$ (as previously) with f_1, g_1 as defined in (3.19) and f_2 given by

$$f_2 = \frac{1}{12}\left(\frac{1}{2}f''(f,f) - f'f'f\right). \tag{3.23}$$

The above term f_2 corrects the deterministic error of size $\mathcal{O}(h^2)$, but the weak order of the integrator (3.22) remains 2. Notice that it would be interesting to search for modified fields $f_{2,h}, g_{2,h}$ to achieve the weak order 3. For comparison, we also consider the integrator of weak order one

$$X_{n+1} = X_n + hf_{h,2}\left(\frac{X_n + X_{n+1}}{2}\right) + g\left(\frac{X_n + X_{n+1}}{2}\right)\Delta W_n, \tag{3.24}$$

where $f_{h,2} = f + hf_1 + h^2 f_2$, with f_1, f_2 given in (3.19),(3.23). Notice that the integrators (3.22) and (3.24) are equivalent to the modifying implicit midpoint rule of order 4 for ODEs (deterministic case) introduced in [10] in the case where the diffusion function g is zero.

We summarize in the following table our theoretical findings. To study the convergence rates in dependence of the noise size, we consider (3.15), where $\circ dW(t)$ is replaced by $\circ\mu\, dW(t)$, with $\mu \in \mathbb{R}^+$ a given scaling parameter.

Table 3.1
Comparison of four integrators conserving exactly all quadratic first integrals for the SDE system $dX = f(X)dt + g(X)\circ\mu\, dW(t)$, $X(0) = X_0$ (1-dimensional noise).

method	weak order of accuracy	order for ODEs (no noise: $\mu = 0$)	weak error $\mathbb{E}(\phi(X_N)) - \mathbb{E}(\phi(X(t_N)))$
IMR, see (3.16)	1	2	$\mathcal{O}(h^2 + \mu^2 h)$
IMR2, see (3.18)	2	2	$\mathcal{O}(h^2)$
IMR(4), see (3.24)	1	4	$\mathcal{O}(h^4 + \mu^2 h)$
IMR2(4), see (3.22)	2	4	$\mathcal{O}(h^4 + \mu h^2)$

Example: a stochastic rigid body model. To illustrate that the integrators previously introduced conserve quadratic first integrals and to compare the performance of the various methods proposed (see Table 3.1), we consider a randomly perturbed rigid body problem that is, the motion of a rigid body in $\mathbb{R}^3$ subject to a scalar white noise perturbation. The equations of motion of an asymmetric rigid body with Stratonovich noise in dimension $m = 1$ are given by [5]

$$\begin{aligned} dX &= \widehat{X}\mathcal{I}^{-1}X dt + \mu\widehat{X}e_1 \circ dW(t), \\ dQ &= Q\widehat{\mathcal{I}^{-1}X} dt + \mu Q\widehat{e_1} \circ dW(t), \end{aligned} \tag{3.25}$$

where $e_1 = (1,0,0)^T$, $\mu \geq 0$ is a parameter and $\mathcal{I} = \mathrm{diag}\,(I_1, I_2, I_3)$. A generalization of equation (3.25) for a 3-dimensional noise is presented in [25, Eq. (6.9)-(6.10)], where one can also find a physical justification for these equations. This model is a variant of the model proposed in [26] with the additional feature that it preserves the spatial angular momentum QX, as detailed below. In the case where $\mu = 0$, we recover the standard deterministic equations of motion of an asymmetric rigid body. We refer to [17, Sect. VI.5] for a survey of geometric and invariant preserving integrators for the rigid body problem in the context of ODEs. The constants $I_1, I_2, I_3 > 0$ are the moments of inertia which characterize the rigid body. The function $X(t)$ represents the angular momentum in $\mathbb{R}^3$ in the body frame. The matrix $Q(t)$ is a rotation matrix in $\mathbb{R}^3$ which gives the orientation of the rigid body in a fixed frame. Notice that the first line in (3.25) can be rewritten simply as

$$\begin{aligned} dX_{[1]} &= \left(\frac{1}{I_3} - \frac{1}{I_2}\right) X_{[2]}X_{[3]}dt, \\ dX_{[2]} &= \left(\frac{1}{I_1} - \frac{1}{I_3}\right) X_{[3]}X_{[1]}dt + \mu X_{[3]} \circ dW(t), \\ dX_{[3]} &= \left(\frac{1}{I_2} - \frac{1}{I_1}\right) X_{[1]}X_{[2]}dt - \mu X_{[2]} \circ dW(t). \end{aligned}$$

