SIAM J. SCI. COMPUT.
Vol. 39, No. 6, pp. A2857–A2878

WEAK SECOND ORDER EXPLICIT EXPONENTIAL RUNGE–KUTTA METHODS FOR STOCHASTIC DIFFERENTIAL EQUATIONS*

YOSHIO KOMORI†, DAVID COHEN‡, AND KEVIN BURRAGE§

Abstract. We propose new explicit exponential Runge–Kutta methods for the weak approximation of solutions of stiff Itô stochastic differential equations (SDEs). We also consider the use of exponential Runge–Kutta methods in combination with splitting methods. These methods have weak order 2 for multidimensional, noncommutative SDEs with a semilinear drift term, whereas they are of order 2 or 3 for semilinear ordinary differential equations. These methods are A-stable in the mean square sense for a scalar linear test equation whose drift and diffusion terms have complex coefficients. We carry out numerical experiments to compare the performance of these methods with an existing explicit stabilized method of weak order 2.

Key words. explicit method, exponential integrator, splitting method, stiffness, noncommutative noise, Itô stochastic differential equation

AMS subject classifications. 60H10, 65L05, 65L06

DOI. 10.1137/15M1041341

1. Introduction.

For stiff ordinary differential equations (ODEs), there are some classes of explicit methods that are well suited. One such class is the Runge–Kutta–Chebyshev (RKC) methods. They are useful for stiff problems whose eigenvalues lie near the negative real axis. Van der Houwen and Sommeijer [38] have constructed a family of first order RKC methods. Abdulle and Medovikov [3] have modified this class and proposed a family of second order RKC methods. Another suitable class of methods is the explicit exponential Runge–Kutta (RK) methods for semilinear problems [12, 17, 18, 19, 26, 32]. Although these methods were proposed many years ago, they have not been regarded as practical until recently because of the cost of calculations for matrix exponentials, especially for large problems. In order to overcome this problem, new methods have been proposed [15, 17, 18, 19]. In addition, the class of splitting methods is also competitive with the classes of numerical methods mentioned above. Exponential RK methods can be used in combination with splitting methods [19, 29].

Similarly, for stochastic differential equations (SDEs) explicit RK methods that have excellent stability properties have been developed. Abdulle and Cirilli [1] have proposed a family of explicit stochastic orthogonal Runge–Kutta–Chebyshev (SROCK)

*Submitted to the journal's Methods and Algorithms for Scientific Computing section September 25, 2015; accepted for publication (in revised form) May 26, 2017; published electronically December 12, 2017.
http://www.siam.org/journals/sisc/39-6/M104134.html
Funding: This work was partially supported by JSPS Grant-in-Aid for Scientific Research 23540143, and partially by an overseas study program of the faculty of Computer Science and Systems Engineering at the Kyushu Institute of Technology.

†Department of Systems Design and Informatics, Kyushu Institute of Technology, 680-4 Kawazu, Iizuka 820-8502, Japan (komori@ces.kyutech.ac.jp, yoshio.kom@gmail.com).
‡Department of Mathematics and Mathematical Statistics, Umeå University, SE-901 87 Umeå, Sweden, and Department of Mathematics, University of Innsbruck, A-6020 Innsbruck, Austria (david.cohen@umu.se, david.cohen@uibk.ac.at).
§ARC Centre of Excellence for Mathematical and Statistical Frontiers, School of Mathematics, Queensland University of Technology (QUT), Australia (kevin.burrage@qut.edu.au, kevin.burrage@gmail.com).

methods with extended mean square (MS) stability regions. Their methods have strong order 1/2 and weak order 1 for noncommutative Stratonovich SDEs, whereas they reduce to the first order RKC methods when applied to ODEs. Abdulle and Li [2] have proposed SROCK methods of the same order for noncommutative Itô SDEs. Komori and Burrage [23] have developed these ideas and have proposed weak second order SROCK methods for noncommutative Stratonovich SDEs. If the methods are applied to ODEs, they reduce to the second order RKC methods of Abdulle and Medovikov [3]. Komori and Burrage [24] have also proposed strong first order SROCK methods for noncommutative Itô and Stratonovich SDEs, which reduce to first or second order RKC methods for ODEs. The weak second order SROCK methods given by Komori and Burrage [23] have the advantage that the stability region is large along the negative real axis, but they still have the drawback that their stability region is not very wide. In order to overcome this, Abdulle, Vilmart, and Zygalakis [5] have proposed a new family of weak second order SROCK methods for noncommutative Itô SDEs, in which another family of second order RKC methods is embedded.

On the other hand, Shi, Xiao, and Zhang [34] have proposed an exponential Euler method for the strong approximation of solutions of SDEs with multiplicative noise driven by a scalar Wiener process. Cohen [9] and Tocino [36] have proposed exponential integrators for second order SDEs with a semilinear drift term and multiplicative noise. Adamu [6], Geiger, Lord, and Tambue [14], and Lord and Tambue [27] have proposed exponential integrators for stochastic partial differential equations with a semilinear drift term and multiplicative noise. Komori and Burrage [25] have proposed another explicit exponential Euler method for noncommutative Itô SDEs with a semilinear drift term, which is of strong order 1/2 and A-stable in the MS.

In the present paper, we derive stochastic exponential Runge–Kutta (SERK) methods and splitting methods for the weak approximation of solutions of noncommutative Itô SDEs with a semilinear drift term. We will achieve this with the help of the derivative-free Milstein–Talay (DFMT) method proposed by Abdulle, Vilmart, and Zygalakis [4, 5] and explicit exponential RK methods for ODEs proposed by Hochbruck and Ostermann [18]. In section 2 we will briefly introduce explicit exponential RK methods and a classical splitting method for ODEs. In section 3 we will derive and analyze novel SERK and splitting methods for SDEs, and in section 4 we will give their stability and error analysis. In section 5 we will present numerical results and in section 6 our conclusions.

2. Explicit exponential RK methods and the Strang splitting method for ODEs. We consider autonomous semilinear ODEs given by

$$\boldsymbol{y}'(t) = A\boldsymbol{y}(t) + \boldsymbol{f}(\boldsymbol{y}(t)), \qquad t \geq 0, \qquad \boldsymbol{y}(0) = \boldsymbol{y}_0, \tag{2.1}$$

where $\boldsymbol{y}$ is an $\mathbb{R}^d$-valued function on $[0, \infty)$, A is a $d \times d$ matrix, and $\boldsymbol{f}$ is an $\mathbb{R}^d$-valued nonlinear function on $\mathbb{R}^d$. In order to introduce some exponential RK methods for (2.1), we make the following assumptions [18].

Assumption 2.1. For a given time $T > 0$, (2.1) satisfies the conditions below.

(1) There exists a constant C such that

$$\left\| \mathrm{e}^{tA} \right\| \leq C$$

for all $t \in [0, T]$.

(2) The nonlinear function $\boldsymbol{f}$ is (locally) Lipschitz continuous in a local region U that contains the exact solution $\boldsymbol{y}$ on $[0, T]$, that is,

$$\{\boldsymbol{y}(t) \mid t \in [0, T]\} \subset U.$$

(3) The solution $\boldsymbol{y}$ is a sufficiently smooth function on $[0, T]$ and $\boldsymbol{f}$ is sufficiently differentiable in U. All derivatives of $\boldsymbol{y}$ and $\boldsymbol{f}$ are uniformly bounded in $[0, T]$ and U, respectively.

Note that the global error estimation of all exponential RK methods introduced in this section can be influenced by the constant C [18].

Let $\boldsymbol{y}_n$ denote a discrete approximation to the solution $\boldsymbol{y}(t_n)$ of (2.1) for an equidistant grid point $t_n \stackrel{\text{def}}{=} nh$ $(n = 1, 2, \ldots, M)$ with step size $h = T/M < 1$ (M is a natural number). By the variation-of-constants formula, the solution of (2.1) is

$$\boldsymbol{y}(t_{n+1}) = \mathrm{e}^{Ah}\boldsymbol{y}(t_n) + \int_{t_n}^{t_{n+1}} \mathrm{e}^{A(t_{n+1}-s)}\boldsymbol{f}(\boldsymbol{y}(s))\mathrm{d}s.$$

By replacing $\boldsymbol{y}(t_n)$ with $\boldsymbol{y}_n$ and interpolating $\boldsymbol{f}(\boldsymbol{y}(s))$ at $\boldsymbol{f}(\boldsymbol{y}_n)$ only, we obtain the simplest exponential method for (2.1) [19]:

$$\boldsymbol{y}_{n+1} = \mathrm{e}^{Ah}\boldsymbol{y}_n + h\varphi_1(Ah)\boldsymbol{f}(\boldsymbol{y}_n),$$

where $\varphi_1(Z) \stackrel{\text{def}}{=} Z^{-1}(\mathrm{e}^Z - I)$ and I stands for the $d \times d$ identity matrix. This is called the explicit exponential Euler method.

Higher order exponential RK methods have been proposed in [18, 19]. For example, the following is a one-parameter family of second order exponential RK methods:

$$\begin{aligned} \boldsymbol{Y}_1 &= \mathrm{e}^{chA}\boldsymbol{y}_n + ch\varphi_1(chA)\boldsymbol{f}(\boldsymbol{y}_n), \\ \boldsymbol{y}_{n+1} &= \mathrm{e}^{hA}\boldsymbol{y}_n + h\left\{\varphi_1(hA) - \frac{1}{c}\varphi_2(hA)\right\}\boldsymbol{f}(\boldsymbol{y}_n) + \frac{1}{c}h\varphi_2(hA)\boldsymbol{f}(\boldsymbol{Y}_1), \end{aligned} \tag{2.2}$$

where c is a parameter and $\varphi_2(Z) \stackrel{\text{def}}{=} Z^{-2}(\mathrm{e}^Z - I - Z)$. In addition to Assumption 2.1, let us assume that there exists a constant C such that

$$\left\| hA \sum_{k=1}^{n-1} \mathrm{e}^{khA} \right\| \leq C$$

for $n = 2, 3, \ldots, M$. (Note that the global error estimation of the following family of exponential RK methods can also be influenced by the above constant C [18].) Then, a two-parameter family of third order exponential RK methods is given by

$$\begin{aligned} \boldsymbol{Y}_1 &= \mathrm{e}^{c_1hA}\boldsymbol{y}_n + c_1h\varphi_1(c_1hA)\boldsymbol{f}(\boldsymbol{y}_n), \\ \boldsymbol{Y}_2 &= \mathrm{e}^{c_2hA}\boldsymbol{y}_n + h\left\{c_2\varphi_1(c_2hA) - \psi(hA)\right\}\boldsymbol{f}(\boldsymbol{y}_n) + h\psi(hA)\boldsymbol{f}(\boldsymbol{Y}_1), \\ \boldsymbol{y}_{n+1} &= \mathrm{e}^{hA}\boldsymbol{y}_n + h\left\{\varphi_1(hA) - \frac{\gamma+1}{\gamma c_1 + c_2}\varphi_2(hA)\right\}\boldsymbol{f}(\boldsymbol{y}_n) \\ &\quad + \frac{h}{\gamma c_1 + c_2}\varphi_2(hA)\left\{\gamma\boldsymbol{f}(\boldsymbol{Y}_1) + \boldsymbol{f}(\boldsymbol{Y}_2)\right\}, \end{aligned} \tag{2.3}$$

where c_1, c_2, and γ are parameters satisfying

$$2(\gamma c_1 + c_2) = 3\left(\gamma c_1^2 + c_2^2\right) \tag{2.4}$$

and $\psi(Z) \stackrel{\text{def}}{=} \gamma c_1\varphi_2(c_1 Z) + \frac{c_2^2}{c_1}\varphi_2(c_2 Z)$.

On the other hand, as a second order splitting method for (2.1), the following Strang splitting method is well known [29, 35]:

$$\boldsymbol{y}_{n+1} = \mathrm{e}^{\frac{h}{2}A}\Phi_h\left(\mathrm{e}^{\frac{h}{2}A}\boldsymbol{y}_n\right),$$

where Φ_h is an integrator of at least order 2 for ODEs given by replacing A with the zero matrix in (2.1).

3. Weak second order SERK methods and splitting methods. We shall now derive SERK methods and splitting methods of weak order 2 by utilizing some results for a well-designed existing stochastic Runge–Kutta (SRK) method. We give a brief introduction to the SRK method in subsection 3.1. We then present SERK methods in subsection 3.2 and splitting methods in subsection 3.3.

3.1. The DFMT method. Similarly to the case of ODEs, we are concerned with autonomous SDEs with a semilinear drift term given by

$$\mathrm{d}\boldsymbol{y}(t) = (A\boldsymbol{y}(t) + \boldsymbol{f}(\boldsymbol{y}(t)))\mathrm{d}t + \sum_{j=1}^{m}\boldsymbol{g}_j(\boldsymbol{y}(t))\mathrm{d}W_j(t), \quad \boldsymbol{y}(0) = \boldsymbol{y}_0, \tag{3.1}$$

where $t \in [0, T]$ and where $\boldsymbol{g}_j$, $j = 1, 2, \ldots, m$, are $\mathbb{R}^d$-valued functions on $\mathbb{R}^d$, the $W_j(t)$, $j = 1, 2, \ldots, m$, are independent Wiener processes, and $\boldsymbol{y}_0$ is independent of $W_j(t) - W_j(0)$ [7, p. 100].

