# Optimal explicit stabilized integrator of weak order one for stiff and ergodic stochastic differential equations

Assyr Abdulle[1], Ibrahim Almuslimani[2] and Gilles Vilmart[2]

June 27, 2017

**Abstract**

A new explicit stabilized scheme of weak order one for stiff and ergodic stochastic differential equations (SDEs) is introduced. In the absence of noise, the new method coincides with the classical deterministic stabilized scheme (or Chebyshev method) for diffusion dominated advection-diffusion problems and it inherits its optimal stability domain size that grow quadratically with the number of internal stages of the method. For mean-square stable stiff stochastic problems, the scheme has an optimal extended mean-square stability domain that grows at the same quadratic rate as the deterministic stability domain size in contrast to known existing methods for stiff SDEs [A. Abdulle and T. Li. Commun. Math. Sci., 6(4), 2008, A. Abdulle, G. Vilmart, and K. C. Zygalakis, SIAM J. Sci. Comput., 35(4)]. Combined with postprocessing techniques, the new methods achieve a convergence rate of order two for sampling the invariant measure of a class of ergodic SDEs, achieving a stabilized version of the non-Markovian scheme introduced in [B. Leimkuhler, C. Matthews, and M. V. Tretyakov, , R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci. A, 470, 2014].

## 1 Introduction

We consider Itô systems of stochastic differential equations of the form

$$dX(t) = f(X(t))dt + \sum_{r=1}^{m} g^r(X(t))dW_r(t), \qquad X(0) = X_0 \tag{1}$$

where $X(t)$ is a stochastic process with values in $\mathbb{R}^d$, $f : \mathbb{R}^d \to \mathbb{R}^d$ is the drift term, $g^r : \mathbb{R}^d \to \mathbb{R}^d$, $r = 1, \ldots, m$ are the diffusion terms, and $W_r(t)$, $r = 1, \ldots, m$, are independent one-dimensional Weiner processes fulfilling the usual assumptions. We assume that the drift and diffusion functions are smooth enough and Lipschitz continuous to ensure the existence and uniqueness of a solution of (1) on a given time interval $(0, T)$. We consider autonomous problems to simplify the presentation, but we emphasise that the scheme can also be extended to non-autonomous SDEs. A one step numerical integrator for the approximation of (1) at time $t = nh$ is a discrete dynamical system of the form

$$X_{n+1} = \Psi(X_n, h, \xi_n) \tag{2}$$

where $h$ denotes the stepsize and $\xi_n$ are independent random vectors. Analogously to the deterministic case, for stiff stochastic problems, standard explicit numerical schemes, such as the

[1]Mathematics Section, École Polytechnique Fédérale de Lausanne, Station 8, 1015 Lausanne, Switzerland, Assyr.Abdulle@epfl.ch

[2]Université de Genève, Section de mathématiques, 2-4 rue du Lièvre, CP 64, CH-1211 Genève 4, Switzerland, Ibrahim.Almuslimani@unige.ch, Gilles.Vilmart@unige.ch

simplest Euler-Maruyama method defined as

$$X_{n+1} = X_n + hf(X_n) + \sum_{r=1}^{m} g^r(X_n)\Delta W_{n,j}, \qquad X(0) = X_0, \tag{3}$$

where $\Delta W_{n,j} = W_j(t_{n+1}) - W_j(t_n)$ are the Brownian increments, face a sever timestep restriction [12, 11, 13], and one can use an implicit or semi-implicit scheme with favourable stability properties. In particular, it is shown in [12] that the implicit $\theta$-method of weak order one is mean-square A-stable if and only if $\theta \geq 1/2$, while weak order two mean-square A-stable are constructed in [6]. An alternative approach is to consider explicit stabilized schemes with extended stability domains, as proposed in [4, 3]. In [4] the deterministic Chebyshev method is extended to the context of mean-square stiff stochastic differential equations with Itô noise, while the Stratonovitch noise case is treated in [3]. In place of a standard small damping, the main idea in [3, 4] is to use a large damping parameter $\eta$ optimized for each number $s$ of stages to stabilize the noise term. This yields a family of Runge-Kutta type schemes with extended stability domain with size $L_s \simeq 0.33s^2$. This stability domain size was improved to $L_s \simeq 0.42s^2$ in [7] where a family of weak second order stabilized schemes (and strong order one under suitable assumptions) is constructed based on the deterministic ROCK2 method [5].

For ergodic SDEs, i.e., when (1) has a unique invariant measure $\mu$ satisfying for each smooth function $\phi$ and for any deterministic initial condition $X_0 = x$,

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \phi(X(s))ds = \int_E \phi(y)d\mu(y), \qquad \text{almost surely}, \tag{4}$$

one is interested in approximating numerically the long-time dynamics and to design numerical scheme with a unique invariant measure such that

$$|\lim_{N\to\infty} \frac{1}{N+1}\sum_{n=0}^{N} \phi(X_n) - \int_E \phi(y)d\mu(y)| \leq Ch^r, \tag{5}$$

where $C$ is independent of $h$ small enough and $X_0$. In such a situation, we say that the numerical scheme as order $r$ with respect to the invariant measure. The Euler-Maruyama method has order 1 with respect to the invariant measure. In [14] the following non-Markovian scheme with the same cost as the Euler-Maruyama method was proposed,

$$X_{n+1} = X_n + hf(X_n) + \sum_{r=1}^{m} g^r(X_n)\frac{\Delta W_{n,j} + \Delta W_{n+1,j}}{2}, \qquad X(0) = X_0, \tag{6}$$

and it was shown in [15] that (6) has order 2 with respect to the invariant measure for Brownian dynamics. However, the admissible stepsizes for such an explicit method to be stable may face sever restriction and alternatively to switching to drift-implicit methods, one may ask if a stabilized version of such an attractive non-Markovian scheme exists.

In this paper we introduce a new family of explicit stabilized schemes with optimal mean-square stability domain of size $L_s = Cs^2$, where $C \geq 2 - \frac{4}{3}\eta$ and $\eta \geq 0$ is a small parameter. We emphasize that in the deterministic case, $L_s = 2s^2$ is the largest, i.e. optimal, stability domain along the negative real axis for an explicit s-stage Runge-Kutta method [11]. We note that the Chebyshev method (8) (with $\eta = 0$) realizes such an optimal stability domain. The new schemes have strong order 1/2 and weak order 1. The main ingredient for the design of the new schemes is to consider second kind Chebyshev polynomials, in addition to the usual first kind Chebyshev polynomials involved in the deterministic Chebychev method and stochastic extensions [4, 3]. For stiff stochastic problems the stability domain sizes are close to the optimal value $2s^2$ and in the deterministic setting the method coincide with the optimal first order explicit stabilized method. Thus these methods are more efficient than previously introduced stochastic stabilized methods [4, 7]. For ergodic dynamical systems, in the context of the ergodic Brownian dynamics, the new family of explicit stabilized schemes allows for a postprocessing [23] to achieve order two of accuracy for sampling the invariant measure. In this context our new methods can be seen as a stabilized version of the non-Markovian scheme (6) introduced in [14, 15].

This paper is organized as follows. In Section 2, we introduce the new family of schemes with optimal stability domain. In Section 3, we recall the main tools for the study of stiff integrators in the mean-square sense. We then analyze its mean-square stability properties (Section 3), and convergence properties (Section 4). In Section 5, using a postprocessor we present a modification with negligible overcost that yields order two of accuracy for the invariant measure of a class of ergodic overdamped Langevin equation. Finally, Section 6 is dedicated to the numerical experiments that confirm our theoretical analysis and illustrate the efficiency of the new schemes.

## 2 New second kind Chebyshev methods

In this section we introduce our new stabilized stochastic method. We first briefly recall the concept of stabilized methods. In the context of ordinary differential equations (ODEs),

$$\frac{dX(t)}{dt} = f(X(t)), \qquad X(0) = X_0, \tag{7}$$

and the Euler method $X_1 = X_0 + hf(X(0))$ a stabilization procedure based on recurrence formula has been introduced in [21]. Its construction relies on Chebyshev polynomials (hence the alternative name "Chebyshev methods"), $T_s(\cos x) = \cos(sx)$ and is based on the explicit $s$-stage method

$$\begin{aligned} K_0 &= X_0, \quad K_1 = X_0 + h\mu_1 f(K_0), \\ K_i &= \mu_i h f(K_{i-1}) + \nu_i K_{i-1} + \kappa_i K_{i-2}, \quad j = 2, \dots, s \\ X_1 &= K_s, \end{aligned} \tag{8}$$

where

$$\omega_0 = 1 + \frac{\eta}{s^2}, \quad \omega_1 = \frac{T_s(\omega_0)}{T_s'(\omega_0)}, \quad \mu_1 = \frac{\omega_1}{\omega_0}, \tag{9}$$

and for all $i = 2, \dots, s$,

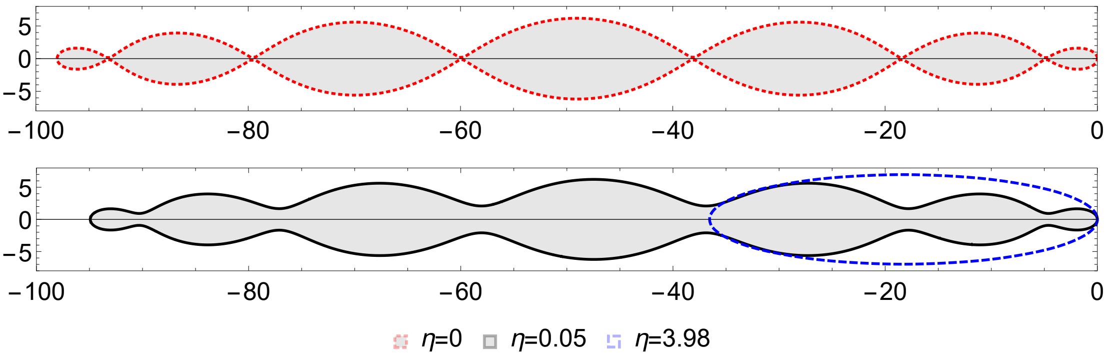
$$\mu_i = \frac{2\omega_1 T_{i-1}(\omega_0)}{T_i(\omega_0)}, \quad \nu_i = \frac{2\omega_0 T_{i-1}(\omega_0)}{T_i(\omega_0)}, \quad \kappa_i = -\frac{T_{i-2}(\omega_0)}{T_i(\omega_0)} = 1 - \nu_i. \tag{10}$$