The system of SDEs (3.25) has several first integrals, all of which are quadratic. It has QX as first integral, which represents the spatial momentum in $\mathbb{R}^3$ with respect to the body frame. It also has $Q^TQ = Id$ as first integral because Q is an orthogonal matrix. Since Q is orthogonal, the Casimir $C(X) = \frac{1}{2}\left(X_{[1]}^2 + X_{[2]}^2 + X_{[3]}^2\right)$ is also conserved. Considering the Hamiltonian $H(X) = \frac{1}{2}\left(X_{[1]}^2/I_1 + X_{[2]}^2/I_2 + X_{[3]}^2/I_3\right)$, we have

$$dH(X) = \mu\frac{X_{[2]}X_{[3]}}{2}\left(\frac{1}{I_2} - \frac{1}{I_3}\right) \circ dW(t),$$

which shows that $H(X)$ is also a first integral if and only if $I_2 = I_3$ (symmetric body) or $\mu = 0$ (the noise is zero).

Using formulas (3.19) where the functions f and g correspond to the right-hand side of (3.25), a straightforward computation yields the modified SDE associated to

[5] We use the standard hat notation for the correspondence between 3×3 skew-symmetric matrices and size 3 vectors, $\widehat{X} = \begin{pmatrix} 0 & -X_{[3]} & X_{[2]} \\ X_{[3]} & 0 & -X_{[1]} \\ -X_{[2]} & X_{[1]} & 0 \end{pmatrix}$, for all $X = \begin{pmatrix} X_{[1]} \\ X_{[2]} \\ X_{[3]} \end{pmatrix}$.

(3.25),

$$dX = \widehat{X}(\mathcal{I}^{-1} + \frac{h\mu^2}{4}\mathcal{J}^{-1})Xdt + \mu\left(1 + \frac{h\mu^2}{4}\right)\widehat{X}e_1 \circ dW(t), \qquad (3.26)$$
$$dQ = Q(\widehat{\mathcal{I}^{-1}X} + \frac{h\mu^2}{4}\widehat{\mathcal{J}^{-1}X})dt + \mu\left(1 + \frac{h\mu^2}{4}\right)Q\widehat{e_1} \circ dW(t),$$

where we define $\mathcal{J} = \operatorname{diag}(I_1, I_3, I_2)$. We obtain from Theorem 3.2 that applying the implicit midpoint rule (3.16) to the Statonovitch SDE (3.26) yields a weak order two approximation of the solution of (3.25) which exactly conserves all quadratic first integrals, i.e. $C(X_{n+1}) = C(X_n)$, $Q_{n+1}X_{n+1} = Q_nX_n$ and $Q_n^TQ_n = Id$ for all n, and in the case $I_2 = I_3$ (symmetric body), we have also $H(X_{n+1}) = H(X_n)$.

Remark 3.6. *Notice that the modified SDE* (3.26) *is of the same form as the original rigid body equations* (3.25) *with modified data parameters. Indeed, replacing* μ *by*

$$\widetilde{\mu} = \mu(1 + h\mu^2/4),$$

and replacing $\mathcal{I} = \operatorname{diag}(I_1, I_2, I_3)$ *in the original SDE* (3.25) *by* $\widetilde{\mathcal{I}} = \operatorname{diag}(\widetilde{I}_1, \widetilde{I}_2, \widetilde{I}_3)$, *where*

$$\frac{1}{\widetilde{I}_1} = \frac{1}{I_1}\left(1 + \frac{h\mu^2}{4}\right), \quad \frac{1}{\widetilde{I}_2} = \frac{1}{I_2} + \frac{h\mu^2}{4I_3}, \quad \frac{1}{\widetilde{I}_3} = \frac{1}{I_3} + \frac{h\mu^2}{4I_2},$$

yields the modified SDE (3.26). *Thus, our modification to high weak order reduces to a perturbation of the parameters and has a negligible overcost.*

Implementation. We now detail the implementation of the standard implicit midpoint rule (3.16) for the stochastic rigid body problem (3.25). The implementation of the modified implicit midpoint rule (3.18) (and similarly for the method (3.22)) is deduced using Remark 3.6 by modifying the moments of inertia I_1, I_2, I_3. We refer to [10] for the implementation of the corrector f_2 in (3.22) and (3.24).