In order to deal with weak approximations for (3.1), let $\boldsymbol{g}_0(\boldsymbol{y})$ denote $A\boldsymbol{y} + \boldsymbol{f}(\boldsymbol{y})$ and let us consider the following DFMT method [4, 5]:

$$\begin{aligned}
&\boldsymbol{K}_1 = \boldsymbol{y}_n + h\boldsymbol{g}_0(\boldsymbol{y}_n), \qquad \boldsymbol{K}_2 = \boldsymbol{K}_1 + \sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j(\boldsymbol{y}_n)\xi_j,\\
&\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + \frac{h}{2}\left\{\boldsymbol{g}_0(\boldsymbol{y}_n) + \boldsymbol{g}_0(\boldsymbol{K}_2)\right\} + \boldsymbol{H}(\boldsymbol{y}_n) + \tilde{\boldsymbol{H}}\left(\frac{\boldsymbol{y}_n + \boldsymbol{K}_1}{2}, \boldsymbol{y}_n\right),
\end{aligned} \tag{3.2}$$

where

$$\begin{aligned}
\boldsymbol{H}(\boldsymbol{y}) &\stackrel{\text{def}}{=} \frac{1}{2}\sum_{j=1}^{m}\left\{\boldsymbol{g}_j\left(\boldsymbol{y} + h\sum_{k=1}^{m}\boldsymbol{g}_k(\boldsymbol{y})\zeta_{kj}\right)\right.\\
&\qquad\qquad \left. -\boldsymbol{g}_j\left(\boldsymbol{y} - h\sum_{k=1}^{m}\boldsymbol{g}_k(\boldsymbol{y})\zeta_{kj}\right)\right\},\\
\tilde{\boldsymbol{H}}(\boldsymbol{y}, \boldsymbol{z}) &\stackrel{\text{def}}{=} \frac{\sqrt{h}}{2}\sum_{j=1}^{m}\left\{\boldsymbol{g}_j\left(\boldsymbol{y} + \sqrt{\frac{h}{2}}\sum_{k=1}^{m}\boldsymbol{g}_k(\boldsymbol{z})\chi_k\right)\right.\\
&\qquad\qquad \left. +\boldsymbol{g}_j\left(\boldsymbol{y} - \sqrt{\frac{h}{2}}\sum_{k=1}^{m}\boldsymbol{g}_k(\boldsymbol{z})\chi_k\right)\right\}\xi_j,
\end{aligned}$$

and where the χ_j and ξ_j, $j = 1, 2, \ldots, m$, are discrete random variables satisfying $P(\chi_j = \pm 1) = 1/2$, $P(\xi_j = \pm\sqrt{3}) = 1/6$, and $P(\xi_j = 0) = 2/3$, and the ζ_{kj}, $j, k = 1, 2, \ldots, m$, are given by

$$\zeta_{kj} \stackrel{\text{def}}{=} \begin{cases} (\xi_j\xi_j - 1)/2 & (j = k),\\ (\xi_k\xi_j - \chi_k)/2 & (j < k),\\ (\xi_k\xi_j + \chi_j)/2 & (j > k). \end{cases}$$

Let $C_P^L(\mathbb{R}^d, \mathbb{R})$ denote the family of L times continuously differentiable real-valued functions on $\mathbb{R}^d$, whose partial derivatives of order less than or equal to L have polynomial growth. Whenever we deal with weak convergence of order q, we will make the following assumption [21, p. 474].

Assumption 3.1. All moments of the initial value $\boldsymbol{y}_0$ exist and $\boldsymbol{g}_j, j = 0, 1, \ldots, m$, are Lipschitz continuous with all their components belonging to $C_P^{2(q+1)}(\mathbb{R}^d, \mathbb{R})$.

Then, we can give the definition of weak convergence of order q [21, p. 327].

DEFINITION 3.2. *When discrete approximations $\boldsymbol{y}_n$ are given by a numerical method, we say that the method is of weak (global) order q if, for all $G \in C_P^{2(q+1)}$ $(\mathbb{R}^d, \mathbb{R})$, constants $C > 0$ (independent of h) and $\delta_0 > 0$ exist such that*

$$|E[G(\boldsymbol{y}(T)] - E[G(\boldsymbol{y}_M)]| \le Ch^q, \quad h \in (0, \delta_0).$$

In order to consider numerical methods of weak order q, the following theorem proposed by Milstein [30] (see [31, p. 100]) is very useful [4, 5, 33].

THEOREM 3.3. *In addition to Assumption* 3.1*, suppose that the following conditions hold:*

(1) *for a sufficiently large $r \in \mathbb{N}$, the moments $E[\|\boldsymbol{y}_n\|^{2r}]$ exist and are uniformly bounded with respect to M and $n = 0, 1, \ldots, M$;*

(2) *for all $G \in C_P^{2(q+1)}(\mathbb{R}^d, \mathbb{R})$, the local error estimation*

$$\left|E[G(\boldsymbol{y}(t_{n+1}))] - E[G(\boldsymbol{y}_{n+1})]\right| \le |K(\boldsymbol{y}_n)| h^{q+1}$$

holds if $\boldsymbol{y}(t_n) = \boldsymbol{y}_n$, where $K \in C_P^0(\mathbb{R}^d, \mathbb{R})$.

Then, the method that gives $\boldsymbol{y}_n, n = 0, 1, \ldots, M$, is of weak (global) order q.

The second condition concerning the local error in the theorem provides us with order conditions for an SRK method to be of weak order q [33]. In addition, the DFMT method is of weak order 2 [4]. These facts give us a way of deriving new SRK methods of weak order 2 [4, 5]. For this, we propose a useful lemma to give a sufficient condition for SRK methods based on the DFMT method in order to satisfy the second condition in Theorem 3.3.

LEMMA 3.4. *For an approximate solution $\boldsymbol{y}_n$, let $\boldsymbol{y}_{n+1}$ be given by* (3.2) *and let $\hat{\boldsymbol{y}}_{n+1}$ be defined by*

$$\hat{\boldsymbol{y}}_{n+1} = \tilde{\boldsymbol{y}}_{n+1} + \frac{h}{2}\boldsymbol{g}_0\left(\boldsymbol{Y}_1 + \sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j\right) + \boldsymbol{H}\left(\boldsymbol{Y}_3\right) + \tilde{\boldsymbol{H}}\left(\boldsymbol{Y}_4, \boldsymbol{Y}_5\right).$$

Here, we assume that the intermediate values $\tilde{\boldsymbol{y}}_{n+1}$ and $\boldsymbol{Y}_i, i = 1, 2, \ldots, 5$, have no random variable and satisfy

$$\tilde{\boldsymbol{y}}_{n+1} + \frac{h}{2}\boldsymbol{g}_0\left(\boldsymbol{Y}_1\right) = \boldsymbol{y}_n + h\boldsymbol{g}_0\left(\boldsymbol{y}_n\right) + \frac{h^2}{2}\boldsymbol{g}_0'\left(\boldsymbol{y}_n\right)\boldsymbol{g}_0\left(\boldsymbol{y}_n\right) + O\left(h^3\right), \tag{3.3}$$

$$\boldsymbol{Y}_i = \boldsymbol{y}_n + O(h) \quad (i = 1, 2, 3, 5), \qquad \boldsymbol{Y}_4 = \boldsymbol{y}_n + \frac{h}{2}\boldsymbol{g}_0\left(\boldsymbol{y}_n\right) + O\left(h^2\right).$$

(Note that the symbol $O(h^p)$ represents terms $\boldsymbol{x}$ such that $\|\boldsymbol{x}\| \le |K(\boldsymbol{y}_n)|h^p$ for $K \in C_P^0(\mathbb{R}^d, \mathbb{R})$ and small $h > 0$.) Then, for all $G \in C_P^r(\mathbb{R}^d, \mathbb{R})$ $(r \ge 3)$,

$$E\left[G\left(\hat{\boldsymbol{y}}_{n+1}\right)\right] - E\left[G\left(\boldsymbol{y}_{n+1}\right)\right] = O\left(h^3\right).$$

For the proof of this lemma, we refer the reader to Appendix A.

3.2. SERK methods. We shall propose weak second order SERK methods for (3.1). As a simple case, let us begin with

$$\begin{aligned} \boldsymbol{y}_{n+1} = \boldsymbol{Y}_1 + h\varphi_2(hA)\left\{\boldsymbol{f}\left(\boldsymbol{Y}_1 + \sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j\right) - \boldsymbol{f}(\boldsymbol{y}_n)\right\} \\ + \mathrm{e}^{\frac{h}{2}A}\left(\boldsymbol{H}(\boldsymbol{Y}_2) + \tilde{\boldsymbol{H}}(\boldsymbol{Y}_2, \boldsymbol{Y}_2)\right). \end{aligned} \tag{3.4}$$

Here and in what follows, we set $\boldsymbol{Y}_1$ and $\boldsymbol{Y}_2$ by

$$\boldsymbol{Y}_1 \stackrel{\text{def}}{=} \mathrm{e}^{hA}\boldsymbol{y}_n + h\varphi_1(hA)\boldsymbol{f}(\boldsymbol{y}_n), \quad \boldsymbol{Y}_2 \stackrel{\text{def}}{=} \mathrm{e}^{\frac{h}{2}A}\boldsymbol{y}_n + \frac{h}{2}\varphi_1\left(\frac{h}{2}A\right)\boldsymbol{f}(\boldsymbol{y}_n). \tag{3.5}$$

Observe that if the diffusion terms vanish, (3.4) is equivalent to (2.2) with $c = 1$.

THEOREM 3.5. *Let $\boldsymbol{g}_0(\boldsymbol{y})$ denote $A\boldsymbol{y} + \boldsymbol{f}(\boldsymbol{y})$ and suppose that* (3.1) *satisfies Assumption* 3.1 *for $q = 2$. Then* (3.4) *is of weak order* 2 *for* (3.1).

Proof. Due to Lemma 3.4, it is clear that the local error of the method

$$\hat{\boldsymbol{y}}_{n+1} = \tilde{\boldsymbol{y}}_{n+1} + \frac{h}{2}\boldsymbol{g}_0\left(\boldsymbol{K}_1 + \sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j\right) + \boldsymbol{H}(\boldsymbol{Y}_2) + \tilde{\boldsymbol{H}}(\boldsymbol{Y}_2, \boldsymbol{Y}_2)$$

is weak order 3, where $\boldsymbol{Y}_1, \boldsymbol{Y}_2$ are given in (3.5), $\tilde{\boldsymbol{y}}_{n+1} = \boldsymbol{y}_n + (h/2)\boldsymbol{g}_0(\boldsymbol{y}_n)$, and $\boldsymbol{K}_1 = \boldsymbol{y}_n + h\boldsymbol{g}_0(\boldsymbol{y}_n)$. Using $\boldsymbol{g}_0(\boldsymbol{y}) = A\boldsymbol{y} + \boldsymbol{f}(\boldsymbol{y})$, we can rewrite this as follows:

$$\begin{aligned} \hat{\boldsymbol{y}}_{n+1} = \boldsymbol{y}_n + \frac{h}{2}\left(A\boldsymbol{y}_n + \boldsymbol{f}\left(\boldsymbol{y}_n\right)\right) + \frac{h}{2}A\boldsymbol{K}_1 \\ + \frac{h}{2}\boldsymbol{f}\left(\boldsymbol{K}_1 + \sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j\right) + \frac{h^{3/2}}{2}A\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j \\ + \boldsymbol{H}(\boldsymbol{Y}_2) + \tilde{\boldsymbol{H}}(\boldsymbol{Y}_2, \boldsymbol{Y}_2). \end{aligned} \tag{3.6}$$

Since we can rewrite $\boldsymbol{H}(\boldsymbol{Y})$ as $\boldsymbol{H}(\boldsymbol{Y}) = h\sum_{j,k=1}^{m}\boldsymbol{g}_j'(\boldsymbol{Y})\boldsymbol{g}_k(\boldsymbol{Y})\zeta_{kj} + O(h^3)$, we have

$$\mathrm{e}^{\frac{h}{2}A}\boldsymbol{H}(\boldsymbol{Y}_2) = \boldsymbol{H}(\boldsymbol{Y}_2) + h^2\boldsymbol{r}_1 + O\left(h^3\right),$$

where $\boldsymbol{r}_1 = (1/2)A\sum_{j,k=1}^{m}\boldsymbol{g}_j'(\boldsymbol{Y}_2)\boldsymbol{g}_k(\boldsymbol{Y}_2)\zeta_{kj}$. In addition, since we can rewrite $\tilde{\boldsymbol{H}}(\boldsymbol{Y}, \boldsymbol{Y})$ as

$$\begin{aligned} \tilde{\boldsymbol{H}}(\boldsymbol{Y}, \boldsymbol{Y}) = \sum_{j=1}^{m}\left\{\sqrt{h}\boldsymbol{g}_j\left(\boldsymbol{Y}\right) + \frac{h^{3/2}}{4}\sum_{k,l=1}^{m}\boldsymbol{g}_j''\left(\boldsymbol{Y}\right)\left[\boldsymbol{g}_k\left(\boldsymbol{Y}\right), \boldsymbol{g}_l\left(\boldsymbol{Y}\right)\right]\chi_k\chi_l\right\}\xi_j \\ + O\left(h^{5/2}\right), \end{aligned}$$

we have

$$\mathrm{e}^{\frac{h}{2}A}\tilde{\boldsymbol{H}}(\boldsymbol{Y}_2, \boldsymbol{Y}_2) = \frac{h^{3/2}}{2}A\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j + \tilde{\boldsymbol{H}}(\boldsymbol{Y}_2, \boldsymbol{Y}_2) + h^{5/2}\boldsymbol{r}_2 + O\left(h^3\right),$$

where

$$\boldsymbol{r}_2 = \frac{1}{8}A\sum_{j=1}^{m}\left\{\sum_{k,l=1}^{m}\boldsymbol{g}_j''\left(\boldsymbol{Y}_2\right)\left[\boldsymbol{g}_k\left(\boldsymbol{Y}_2\right), \boldsymbol{g}_l\left(\boldsymbol{Y}_2\right)\right]\chi_k\chi_l + A\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\right\}\xi_j.$$