One can easily check that the (family) of methods (8) has the same first order accuracy as the Euler method. In addition, the scheme (8) has a low memory requirement (only two stages should be stored when applying the recurrence formula) and it has a good internal stability with respect to round-off errors [21]. The attractive feature of such a scheme comes from its stability behavior. Indeed applied to the linear test problem $dX(t)/dt = \lambda X(t)$ the method (8), using the recurrence relation

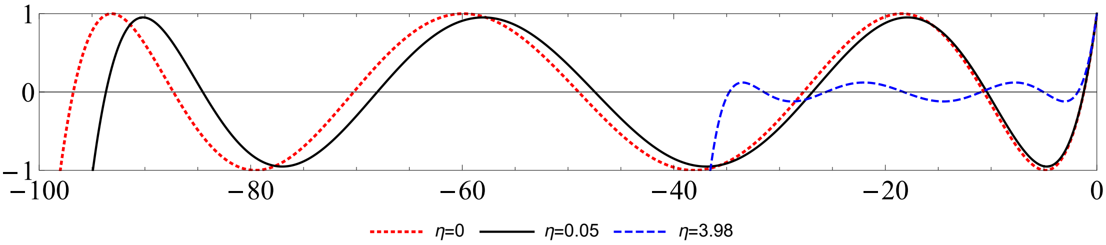
$$T_j(p) = 2pT_{j-1}(p) - T_{j-2}(p), \tag{11}$$

where $T_0(p) = 1, T_1(p) = p$, gives for $p = \lambda h$

$$X_1 = R_{s,\eta}(p) X_0 = \frac{T_s(\omega_0 + \omega_1 p)}{T_s(\omega_0)} X_0, \tag{12}$$




(a) Complex stability domain for different damping parameters $\eta$.




(b) Corresponding stability functions $R_{7,\eta}(p)$.

Figure 1: Stability domains and stability functions of the deterministic Chebyshev method for $s = 7$ and different damping values $\eta = 0, 0.05, 3.98$.

where the dependence of the stability function on the parameters $s$ and $\eta$ is emphasized with a corresponding subscript. The real negative interval $(-C_s(\eta) \cdot s^2, 0)$ is included in the stability domain of the method

$$\mathcal{S} := \{p \in \mathbb{C}; |R(z)| \leq 1\}. \tag{13}$$

The constant $C_s(\eta) \simeq 2 - 4/3\,\eta$ depends on the so-called damping parameter $\eta$ and for $\eta = 0$, it reaches the maximal value $C_s = 2$. Hence, given the stepsize $h$, for systems with a Jacobian having large real negative eigenvalues (such as diffusion problems) with spectral radius $\lambda_{\max}$ at $X_n$, the parameter $s$ for the next step $X_{n+1}$ can be chosen adaptively as[1]

$$s = \left[\sqrt{\frac{h\lambda_{\max}}{2 - 4/3\,\eta}}\right] + 1,$$

see [1] in the context of deterministically stabilized schemes of order two with adaptative stepsizes.

The method (8) is much more efficient as its stability domain increase *quadratically* with the number $s$ of function evaluations while a composition of $s$ explicit Euler steps (same cost) has a stability domain that only increases *linearly* with $s$. In Figure 1(a) we plot the complex stability domain $\{p \in \mathcal{C}\,; |R_{s,\eta}(p)| \leq 1\}$ for $s = 7$ stages and different values $\eta = 0$, $\eta = 0.05$ and $\eta = 3.98$, respectively. We also plot in Figure 1(b) the corresponding stability function $R_{s,\eta}(p)$ as a function of $p$ real, to illustrate that the stability domain along the negative real axis corresponds to the values for which $|R_{s,\eta}(p)| \leq 1$. We observe that in the absence of damping ($\eta = 0$), the stability domain includes the large real interval $[-2 \cdot s^2, 0]$ of width $2 \cdot 7^2 = 98$. However for all $p$ that are a local extrema of the stability function, where $|R_{s,\eta}(p)| = 1$, the stability domain is very thin and does not include a neighbourhood close to the negative real axis. To make the scheme robust with respect to small perturbations of the eigenvalues, it is therefore needed to add some

[1] The notation $[x]$ standard for the integer rounding of real numbers.

damping is necessary and a typical values is $\eta = 0.05$, see for instance the reviews [22, 2]. The advantage is that the stability domain now includes a neighbourhood of the negative real axis portion. The price of this improvement is a slight reduction of the stability domain size $C_\eta s^2$, where $C_\eta \simeq 2 - \frac{4}{3}\eta$. Chebyshev methods have been first generalized for Itô SDEs in [4] with the scheme[2]

$$
\begin{aligned}
K_0 &= X_0 \\
K_1 &= X_0 + \mu_1 h f(X_0) \\
K_i &= \mu_i h f(K_{i-1}) + \nu_i K_{i-1} + \kappa_i K_{i-2}, \quad i = 2, \dots, s, \\
X_1 &= K_s + \sum_{r=1}^{m} g^r(K_s)\Delta W_j,
\end{aligned}
\tag{14}
$$

where the coefficients $\mu_i, \nu_i, \kappa_i$ are defined in (9),(10). In contrast to the deterministic method (8) where $\eta$ is chosen small and fixed (typically $\eta = 0.05$), in stochastic case for the classical S-ROCK method [4], the damping $\eta = \eta_s$ is not small and chosen as an increasing function of $s$ that plays a crucial in stabilizing the noise and obtain an increasing portion of the true stability domain (18) as $s$ increases.

In the context of stiff SDEs, a relevant stability concept is that of mean-square stability. A test problem widely used in the literature is [19, 12, 9, 20] ,

$$
dX(t) = \lambda X(t)dt + \mu X(t)dW(t), \qquad X(0) = 1, \tag{15}
$$

in dimensions $d = m = 1$ with fixed complex parameters $\lambda, \mu$. Note that other stability test problem in multiple dimensions are also be considered in [8] and references therein. The exact solution of (15) is called mean-square stable if $\lim_{t\to\infty} \mathbb{E}\big(|X(t)|^2\big) = 0$ and this holds if and only if $(\lambda, \mu) \in \mathcal{S}^{MS}$, where

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

Indeed, the exact solution of (15) is given by $X(t) = \exp((\lambda + \frac{1}{2}\mu^2)t + \mu W(t))$, and an application of the the Itô formula yields $\mathbb{E}(|X(t)|^2) = \exp((\Re(\lambda) + \frac{1}{2}\mu^2)t)$ which tends to zero at infinity if and only if $\Re\lambda + \frac{1}{2}\mu^2 < 0$. We say that a numerical scheme $\{X_n\}$ for the test problem 15 is mean-square stable if and only if $\lim_{n\to\infty} \mathbb{E}(|X_n|^2) = 0$. For a one-step integrator applied to the test SDE (15), we obtain in general a induction of the form

$$
X_{n+1} = R(p, q, \xi_n)X_n, \tag{16}
$$

where $p = \lambda h, q = \mu\sqrt{h}$, and $\xi_n$ is a random variable (e.g. a Gaussian $\xi_n \sim \mathcal{N}(0, 1)$ or a discrete random variable). Using $\mathbb{E}(|X_{n+1}|^2) = \mathbb{E}(|R(p, q, \xi_n)|^2)\mathbb{E}(|X_n|^2)$, we obtain the mean-square stability condition [12]

$$
\lim_{n\to\infty} \mathbb{E}(|X_n|^2) = 0 \iff (p, q) \in \mathcal{S}_{num}, \tag{17}
$$

where we define $\mathcal{S}_{num} = \big\{(p, q) \in \mathbb{C}^2 \; ; \mathbb{E}|R(p, q, \xi)|^2 < 1\big\}$. For instance, the stability function of the Euler-Maruyama method (3) reads $R(p, q, \xi) = 1 + p + q\xi$ and we have

$$
\mathbb{E}(|R(p, q, \xi)|^2) = (1 + p)^2 + q^2.
$$

We say that a numerical integrator is mean-square $A$-stable if $\mathcal{S} \subset \mathcal{S}_{num}$. This means that the numerical scheme applied to (15) is mean-square stable for all all $h > 0$ and all $(\lambda, \mu) \in \mathcal{S}$ for which the exact solution of (15) is mean-square stable. An explicit Runge-Kutta type scheme

[2] A variant with analogous stability properties is proposed in [4] with $g^r(K_s)$ replaced by $g^r(K_{s-1})$ in (14).

cannot however be mean-square stable because its stability domain $\mathcal{S}_{num}$ is necessary bounded along the $p$-axis. Following [3, 4], we consider the following portion of the true mean-square stability domain

$$\mathcal{S}_a = \{(p,q) \in (-a,0) \times \mathbb{R} \ ; \ p + \frac{1}{2}|q|^2 < 0\}, \tag{18}$$

and define for a given method

$$L = \sup\{a > 0 \ ; \ \mathcal{S}_a \subset \mathcal{S}_{num}\}. \tag{19}$$

and we search for explicit schemes for which the length $L$ of the stability domain is large. For example, for the classical S-ROCK method [4], the value $\eta = 3.98$ is the optimal damping maximising $L$ for $s = 7$ stages and we can see in Figure 1 that this damping reduces significantly the stability domain compared to the optimal deterministic domain.

The new S-ROCK method, called SK-ROCK (for second kind orthogonal Runge-Kutta-Chebyshev method) introduced in this paper is defined as

$$\begin{aligned} K_0 &= X_0 && (20)\\ K_1 &= X_0 + \mu_1 h f(X_0 + \nu_1 Q) + \kappa_1 Q && (21)\\ K_i &= \mu_i h f(K_{i-1}) + \nu_i K_{i-1} + \kappa_i K_{i-2}, \quad i = 2, \ldots, s. && (22)\\ X_1 &= K_s, && (23) \end{aligned}$$

where

$$Q = \sum_{r=1}^{m} g^r(X_0)\Delta W_j \tag{24}$$

and $\mu_1 = \omega_1/\omega_0, \nu_1 = s\omega_1/2, \kappa_1 = s\omega_1/\omega_0$ and $\mu_i, \nu_i, \kappa_i,\ i = 2, \ldots, s$ are given by (10), with a fixed small damping parameter $\eta$. In the absence of noise ($g^r = 0, r = 1, \ldots, m$ deterministic case), this method coincides with the standard deterministic order 1 Chebychev method, see the review [2].