It is a standard approach to use quaternions q_n to represent the orthogonal matrices Q_n (see [17] in the context of rigid body integrator implementations). The implicit midpoint rule (3.16) for the angular momentum $X(t)$ can be written as

$$X_{n+1} = X_n + h\widehat{Y}\mathcal{I}^{-1}Y + \mu\widehat{Y}e_1\Delta W_n,$$

where we denote $Y = (X_{n+1} + X_n)/2$. This implicit system can be solved by a few fixed point iterations. Next, the configuration update

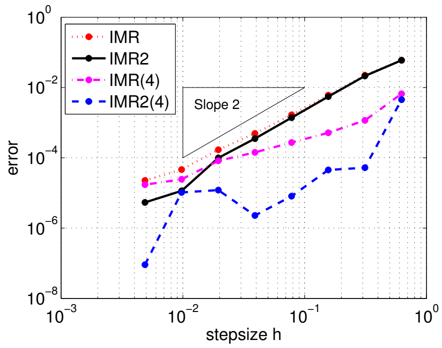
$$Q_{n+1} = Q_n + h\left(\frac{Q_n + Q_{n+1}}{2}\right)\widehat{\mathcal{I}^{-1}Y} + \mu\left(\frac{Q_n + Q_{n+1}}{2}\right)\widehat{e_1}\Delta W_n,$$

is equivalent to $Q_{n+1} = Q_n\Omega$, where Ω is the orthogonal matrix defined by the Cayley transform $\Omega = \big(Id + \widehat{Z}\big)\big(Id - \widehat{Z}\big)^{-1}$ with $Z = \frac{h}{2}\mathcal{I}^{-1}Y + \frac{\Delta W_n\mu}{2}e_1$. Thus, the configuration update for the rotation matrix Q_n reduces to a multiplication of quaternions[6]

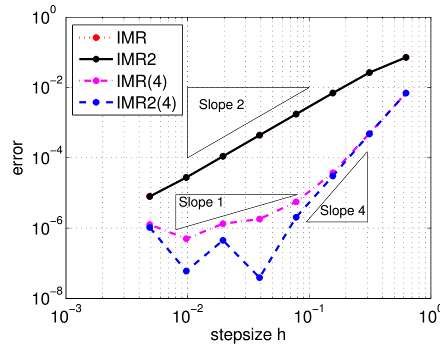
$$q_{n+1} = q_n \cdot \frac{\omega}{\|\omega\|}, \qquad \text{with } \omega = 1 + \frac{h}{2}\left(i\frac{Y_{[1]}}{I_1} + j\frac{Y_{[2]}}{I_2} + k\frac{Y_{[3]}}{I_3}\right) + i\frac{\Delta W_n}{2}\mu,$$

where the matrices Q_n, Q_{n+1} are represented by the quaternions q_{n+1}, q_n, respectively.

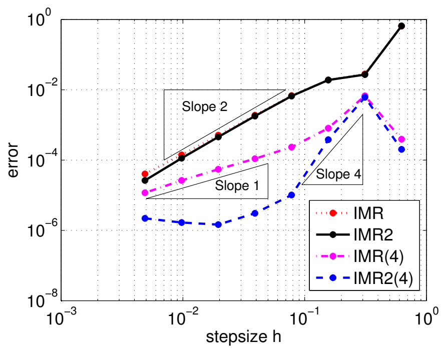
[6]Notice that $\|\omega\|$ denotes the norm of the quaternion ω, thus $\omega/\|\omega\|$ is a quaternion of norm 1.



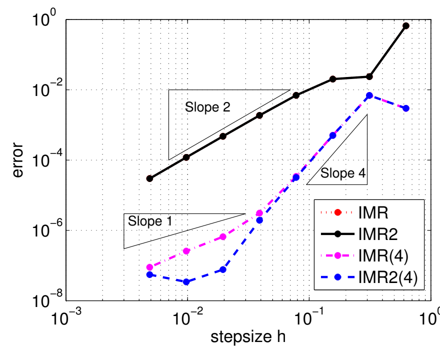
(a) Error for $\mathbb{E}(X_{[1]}^2)$. Noise size $\mu = 0.1$.