If $\boldsymbol{g}_j \equiv \boldsymbol{0}$ for $j = 1, 2, \ldots, m$, then $\boldsymbol{y}_{n+1} - \hat{\boldsymbol{y}}_{n+1} = O\left(h^3\right)$ as (3.4) and (3.6) are of order 2 for semilinear ODEs. Hence, all that remains concerning the local error is to check the difference between $\boldsymbol{u}$ and $\hat{\boldsymbol{u}}$ given by

$$\boldsymbol{u} = h\varphi_2(hA)\left\{\boldsymbol{f}\left(\boldsymbol{Y}_1 + \sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j\right) - \boldsymbol{f}\left(\boldsymbol{Y}_1\right)\right\}$$

and

$$\hat{\boldsymbol{u}} = \frac{h}{2}\left\{\boldsymbol{f}\left(\boldsymbol{K}_1 + \sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j\right) - \boldsymbol{f}\left(\boldsymbol{K}_1\right)\right\}. \tag{3.7}$$

As $\boldsymbol{Y}_1 = \boldsymbol{K}_1 + O\left(h^2\right)$, we have

$$\boldsymbol{u} - \hat{\boldsymbol{u}} = h^{5/2}\boldsymbol{r}_3 + O\left(h^3\right),$$

where $\boldsymbol{r}_3 = (1/6)\sum_{j=1}^{m} A\boldsymbol{f}'\left(\boldsymbol{K}_1\right)\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j$. Since $E[\boldsymbol{r}_1] = E[\boldsymbol{r}_2] = E[\boldsymbol{r}_3] = E[\xi_j\boldsymbol{r}_1] = \boldsymbol{0}$ $(j = 1, 2, \ldots, m)$, the local error of (3.4) is also weak order 3.

As a sufficient condition for (1) in Theorem 3.3, it is known that the following two inequalities hold for all sufficiently small $h > 0$:

$$\left\|E\left[\boldsymbol{y}_{n+1} - \boldsymbol{y}_n \mid \boldsymbol{y}_n\right]\right\| \le C\left(1 + \|\boldsymbol{y}_n\|\right)h, \qquad \left\|\boldsymbol{y}_{n+1} - \boldsymbol{y}_n\right\| \le X_n\left(1 + \|\boldsymbol{y}_n\|\right)\sqrt{h},$$

where C is a positive constant and X_n is a random variable which has moments of all orders [31, p. 102]. From the definition of $\boldsymbol{Y}_2$ in (3.4) and Assumption 3.1,

$$\begin{aligned}&\frac{1}{2}\left\|\boldsymbol{g}_j\left(\boldsymbol{Y}_2 + h\sum_{k=1}^{m}\boldsymbol{g}_k\left(\boldsymbol{Y}_2\right)\zeta_{kj}\right) - \boldsymbol{g}_j\left(\boldsymbol{Y}_2 - h\sum_{k=1}^{m}\boldsymbol{g}_k\left(\boldsymbol{Y}_2\right)\zeta_{kj}\right)\right\| \\ &\le C_1\left\|\boldsymbol{g}_j'\left(\boldsymbol{y}_n\right)h\sum_{k=1}^{m}\boldsymbol{g}_k\left(\boldsymbol{y}_n\right)\zeta_{kj}\right\| \le C_2\left(1 + \|\boldsymbol{y}_n\|\right)h\end{aligned}$$

for constants $C_1, C_2 > 0$. Here, note that the smoothness and global Lipschitzness of $\boldsymbol{g}_j, j = 1, 2, \ldots, m$, imply $\|\boldsymbol{g}_j'(\boldsymbol{y})\boldsymbol{g}_k(\boldsymbol{y})\| \le C(1 + \|\boldsymbol{y}\|)$ for a constant $C > 0$, whereas the global Lipschitzness implies $\|\boldsymbol{g}_j(\boldsymbol{y})\| \le C(1 + \|\boldsymbol{y}\|)$. From these facts, we can see that the two inequalities requested above hold for (3.4). Consequently, (3.4) is of weak order 2 by Theorem 3.3. ◻

If an SRK method is of higher deterministic order, it can be expected to have a better approximation to the expectation of a solution for SDEs with small noise [22]. For this, as a slightly complicated case, let us consider the SERK methods given by

$$\begin{aligned}\boldsymbol{y}_{n+1} = \boldsymbol{Y}_1 &+ \frac{h}{\gamma c_1 + c_2}\varphi_2(hA)\left\{\gamma\boldsymbol{f}\left(\boldsymbol{Y}_3 + b_1\sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j\right)\right. \\ &\left. + \boldsymbol{f}\left(\boldsymbol{Y}_4 + b_2\sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j\right) - (\gamma + 1)\boldsymbol{f}(\boldsymbol{y}_n)\right\} \\ &+ \mathrm{e}^{\frac{h}{2}A}\left(\boldsymbol{H}(\boldsymbol{Y}_2) + \tilde{\boldsymbol{H}}(\boldsymbol{Y}_2, \boldsymbol{Y}_2)\right),\end{aligned} \tag{3.8}$$

where $\boldsymbol{Y}_1, \boldsymbol{Y}_2$ are given in (3.5),

$$\begin{aligned}\boldsymbol{Y}_3 &= \mathrm{e}^{c_1hA}\boldsymbol{y}_n + c_1h\varphi_1(c_1hA)\boldsymbol{f}(\boldsymbol{y}_n), \\ \boldsymbol{Y}_4 &= \mathrm{e}^{c_2hA}\boldsymbol{y}_n + h\left\{c_2\varphi_1(c_2hA) - \psi(hA)\right\}\boldsymbol{f}(\boldsymbol{y}_n) + h\psi(hA)\boldsymbol{f}(\boldsymbol{Y}_3),\end{aligned}$$

and b_1 and b_2 are parameters along with c_1, c_2, and γ satisfying (2.4). Observe also that if the diffusion terms vanish, (3.8) is equivalent to (2.3).

THEOREM 3.6. *Let $\boldsymbol{g}_0(\boldsymbol{y})$ denote $A\boldsymbol{y} + \boldsymbol{f}(\boldsymbol{y})$ and suppose that* (3.1) *satisfies Assumption* 3.1 *for $q = 2$. Then,* (3.8) *is of weak order* 2 *for* (3.1) *if the parameters satisfy*

$$\frac{\gamma b_1 + b_2}{\gamma c_1 + c_2} = 1, \qquad \frac{\gamma b_1^2 + b_2^2}{\gamma c_1 + c_2} = 1 \tag{3.9}$$

as well as (2.4).

For the proof of this theorem, we refer the reader to Appendix B.

Remark 3.7. As a simple solution of (2.4) and (3.9), we can find

$$c_1 = \frac{1}{2}, \quad c_2 = 1, \quad \gamma = 4, \quad b_1 = \frac{6 \pm \sqrt{6}}{10}, \quad b_2 = \frac{3 \mp 2\sqrt{6}}{5}$$

(double sign in order). For this solution, the intermediate values $\boldsymbol{Y}_3$ and $\boldsymbol{Y}_4$ satisfy

$$\boldsymbol{Y}_3 = \boldsymbol{Y}_2, \qquad \boldsymbol{Y}_4 = \boldsymbol{Y}_1 + h\psi(hA)\left\{\boldsymbol{f}\left(\boldsymbol{Y}_2\right) - \boldsymbol{f}\left(\boldsymbol{y}_n\right)\right\},$$

where $\boldsymbol{Y}_1, \boldsymbol{Y}_2$ are given in (3.5).

3.3. Splitting methods. In the previous subsection we derived our SERK methods by setting terms including exponentials in the intermediate values and by utilizing Lemma 3.4. Taking this into account, let us consider the following Strang splitting method for the weak second order approximation to the solution of (3.1):

$$\boldsymbol{y}_{n+1} = \mathrm{e}^{\frac{h}{2}A}\Phi_h\left(\mathrm{e}^{\frac{h}{2}A}\boldsymbol{y}_n\right), \tag{3.10}$$

where Φ_h is the DFMT method for SDEs given by replacing A with the zero matrix in (3.1). We name (3.10) the Strang splitting method based on the DFMT method, and call it the SSDFMT method.

THEOREM 3.8. *Let $\boldsymbol{g}_0(\boldsymbol{y})$ denote $A\boldsymbol{y} + \boldsymbol{f}(\boldsymbol{y})$ and suppose that* (3.1) *satisfies Assumption* 3.1 *for $q = 2$. Then,* (3.10) *is of weak order* 2 *for* (3.1).

In order to prove this theorem, we can utilize parts of the proof of Lemma 3.4. For details, refer to Appendix C.

The SSDFMT method is of order 2 for semilinear ODEs. In order to make it possible for splitting methods to have a higher deterministic order, let us consider splitting methods given by the following formulation:

$$\boldsymbol{y}_{n+1} = \Psi_{\frac{h}{2}}\left(\hat{\Phi}_h\left(\Psi_{\frac{h}{2}}(\boldsymbol{y}_n)\right)\right), \tag{3.11}$$

where $\hat{\Phi}_h$ is the DFMT method for SDEs given by making the drift term zero and where Ψ_h denotes an exponential integrator which at least satisfies

$$\Psi_h(\boldsymbol{y}_n) = \boldsymbol{y}_n + h\boldsymbol{g}_0(\boldsymbol{y}_n) + \frac{h^2}{2}\boldsymbol{g}_0'(\boldsymbol{y}_n)\boldsymbol{g}_0(\boldsymbol{y}_n) + O\left(h^3\right).$$

THEOREM 3.9. *Let $\boldsymbol{g}_0(\boldsymbol{y})$ denote $A\boldsymbol{y} + \boldsymbol{f}(\boldsymbol{y})$ and suppose that* (3.1) *satisfies Assumption* 3.1 *for $q = 2$. Then,* (3.11) *is of weak order* 2 *for* (3.1).

Proof. Due to Lemma 3.4, it is clear that the local error of the method

$$\hat{\boldsymbol{y}}_{n+1} = \tilde{\boldsymbol{y}}_{n+1} + \frac{h}{2}\boldsymbol{g}_0\left(\tilde{\boldsymbol{Y}} + \sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\tilde{\boldsymbol{Y}}\right)\xi_j\right) + \boldsymbol{H}\left(\tilde{\boldsymbol{Y}}\right) + \tilde{\boldsymbol{H}}\left(\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{Y}}\right) \tag{3.12}$$

is weak order 3, where $\tilde{\boldsymbol{y}}_{n+1} = \boldsymbol{y}_n + (h/2)\boldsymbol{g}_0(\boldsymbol{y}_n) + (h^2/4)\boldsymbol{g}_0'(\boldsymbol{y}_n)\boldsymbol{g}_0(\boldsymbol{y}_n)$ and $\tilde{\boldsymbol{Y}} = \Psi_{h/2}(\boldsymbol{y}_n)$. Incidentally, (3.11) can be rewritten as follows:

$$\begin{aligned}\boldsymbol{y}_{n+1} = \tilde{\boldsymbol{Y}} &+ \frac{h}{2}\boldsymbol{g}_0\left(\hat{\Psi}_h\left(\tilde{\boldsymbol{Y}}\right)\right) + \frac{h^2}{8}\boldsymbol{g}_0'\left(\hat{\Psi}_h\left(\tilde{\boldsymbol{Y}}\right)\right)\boldsymbol{g}_0\left(\hat{\Psi}_h\left(\tilde{\boldsymbol{Y}}\right)\right)\\ &+ \boldsymbol{H}\left(\tilde{\boldsymbol{Y}}\right) + \tilde{\boldsymbol{H}}\left(\tilde{\boldsymbol{Y}},\tilde{\boldsymbol{Y}}\right) + O\left(h^3\right).\end{aligned} \tag{3.13}$$