In addition to the recurrence relation for the first kind (11), using the similar recurrence relation for the second kind Chebyshev polynomials

$$U_j(p) = 2pU_{j-1}(p) - U_{j-2}(p), \tag{25}$$

where $U_0(p) = 1, U_1(p) = 2p$, we obtain (see Lemma 3.1 in Section 3)

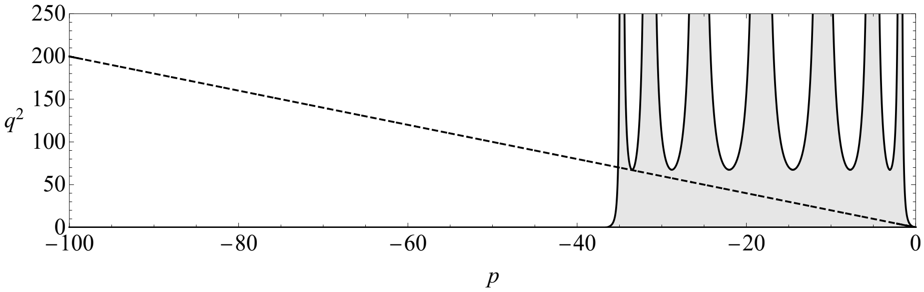
$$\mathbb{E}(|R(p,q,\xi)|^2) = A(p)^2 + B(p)^2 q^2 \tag{26}$$

where

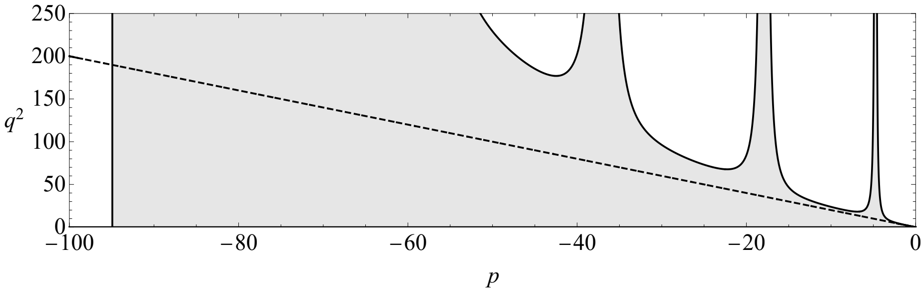
$$A(p) = \frac{T_s(\omega_0 + \omega_1 p)}{T_s(\omega_0)} \qquad B(p) = \frac{U_{s-1}(\omega_0 + \omega_1 p)}{U_{s-1}(\omega_0)}(1 + \frac{\omega_1}{2}p),$$

correspond to the drift and diffusion contributions, respectively. The relation $T_s'(p) = sU_{s-1}(p)$ between first and second kind Chebyshev polynomials will be repeatedly used in our analysis.

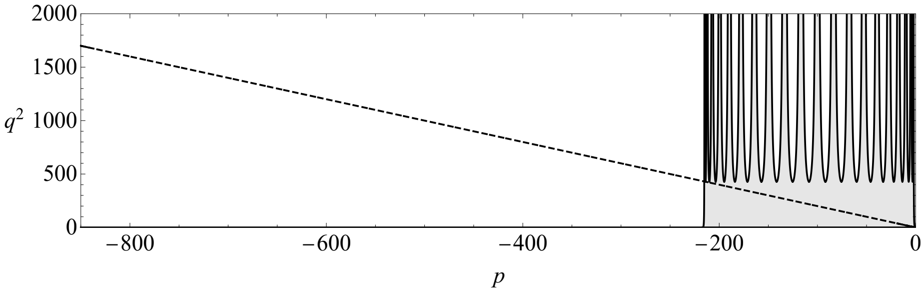
We observe that the new class of methods (23) is closely related to the standard S-ROCK method (14). Comparing the two schemes (23) and (14), the two differences are on the one hand that the noise term is computed at the first internal stage $K_1$ for (23), whereas it is computed at the final stage in (14), and on the other hand, for the new method (23) the damping parameter $\eta$ involved in (9) is fixed and small independently of $s$ (typically $\eta = 0.05$), whereas for the standard method (23), the damping $\eta$ is an increasing function of $s$, optimized numerically for each number of stages $s$.



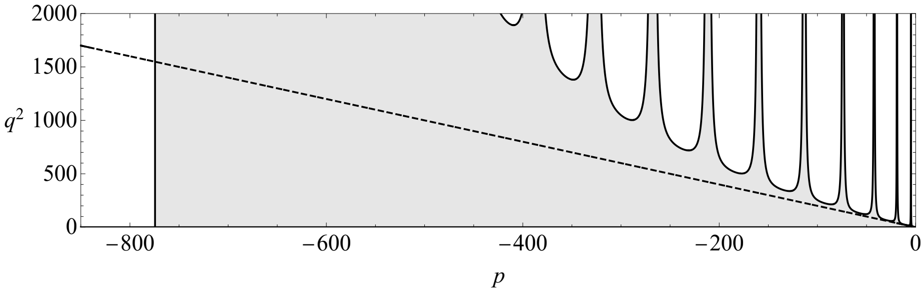

(a) Standard S-ROCK method ($s = 7$, $\eta = 3.98$).




(b) New S-ROCK method ($s = 7$, $\eta = 0.05$).




(c) Standard S-ROCK method ($s = 20$, $\eta = 6.95$).

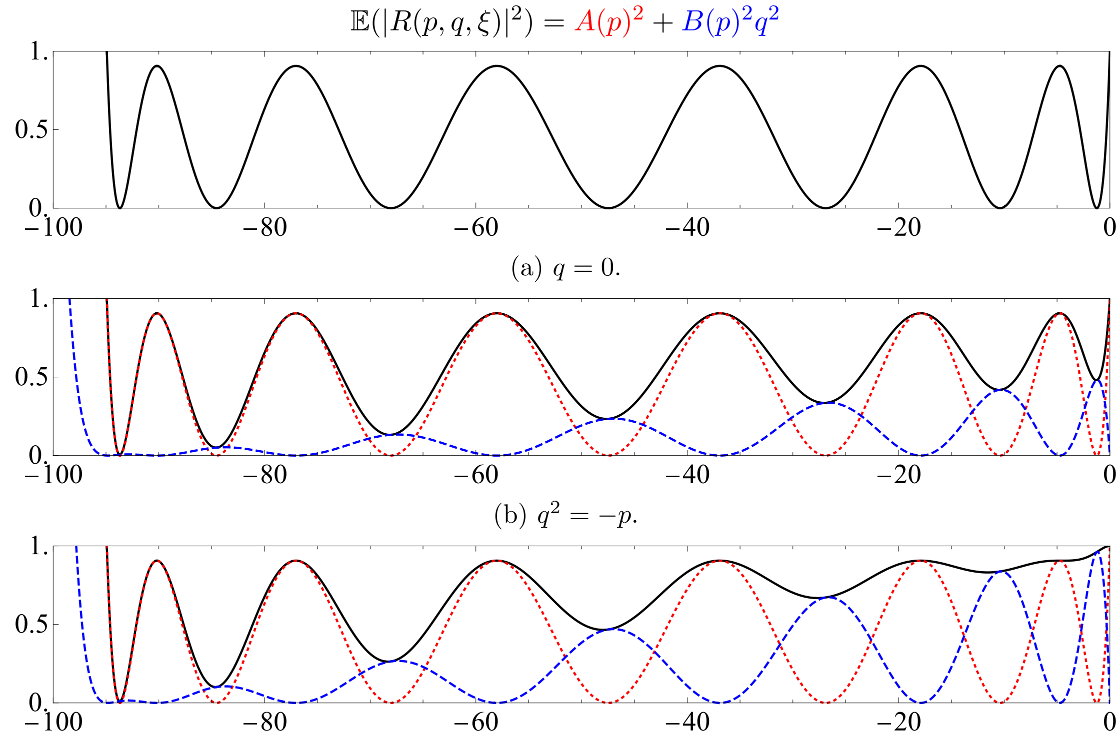



(d) New S-ROCK method ($s = 20$, $\eta = 0.05$).

Figure 2: Mean-square stability domains of the standard and new stochastic Chebyshev methods in the $p$–$q^2$ plane for $s = 7, 20$ stages, respectively. The dashed lines corresponds to the upper boundary $q^2 = -2p$ of the real mean-square stability domain $\mathcal{S} \cap \mathbb{R}^2$ of the exact solution.

In Figure 2(b)(d), we plot the mean-square stability domain of the SK-ROCK method for $s = 7$ and $s = 20$ stages, respectively and the same small damping $\eta = 0.05$ as for the deterministic Chebyshev method. We observe that the stability domain has length $\simeq (2 - \frac{4}{3}\eta)s^2$. For comparison, we also include in Figure 2(a)(c) the mean-square stability domain of the standard S-ROCK method with smaller stability domain size $L_s \simeq 0.33 \cdot s^2$.

In Figure 3, we plot the stability function $\mathbb{E}(|R(p, q, \xi)|^2)$ in (26) as a function of $p$ for various scaling of the noise for $s = 7$ stages and damping $\eta = 0.05$. We see that it is bounded by 1 for




(a) $q = 0$.

(b) $q^2 = -p$.

(c) $q^2 = -2p$.

Figure 3: Stability function $\mathbb{E}(|R(p,q,\xi)|^2) = A(p)^2 + B(p)^2q^2$ as a function of $p$ in (26) (solid black lines) of the new SK-ROCK method, $s = 7, \eta = 0.05$, for various noise scalings $q^2 = 0, -p, -2p$, respectively. We also include the drift contribution $A(p)^2$ (red dotted lines) and diffusion contribution $B(p)^2q^2$ (blue dashed lines).