(b) Error for $\mathbb{E}(X_{[1]}^2)$. Noise size $\mu = 0.01$.



(c) Error for $\mathbb{E}(\sin(e_1^T Q e_1))$. Noise size $\mu = 0.1$.



(d) Error for $\mathbb{E}(\sin(e_1^T Q e_1))$. Noise size $\mu = 0.01$.

FIG. 3.3. *Rigid body problem* (3.25). *Comparison of weak convergence rates for* IMR, *see* (3.16) *(dotted lines),* IMR2, *see* (3.18) *(solid lines),* IMR2(4), *see* (3.22) *(dashed lines), and* IMR(4), *see* (3.24) *(dashed-dotted lines).*

Convergence rates. We consider the system (3.25) on the time interval $[0, 10]$. We take the moments of inertia $I_1 = 0.345$, $I_2 = 0.653$, $I_3 = 1.0$, which correspond to the water molecule (nearly flat body). Initial values are $X(0) = (0.8, 0.6, 0)^T$ and $Q(0)$ is the identity matrix. We have carefully implemented the above integrators in FORTRAN, using quaternions for the rotation matrices. In Figure 3.3, we plot the errors for $\mathbb{E}(X_{[1]}^2)$, $\mathbb{E}(\sin(e_1^T Q e_1))$ at final time $t = 10$ versus the timestep $h = 2^{-i}$, $i = 1, \dots, 8$. The reference solution is computed using the small timestep $h = 2^{-14}$. To check carefully the accuracy of the methods, we compute numerically $\mathbb{E}(X_{[1]}^2)$ and $\mathbb{E}(\sin(e_1^T Q e_1))$ using the averages over 300 millions of trajectories. We consider two values of the noise parameter: $\mu = 0.1$ and $\mu = 0.01$. We observe in all cases lines of slope two for the modified midpoint rule IMR2 (3.18) which confirms its weak order two of accuracy. For the standard midpoint rule IMR in (3.16) and the modified version IMR(4) in (3.24) which both have weak order one, we observe for large stepsizes h, lines of slope four and two respectively in the case where the deterministic error (h^2 or h^4) is dominant compared to $\mu^2 h$. For smaller timesteps, we retrieve lines of slope one, the weak order of these two methods. Similarly, for the improved modified midpoint rule IMR2(4) in (3.22), we observe lines of slope four for large timesteps and only two for small timestep. This corroborates the theoretical results collected in Table 3.1.

4. Conclusion. In this paper, we introduced a new framework for increasing the weak order of accuracy of a given numerical method for SDE by considering the numerical integration of a suitably modified problem. Our methodology, which uses tools developed for backward error analysis for stochastic problems [46, 14, 41], generalizes to SDEs the framework of numerical integrators based on modified equations introduced in [10] for deterministic problems. This approach permits to fulfill automatically the numerous order conditions for high weak order schemes. We illustrated our approach at the example of the θ-Milstein-Talay method, and obtained for $\theta = 1$ a mean-square stable integrator of weak order two. The numerical experiments conducted for a stiff problem in economy show an improvement in accuracy of two orders of magnitude over the classical θ-method of weak order one.

In the spirit of backward error analysis for the study of geometric integrators for ODEs, where the modified equations inherit the geometric properties of the integrators, we also derived new high weak order integrators based on the implicit midpoint rule, that automatically conserve all quadratic first integrals. The efficiency of the approach is illustrated at the example of a stochastic rigid body model which possesses several quadratic first integrals. A natural extension of this work would be to search for modified equations to construct new integrators of weak order three or more with good stability or geometric properties.

We note that this new approach also allows to construct higher order Chebyshev methods for stiff SDEs. An attempt to generalize such methods, introduced in [1, 2, 3], to higher order has been proposed in [9, 24]. This generalization involves the solution of a large number of order conditions and the resulting methods appear to have less favorable stability properties than the method proposed in [1, 2, 3]. In [4], we show that using techniques based on modifying equations as proposed in this paper, it is possible to construct high weak order Chebyshev method in an efficient way with better stability properties than the method given in [9, 24].