The last two terms in the right-hand side of (3.12) are the same as the fourth and fifth terms in the right-hand side of (3.13), respectively. In addition, as $\boldsymbol{y}_{n+1} - \hat{\boldsymbol{y}}_{n+1} = O(h^3)$ if $\boldsymbol{g}_j \equiv \boldsymbol{0}$ for $j = 1, 2, \ldots, m$, then all that remains concerning the local error is to check the difference between $\boldsymbol{u}$ and $\hat{\boldsymbol{u}}$ given by

$$\begin{aligned}\boldsymbol{u} = \frac{h}{2}&\left\{\boldsymbol{g}_0\left(\hat{\Psi}_h\left(\tilde{\boldsymbol{Y}}\right)\right) - \boldsymbol{g}_0\left(\tilde{\boldsymbol{Y}}\right)\right\} + \frac{h^2}{8}\left\{\boldsymbol{g}_0'\left(\hat{\Psi}_h\left(\tilde{\boldsymbol{Y}}\right)\right)\boldsymbol{g}_0\left(\hat{\Psi}_h\left(\tilde{\boldsymbol{Y}}\right)\right)\right.\\ &\left. - \boldsymbol{g}_0'\left(\tilde{\boldsymbol{Y}}\right)\boldsymbol{g}_0\left(\tilde{\boldsymbol{Y}}\right)\right\}\end{aligned}$$

and $\hat{\boldsymbol{u}} = (h/2)\{\boldsymbol{g}_0(\tilde{\boldsymbol{Y}} + \sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j(\tilde{\boldsymbol{Y}})\xi_j) - \boldsymbol{g}_0(\tilde{\boldsymbol{Y}})\}$. By seeking a Taylor expansion of $\hat{\Phi}_h(\tilde{\boldsymbol{Y}})$ centered at $\tilde{\boldsymbol{Y}}$ and by utilizing this, we obtain $\boldsymbol{u} - \hat{\boldsymbol{u}} = h^2\tilde{\boldsymbol{r}}_1 + h^{5/2}\tilde{\boldsymbol{r}}_2 + h^{5/2}\tilde{\boldsymbol{r}}_3 + O(h^3)$, where $\tilde{\boldsymbol{r}}_1 = \frac{1}{2}\sum_{j,k=1}^{m}\boldsymbol{g}_0'(\tilde{\boldsymbol{Y}})\boldsymbol{g}_j'(\tilde{\boldsymbol{Y}})\boldsymbol{g}_k(\tilde{\boldsymbol{Y}})\zeta_{kj}$,

$$\begin{aligned}\tilde{\boldsymbol{r}}_2 = \frac{1}{8}\sum_{j=1}^{m}&\left\{\sum_{k,l=1}^{m}\boldsymbol{g}_0'\left(\tilde{\boldsymbol{Y}}\right)\boldsymbol{g}_j''\left(\tilde{\boldsymbol{Y}}\right)\left[\boldsymbol{g}_k\left(\tilde{\boldsymbol{Y}}\right),\boldsymbol{g}_l\left(\tilde{\boldsymbol{Y}}\right)\right]\chi_k\chi_l\right.\\ &\left. + \boldsymbol{g}_0'\left(\tilde{\boldsymbol{Y}}\right)\boldsymbol{g}_0'\left(\tilde{\boldsymbol{Y}}\right)\boldsymbol{g}_j\left(\tilde{\boldsymbol{Y}}\right)\right\}\xi_j,\end{aligned}$$

and

$$\begin{aligned}\tilde{\boldsymbol{r}}_3 = \frac{1}{2}\sum_{j=1}^{m}&\left\{\frac{1}{4}\boldsymbol{g}_0''\left(\tilde{\boldsymbol{Y}}\right)\left[\boldsymbol{g}_0\left(\tilde{\boldsymbol{Y}}\right),\boldsymbol{g}_j\left(\tilde{\boldsymbol{Y}}\right)\right]\right.\\ &\left. + \sum_{k,l=1}^{m}\boldsymbol{g}_0''\left(\tilde{\boldsymbol{Y}}\right)\left[\boldsymbol{g}_j\left(\tilde{\boldsymbol{Y}}\right),\boldsymbol{g}_k'\left(\tilde{\boldsymbol{Y}}\right)\boldsymbol{g}_l\left(\tilde{\boldsymbol{Y}}\right)\right]\zeta_{lk}\right\}\xi_j.\end{aligned}$$

Since $E[\tilde{\boldsymbol{r}}_1] = E[\tilde{\boldsymbol{r}}_2] = E[\tilde{\boldsymbol{r}}_3] = E[\xi_j\tilde{\boldsymbol{r}}_1] = \boldsymbol{0}$ $(j = 1, 2, \ldots, m)$, the local error of (3.11) is also weak order 3. Thus we conclude this proof similarly to the end of the proof of Theorem 3.5. □

Remark 3.10. As $\hat{\Phi}_h(\boldsymbol{y}) = \boldsymbol{y}$ if $\boldsymbol{g}_j \equiv \boldsymbol{0}$ for $j = 1, 2, \ldots, m$, (3.11) can have the same deterministic order as Ψ_h.

4. MS stability analysis and error analysis. We investigate the stability properties of our SERK and splitting methods. In addition, we analyze their error in a small interval for a linear scalar SDE.

4.1. MS stability analysis for SERK methods and splitting methods. Let us consider the following scalar test SDE [16]:

$$\mathrm{d}y(t) = \lambda y(t)\mathrm{d}t + \sum_{j=1}^{m} \sigma_j y(t)\mathrm{d}W_j(t), \qquad t \geq 0, \qquad y(0) = y_0, \tag{4.1}$$

where $y_0 \neq 0$ with probability 1 (w.p.1) and where $\lambda, \sigma_j \in \mathbb{C}, 1 \leq j \leq m$, satisfy

$$2\Re(\lambda) + \sum_{j=1}^{m} \left|\sigma_j\right|^2 < 0. \tag{4.2}$$

Due to (4.2), the solution of (4.1) is MS-stable ($\lim_{t\to\infty} E[|y(t)|^2] = 0$).

When an SRK method is applied to (4.1), it is generally expressed by

$$y_{n+1} = R\left(h, \lambda, \{\sigma_j\}_{j=1}^{m}, \boldsymbol{\eta}\right) y_n,$$

where $\boldsymbol{\eta}$ is a vector whose components are random variables appearing in the method. The method is said to be MS-stable for particular h, λ, σ_j, $j = 1, 2, \ldots, m$, if

$$E\left[\left|R\left(h, \lambda, \{\sigma_j\}_{j=1}^{m}, \boldsymbol{\eta}\right)\right|^2\right] < 1,$$

which means that $E[|y_n|^2] \to 0$ $(n \to \infty)$ for the given h, λ, σ_j. Further, the method is said to be A-stable in the MS if it is MS-stable for any $h > 0$ when (4.2) holds [16].

THEOREM 4.1. *The SERK method* (3.4) *is A-stable in the MS for* (4.1).

Proof. If we apply (3.4) to (4.1), then we have

$$y_{n+1} = R\left(h, \lambda, \{\sigma_j\}_{j=1}^{m}, \{\xi_j\}_{j=1}^{m}, \{\zeta_{jk}\}_{j,k=1}^{m}\right) y_n,$$

where

$$R\left(h, \lambda, \{\sigma_j\}_{j=1}^{m}, \{\xi_j\}_{j=1}^{m}, \{\zeta_{jk}\}_{j,k=1}^{m}\right) = \mathrm{e}^{h\lambda}\left\{1 + \sqrt{h}\sum_{j=1}^{m}\sigma_j\xi_j + h\sum_{j,k=1}^{m}\sigma_j\sigma_k\zeta_{kj}\right\}.$$

From this, the MS stability function $\hat{R}$ of (3.4) is given by

$$\hat{R}(p_r, q) \overset{\mathrm{def}}{=} E\left[|R|^2\right] = \mathrm{e}^{2p_r}\left(1 + q + \frac{q^2}{2}\right),$$

where $p_r \overset{\mathrm{def}}{=} \Re(\lambda)h$ and $q \overset{\mathrm{def}}{=} \sum_{j=1}^{m} |\sigma_j|^2 h$. As we can rewrite (4.2) by $2p_r + q < 0$,

$$\hat{R}(p_r, q) < \mathrm{e}^{2p_r}(1 - 2p_r + 2p_r^2).$$

The function in the right-hand side is less than 1 for any $p_r < 0$. Thus, $\hat{R}(p_r, q) < 1$ whenever $2p_r + q < 0$. Consequently, (3.4) is A-stable in the MS. ∎

As a comparison, let us look at stability properties of the SROCK2 method [5]. When $m = 1$, its MS stability function is given by

$$\hat{R}(p, q) = |A(p)|^2 + |B(p)|^2 q + |C(p)|^2 \frac{q^2}{2},$$

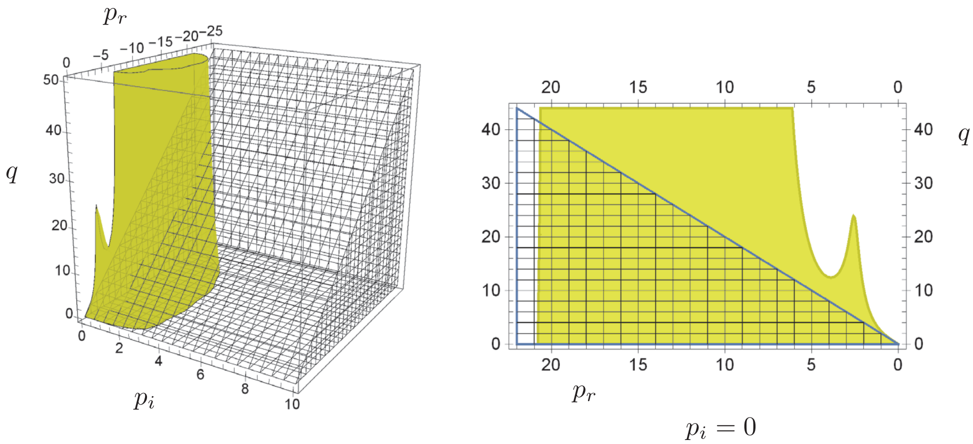

FIG. 1. *MS stability domain (left) and its profile (right) for the SROCK2 method with six stages.*

where $p \stackrel{\text{def}}{=} \lambda h$ and $A(p), B(p)$, and $C(p)$ are polynomial functions of p. For details, see [5]. Now, we can plot the MS stability domain, that is, $\{(p, q) \mid \hat{R}(p, q) < 1\}$. For the SROCK2 method with six stages, the MS stability domain and its profile are given in Figure 1. The MS stability domain is indicated by the colored part in the left-hand plot, and p_i denotes $\Im(\lambda)h$. The other part enclosed by the mesh indicates the domain in which the solution of the test SDE is MS-stable. In the right-hand plot, the colored area indicates the profile of the MS stability domain when $p_i = 0$. We can see that the MS stability domain is large along the negative axis of p_r, but it is thin in the axis of p_i. This stems from the fact that the SROCK2 method has been designed to achieve a large stability domain when the eigenvalues of the drift term lie near the negative real axis.

Remark 4.2. Let us consider d-dimensional SDEs of the following form:

$$\mathrm{d}\boldsymbol{y}(t) = A\boldsymbol{y}(t)\mathrm{d}t + \sum_{j=1}^{m} B_j \boldsymbol{y}(t)\mathrm{d}W_j(t), \qquad t \geq 0, \qquad \boldsymbol{y}(0) = \boldsymbol{y}_0, \tag{4.3}$$

where $\boldsymbol{y}_0 \neq \mathbf{0}$ (w.p.1) and where A, B_j $(j = 1, 2, \ldots, m) \in \mathbb{R}^{d\times d}$. When applied to this, the methods (3.4), (3.8), (3.10), and (3.11) with (2.2) or (2.3) as Ψ_h lead to

$$\boldsymbol{y}_{n+1} = \left\{ \mathrm{e}^{hA} + \sqrt{h}\mathrm{e}^{\frac{h}{2}A} \sum_{j=1}^{m} B_j \xi_j \mathrm{e}^{\frac{h}{2}A} + h\mathrm{e}^{\frac{h}{2}A} \sum_{j,k=1}^{m} B_j B_k \zeta_{kj} \mathrm{e}^{\frac{h}{2}A} \right\} \boldsymbol{y}_n.$$

For this, they have the same stability properties not only for (4.1), but also for (4.3). Thus, for example, the stability properties of these methods are equivalent for all the two-dimensional test SDEs in [8, 37].