$p \in (-2(1-\frac{2}{3}\eta)s^2, 0)$ which is proved asymptotically in Theorem 3.3. The case $q = 0$ corresponds to the deterministic case, and we see in Figure 3(a), the polynomial $\mathbb{E}(|R(p,0,\xi)|^2) = A(p)^2$. Noticing that $\mathbb{E}(|R(p,q,\xi)|^2)$ is an increasing function of $q$, the case $q^2 = \sqrt{-2p}$ represented in Figure 3(c) corresponds to the upper border of the stability domain $\mathcal{S}_L$ defined in (18) (note that this is the stability function value along the dashed boundary in Figure 2), while the scaling $q^2 = -p$ in Figure 3(c) is an intermediate regime. In Figures 3(b)(c), we also include the drift function $A(p)^2$ (red dotted lines) and diffusion function $B(p)^2q^2$ (blue dashed lines), and it can be observed that their oscillations alternate, which means that any local maxima of one function is close to a zero of the other function. This is note surprising because $A(p)$ and $B(p)$ are related to the first kind and second kind Chebyshev polynomials, respectively, corresponding to the cosinus and sinus functions. This also explains how a large mean-square stability domain can be achieved by the new SK-ROCK method (23) with a small damping parameter $\eta$, in contrast to the standard S-ROCK method (14) from [4] that uses a large and $s$-dependent damping parameter $\eta$ with smaller stability domain size $L_s \simeq 0.33 \cdot s^2$(see Figure 3(a)(c)).

## 3 Mean-square stability analysis

**Lemma 3.1.** *Let $s \geq 1$ and $\eta \geq 0$. Applied to the linear test equation $dX = \lambda X dt + \mu X dW$, the scheme* (23) *yields*

$$X_{n+1} = R(\lambda h, \mu\sqrt{h}, \xi_n)X_n$$

*where $p = \lambda h, q = \mu\sqrt{h}$, $\xi_n \sim \mathcal{N}(0,1)$ is a Gaussian variable and the stability function given by*

$$R(p,q,\xi) = \frac{T_s(\omega_0+\omega_1 p)}{T_s(\omega_0)} + \frac{U_{s-1}(\omega_0+\omega_1 p)}{U_{s-1}(\omega_0)}(1+\frac{\omega_1}{2}p)q\xi. \tag{27}$$

*Proof.* Indeed, we advantage that $T_j$ and $U_j$ have the same recurrence relations (11),(25), and only the initialization changes with $T_1(x) = x$ and $U_1(x) = 2x$, we deduce $Q = X_0\mu\sqrt{h}\xi$, and we obtain by induction on $i \geq 1$,

$$K_i = \frac{T_i(\omega_0+\omega_1 p)}{T_i(\omega_0)}X_0 + \frac{U_{i-1}(\omega_0+\omega_1 p)}{T_i(\omega_0)}(1+\frac{\omega_1 p}{2})s\omega_1 Q$$

and we use $T_s'(x) = xU_{s-1}(x)$ and $s\omega_1/T_s(\omega_0) = 1/U_{s-1}(\omega_0)$, which yields the result for $X_1 = K_s$. $\square$

**Remark 3.2.** *In the special case $\eta = 0$, the size* (19) *of the mean-square stability domain maximizes to $L = 2s^2$. Indeed, the stability function* (27) *reduces to*

$$R(p,q,\xi) = T_s(1+\frac{p}{s^2}) + s^{-1}U_{s-1}(1+\frac{p}{s^2})(1+\frac{p}{2s^2})q\xi.$$

*and we show $\mathbb{E}(|R(p,q,\xi)|^2) \leq 1$, for all $p \in [-2s^2, 0]$ and all $q \in \mathbb{C}$ such that $p + |q|^2/2 \leq 0$. For $p \in [-2s^2, 0]$, we denote $\cos\theta = 1 + \frac{p}{s^2} \in [-1,1]$ and using $T_s(\cos(\theta)) = \cos(s\theta)$ and $\sin(\theta)U_{s-1}(\sin(\theta)) = \sin(s\theta)$, we obtain*

$$\mathbb{E}(|R(p,q,\xi)|^2) \leq \mathbb{E}(|R(p,\sqrt{-2p}) = \cos(s\theta)^2 + \sin(s\theta)^2\frac{1+\cos\theta}{2} \leq 1,$$

*where we used $-2p = 2s^2(1-\cos\theta)$, $1 + \frac{p}{2s^2} = \frac{1+\cos\theta}{2}$ and $\sin\theta = \frac{1+\cos\theta}{2}(1-\cos\theta)$.*

For a positive damping $\eta$, we prove the following main result of this section, showing that a quadratic growth $L \geq (2 - 4/3\,\eta)s^2$ of the mean-square stability domain is achieved for all $\eta$ small enough, all stage number $s$ large enough.

**Theorem 3.3.** *There exists $\eta_0 > 0$ and $s_0$ such that for all $\eta \in [0,\eta_0]$ and all $s \geq s_0$ for all $p \in [-2\omega_1^{-1}, 0]$ and $p + \frac{1}{2}|q|^2 \leq 0$, we have $\mathbb{E}(|R(p,q,\xi)|^2) \leq 1$.*

**Remark 3.4.** *Notice that for $s \to \infty$, we have $2\omega_1^{-1}s^{-2} \to 2\frac{\tanh(\sqrt{2\eta})}{\sqrt{2\eta}} = 2 - \frac{4}{3}\eta + \mathcal{O}(\eta^2)$ for all $\eta \leq \eta_0$. We recover from Theorem 3.3, that the mean-square stability domain size* (19) *of SK-ROCK grows as $(2-\frac{4}{3}\eta)s^2$ which is arbitrarity close to the optimal stability domain size $2s^2$ for no damping.*

The following lemma with proof postponed to appendix will be useful, see [21] for analogous results.

**Lemma 3.5.** *We have the following convergences as $s \to \infty$ to analytic functions*[3] *uniformly for $z$ in any bounded set of the complex plan,*

$$T_s(1+z/s^2) \to \cosh\sqrt{2z},$$
$$s^{-1}U_{s-1}(1+z/s^2) \to \alpha(z) := \frac{\sinh\sqrt{2z}}{\sqrt{2z}},$$
$$\omega_1^{-1}s^{-2} \to C := \frac{\tanh\sqrt{2\eta}}{\sqrt{2\eta}}.$$

[3]Note that for $z < 0$, we can use $\sqrt{2z} = i\sqrt{-2z}$ and obtain $T_s(1+z/s^2) \to \cos(\sqrt{-2z})$ for $s \to \infty$ and similarly $\alpha(z) = \mathrm{sinc}(\sqrt{-2z})$.

**Lemma 3.6.** *For all $\eta$ small enough and all $s$ large enough, we have the following estimate:*

$$\frac{s^2\omega_1}{T_s(w_0)^2}\frac{1-(1-\omega_1)^2}{1-(\omega_0-\omega_1)^2} \leq 1 \tag{28}$$

*Proof of Lemma 3.6.* Using the Lemma 3.5 we we have for $s \to \infty$, uniformly for all $\eta \in [0, \eta_0]$,

$$\frac{s^2\omega_1}{T_s(\omega_0)^2} \to \frac{2\sqrt{2\eta}}{\sinh(2\sqrt{2\eta})} \; and \; \frac{1-(1-\omega_1)^2}{1-(\omega_0-\omega_1)^2} \to \frac{1}{1-C\eta}.$$

Now if we expand both functions in Taylor series we get:

$$\frac{2\sqrt{2\eta}}{\sinh(2\sqrt{2\eta})} = 1 - \frac{4}{3}\eta + \mathcal{O}(\eta^2) \; and \; \frac{1}{1-C\eta} = 1 + \eta + \mathcal{O}(\eta^2), \tag{29}$$

and this implies that as $s \to \infty$

$$\frac{s^2\omega_1}{T_s(w_0)^2}\frac{1-(1-\omega_1)^2}{1-(\omega_0-\omega_1)^2} \to (1-\frac{4}{3}\eta+\mathcal{O}(\eta^2))(1+\eta+\mathcal{O}(\eta^2)) = 1-\frac{1}{3}\eta+\mathcal{O}(\eta^2), \tag{30}$$

which is less than 1 for $\eta$ small enough. □

**Remark 3.7.** *Numerical evidence suggests that the result of Theorem 3.3 holds for all $s \geq 1$ and all $\eta \geq 0$. Indeed, it can be checked numerically that (28) holds for all $\eta \in (0,1)$ and all $s \geq 1$.*

*Proof of Theorem 3.3.* Setting $x = w_0 + w_1 p$, a calculation yields

$$\begin{aligned}\mathbb{E}(|R(p,q,\xi)|^2) &\leq \mathbb{E}(|R(p,\sqrt{-2p})|^2) \\ &= \frac{T_s(x)^2}{T_s(w_0)^2} + \frac{U_{s-1}(x)^2}{U_{s-1}(w_0)^2}(1+\frac{w_1}{2}p)^2(-2p)\end{aligned}$$

The proof is conducted in two steps, where we treat separately the cases $p \in [-2\omega_1^{-1}, -1]$ and $p \in [-1, 0]$. For the first case $p \in [-2\omega_1^{-1}, -1]$, which corresponds to $x \in [-1+\eta/s^2, \omega_0-\omega_1]$, we have

$$\begin{aligned}\mathbb{E}(|R(p,q,\xi)|^2) &= \frac{T_s(x)^2}{T_s(w_0)^2} + \frac{U_{s-1}(x)^2}{U_{s-1}(w_0)^2}\left(1-\frac{w_0-x}{2}\right)^2 2\frac{w_0-x}{w_1} \\ &= \frac{T_s(x)^2}{T_s(w_0)^2} + U_{s-1}(x)^2(1-x^2)Q_s(x)\end{aligned}$$

where we denote

$$Q_s(x) = \frac{s^2\omega_1}{T_s(w_0)^2}\left(\frac{1+x-\frac{\eta}{s^2}}{2}\right)\frac{1-(x-\frac{\eta}{s^2})^2}{1-x^2}$$

First, we note that $\frac{1+x-\frac{\eta}{s^2}}{2} \in [0, 1-\frac{\omega_1}{2}]$. Next, using $\frac{\eta}{s^2} \leq 2$, we deduce $\frac{d}{dx}\left(\frac{1-(x-\frac{\eta}{s^2})^2}{1-x^2}\right) = \frac{2\eta}{s^2}\frac{1+x^2-\eta/s^2 x}{(1-x^2)^2} \geq \frac{2\eta}{s^2}\frac{(1-x)^2}{(1-x^2)^2} \geq 0$. Thus, $\frac{1-(x-\frac{\eta}{s^2})^2}{1-x^2}$ is an increasing function of $x$, smaller than its value at $x = \omega_0 - \omega_1$,

$$\frac{1-(x-\frac{\eta}{s^2})^2}{1-x^2} \leq \frac{1-(1-\omega_1)^2}{1-(\omega_0-\omega_1)^2}$$

Using Lemma 3.6 we obtain $|Q_s(x)| \leq 1$. This yields $\mathbb{E}(|R(p,q,\xi)|^2) \leq T_s(x)^2 + U_{s-1}(x)^2(1-x^2) = 1$.