Acknowledgments. The research of A. A. is partially supported under Swiss National Foundation Grant 200021_140692. K. C. Z. was supported by Award No. KUK-C1-013-04 of the King Abdullah University of Science and Technology (KAUST).

REFERENCES

[1] A. Abdulle and S. Cirilli. Stabilized methods for stiff stochastic systems. *C. R. Math. Acad. Sci. Paris*, 345(10):593–598, 2007.
[2] A. Abdulle and S. Cirilli. S-ROCK: Chebyshev methods for stiff stochastic differential equations. *SIAM J. Sci. Comput.*, 30:997–1014, 2008.
[3] A. Abdulle and T. Li. S-ROCK methods for stiff Itô SDEs. *Commun. Math. Sci.*, 6(4):845–868, 2008.
[4] A. Abdulle, G. Vilmart, and K. C. Zygalakis. Second weak order explicit stabilized methods for stiff stochastic differential equations. *Submitted*, 2012.
[5] L. Arnold. *Stochastic Differential Equations, Theory and applications.* Wiley-Interscience Publication, 1974.
[6] E. Buckwar and C. Kelly. Towards a systematic linear stability analysis of numerical methods for systems of stochastic differential equations. *SIAM J. Numer. Anal.*, 48(1):298–321, 2010.
[7] E. Buckwar and T. Sickenberger. A comparative linear mean-square stability analysis of Maruyama- and Milstein-type methods. *Math. Comput. Simulation*, 81(6):1110–1127, 2011.
[8] K. Burrage and P.M. Burrage. General order conditions for stochastic Runge-Kutta methods for both commuting and non-commuting stochastic ordinary differential equation systems. *Appl. Numer. Math.*, 28(2-4):161–177, 1998. Eighth Conference on the Numerical Treatment of Differential Equations (Alexisbad, 1997).