4.2. Error analysis in a small interval. Let us consider the linear scalar SDE given by

$$\mathrm{d}y(t) = \lambda y(t)\mathrm{d}t + \sigma y(t)\mathrm{d}W(t), \qquad 0 \leq t \leq T_0 < 1, \qquad y(0) = y_0, \tag{4.4}$$

where $y_0 \neq 0$ (w.p.1) and $\lambda, \sigma \in \mathbb{R}$. For the solution of this SDE, we have

$$E\left[y(T_0)\right] = E[y_0]\mathrm{e}^{\lambda T_0}, \qquad E\left[(y(T_0))^2\right] = E[(y_0)^2]\mathrm{e}^{(2\lambda+\sigma^2)T_0}.$$

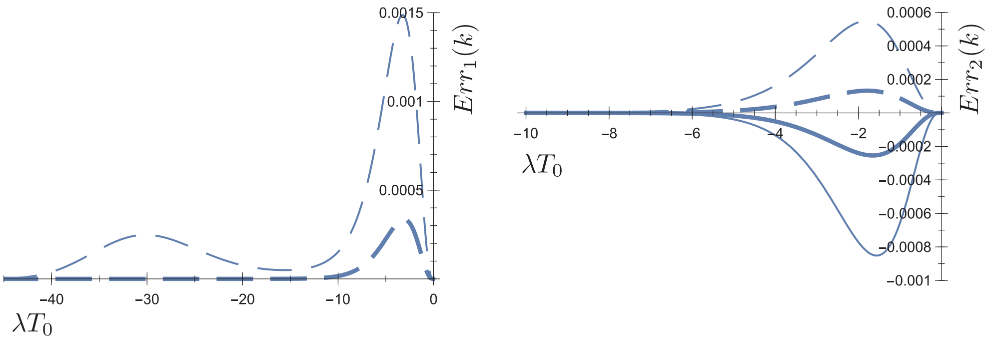

FIG. 2. *Error versus λT_0 for the expectation and the second moment of the amplification factor of methods with $h = T_0/2^k$. Solid lines: SERK, splitting methods ($k = 0$); thick lines: SERK, splitting methods ($k = 1$); dashed lines: SROCK2 ($k = 3$); thick dashed lines: SROCK2 ($k = 4$).*

On the other hand, if we set $h = T_0$ and apply our methods or the SROCK2 method to (4.4), then for an approximation to $y(T_0)$ we have

$$E[y_1] = E[y_0]\bar{R}(\lambda T_0), \qquad E\left[(y_1)^2\right] = E\left[(y_0)^2\right]\hat{R}(\lambda T_0, \sigma^2 T_0),$$

where $\bar{R}(\lambda T_0) \stackrel{\text{def}}{=} E[R] = \mathrm{e}^{\lambda T_0}$ for our methods or $A(\lambda T_0)$ for the SROCK2 method. Thus, as errors of an amplification factor R of a method we can consider

$$\bar{R}(\lambda T_0) - \mathrm{e}^{\lambda T_0}, \qquad \hat{R}(\lambda T_0, \sigma^2 T_0) - \mathrm{e}^{(2\lambda+\sigma^2)T_0}.$$

In general, for $h = T_0/2^k$ (k is a positive integer) we have

$$Err_1(k) \stackrel{\text{def}}{=} \left\{\bar{R}(\lambda T_0/2^k)\right\}^{2^k} - \mathrm{e}^{\lambda T_0},$$
$$Err_2(k) \stackrel{\text{def}}{=} \left\{\hat{R}(\lambda T_0/2^k, \sigma^2 T_0/2^k)\right\}^{2^k} - \mathrm{e}^{(2\lambda+\sigma^2)T_0}.$$

In Figure 2, these errors are indicated for our methods and for the SROCK2 method with six stages. The solid and thick lines denote our methods for $k = 0$ and $k = 1$, whereas the dashed and thick dashed lines denote the SROCK2 method with six stages for $k = 3$ and $k = 4$. In the right-hand plot, we deal with a case in which $\sigma^2/\lambda = -3/10$. In this case, our methods with step size h can be expected to have similar precision to the SROCK2 method with $h/8$ in the approximation to the second moment of the solution. This will be numerically checked later.

Incidentally, although there are no differences between our methods in the linear case as we have seen, the situation is quite different for nonlinear cases. For example, when diffusion terms vanish, our SERK methods lead to the exponential RK method (2.2) or (2.3) with step size h. On the other hand, our splitting methods with h lead to methods that twice apply (2.2) or (2.3) with $h/2$ if Ψ_h is (2.2) or (2.3) in (3.11). Taking these facts into account, we will use one of our SERK methods with $h_{min}/4$ to obtain a reference solution in a numerical experiment, where h_{min} is the minimum step size used for our methods in the numerical experiment.

5. Numerical experiments. In section 3, we derived our SERK and splitting methods. For example, (3.4) is an SERK method of weak order 2 and deterministic order 2. In what follows, let us call this the SERKW2D2 method. As we have seen in Remark 3.7, (3.8) with $c_1 = 1/2, c_2 = 1, \gamma = 4, b_1 = (6+\sqrt{6})/10$, and $b_2 = (3-2\sqrt{6})/5$ is an SERK method of weak order 2 and deterministic order 3. Let us call this the SERKW2D3 method. If Ψ_h is (2.2) with $c = 1$, then (3.11) is a splitting method of weak order 2 and deterministic order 2. We call this the SPLITW2D2 method. Similarly, we call (3.11) the SPLITW2D3 method if Ψ_h is (2.3) with the parameter values mentioned above. As an implementation of the SROCK2 method, we do not directly use the Fortran codes from http://anmc.epfl.ch/Pdf/srock2.zip, but have implemented C codes by including rectp.f from the Fortran codes. Thus, the SROCK2 method in our C codes has the same parameter values as in the Fortran codes.

In order to confirm the performance of the methods, we investigate the expectation and/or the second moment of the solution of SDEs in our numerical experiments. As a first example, let us consider the stochastic Verhulst equation

$$\mathrm{d}y(t) = \left\{\alpha y(t) - \gamma (y(t))^2\right\} \mathrm{d}t + \beta y(t) \mathrm{d}W(t), \quad y(0) = y_0 \text{ (w.p.1)}, \tag{5.1}$$

where $t \in [0, 1/2]$ and $\alpha, \beta, \gamma, y_0 \in \mathbb{R}$. The solution is given as [21, p. 125]

$$y(t) = \frac{y_0 \exp\left(\left(\alpha - \frac{\beta^2}{2}\right) t + \beta W(t)\right)}{1 + y_0 \gamma \int_0^t \exp\left(\left(\alpha - \frac{\beta^2}{2}\right) s + \beta W(s)\right) \mathrm{d}s},$$

but we do not know the exact value of its expectation and second moment. For this, we seek approximations to them by the SERKW2D3 method with $h = 2^{-7}$ and refer to these approximations instead of the exact expectation and second moment. We use the Mersenne twister algorithm [28] to generate pseudorandom numbers.

Let us set $y_0 = 1$, $\alpha = -3$, $\beta = 3/\sqrt{10}$, and $\gamma = 1/25$. In this example, we simulate 16384×10^6 independent trajectories for a given h, and seek a numerical approximation to the expectation or the second moment of $y(1/2)$. The results are indicated in Figure 3. The solid, dash-dotted, long-dashed, thick, thick long-dashed, thick dashed, and dashed lines denote the SERKW2D2 method, the SERKW2D3 method, the DFMT method, the SSDFMT method, the SPLITW2D2 method, the SPLITW2D3 method, and the SROCK2 method with six stages, respectively. Here and in what follows, the dotted line is a reference line with slope 2. All the methods show the theoretical order of convergence. In the linear case (4.4), our exponential methods have the same error in the approximation to the second moment, but the existence of a nonlinear term in (5.1) makes a difference. The SSDFMT method is

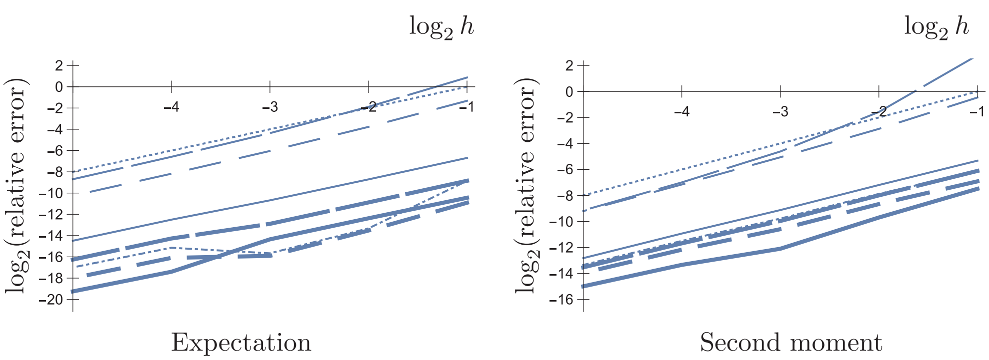

Fig. 3. *Log-log plots of the relative error versus h for the expectation and the second moment in* (5.1). *Solid lines: SERKW2D2; dash-dotted lines: SERKW2D3; long-dashed lines: DFMT; thick lines: SSDFMT; thick long-dashed lines: SPLITW2D2; thick dashed lines: SPLITW2D3; dashed lines: SROCK2; dotted lines: reference line with slope 2.*

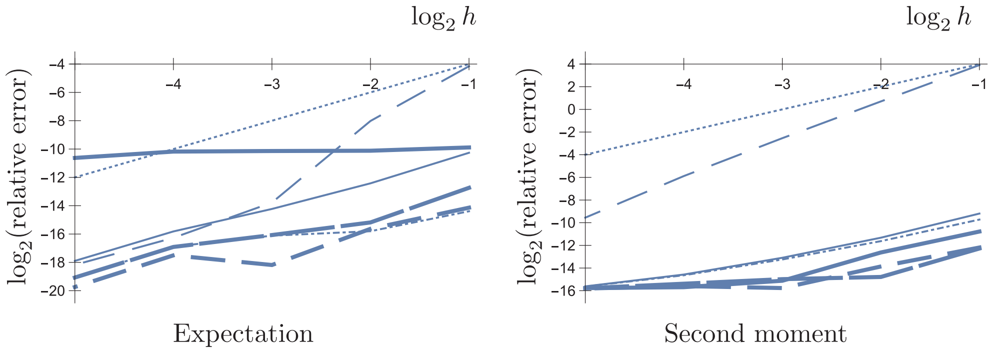

FIG. 4. *Log-log plots of the relative error versus h for the expectation and the second moment in* (5.2). *Solid lines: SERKW2D2; dash-dotted lines: SERKW2D3; thick lines: SSDFMT; thick long-dashed lines: SPLITW2D2; thick dashed lines: SPLITW2D3; dashed lines: SROCK2; dotted lines: reference line with slope* 2.

the best. In addition, as we saw in subsection 4.2, we can observe that the SROCK2 method needs a small step size $h/8$ to achieve a similar precision to our exponential methods with h.

As a second example, let us consider the following mildly stiff noncommutative SDE, which is obtained by adding a nonlinear term to (36) in [11] and by making small changes:

$$\begin{aligned} \mathrm{d}\boldsymbol{y}(t) &= \left\{ \begin{bmatrix} -\frac{273}{1024} & 0 \\ -\frac{1}{160} & -\frac{785}{4} + \frac{\sqrt{2}}{8} \end{bmatrix} \boldsymbol{y}(t) - \begin{bmatrix} \frac{1}{10}\mathrm{e}^{y_1(t)} \\ \frac{1}{10}\mathrm{e}^{y_2(t)} \end{bmatrix} \right\} \mathrm{d}t \\ &\quad + \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1-2\sqrt{2}}{4} \end{bmatrix} \boldsymbol{y}(t)\mathrm{d}W_1(t) + \begin{bmatrix} \frac{1}{16} & 0 \\ \frac{1}{10} & \frac{1}{16} \end{bmatrix} \boldsymbol{y}(t)\mathrm{d}W_2(t), \\ \boldsymbol{y}(0) &= [1\ 1]^\top \text{ (w.p.1)}, \end{aligned} \tag{5.2}$$

where $t \in [0, 1]$. We seek an approximation to the expectation of $\boldsymbol{y}(1)$ or to the second moment of each element of $\boldsymbol{y}(1)$, that is, $[E[(y_1(1))^2] \;\; E[(y_2(1))^2]]^\top$. As we do not know the exact solution of this SDE, we seek numerical approximations by the SERKW2D3 method with $h = 2^{-7}$ and use them instead of the exact expectation and second moment. In this example, we simulate 4096×10^6 independent trajectories for a given h. The results are indicated in Figure 4. As the solution is a vector, the Euclidean norm is used. Similarly to the first example, the solid, dash-dotted, thick, thick long-dashed, thick dashed, and dashed lines denote the SERKW2D2 method, the SERKW2D3 method, the SSDFMT method, the SPLITW2D2 method, the SPLITW2D3 method, and the SROCK2 method with 15 stages, respectively. Incidentally, the DFMT method needs a smaller step size than 2^{-5} to solve (5.2) numerically stably. The SPLITW2D3 method shows the best accuracy in approximations to the expectation. On the other hand, the SSDFMT method suffers from order reduction in approximations to the expectation, and the SROCK2 method shows large relative errors in approximations to the second moment.

In order to deal with the stiff case, let us consider the following SDE:

$$\begin{aligned} \mathrm{d}\boldsymbol{y}(t) &= \begin{bmatrix} \alpha & 1 \\ -\omega^2 & \alpha \end{bmatrix} \boldsymbol{y}(t)\mathrm{d}t + \begin{bmatrix} \beta & 0 \\ 0 & \beta \end{bmatrix} \boldsymbol{y}(t)\mathrm{d}W(t), \quad t \geq 0, \\ \boldsymbol{y}(0) &= [1\ 1]^\top \text{ (w.p.1)} \end{aligned} \tag{5.3}$$

TABLE 1
Step size for numerical stability in (5.3).