For the second case $p \in [-1,0]$ wich corresponds to $x \in [\omega_0 - \omega_1, \omega_0]$, we deduce from $T_s(x)^2 + U_{s-1}(x)^2(1-x^2) = 1$ that

$$\mathbb{E}(|R(p,q,\xi)|^2) \leq \frac{1}{T_s(w_0)^2} + \frac{U_{s-1}(x)^2}{U_{s-1}(w_0)^2}\left((1+\frac{w_1}{2}p)^2(-2p) - \frac{(1-x^2)U_{s-1}(\omega_0)^2}{T_s(w_0)^2}\right)$$

Using Lemma 3.5, we get

$$\begin{aligned}\mathbb{E}(|R(p,q,\xi)|^2) &\leq \frac{1}{T_s(w_0)^2} + \frac{U_{s-1}(x)^2}{U_{s-1}(w_0)^2}\left((1+\frac{w_1}{2}p)^2(-2p) - \frac{(1-x^2)U_{s-1}(\omega_0)^2}{T_s(w_0)^2}\right) \\ &\to l(\eta, p) := \frac{1}{\cosh^2\sqrt{2\eta}} + \frac{\alpha(\eta+p/c)^2}{\alpha(\eta)^2}(-2p(C-1)+2C^2\eta) \quad as \quad s \to \infty.\end{aligned}$$

Using the fact that $C = 1 - \frac{2}{3}\eta + O(\eta^2)$ one can easily show that

$$\frac{dl}{d\eta}|_{\eta=0} = -2 + \alpha(p)^2(-\frac{4}{3}p + 2).$$

By Taylor series in the neighborhood of zero we have

$$\alpha(p)^2 = 1 + \frac{2}{3}p + \frac{8}{45}p^2 + O(p^3),$$

and for $p \in [-1,0]$, $\alpha(p)^2 \leq 1 + \frac{2}{3}p + \frac{8}{45}p^2$, thus

$$\begin{aligned}\frac{dl}{d\eta}|_{\eta=0} &\leq -2 + (1 + \frac{2}{3}p + \frac{8}{45}p^2)(-\frac{4}{3}p + 2) \\ &= -\frac{8}{135}p^2(4p+9) \leq 0 \quad for \quad p \in [-1,0].\end{aligned}$$

Therefore, there exists $\eta_0$ small enough such that for all $\eta \leq \eta_0$, $l(\eta) \leq l(0) = 1$. This concludes the proof of Theorem 3.3. □

# 4 Convergence analysis

We show in this section that the proposed scheme (23) has strong order 1/2 and weak order 1 for general systems of SDEs of the form (1) with Lipschitz and smooth vector fields, analogously to the simplest Euler-Maruyama method.

We denote by $C_P^4(\mathbb{R}^d, \mathbb{R}^d)$ the set of functions from $\mathbb{R}^d$ to $\mathbb{R}^d$ that are 4 times continuously differentiable with all derivatives with at most polynomial growth. The following theorem shows that to proposed SK-ROCK has strong order 1/2 and weak order 1 for general SDEs.

**Theorem 4.1.** *Consider the system of SDEs* (1) *on a time interval of length* $T > 0$, *with* $f, g \in C_P^4(\mathbb{R}^d, \mathbb{R}^d)$, *Lipschitz continuous. Then the scheme* (23) *has strong order* 1/2 *and weak order* 1,

$$\mathbb{E}|(\|X(t_n) - X_n\|) \leq Ch^{1/2}, \qquad t_n = nh \leq T, \tag{31}$$

$$|\mathbb{E}(\phi(X(t_n))) - \mathbb{E}(\phi(X_n))| \leq Ch, \qquad t_n = nh \leq T, \tag{32}$$

*for all* $\phi \in C_P^4(\mathbb{R}^d, \mathbb{R})$, *where* $C$ *is independent of* $n, h$.

For the proof the Theorem4.1, the following lemma will be useful. It relies on the linear stability analysis of Lemma 3.1.

**Lemma 4.2.** *The scheme* (23) *has the following Taylor expansion after one timestep,*

$$X_1 = X_0 + hf(X_0) + \sum_{r=1}^{m} g^r(X_0)\Delta W_j + h\left(\frac{T_s''(\omega_0)\omega_1^2}{T_s(\omega_0)} + \frac{\omega_1}{2}\right) f'(X_0) \sum_{r=1}^{m} g^r(X_0)\Delta W_j + h^2 R_h(X_0),$$

*where all the moments of $R_h(X_0)$ are bounded uniformly with respect to $h$ assumed small enough, with a polynomial growth with respect to $X_0$.*

*Proof.* Using the definition (23) of the scheme and the recurrence relations (11),(25), we obtain by induction on $i = 1, \ldots, s$,

$$K_i = X_0 + h\frac{T_i'(\omega_0)\omega_1}{T_i(\omega_0)} f(X_0) + \frac{sT_i'(\omega_0)\omega_1}{iT_i(\omega_0)} \sum_{r=1}^{m} g^r(X_0)\Delta W_j$$
$$+ h\left(\frac{sT_i''(\omega_0)\omega_1^2}{iT_i(\omega_0)} + \frac{sT_i'(\omega_0)\omega_1^2}{2iT_i(\omega_0)}\right) f'(X_0) \sum_{r=1}^{m} g^r(X_0)\Delta W_j + h^2 R_{i,h}(X_0),$$

and $R_{i,h}(X_0)$ has the properties claimed on $R_h(X_0)$. Using $\omega_1 = T_s(\omega_0)/T_s'(\omega_0)$, this yields the result for $X_1 = K_s$. □

*Proof of Theorem 4.1.* A well-known theorem of Milstein [17] (see [18, Chap. 2.2]) allows to infer the global orders of convergence from the error after one step. We first show that for all $r \in \mathbb{N}$ the moments $\mathbb{E}(|X_n|^{2r})$ are bounded for all $n, h$ with $0 \le nh \le T$ uniformly with respect to all $h$ sufficiently small. Then, it is sufficient to show the local error estimate

$$|\mathbb{E}(\phi(X(t_1))) - \mathbb{E}(\phi(X_1))| \le Ch^2,$$

for all initial value $X(0) = X_0$ and where $C$ has at most polynomial growth with respect to $X_0$, to deduce the weak convergence estimate (32). For the strong convergence (32), using the classical result from [16], it is sufficient to show in addition the local error estimate

$$\mathbb{E}(\|X(t_1) - X_1\|) \le Ch$$

for all initial value $X(0) = X_0$ and where $C$ has at most polynomial growth with respect to $X_0$. These later two local estimates are an immediate consequence of Lemma 4.2.

To conclude the proof of the global error estimates, it remains to check that for all $r \in \mathbb{N}$ the moments $\mathbb{E}(|X_n|^{2r})$ are bounded uniformly with respect to all $h$ small enough for all $0 \le nh \le T$. We use here the approach of [18, Lemma 2.2, p. 102] which states that it is sufficient to show

$$|\mathbb{E}(X_{n+1} - X_n|X_n)| \le C(1 + |X_n|)h, \qquad |X_{n+1} - X_n| \le M_n(1 + |X_n|)\sqrt{h}, \tag{33}$$

where $C$ is independent of $h$ and $M_n$ is a random variable with moments of all orders bounded uniformly with respect to all $h$ small enough. These estimates are a straightforward consequence of the definition (23) of the scheme and the linear growth of $f, g$ (a consequence of their Lischitzness). This concludes the proof of Theorem 4.1. □

**Remark 4.3.** *In the case of additive noise, i.e. $g^r, r = 1, \ldots, m$ are constant fonctions, one can show that the order of strong convergence* (31) *become* 1*, analogously to the case of the Euler-Maruyama method. For a general multiplicative noise, a scheme of strong order one can also be constructed with $\mathbb{E}(|R(p,q,\xi)|^2) \le 1$ for all $p \in [-2\omega_1^{-1}, 0]$ and all $q$ with $p + \frac{|q|^2}{2} \le 0$, as it can be check numerically. The idea is to modify the first stages of the scheme such that the stability function* (27) *becomes*

$$R(p,q,\xi) = \frac{T_s(\omega_0 + \omega_1 p)}{T_s(\omega_0)} + \frac{U_{s-1}(\omega_0 + \omega_1 p)^2}{U_{s-1}(\omega_0)^2}(1 + \frac{w_1}{2}p - \frac{\omega_1^4}{2}p^2)\left(q\xi + q^2\frac{\xi^2 - 1}{2}\right).$$

*We refer to [7, Remark 3.2] for details.*

# 5 Long term accuracy for Brownian dynamics

In this section we discuss the long-time accuracy of the SK-ROCK for Brownian dynamics (also called overdamped Langevin dynamics). We will see that using postprocessing techniques we can derive an SK-ROCK method that capture the invariant measure of Brownian dynamics with second order accuracy. In doing so, we do not need our stabilized method to be of weak order 2 and obtain a method that is cheaper as the second order weak S-ROCK method proposed [7] that has many more function evaluations per time-step and a smaller stability domain.