[9] K. Burrage and Y. Komori. Explicit stochastic Runge-Kutta methods with large stability regions. volume 1281, pages 2057–2060, 2010. International Conference on Numerical Analysis and Applied Mathematics (Rhodes, 2010).
[10] P. Chartier, E. Hairer, and G. Vilmart. Numerical integrators based on modified differential equations. *Math. Comp.*, 76(260):1941–1953 (electronic), 2007.
[11] D. Crisan and B. Rozovskii (e.d.). *The Oxford handbook of nonlinear filtering.* Oxford University Press, USA, 2011.
[12] K. Debrabant and A. Rößler. Classification of stochastic Runge-Kutta methods for the weak approximation of stochastic differential equations. *Math. Comput. Simulation*, 77(4):408–420, 2008.
[13] K. Debrabant and A. Rößler. Families of efficient second order Runge-Kutta methods for the weak approximation of Itô stochastic differential equations. *Appl. Numer. Math.*, 59(3-4):582–594, 2009.
[14] A. Debussche and E. Faou. Weak backward error analysis for SDEs. *preprint*, 2011. `http://fr.arxiv.org/abs/1105.0489`.
[15] G.W. Gardiner. *Handbook of stochastic processes for physics, chemistry and natural sciences.* Springer Verlag, 2 edition, 1985.
[16] M.B. Giles. Multilevel Monte Carlo path simulation. *Operations Research*, 56(3):607–617, 2008.
[17] E. Hairer, C. Lubich, and G. Wanner. *Geometric numerical integration. Structure-preserving algorithms for ordinary differential equations.* Springer Series in Computational Mathematics 31. Springer-Verlag, Berlin, second edition, 2006.
[18] D.J. Higham. *A*-stability and stochastic mean-square stability. *BIT*, 40(2):404–409, 2000.
[19] D.J. Higham. Mean-square and asymptotic stability of the stochastic theta method. *SIAM J. Numer. Anal.*, 38:753–769, 2000.
[20] D.J. Higham and P.E. Kloeden. MAPLE and MATLAB for stochastic differential equations in finance. In Soren Bo Nielsen, editor, *Programming Languages and Systems in Computational Economics and Finance*, volume 18 of *Advances in Computational Economics*, pages 233–270. Springer, Kluwer, 2002.
[21] N. Hofmann, E. Platen, and M. Schweizer. Option pricing under incompleteness and stochastic volatility. *Mathematical Finance*, 2(3):153–187, 1992.
[22] P.E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1992.
[23] P.E. Kloeden, E. Platen, and N. Hofmann. Extrapolation methods for the weak approximation of Itô diffusions. *SIAM J. Numer. Anal.*, 32(5):1519–1534, 1995.
[24] Y. Komori and K. Burrage. Weak second order s-rock methods for Stratonovich stochastic differential equations. *J. Comput. Appl. Math.*, 236(11):2895–2908, May 2012.
[25] J.-A. Lázaro-Camí and J.-P. Ortega. Reduction, reconstruction, and skew-product decomposition of symmetric stochastic differential equations. *Stoch. Dyn.*, 9(1):1–46, 2009.
[26] M. Liao. Random motion of a rigid body. *J. Theoret. Probab.*, 10(1):201–211, 1997.
[27] V. Mackevicius. Second order weak approximation for Stratonovich stochastic differential equations. *Liet. Mat. Rink.*, 34(2):183–200, 1994.
[28] G. Maruyama. Continuous markov processes and stochastic equations. *Rend. Circ. Mat. Palermo*, 4(0):48–90, 1955.
[29] G.N. Milstein. A method of second order accuracy integration of stochastic differential equation. *Theory Probab. Appl.*, 23:396–401, 1978.
[30] G.N. Milstein. Weak approximation of solutions of systems of stochastic differential equations. *Theory Probab. Appl.*, 30(4):750–766, 1986.
[31] G.N. Milstein, Y.M. Repin, and M.V. Tretyakov. Numerical methods for stochastic systems preserving symplectic structure. *SIAM J. Numer. Anal.*, 40(4):1583–1604 (electronic), 2002.
[32] G.N. Milstein and M.V. Tretyakov. *Stochastic numerics for mathematical physics.* Scientific Computation. Springer-Verlag, Berlin, 2004.
[33] B. Oksendal. *Stochastic Differential Equations.* Springer, 6th edition, 2003.
[34] G.A. Pavliotis and A.M. Stuart. *Multiscale methods: averaging and homogenization.* Springer-Verlag, New York, 2008.
[35] W.P. Petersen. Lagged Fibonacci series random number generators for the NEC SX-3. *Internat. J. High Speed Computing*, 6(3):387–398, 1994.
[36] E. Platen. High-order weak approximation of ito diffusions by markov chains. *Probab. Engrg. Inform. Sci.*, 6:391–408, 1992.
[37] S.T. Rachev (e.d.). *Handbook of computational and numerical methods in finance.* Birkhauser, Boston, 2004.
[38] H. Risken. *The Fokker-Planck equation*, volume 18 of *Springer Series in Synergetics*. Springer-

Verlag, Berlin, 1989.

[39] A. Rößler. Second order Runge-Kutta methods for Itô stochastic differential equations. *SIAM J. Numer. Anal.*, 47(3):1713–1738, 2009.

[40] Y. Saito and T. Mitsui. Stability analysis of numerical schemes for stochastic differential equations. *SIAM J. Numer. Anal.*, 33(6):2254–2267, 1996.

[41] T. Shardlow. Modified equations for stochastic differential equations. *BIT*, 46(1):111–125, 2006.

[42] P.D. Spanos (e.d.). *Computational stochastic mechanics.* Balkema, Rotterdam, Netherlands, 1999.

[43] D. Talay. Efficient numerical schemes for the approximation of expectations of functionals of the solution of a SDE and applications. In *Filtering and control of random processes (Paris, 1983)*, volume 61 of *Lecture Notes in Control and Inform. Sci.*, pages 294–313. Springer, Berlin, 1984.

[44] D. Talay and L. Tubaro. Expansion of the global error for numerical schemes solving stochastic differential equations. *Stochastic Anal. Appl.*, 8(4):483–509 (1991), 1990.

[45] A. Tocino and J. Vigo-Aguiar. Weak second order conditions for stochastic Runge-Kutta methods. *SIAM J. Sci. Comput.*, 24(2):507–523 (electronic), 2002.

[46] K.C. Zygalakis. On the existence and the applications of modified equations for stochastic differential equations. *SIAM J. Sci. Comput.*, 33(1):102–130, 2011.