	Method	Step size	Absolute errors
Case 1	SROCK2 (10 stages)	$h = 2^{-1}$	9.1×10^{-5} (stable)
	SERKW2D2	$h = 2^{-1}$	9.1×10^{-5} (stable)
Case 2	SROCK2 (3 stages)	$h = 2^{-9}$	1.9 (stable)
	SROCK2 (all stages)	$h = 2^{-8}$	∞ (unstable)
	SERKW2D2	$h = 2^{-1}$	4.1×10^{-5} (stable)
Case 3	SROCK2 (3 stages)	$h = 2^{-7}$	9.1×10^{-5} (stable)
	SROCK2 (5 stages)	$h = 2^{-6}$	9.1×10^{-5} (stable)
	SROCK2 (all stages)	$h = 2^{-5}$	∞ (unstable)
	SERKW2D2	$h = 2^{-1}$	9.1×10^{-5} (stable)

for $\alpha, \omega, \beta \in \mathbb{R}$. Since the eigenvalues of the matrix in the drift term are $\alpha \pm i\omega$, $\lim_{t\to\infty} E[\|\boldsymbol{y}(t)\|^2] = 0$ holds if $2\alpha + \beta^2 < 0$. We investigate three cases:

$$\begin{aligned} &\text{Case 1:}\quad \alpha = -100,\ \omega = 1,\ \beta = \sqrt{199}, \qquad \text{Case 2:}\quad \alpha = -\tfrac{1}{4},\ \omega = 30\pi,\ \beta = \tfrac{1}{4},\\ &\text{Case 3:}\quad \alpha = -100,\ \omega = 30\pi,\ \beta = \sqrt{199}. \end{aligned}$$

In this example, we simulate 1×10^6 independent trajectories for a given h until $t = 10$ and we seek numerical solutions to $E[\|\boldsymbol{y}(10)\|^2]$ by the SROCK2 and SERKW2D2 methods. Note that the SERKW2D2 method and our other methods are equivalent for (5.3) due to Remark 4.2. For the solution $\boldsymbol{y}(t)$ in each case, we have

$$\begin{aligned} &\text{Case 1:}\quad E\left[(y_1(10))^2\right] = \{1 + \sin(20)\}\mathrm{e}^{-20}, \quad E\left[(y_2(10))^2\right] = \{1 - \sin(20)\}\mathrm{e}^{-20},\\ &\text{Case 2:}\quad E\left[(y_1(10))^2\right] = E\left[(y_2(10))^2\right] = \mathrm{e}^{-35/8},\\ &\text{Case 3:}\quad E\left[(y_1(10))^2\right] = E\left[(y_2(10))^2\right] = \mathrm{e}^{-10}. \end{aligned}$$

Table 1 gives numerical results, which indicate how the small step size is necessary for each method to solve (5.3) numerically stably. In Case 1, the SROCK2 method with 10 stages can solve it for $h = 2^{-1}$, but with fewer than 10 stages it cannot. In Case 2, the SROCK2 method cannot solve the SDE for $h = 2^{-8}$ even if we make the stage number large. This is understandable, since increasing the stage number does not result in a large enough MS stability domain in the axis of p_i. In Case 3, the SROCK2 method with three stages cannot solve the SDE for $h = 2^{-6}$, but with five stages it can. However, for $h = 2^{-5}$ the SROCK2 method cannot solve it by making the stage number large. The results of the SROCK2 method in Cases 2 and 3 can be explained by the argument we presented before Remark 4.2. On the other hand, the SERKW2D2 method can solve for $h = 2^{-1}$ in all cases.

The fourth example comes from a stochastic Burgers equation with white noise in time only. Da Prato and Gatarek [10] have proved the existence and uniqueness of the global solution of a scalar Burgers equation with multiplicative noise driven by a scalar Wiener process. Now, we consider

(5.4)

$$\begin{aligned} &\mathrm{d}u(t,x) = \left(\frac{\partial^2 u}{\partial x^2}(t,x) + u(t,x)\frac{\partial u}{\partial x}(t,x)\right)\mathrm{d}t + ku(t,x)\mathrm{d}W(t), \quad (t,x) \in [0,1/2] \times [0,1],\\ &u(t,0) = u(t,1) = \alpha \ (\text{w.p.1}), \qquad t \in [0,1/2],\\ &u(0,x) = 2\sin(\pi x) + \alpha \ (\text{w.p.1}), \qquad x \in [0,1], \end{aligned}$$

where $k, \alpha \in \mathbb{R}$. If we discretize the space interval by $N+2$ equidistant points x_i $(0 \le i \le N+1)$ and define a vector-valued function by $\boldsymbol{y}(t) \stackrel{\text{def}}{=} [u(t,x_1)\ u(t,x_2), \ldots, u(t,x_N)]^\top$, then we obtain the SDE

$$
\begin{aligned}
&\mathrm{d}\boldsymbol{y}(t) = \big(A\boldsymbol{y}(t) + \boldsymbol{f}(\boldsymbol{y}(t))\big)\mathrm{d}t + k\boldsymbol{y}(t)\mathrm{d}W(t),\\
&\boldsymbol{y}(0) = [2\sin(\pi x_1) + \alpha\ 2\sin(\pi x_2) + \alpha, \ldots,\ 2\sin(\pi x_N) + \alpha]^\top \text{ (w.p.1)}
\end{aligned}
\tag{5.5}
$$

by applying the central difference scheme to (5.4), where

$$
A \stackrel{\text{def}}{=} (N+1)^2 \operatorname{tridiag}(1, -2, 1),
$$

$$
\boldsymbol{f}(\boldsymbol{y}) \stackrel{\text{def}}{=} \frac{N+1}{2} \begin{bmatrix} y_1(y_2 - \alpha) \\ y_2(y_3 - y_1) \\ \vdots \\ y_{N-1}(y_N - y_{N-2}) \\ y_N(\alpha - y_{N-1}) \end{bmatrix} + (N+1)^2 \begin{bmatrix} \alpha \\ 0 \\ \vdots \\ 0 \\ \alpha \end{bmatrix}.
$$

In this example, we set $N = 127$ and $k = 1/10$, and simulate 256×10^6 independent trajectories for a given h. For $\alpha = 1/10$ we seek an approximation to the expectation or the second moment of $\boldsymbol{y}(1/2)$. As we do not know the exact solution of the SDE, we seek numerical approximations by the SERKW2D3 method with $h = 2^{-8}$ and use them instead of the exact expectation and second moment. In order to solve the SDE numerically stably with reasonable cost by the SROCK2 method, we set the stage numbers of the method as 200, 150, 104, 74, 49, 35, 24, 17, and 12 corresponding to the step sizes $2^{-2}, 2^{-3}, 2^{-4}, 2^{-5}, 2^{-6}, 2^{-7}, 2^{-8}, 2^{-9}$, and 2^{-10}, respectively. The results are indicated in Figure 5. The solid, dash-dotted, thick, thick long-dashed, thick dashed, and dashed lines denote the SERKW2D2 method, the SERKW2D3 method, the SSDFMT method, the SPLITW2D2 method, the SPLITW2D3 method, and the SROCK2 method, respectively. The figure indicates that the SSDFMT method suffers from order reduction [13], and the SROCK2 method has large relative errors especially in approximations to the second moment. The SERKW2D3 method and the SPLITW2D3 method show higher convergence rates in approximations to the expectation.

In order to avoid the order reduction of the SSDFMT method, we can use the change of variable in this example. When we define $v(t,x) \stackrel{\text{def}}{=} u(t,x) - \alpha$ and solve

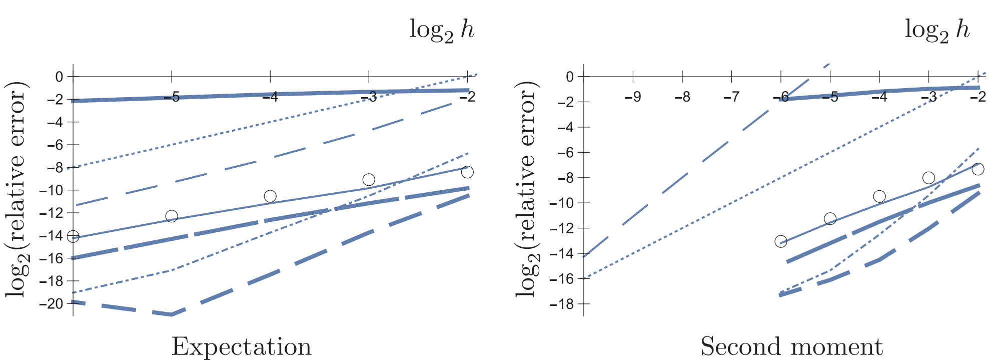

FIG. 5. *Log-log plots of the relative error versus h for the expectation and the second moment in* (5.5) *when $\alpha = 1/10$. Solid lines: SERKW2D2; dash-dotted lines: SERKW2D3; thick lines: SSDFMT; thick long-dashed lines: SPLITW2D2; thick dashed lines: SPLITW2D3; dashed lines: SROCK2; dotted lines: reference line with slope 2; white circles: modified SSDFMT.*

TABLE 2
Comparisons of computational cost in one step and one trajectory.

Method	n_e	n_r	n_m
SROCK2 with s stages	$s + 5m + 2$	$2m$	0
SERKW2D2	$5m + 2$	$2m$	6
SERKW2D3	$5m + 4$	$2m$	7
SSDFMT	$5m + 2$	$2m$	2
SPLITW2D2	$5m + 4$	$2m$	6
SPLITW2D3	$5m + 6$	$2m$	12

another SDE by the SSDFMT method, the order reduction does not occur. The results are indicated by white circles in the figure. However, such a remedy is not always available if the SDE comes from a multidimensional stochastic partial differential equation instead of a scalar equation such as (5.4) [13].

Finally, Table 2 indicates comparisons of computational cost for each method in one step and one trajectory. In the table, n_e, n_r, and n_m stand for the number of evaluations on the drift or diffusion terms, the number of generated pseudorandom numbers, and the number of products of a matrix exponential function with a vector, respectively. For stiff problems whose drift term has eigenvalues lying near the negative real axis such as the fourth experiment, the computational cost of the SROCK2 method increases as the stage number s needs to increase for a large step size. On the other hand, in high dimensional problems, the computational cost of our SERK and splitting methods increases due to the products of matrix exponential functions with vectors. Their computational cost for the products can be significant if SDEs come from multidimensional stochastic partial differential equations. The column for n_m in the table indicates how much each method owes the computational cost for the products of matrix exponential functions with vectors. Note especially that the SROCK2 method has no cost assigned. We will mention this again in the final section.

6. Concluding remarks. We have derived explicit SERK methods and splitting methods that achieve weak order 2 for noncommutative Itô SDEs with a semilinear drift term, and simultaneously achieve order 2 or 3 for ODEs. Using a scalar test SDE with complex coefficients, we have investigated the stability properties of the methods. As a result, we have proved that they are A-stable in the MS for the test SDE. To the best of our knowledge, there seems to be no weak second order method for which the A-stability in the MS is proven using the test SDE with complex coefficients, except for a drift-implicit method of weak order 2 and deterministic order 2 in [4]. In addition, as an example of an explicit stabilized method, we have selected the SROCK2 method and plotted its MS stability domain.

In order to check the numerical performance of the methods as well as their stability properties, we performed four numerical experiments. In the first experiment, the stochastic Verhulst equation was considered. The experiment confirmed the theoretical convergence of weak order 2 for our methods as well as for the DFMT method and the SROCK2 method. In the second experiment, a mildly stiff noncommutative SDE was considered. The experiment showed the superiority of our SERK methods and our splitting methods, combined with exponential RK methods, over the SSDFMT method and the SROCK2 method in approximations to the expectation and the second moment, respectively. In the third experiment, we dealt with three types of the stiff case. The experiment indicates that if the imaginary part of the eigenvalues in the drift term is large, then the SROCK2 method needs a very small step size for stability, whereas the SERK and splitting methods do not need such a step size restriction. In the final experiment, we considered a stochastic Burgers equation with white noise, and compared our methods with the SROCK2 method with several stages. This experiment showed the superiority of our SERK and splitting methods over the SROCK2 method in terms of computational accuracy for relatively large step size. It also indicated that the SSDFMT method suffers from order reduction for nonhomogeneous boundary value problems. In this example, the order reduction was successfully removed by a change of variable, but such a remedy is not always available if the SDE comes from a multidimensional stochastic partial differential equation. In the deterministic case, some techniques have been proposed to avoid order reduction [13]. In the stochastic case, a new analysis may be necessary, but it is outside the scope of the present paper.

Finally, we should make the following remarks. As we have seen, we can apply our methods to SDEs with a semilinear drift term and they have very good performance if the stiffness of the problem is in the matrix A as opposed to the nonlinear function $\boldsymbol{f}$. The SROCK2 method is applicable to more general SDEs without such a restriction and it can also cope with stiff problems by increasing the stage number. When the dimension of a system of SDEs is not large and the stiffness is very strong, our methods have a significant advantage over the SROCK2 method. This is because the method has to increase the stage number significantly, which leads to high computational cost. On the other hand, when the dimension of the SDEs is very large, the SROCK2 method can still cope with high-dimensional stiff SDEs by just increasing the stage number, but our methods need techniques in order to calculate matrix exponentials efficiently, such as Krylov methods and other methods for their computation (see [20] and the references therein). Although we have not used such approaches for matrix exponentials in this paper, the application of these techniques will have a considerable impact on our methods to challenge very high-dimensional SDEs with a semilinear drift term.