## 5.1 An exact SK-ROCK method for the Orstein-Uhlenbeck process.

We consider the 1-dimensional Orstein-Uhlenbeck problem

$$dX(t) = -\delta X(t)dt + \sigma dW(t), \tag{34}$$

that is ergodic and has a Gaussian invariant measure with mean zero and variance given by $\lim_{t\to\infty} \mathbb{E}(X(t)^2) = \sigma^2/(2\delta)$. Applying the SK-ROCK method to the above system we obtain

$$X_{n+1} = A(p)X_n + B(p)\sigma\sqrt{h}\xi_n \tag{35}$$

where $p = -\delta h$ $\xi_n \sim \mathcal{N}(0,1)$ is a Gaussian variable and similarly as for (27) we have

$$A(p) = \frac{T_s(\omega_0+\omega_1 p)}{T_s(\omega_0)}, \quad B(p) = \frac{U_{s-1}(\omega_0+\omega_1 p)}{U_{s-1}(\omega_0)}(1+\frac{\omega_1}{2}p). \tag{36}$$

A simple calculation (using that $|A(p)| < 1$) gives

$$\lim_{n\to\infty} \mathbb{E}(X_n^2) = \frac{\sigma}{2\delta}\frac{2pB(p)^2}{A(p)^2-1} = \frac{\sigma}{2\delta}\tilde{R}(p). \tag{37}$$

From the above equation, we see that the SK-ROCK method will have order $r$ for the invariant measure of (34) if and only if $\tilde{R}(p) = 1 + \mathcal{O}(p^r)$ an a short calculation using (36) reveals that $\tilde{R}(p) = 1 + \mathcal{O}(p)$, it has order one for the invariant measure (we of course already know that since the SK-ROCK has weak order one). We next apply the techniques of postprocessed integrator derived in [23] [cite Butcher case ODE ??]. The idea is to consider a postprocessed dynamics $\bar{X}_n = G_n(X_n)$ (of negligible cost) such that the process $\bar{X}_n$ approximate the invariant measure of dynamical system with higher order. For the process (34), we consider the postprocessor

$$G_n(x) = x + c\sigma\sqrt{h}\xi_n. \tag{38}$$

In the case of SK-ROCK method with $\eta = 0$ (zero damping), we have $A(p) = T_s(1+p/s^2), B(p) = U_{s-1}(1+p/s^2)(1+p/(2s^2))/s$. Setting $c = 1/(2s)$ using the identity

$$(1-x^2)U_{s-1}^2(x) = 1 - T_s^2(x),$$

with $x = 1 + p/s^2$, reveals that $\tilde{R}(p) = 1$, hence the SK-ROCK method captures exactly the invariant measure of the 1-dimensional Orstein-Uhlenbeck problem (34). Such a behavior is known for the drift-implicit $\theta$ method and has recently also been shown for the non-Markovian Euler scheme [14]. In [23] an interpretation of the scheme [14] as an Euler-Maruyama method with postprocessing (38) $c = 1/2$ has been proposed and we observe that this is exactly the same postprocessor as for the SK-ROCK method (with $s = 1$). As the SK-ROCK method with zero damping is mean-square stable, it can be seen as a stabilized version of the scheme [14]. However, the SK-ROCK method with zero damping is dangerous to use as the stability domain along the drift axis does not allow for any imaginary perturbation at the points where $T_s(1 + p/s^2) = 1$.

**Stability analysis for Orstein-Uhlenbeck** Let $A \in \mathbb{R}^{d\times d}$ denote a symmetric matrix with eigenvalues $-\lambda_d \leq \ldots \leq -\lambda_1 < 0$, and consider the Orstein-Uhlenbeck problem

$$dX(t) = AX(t)dt + \sigma dW(t) \tag{39}$$

where $W(t)$ denotes a $d$-dimensional standard Wiener process.

The following theorem shows that the damping parameter $\eta > 0$ plays an essential role to warranty the convergence to the invariant measure at an exponentially fast rate.

**Theorem 5.1.** *Let $\eta > 0$. Consider the scheme (23) applied to (41) with stepsize $h$ and stage parameter $s$ such that $L_s \geq h\lambda_d$. Then, for all $h \leq \eta/\lambda_1$,*

$$|\mathbb{E}(\phi(X_n)) - \int_{\mathbb{R}^d} \phi(x)\rho_\infty^h(x)dx| \leq C\exp(-\lambda_1(1+\eta)^{-2}t_n)$$

*where $C$ is independent of $h, n, s, \lambda_1, \ldots, \lambda_d$.*

*Proof.* It is sufficient to show the estimate

$$|R_s(-\lambda_j h)| \leq \exp(-\lambda_1(1+\eta)^{-2}h) \tag{40}$$

for all $h \leq h_0$, where we denote $R_s(z) = \frac{T_s(\omega_0+\omega_1 z)}{T_s(\omega_0)}$. Indeed, considering two initial conditions $X_0^1, X_0^2$ and the corresponding numerical solutions $X_n^1, X_n^2$ obtained with (23), we obtain $X_n^1 - X_n^2 = R_s(hA)(X_{n-1}^1 - X_{n-1}^2)$ and using $\|R_s(hA)\| = \max_j |R_s(-\lambda_j h)|$ , we deduce by induction on $n$,

$$\|X_n^1 - X_n^2\| \leq \exp(-\lambda_1(1+\eta)^{-2}t_n),$$

and taking $X_0^2$ distributed according to the numerical invariant measure yields the result.

For the proof of (40), let $z = -\lambda_j h$. Consider first the case $z \in (-\eta\omega_1^{-1}s^{-2}, 0)$. Using the convexity of $R_s(z)$ on $[-\eta\omega_1^{-1}s^{-2}, 0]$ (note that $T_s'(x)$ is increasing on $[1,\infty)$), we can bound $R_s(z)$ by the affine function passing by the points $(x_1, R(x_1)), (x_2, R(x_2))$ with $x_1 = -\eta\omega_1^{-1}s^{-2}$, $x_2 = 0$,

$$R_s(z) \leq 1 + z(1 - 1/T_s(\omega_0))\eta^{-1}\omega_1^{-1}s^{-2}$$

We note that $\omega_1^{-1}s^{-2}$ and $T_s(\omega_0)$ are increasing functions of $s$ bounded from below by there values at $s = 1$, which yields $\omega_1^{-1}s^{-2} \geq (1+\eta)^{-1}$ and $T_s(\omega_0) \geq 1+\eta$. We obtain

$$(1 - 1/T_s(\omega_0))\eta^{-1}\omega_1^{-1}s^{-2} \geq (1 - (1+\eta)^{-1})\eta^{-1}(1+\eta)^{-1} = (1+\eta)^{-2}.$$

This yields for all $z \in [-\eta\omega_1^{-1}s^{-2}, 0]$,

$$R_s(z) \leq 1 + z(1+\eta)^{-2} \leq \exp(z(1+\eta)^{-2})$$

where we used the convexity of $\exp(z(1+\eta)^{-2})$ bounded from below by its tangent at $z = 0$. We obtain

$$R_s(-\lambda_j h) \leq e^{-\lambda_j h(1+\eta)^{-2}} \leq e^{-\lambda_1 h(1+\eta)^{-2}}.$$

We now consider the case $z \in [-L_s, -\eta\omega_1^{-1}s^{-2}]$. We have $|\omega_0 + \omega_1 z| \leq 1$, thus $|T_s(\omega_0 + \omega_1 z)| \leq 1$ and

$$|R_s(z)| \leq 1/T_s(\omega_0) \leq \exp(-\lambda_1(1+\eta)^{-2}h)$$

for all $h \leq (1+\eta)^2\log(T_s(\omega_0))/\lambda_1$, and thus also for $h \leq \eta/\lambda_1$ where we use $T_s(\omega_0) \geq 1+\eta$ and $(1+\eta)^2\log(1+\eta) \geq \eta$. This concludes the proof. □

**Remark 5.2.** *Notice that Theorem 5.1 allows to use an $h$-dependent value of $\eta$ such as $\eta = h\widetilde{\lambda}_1$ where $\widetilde{\lambda}_1 \geq \lambda_1$ is an upper bound for $\lambda_1$. Then, the exponential convergence of Theorem 5.1 holds for all $h \leq 1$.*

We end this section by noting that being exact for the invariant measure of Brownian dynamics is only true for the SK-ROCK method (or the method in [14]) for linear processes with a symmetric drift. But second order accuracy for the invariant measure has been shown for the method [14]) in [15] for general nonlinear Brownian dynamics. This will also be shown for the SK-ROCK method in the next section.

### 5.2 A second order SK-ROCK method for nonlinear Brownian dynamics

We consider the overdamped Langevin equation,

$$dX(t) = -\nabla V(X(t))dt + \sigma dW(t), \tag{41}$$

where the stochastic process $X(t)$ takes values in $\mathbb{R}^d$ and $W(t)$ is a $d$-dimensional Wiener process. We assume that the potential $V : \mathbb{R}^d \to \mathbb{R}$ has class $C^\infty$ and satisfies the at least quadratic growth assumption

$$x^T \nabla V(x) \geq C_1 x^T x - C_2$$

for two constants $C_1, C_2 > 0$ independent of $x \in \mathbb{R}^d$. The above assumptions warranty that the system (42) is ergodic with exponential convergence to a unique invariant measure with Gibbs density $\rho_\infty = Z\exp(-2\sigma^{-2}V(x))$,

$$|\mathbb{E}(\phi(X(t)) - \int_{\mathbb{R}^d} \phi(x)\rho_\infty(x)dx| \leq Ce^{-\lambda t},$$

for test function $\phi$ and all initial condition $X_0$, where $C, \lambda$ are independent of $t$.

We propose to modify the internal stage $R_1 = X_0 + \mu_1 hf(X_0 + \nu_1 G) + \kappa_1 G$ of the method (23) as follows:

$$R_1 = X_0 + \mu_1 hf(X_0 + \nu_1 G) + \kappa_1 G + \alpha h\big(f(X_0 + \nu_1 G) - 2hf(X_0) + f(X_0 - \nu_1 G)\big), \tag{42}$$

where $\alpha$ is a parameter whose optimal value depends on $s$ and $\eta$ is discussed below.

**Remark 5.3.** *Notice that for $\alpha = 0$, we recover the original definition from* (23). *We note that the parameter $\alpha$ does not modify the stability function of Lemma 3.1, and yields a perturbation of order $\mathcal{O}(h^2)$ in the definition of $X_1$. Thus, the results of Theorem 3.3 and Theorem 4.1 remain valid for any values of $\alpha$ for the scheme* (23) *with modified internal stage* (43).