Appendix A. Proof of Lemma 3.4. From the assumption, we can write $\boldsymbol{Y}_i, i = 1, 2, \ldots, 5$, as $\boldsymbol{Y}_i = \boldsymbol{y}_n + h\boldsymbol{a}_i + O\left(h^2\right)$, $i = 1, 2, 3, 5$, and $\boldsymbol{Y}_4 = \boldsymbol{y}_n + (h/2)\boldsymbol{g}_0\left(\boldsymbol{y}_n\right) + h^2\boldsymbol{a}_4 + O\left(h^3\right)$, where $\boldsymbol{a}_i, i = 1, 2, \ldots, 5$, are vectors independent of h. As

$$
\begin{aligned}
&\boldsymbol{g}_0\left(\boldsymbol{Y}_1 + \sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j\right)\\
&= \boldsymbol{g}_0\left(\boldsymbol{Y}_1\right) + \sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_0'\left(\boldsymbol{y}_n\right)\boldsymbol{g}_j\left(\boldsymbol{y}_n\right)\xi_j + \frac{h}{2}\sum_{j,k=1}^{m}\boldsymbol{g}_0''\left(\boldsymbol{y}_n\right)\left[\boldsymbol{g}_j(\boldsymbol{y}_n), \boldsymbol{g}_k(\boldsymbol{y}_n)\right]\xi_j\xi_k\\
&\quad + \frac{h^{3/2}}{6}\sum_{j,k,l=1}^{m}\boldsymbol{g}_0'''\left(\boldsymbol{y}_n\right)\left[\boldsymbol{g}_j(\boldsymbol{y}_n), \boldsymbol{g}_k(\boldsymbol{y}_n), \boldsymbol{g}_l(\boldsymbol{y}_n)\right]\xi_j\xi_k\xi_l\\
&\quad + h^{3/2}\sum_{j=1}^{m}\boldsymbol{g}_0'\left(\boldsymbol{y}_n\right)\boldsymbol{g}_j'\left(\boldsymbol{y}_n\right)\boldsymbol{a}_2\xi_j + h^{3/2}\sum_{j=1}^{m}\boldsymbol{g}_0''\left(\boldsymbol{y}_n\right)\left[\boldsymbol{a}_1, \boldsymbol{g}_j(\boldsymbol{y}_n)\right]\xi_j + O\left(h^2\right),
\end{aligned}
$$

we have, from (3.3),

$$
\begin{aligned}
&\tilde{\boldsymbol{y}}_{n+1} + \frac{h}{2}\boldsymbol{g}_0\left(\boldsymbol{Y}_1 + \sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j\right)\\
&= \boldsymbol{y}_n + \frac{h}{2}\left(\boldsymbol{g}_0\left(\boldsymbol{y}_n\right) + \boldsymbol{g}_0\left(\boldsymbol{K}_2\right)\right) + h^{5/2}\boldsymbol{r}_1 + O\left(h^3\right),
\end{aligned}
\tag{A.1}
$$

where $\boldsymbol{r}_1 = (1/2)\sum_{j=1}^{m}\left\{\boldsymbol{g}_0'\left(\boldsymbol{y}_n\right)\boldsymbol{g}_j'\left(\boldsymbol{y}_n\right)\boldsymbol{a}_2 + \boldsymbol{g}_0''\left(\boldsymbol{y}_n\right)\left[\boldsymbol{a}_1 - \boldsymbol{g}_0(\boldsymbol{y}_n), \boldsymbol{g}_j(\boldsymbol{y}_n)\right]\right\}\xi_j$. Similarly, we have

$$\boldsymbol{H}\left(\boldsymbol{Y}_3\right) - \boldsymbol{H}\left(\boldsymbol{y}_n\right) = h^2\boldsymbol{r}_2 + O\left(h^3\right), \tag{A.2}$$

where $\boldsymbol{r}_2 = \sum_{j,k=1}^{m}\left\{\boldsymbol{g}_j''\left(\boldsymbol{y}_n\right)\left[\boldsymbol{a}_3, \boldsymbol{g}_k\left(\boldsymbol{y}_n\right)\right] + \boldsymbol{g}_j'\left(\boldsymbol{y}_n\right)\boldsymbol{g}_k'\left(\boldsymbol{y}_n\right)\boldsymbol{a}_3\right\}\zeta_{kj}$. As

$$\begin{aligned}
&\sum_{j=1}^{m}\left\{\boldsymbol{g}_j\left(\boldsymbol{Y}_4 + \sqrt{\frac{h}{2}}\sum_{k=1}^{m}\boldsymbol{g}_k\left(\boldsymbol{Y}_5\right)\chi_k\right) + \boldsymbol{g}_j\left(\boldsymbol{Y}_4 - \sqrt{\frac{h}{2}}\sum_{k=1}^{m}\boldsymbol{g}_k\left(\boldsymbol{Y}_5\right)\chi_k\right)\right\}\xi_j \\
&= \sum_{j=1}^{m}\left\{\boldsymbol{g}_j\left(\boldsymbol{y}_n + \frac{h}{2}\boldsymbol{g}_0\left(\boldsymbol{y}_n\right) + \sqrt{\frac{h}{2}}\sum_{k=1}^{m}\boldsymbol{g}_k\left(\boldsymbol{y}_n\right)\chi_k\right)\right. \\
&\quad \left. + \boldsymbol{g}_j\left(\boldsymbol{y}_n + \frac{h}{2}\boldsymbol{g}_0\left(\boldsymbol{y}_n\right) - \sqrt{\frac{h}{2}}\sum_{k=1}^{m}\boldsymbol{g}_k\left(\boldsymbol{y}_n\right)\chi_k\right)\right\}\xi_j + 2h^2\boldsymbol{r}_3 + O\left(h^{5/2}\right)
\end{aligned}$$

where

$$\boldsymbol{r}_3 = \sum_{j=1}^{m}\left\{\boldsymbol{g}_j'\left(\boldsymbol{y}_n\right)\boldsymbol{a}_4 + \frac{1}{2}\sum_{k,l=1}^{m}\boldsymbol{g}_j''\left(\boldsymbol{y}_n\right)\left[\boldsymbol{g}_k\left(\boldsymbol{y}_n\right),\ \boldsymbol{g}_l'\left(\boldsymbol{y}_n\right)\boldsymbol{a}_5\right]\chi_k\chi_l\right\}\xi_j,$$

we have

$$\tilde{\boldsymbol{H}}\left(\boldsymbol{Y}_4, \boldsymbol{Y}_5\right) - \tilde{\boldsymbol{H}}\left(\frac{\boldsymbol{y}_n + \boldsymbol{K}_1}{2}, \boldsymbol{y}_n\right) = h^{5/2}\boldsymbol{r}_3 + O\left(h^3\right). \tag{A.3}$$

Thus, from (3.2) and $\hat{\boldsymbol{y}}_{n+1} - \boldsymbol{y}_{n+1} = h^2\boldsymbol{r}_2 + h^{5/2}\boldsymbol{r}_1 + h^{5/2}\boldsymbol{r}_3 + O\left(h^3\right)$, we have

$$\begin{aligned}
&G\left(\hat{\boldsymbol{y}}_{n+1}\right) - G\left(\boldsymbol{y}_{n+1}\right) \\
&= G'\left(\boldsymbol{y}_n\right)\left(\hat{\boldsymbol{y}}_{n+1} - \boldsymbol{y}_{n+1}\right) + h^{5/2}G''\left(\boldsymbol{y}_n\right)\left[\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{y}_n\right)\xi_j, \boldsymbol{r}_2\right] + O\left(h^3\right).
\end{aligned}$$

Consequently, we obtain $E\left[G\left(\hat{\boldsymbol{y}}_{n+1}\right)\right] - E[G(\boldsymbol{y}_{n+1})] = O(h^3)$ since $E[\boldsymbol{r}_1] = E[\boldsymbol{r}_2] = E[\boldsymbol{r}_3] = E[\xi_j\boldsymbol{r}_2] = 0 \quad (j = 1, 2, \ldots, m)$. ∎

Appendix B. Proof of Theorem 3.6. As methods (3.4) and (3.8) are the same except for the second term in the right-hand side, all that remains concerning the local error is to check the difference between $\boldsymbol{u}$ and $\hat{\boldsymbol{u}}$ given by

$$\begin{aligned}
\boldsymbol{u} = \frac{h}{\gamma c_1 + c_2}\varphi_2(hA)&\left\{\gamma\boldsymbol{f}\left(\boldsymbol{Y}_3 + b_1\sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j\right) - \gamma\boldsymbol{f}\left(\boldsymbol{Y}_3\right)\right. \\
&\left. + \boldsymbol{f}\left(\boldsymbol{Y}_4 + b_2\sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j\right) - \boldsymbol{f}\left(\boldsymbol{Y}_4\right)\right\}
\end{aligned}$$

and (3.7). As $\boldsymbol{Y}_3 = \boldsymbol{K}_1 + (c_1 - 1)h\boldsymbol{g}_0\left(\boldsymbol{y}_n\right) + O\left(h^2\right)$ and $\boldsymbol{Y}_4 = \boldsymbol{K}_1 + (c_2 - 1)h\boldsymbol{g}_0\left(\boldsymbol{y}_n\right) + O\left(h^2\right)$ where $\boldsymbol{K}_1 = \boldsymbol{y}_n + h\boldsymbol{g}_0(\boldsymbol{y}_n)$, using (3.9) we have

$$\frac{1}{\gamma c_1 + c_2}\left\{\gamma \boldsymbol{f}\left(\boldsymbol{Y}_3 + b_1\sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j\right) - \gamma\boldsymbol{f}\left(\boldsymbol{Y}_3\right)\right.$$
$$\left. + \boldsymbol{f}\left(\boldsymbol{Y}_4 + b_2\sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j\right) - \boldsymbol{f}\left(\boldsymbol{Y}_4\right)\right\}$$
$$= \sqrt{h}\sum_{j=1}^{m}\boldsymbol{f}'\left(\boldsymbol{K}_1\right)\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\xi_j + \frac{h}{2}\sum_{j,k=1}^{m}\boldsymbol{f}''\left(\boldsymbol{K}_1\right)\left[\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right), \boldsymbol{g}_k\left(\boldsymbol{Y}_2\right)\right]\xi_j\xi_k + h^{3/2}\tilde{\boldsymbol{r}}_1$$
$$+ \frac{\gamma b_1^3 + b_2^3}{6(\gamma c_1 + c_2)}h^{3/2}\sum_{j,k,l=1}^{m}\boldsymbol{f}'''\left(\boldsymbol{K}_1\right)\left[\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right), \boldsymbol{g}_k\left(\boldsymbol{Y}_2\right), \boldsymbol{g}_l\left(\boldsymbol{Y}_2\right)\right]\xi_j\xi_k\xi_l + O\left(h^2\right),$$

where

$$\tilde{\boldsymbol{r}}_1 = \left(\frac{\gamma b_1 c_1 + b_2 c_2}{\gamma c_1 + c_2} - 1\right)\sum_{j=1}^{m}\boldsymbol{f}''\left(\boldsymbol{K}_1\right)\left[\boldsymbol{g}_0\left(\boldsymbol{y}_n\right), \boldsymbol{g}_j\left(\boldsymbol{Y}_2\right)\right]\xi_j.$$

By seeking a similar Taylor expansion of $\boldsymbol{f}(\boldsymbol{K}_1 + \sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j(\boldsymbol{Y}_2)\xi_j) - \boldsymbol{f}(\boldsymbol{K}_1)$ centered at $\boldsymbol{K}_1$, and by utilizing these results and $h\varphi_2(hA) = (h/2)I + (h^2/6)A + O\left(h^3\right)$, we obtain $\boldsymbol{u} - \hat{\boldsymbol{u}} = (h^{5/2}/2)\tilde{\boldsymbol{r}}_1 + h^{5/2}\tilde{\boldsymbol{r}}_2 + h^{5/2}\boldsymbol{r}_3 + O(h^3)$, where $\boldsymbol{r}_3$ is given in the proof of Theorem 3.5 and

$$\tilde{\boldsymbol{r}}_2 = \frac{1}{12}\left(\frac{\gamma b_1^3 + b_2^3}{\gamma c_1 + c_2} - 1\right)\sum_{j,k,l=1}^{m}\boldsymbol{f}'''\left(\boldsymbol{K}_1\right)\left[\boldsymbol{g}_j\left(\boldsymbol{Y}_2\right), \boldsymbol{g}_k\left(\boldsymbol{Y}_2\right), \boldsymbol{g}_l\left(\boldsymbol{Y}_2\right)\right]\xi_j\xi_k\xi_l.$$

Since $E[\tilde{\boldsymbol{r}}_1] = E[\tilde{\boldsymbol{r}}_2] = E[\boldsymbol{r}_3] = \boldsymbol{0}$, the local error of (3.8) is weak order 3. Thus we conclude this proof similarly to the end of the proof of Theorem 3.5. ∎

Appendix C. Proof of Theorem 3.8. If we denote $\mathrm{e}^{hA/2}\boldsymbol{y}_n + h\boldsymbol{f}(\mathrm{e}^{hA/2}\boldsymbol{y}_n)$ by $\tilde{\boldsymbol{K}}_1$, similarly to (A.1), we have

$$\mathrm{e}^{\frac{h}{2}A}\left\{\mathrm{e}^{\frac{h}{2}A}\boldsymbol{y}_n + \frac{h}{2}\boldsymbol{f}\left(\mathrm{e}^{\frac{h}{2}A}\boldsymbol{y}_n\right) + \frac{h}{2}\boldsymbol{f}\left(\tilde{\boldsymbol{K}}_1 + \sqrt{h}\sum_{j=1}^{m}\boldsymbol{g}_j\left(\mathrm{e}^{\frac{h}{2}A}\boldsymbol{y}_n\right)\xi_j\right)\right\}$$
$$= \boldsymbol{y}_n + \frac{h}{2}\left(\boldsymbol{g}_0\left(\boldsymbol{y}_n\right) + \boldsymbol{g}_0\left(\boldsymbol{K}_2\right)\right) - \frac{h^{3/2}}{2}A\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{y}_n\right)\xi_j + h^{5/2}\tilde{\boldsymbol{r}}_1 + O\left(h^3\right), \quad \text{(C.1)}$$

where $\tilde{\boldsymbol{a}}_1 = (A/2)\boldsymbol{y}_n + \boldsymbol{f}\left(\boldsymbol{y}_n\right)$, $\tilde{\boldsymbol{a}}_2 = (A/2)\boldsymbol{y}_n$, and

$$\tilde{\boldsymbol{r}}_1 = \frac{1}{2}\sum_{j=1}^{m}\left\{\boldsymbol{f}'\left(\boldsymbol{y}_n\right)\boldsymbol{g}_j'\left(\boldsymbol{y}_n\right)\tilde{\boldsymbol{a}}_2 + \boldsymbol{g}_0''\left(\boldsymbol{y}_n\right)\left[\tilde{\boldsymbol{a}}_1 - \boldsymbol{g}_0(\boldsymbol{y}_n), \boldsymbol{g}_j(\boldsymbol{y}_n)\right]\right.$$
$$\left. + \frac{1}{2}A\boldsymbol{f}'\left(\boldsymbol{y}_n\right)\boldsymbol{g}_j\left(\boldsymbol{y}_n\right)\right\}\xi_j.$$

Similarly to (A.2),

$$\mathrm{e}^{\frac{h}{2}A}\boldsymbol{H}\left(\mathrm{e}^{\frac{h}{2}A}\boldsymbol{y}_n\right) - \boldsymbol{H}(\boldsymbol{y}_n) = h^2\tilde{\boldsymbol{r}}_2 + O\left(h^3\right), \quad \text{(C.2)}$$

where $\tilde{\boldsymbol{a}}_3 = (A/2)\boldsymbol{y}_n$ and

$$\tilde{\boldsymbol{r}}_2 = \sum_{j,k=1}^{m}\left\{\boldsymbol{g}_j''\left(\boldsymbol{y}_n\right)\left[\tilde{\boldsymbol{a}}_3, \boldsymbol{g}_k\left(\boldsymbol{y}_n\right)\right] + \boldsymbol{g}_j'\left(\boldsymbol{y}_n\right)\boldsymbol{g}_k'\left(\boldsymbol{y}_n\right)\tilde{\boldsymbol{a}}_3 + \frac{1}{2}A\boldsymbol{g}_j'\left(\boldsymbol{y}_n\right)\boldsymbol{g}_k\left(\boldsymbol{y}_n\right)\right\}\zeta_{kj}.$$

Finally, similarly to (A.3), we have

$$
\begin{aligned}
&\mathrm{e}^{\frac{h}{2}A}\tilde{\boldsymbol{H}}\left(\frac{\mathrm{e}^{\frac{h}{2}A}\boldsymbol{y}_n+\tilde{\boldsymbol{K}}_1}{2},\mathrm{e}^{\frac{h}{2}A}\boldsymbol{y}_n\right)-\tilde{\boldsymbol{H}}\left(\frac{\boldsymbol{y}_n+\boldsymbol{K}_1}{2},\boldsymbol{y}_n\right)\\
&=\frac{h^{3/2}}{2}A\sum_{j=1}^{m}\boldsymbol{g}_j\left(\boldsymbol{y}_n\right)\xi_j+h^{5/2}\tilde{\boldsymbol{r}}_3+O\left(h^3\right),
\end{aligned}
\tag{C.3}
$$

where $\tilde{\boldsymbol{a}}_4=(A^2/8)\boldsymbol{y}_n+\boldsymbol{f}'(\boldsymbol{y}_n)(A/4)\boldsymbol{y}_n$, $\tilde{\boldsymbol{a}}_5=(A/2)\boldsymbol{y}_n$, and

$$
\begin{aligned}
\tilde{\boldsymbol{r}}_3=\sum_{j=1}^{m}\Bigg\{&\boldsymbol{g}_j'\left(\boldsymbol{y}_n\right)\tilde{\boldsymbol{a}}_4+\frac{1}{2}\sum_{k,l=1}^{m}\boldsymbol{g}_j''\left(\boldsymbol{y}_n\right)\left[\boldsymbol{g}_k\left(\boldsymbol{y}_n\right),\boldsymbol{g}_l'\left(\boldsymbol{y}_n\right)\tilde{\boldsymbol{a}}_5\right]\chi_k\chi_l\\
&+\frac{A}{4}\boldsymbol{g}_j'\left(\boldsymbol{y}_n\right)\boldsymbol{g}_0\left(\boldsymbol{y}_n\right)+\frac{A}{8}\sum_{k,l=1}^{m}\boldsymbol{g}_j''\left[\boldsymbol{g}_k\left(\boldsymbol{y}_n\right),\boldsymbol{g}_l\left(\boldsymbol{y}_n\right)\right]\chi_k\chi_l+\frac{A^2}{8}\boldsymbol{g}_j\left(\boldsymbol{y}_n\right)\Bigg\}\xi_j.
\end{aligned}
$$

From (C.1), (C.2), and (C.3), the difference between (3.2) and (3.10) is $h^2\tilde{\boldsymbol{r}}_2+h^{5/2}\tilde{\boldsymbol{r}}_1+h^{5/2}\tilde{\boldsymbol{r}}_3+O\left(h^3\right)$. Since $E[\tilde{\boldsymbol{r}}_1]=E[\tilde{\boldsymbol{r}}_2]=E[\tilde{\boldsymbol{r}}_3]=E[\xi_j\tilde{\boldsymbol{r}}_2]=0\quad(j=1,2,\ldots,m)$, the local error of (3.10) is weak order 3. Thus we conclude this proof similarly to the end of the proof of Theorem 3.5. ◻

Acknowledgments. The authors would like to thank the referees for their valuable comments which helped to improve the earlier versions of this paper. It should be mentioned that suggestions made by one of the referees were very helpful for deriving the splitting methods and combining the exponential RK methods with them.

REFERENCES

[1] A. Abdulle and S. Cirilli, *S-ROCK: Chebyshev methods for stiff stochastic differential equations*, SIAM J. Sci. Comput., 30 (2008), pp. 997–1014.

[2] A. Abdulle and T. Li, *S-ROCK methods for stiff Itô SDEs*, Commun. Math. Sci., 6 (2008), pp. 845–868.

[3] A. Abdulle and A. A. Medovikov, *Second order Chebyshev methods based on orthogonal polynomials*, Numer. Math., 90 (2001), pp. 1–18.

[4] A. Abdulle, G. Vilmart, and K. C. Zygalakis, *Mean-square A-stable diagonally drift-implicit integrators of weak second order for stiff Itô stochastic differential equations*, BIT, 53 (2013), pp. 827–840.

[5] A. Abdulle, G. Vilmart, and K. C. Zygalakis, *Weak second order explicit stabilized methods for stiff stochastic differential equations*, SIAM J. Sci. Comput., 35 (2013), pp. A1792–A1814.

[6] I. A. Adamu, *Numerical Approximation of SDEs and Stochastic Swift-Hohenberg Equation*, Ph.D. thesis, Heriot-Watt University, Edinburgh, UK, 2011.

[7] L. Arnold, *Stochastic Differential Equations: Theory and Applications*, John Wiley & Sons, New York, 1974.

[8] E. Buckwar and T. Sickenberger, *A structural analysis of mean-square stability for multi-dimensional linear stochastic differential systems*, Appl. Numer. Math., 62 (2012), pp. 842–859.

[9] D. Cohen and M. Sigg, *Convergence analysis of trigonometric methods for stiff second-order stochastic differential equations*, Numer. Math., 121 (2012), pp. 1–29.

[10] G. Da Prato and D. Gatarek, *Stochastic Burgers equation with correlated noise*, Stochastics Stochastics Rep., 52 (1995), pp. 29–41.

[11] 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 (2009), pp. 582–594.

[12] B. L. Ehle and J. D. Lawson, *Generalized Runge-Kutta processes for stiff initial-value problems*, IMA J. Appl. Math., 16 (1975), pp. 11–21.

[13] L. Einkemmer and A. Ostermann, *Overcoming order reduction in diffusion-reaction splitting. Part* 1*: Dirichlet boundary conditions*, SIAM J. Sci. Comput., 37 (2015), pp. A1577–A1592.

[14] S. Geiger, G. Lord, and A. Tambue, *Exponential time integrators for stochastic partial differential equations in* 3*D reservoir simulation*, Comput. Geosci., 16 (2012), pp. 323–334.

[15] N. Hale, N. J. Higham, and N. Trefethen, *Computing* A^α*,* $\log(A)$*, and related matrix functions by contour integrals*, SIAM J. Numer. Anal., 46 (2008), pp. 2505–2523.

[16] D. J. Higham, *A-stability and stochastic mean-square stability*, BIT, 40 (2000), pp. 404–409.

[17] M. Hochbruck, C. Lubich, and H. Selhofer, *Exponential integrators for large systems of differential equations*, SIAM J. Sci. Comput., 19 (1998), pp. 1552–1574.

[18] M. Hochbruck and A. Ostermann, *Explicit exponential Runge–Kutta methods for semilinear parabolic problems*, SIAM J. Numer. Anal., 43 (2005), pp. 1069–1090.

[19] M. Hochbruck and A. Ostermann, *Exponential integrators*, Acta Numer., 19 (2010), pp. 209–286.

[20] P. Kandolf, *Computationally Efficient Exponential Integrators*, Ph.D. thesis, University of Innsbruck, Innsbruck, Austria, 2016.

[21] P. E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin, 1999; corrected third printing.

[22] Y. Komori, *Weak second-order stochastic Runge–Kutta methods for non-commutative stochastic differential equations*, J. Comput. Appl. Math., 206 (2007), pp. 158–173.

[23] Y. Komori and K. Burrage, *Weak second order S-ROCK methods for Stratonovich stochastic differential equations*, J. Comput. Appl. Math., 236 (2012), pp. 2895–2908.

[24] Y. Komori and K. Burrage, *Strong first order S-ROCK methods for stochastic differential equations*, J. Comput. Appl. Math., 242 (2013), pp. 261–274.

[25] Y. Komori and K. Burrage, *A stochastic exponential Euler scheme for simulation of stiff biochemical reaction systems*, BIT, 54 (2014), pp. 1067–1085.

[26] J. D. Lawson, *Generalized Runge-Kutta processes for stable systems with large Lipschitz constants*, SIAM J. Numer. Anal., 4 (1967), pp. 372–380.

[27] G. J. Lord and A. Tambue, *Stochastic exponential integrators for the finite element discretization of SPDEs for multiplicative and additive noise*, IMA J. Numer. Anal., 33 (2013), pp. 515–543.

[28] M. Matsumoto and T. Nishimura, *Mersenne twister: A* 623*-dimensionally equidistributed uniform pseudo-random number generator*, ACM Trans. Model. Comput. Simul., 8 (1998), pp. 3–30.

[29] R. I. McLachlan and G. R. W. Quispel, *Splitting methods*, Acta Numer., 11 (2002), pp. 341–434.

[30] G. N. Milstein, *Weak approximation of solutions of systems of stochastic differential equations*, Theory Probab. Appl., 30 (1986), pp. 750–766.

[31] G. N. Milstein and M. V. Tretyakov, *Stochastic Numerics for Mathematical Physics*, Springer-Verlag, Berlin, 2004.

[32] D. Pope, *An exponential method of numerical integration of ordinary differential equations*, Commun. ACM, 6 (1963), pp. 491–493.

[33] A. Rössler, *Rooted tree analysis for order conditions of stochastic Runge–Kutta methods for the weak approximation of stochastic differential equations*, Stochastic Anal. Appl., 24 (2006), pp. 97–134.

[34] C. Shi, Y. Xiao, and C. Zhang, *The convergence and MS stability of exponential Euler method for semilinear stochastic differential equations*, Abstr. Appl. Anal., 2012 (2012), 350407.

[35] G. Strang, *On the construction and comparison of difference schemes*, SIAM J. Numer. Anal., 5 (1968), pp. 506–517.

[36] A. Tocino, *On preserving long-time features of a linear stochastic oscillator*, BIT, 47 (2007), pp. 189–196.

[37] A. Tocino and M. J. Senosiain, *Mean-square stability analysis of numerical schemes for stochastic differential systems*, J. Comput. Appl. Math., 236 (2012), pp. 2660–2672.

[38] P. J. van der Houwen and B. P. Sommeijer, *On the internal stability of explicit,* m*-stage Runge–Kutta methods for large* m*-values*, ZAMM Z. Angew. Math. Mech., 60 (1980), pp. 479–485.