**Theorem 5.4.** *Under the above assumptions, consider the scheme* (23) *with modified internal stage $R_1$ using* (43). *Consider in addition a postprocessor defined as*

$$\overline{X}_n = X_n + c\sigma\sqrt{h}\xi.$$

*where*

$$\alpha = \frac{2}{s\omega_0\omega_1}(c^2 + \frac{\omega_1^2 T_s''(\omega_0)}{2T_s(\omega_0)} - r_s)$$
$$c^2 = -\frac{1}{4} + \frac{\omega_1}{2} + \frac{\omega_1 T_s''(\omega_0)}{T_s'(\omega_0)} - \frac{\omega_1^2 T_s''(\omega_0)}{4T_s(\omega_0)}$$

*and $r_s$ can be obtained as follows:*

$$r_0 = 0 \tag{43}$$

$$r_1 = \frac{s^2\omega_1^3}{4\omega_0} := \Delta_1 \tag{44}$$

$$r_i = \nu_i r_{i-1} + \kappa_i r_{i-2} + \Delta_i, \tag{45}$$

*where,*

$$\Delta_i = \mu_i \frac{sT'_{i-1}(\omega_0)\omega_1}{(i-1)T_{i-1}(\omega_0)}.$$

*Then, $\overline{X}_n$ yields order two for the invariant measure,*

$$|\mathbb{E}(\phi(\overline{X}_n) - \int_{\mathbb{R}^d} \phi(x)\rho_\infty(x)dx| \leq C_1 e^{-\lambda t_n} + C_2 h^2, \tag{46}$$

*for all $t_n = nh$, where $C_1, C_2$ are independent of $h$ assumed small enough, and $C_2$ is independent of the initial condition $X_0$.*

*Proof.* A weak Taylor expansion yields $\mathbb{E}(\phi(X_1)|X_0 = x) = \phi(x) + h\mathcal{L}\phi(x) + h^2\mathcal{A}_1\phi(x) + \mathcal{O}(h^3)$, where for $f = -\nabla V(x)$, we have:

$$\mathcal{A}_1\phi = \frac{1}{2}\phi''(f,f) + \frac{\sigma^2}{2}\sum_{i=1}^{d}\phi'''(e_i,e_i,f) + \frac{\sigma^4}{8}\sum_{i,j=1}^{d}\phi^{(4)}(e_ie_i,e_j,e_j) + c_2\phi' f' f \tag{47}$$

$$+ \ c_3\frac{\sigma^2}{2}\phi'\sum_{i=1}^{d} f''(e_i,e_i) + c_4\sigma^2\sum_{i=1}^{d}\phi''(f'e_i,e_i). \tag{48}$$

where,

$$c_2 = \frac{\omega_1^2 T''_s(\omega_0)}{2T_s(\omega_0)} \tag{49}$$

$$c_3 = r_s + \frac{\omega_0}{s\omega_1}\alpha \tag{50}$$

$$c_4 = \frac{T''_s(\omega_0)\omega_1}{T'_s(\omega_0)} + \frac{\omega_1}{2}. \tag{51}$$

Indeed, We set:

$$\tilde{r}_0 = 0 \tag{52}$$

$$\tilde{r}_1 = \Delta_1 + \alpha \tag{53}$$

$$\tilde{r}_i = \nu_i\tilde{r}_{i-1} + \kappa_i\tilde{r}_{i-2} + \Delta_i, \tag{54}$$

and

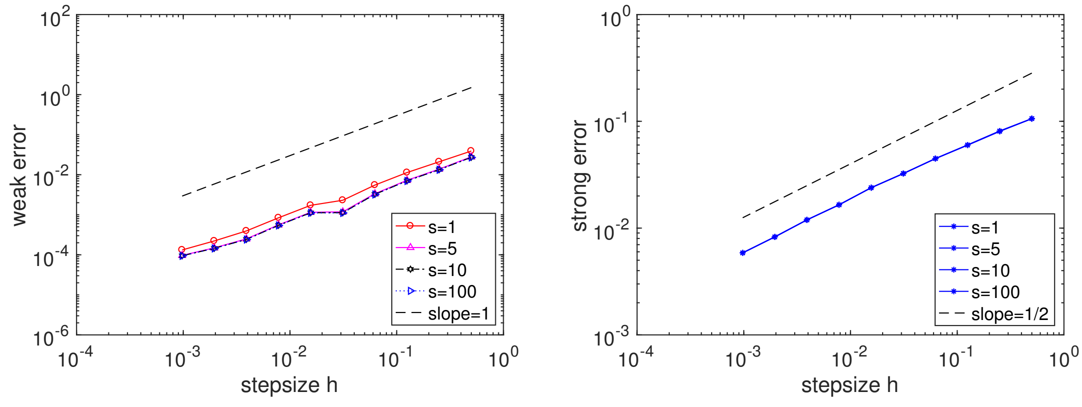
$$d_i = \tilde{r}_i - r_i.$$

It can be easily shown that

$$d_i = \frac{U_{i-1}(\omega_0)}{T_i(\omega_0)}\omega_0\alpha \quad \forall \quad i = 0,..,s.$$

Taking $c_3 = \tilde{r}_s$ and using $\tilde{r}_s = r_s + d_s$ we get the desired value for $c_3$.
Using [23, Theorem 4.1], the order two convergence estimate (47) holds if and only if $(A_1 + [\overline{A}_1, \mathcal{L}])^*\rho_\infty = 0$, which is equivalent to

$$\langle \mathcal{A}_1\phi + [\overline{A}_1, \mathcal{L}]\phi \rangle = 0$$




(a) Weak error $|\mathbb{E}(\operatorname{arcsinh}(X(T)) - \mathbb{E}(\operatorname{arcsinh}(X_N))|$.

(b) Strong error $\mathbb{E}(|X(T) - X_N|)$.

Figure 4: Nonlinear problem (60). Strong and weak convergence plots using SK-ROCK with final time $T = 1$, stepsizes $h = 2^{-p}, p = 1..10$, $10^4$ samples and number of stages $s = 1, 5, 10, 100$.

for all test function $\phi$, where we define $\langle\phi\rangle = \int_{\mathbb{R}^d} \phi\rho_\infty dx$. Following the proof of [23, Theorem 4.2], this quantity can be computed as

$$\langle \mathcal{A}_1\phi + [\overline{A}_1, \mathcal{L}]\phi \rangle = \left\langle (c_3 - c_2 - c^2)\frac{\sigma^2}{2}\phi' \sum_{i=1}^{d} f''(e_i, e_i) + (c_4 - \frac{1}{4} - \frac{c_2}{2} - c^2)\sigma^2 \sum_{i=1}^{d} \phi''(f'e_i, e_i) \right\rangle \tag{55}$$

The condition to achieve an order two convergence rate (47) for the invariant measure is thus

$$c_3 - c_2 = c_4 - \frac{1}{4} - \frac{c_2}{2} = c^2. \tag{56}$$

We observe that the above condition can be satisfied for a unique value of $\alpha$ and $c \geq 0$ which concludes the proof.

□

**Remark 5.5.** *TO BE COMPLETED (not accurate)*

$$c_3 - c_2 = \alpha - \frac{1}{6} + \frac{2}{3s^2} - \frac{4}{45}\eta + \mathcal{O}(\eta^2 + s^{-4}) \tag{57}$$

$$c_4 - \frac{1}{4} - \frac{c_2}{2} = \frac{1}{4s^2} + \frac{4}{30}\eta + \mathcal{O}(\eta^2 + s^{-4}) \tag{58}$$

*Using* $c_3 = \frac{c_4}{2} - \frac{1}{8} + \frac{c_2}{4}$

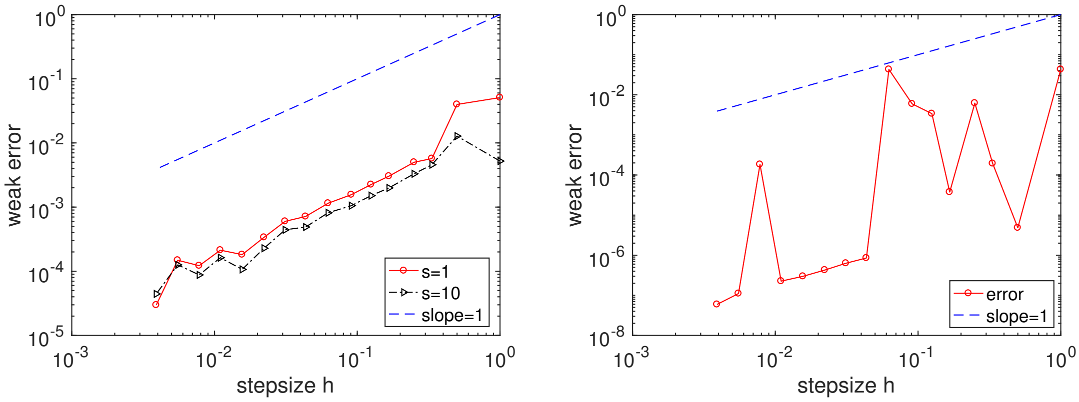
# 6 Numerical experiments

## 6.1 A nonlinear nonstiff problem

We first consider the following non-stiff non-linear SDE,

$$dX = \left(\frac{1}{4}X + \frac{1}{2}\sqrt{X^2+1}\right)dt + \sqrt{\frac{1}{2}(X^2+1)}dW, \quad X(0) = 0. \tag{59}$$

whose exact solution is $X(t) = \sinh(\frac{t}{2} + \frac{W(t)}{\sqrt{2}})$. In Figure 4, we consider the SK-ROCK method (23) and plot the strong error $\mathbb{E}(|X(T) - X_N|$ and second moment errors $|\mathbb{E}(X(T)^2 - \mathbb{E}(X_N^2|$




(a) Non-stiff case $\lambda_1 = \mu_1 = 1$.  (b) Stiff case $\lambda_1 = \mu_1^2 = 500$.

Figure 5: Nonlinear problem (61) with $\alpha = \mu_2 = 1, \lambda_2 = 4$. Weak convergence plots using SK-ROCK for $\mathbb{E}(X_1(T)^2)$ where $T = 1$, $h = T/[2^{i/2} + 0.5], i = 1, \ldots, 16$, and $10^6$ samples. For the stiff case (Fig. 5b), the method uses the following number of stages respectively: $s = 17, 12, 10, 9, 7, 6, 5, 5, 4, 3, 3, 3, 2, 2, 2, 2$.

at the final time $T = Nh = 1$ using $10^4$ samples and number of stages $s = 1, 5, 10, 100$. We obtain convergence slopes 1 and 1/2, respectively, which confirms Theorem 4.1 stating the strong order 1/2 and weak order 1 of the proposed scheme. Note that $s = 1$ stage is sufficient for the stability of the scheme in the non-stiff case. The results for $s = 5, 10, 100$ yield nearly identical curves which illustrates that the error constants of the method are nearly independent of the stage number of the scheme.

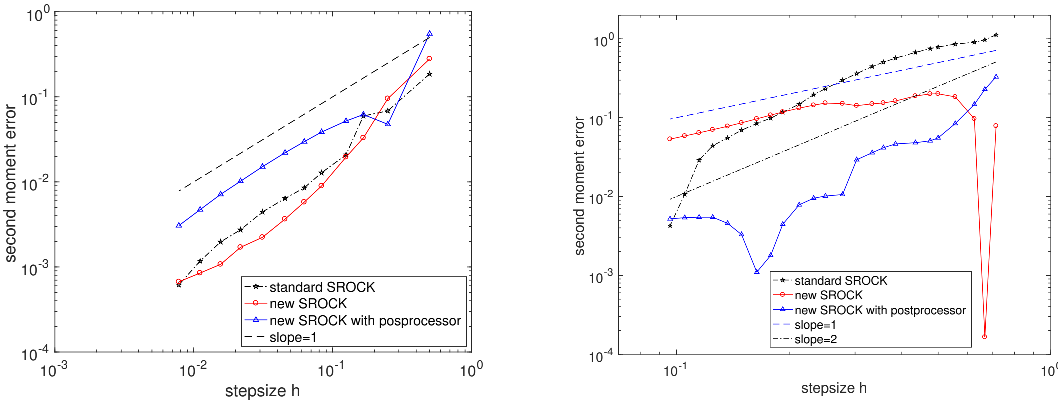
### 6.2 Nonlinear nonglobally Lipschitz stiff problems

Consider the following nonlinear SDE in dimensions $d = 2$ with a one-dimensional noise ($d = 2, m = 1$).This is a modification of a one-dimensional population dynamics model from [10, Chap. 6.2] considered in [7, 6] for testing stiff integrator performances.

$$\begin{aligned} dX &= (\alpha(Y-1) - \lambda_1 X(1-X))dt - \mu_1 X(1-X)dW, \\ dY &= -\lambda_2 Y(1-Y)dt - \mu_2 Y(1-Y)dW, \\ X(0) &= Y(0) = 0.95. \end{aligned} \tag{60}$$

Observe that linearizing (61) close to the equilibrium $(X, Y) = (1, 1)$, we recover for $\alpha = 0$ the scalar test problem (15). In Figure 5 we consider the SK-ROCK method applied to (61) with the initial condition $X(0) = Y(0) = 0.95$ close to this steady state and use the parameters $\alpha = \mu_2 = 1, \lambda_2 = 4$. In a nonstiff regime ($\lambda_1 = \mu_1 = 1$ in Figure 5(a)), we observe a convergence slope 1 which illustrates the weak order one of the scheme, although our analysis in Theorem 4.1 applies only for globally Lipschitz vector fields. The stage number $s = 1$ is sufficient for stability, but we also include for comparison the results for $s = 10$. The convergence curves are obtained as an average over $10^6$ samples. In a stiff regime ($\lambda_1 = \mu_1 = 500$), both the S-ROCK and SK-ROCK methods reveal unstable, we thus consider a modification of SK-ROCK described in the following remark.

**Remark 6.1.** *For severely nonlinear stiff problems, we thus introduce an additional stabilization in the definition of the noise term and replace $g^r(X_0)$ by $g^r(\widetilde{X}_0)$ in equation (24) where we define*




(a) Final time $T = 0.5$, $h = T/[2^{i/2}]$, $i = 1, \dots, 12$.

(b) Final time $T = 10$, $h = T/[13 \times 2^{i/8}]$, $i = 1, \dots, 24$.

Figure 6: Second moment errors versus the average number of drift function evaluations for problem (64) using S-ROCK and the new methods SK-ROCK and its postprocessed version PSK-ROCK. We use discrete random increments and $10^8$ samples.

*$\widetilde{X}_0$ as $X_1$ in (8) where $\mu_1 = \omega_1/\omega_0$ is unchanged and $\mu_i, \nu_i, \kappa_i, i = 2, \dots, s$ are replaced by*

$$\widetilde{\mu}_i = \frac{2\omega_1 U_{i-1}(\omega_0)}{U_i(\omega_0)}, \quad \widetilde{\nu}_i = \frac{2\omega_0 U_{i-1}(\omega_0)}{U_i(\omega_0)}, \quad \widetilde{\kappa}_i = -\frac{U_{i-2}(\omega_0)}{U_i(\omega_0)} = 1 - \widetilde{\nu}_i. \tag{61}$$

*For linear problems ($f(x) = \lambda x$ and $p$=$h\lambda$), notice that*

$$\widetilde{X}_0 = \frac{U_{s-1}(w_0 + w_1 p)}{U_{s-1}(w_0)} X_0. \tag{62}$$

*In practice, it is sufficient to apply the above stabilization only for the first few five steps of the integration. Alternatively, one could consider a large value of the damping $\eta$, or switch to drift-implicit schemes [12, 6].*

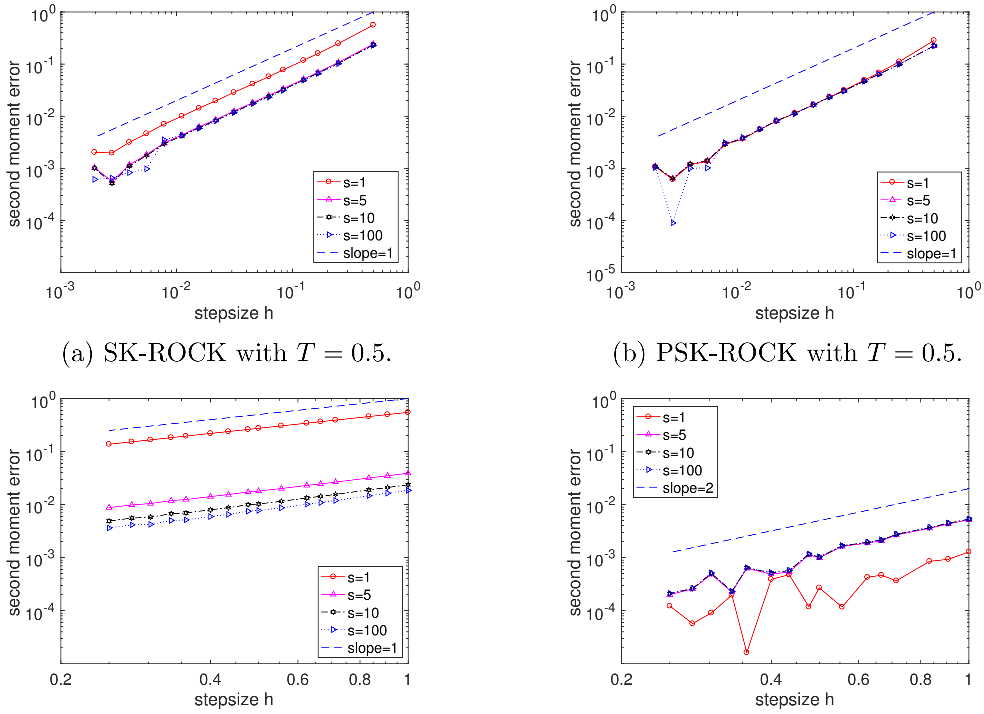
We plot in Figure 5(b) the convergence curves of the SK-ROCK method, where $s$ is adjusted automatically such that $Cs^2 \geq \rho h$ where $\rho = |\lambda_1| = 500$ which allows using large timesteps $h$. For the largest stepsize $h = 1$ and at $t = 0$, the SK-ROCK uses $s = 17$ stages for stabilizations, whereas the standard S-ROCK would need 35 stages.

### 6.3 Nonglobally Lipschitz Brownian dynamics SECTION TO BE UPDATED/COMPLETED

To illustrate the advantage of the PSK-ROCK method applied to nonglobally Lipschitz ergodic Brownian dynamics, we next consider the following double well potential $V(x) = (1 - X^2)^2/4$ and the corresponding one-dimensional Brownian dynamics problem

$$dX = (-X^3 + X)dt + \sqrt{2}dW, \qquad X(0) = 0, \tag{63}$$

In Figure 6, we compare the performances of S-ROCK and the new PSK-ROCK and PSK-ROCK method at short and long time (figure 6a and figure 6b respectively). Both methods exhibit similar accuracies TO BE UPDATED, with an advantage to the new ROCK method for large time steps at long time simulation. Also for long time, PSK-ROCK exhibits a slope two of accuracy with errors reduced by a factor $10 - 100$. For short time we can see order one with




(a) SK-ROCK with $T = 0.5$. (b) PSK-ROCK with $T = 0.5$.

(c) SK-ROCK with $T = 10$. (d) PSK-ROCK with $T = 10$.

Figure 7: Linear additive problem (65). Second moment error $\mathbb{E}(X(T)^2)$ for for short time $T = 0.5$ (top pictures) and long time $T = 10$ (bottom pictures) without (SK-ROCK) or with a postprocessor (PSK-ROCK). where, $h = T/[10 \times 2^{i/8}]$ for $T = 10$, and $h = T/[2^{i/2} + 0.5]$ for $T = 0.5$ with $i = 1, \ldots, 16$, and $10^8$ samples.

small errors for both methods even for PSK-ROCK. The SK-ROCK method used $s = 3, 2, 1$ stages for the different values of the stepsize $h$ while the standard S-ROCK used $s = 7, 6, 5, 4, 3, 2$ stages . Note that for the postprocessed SORCK, we used discrete random increments (with $\mathbb{P}(\xi = \pm\sqrt{3}) = 1/6, \mathbb{P}(\xi = 0) = 2/3)$ which exhibits in our numerical tests a better accuracy compared to Gaussian increments (see Remark TO BE COMPLETED). WE USED DISCRETE RANDOM INCREMENTS FOR POSTPROCESSED AND NON POSTPROCESSED SROCK, AND FOR STANDARD SROCK.

### 6.4 Linear case: Orstein-Uhlenbeck process

We now illustrate numerically the role of the postprocessor introduced in Theorem 5.4 for the scalar linear case of the Orstein-Uhlenbeck process in dimension $d = m = 1$,

$$dX = -\lambda X dt + \sigma dW, \quad X(0) = 2 \tag{64}$$

where we choose $\lambda = 1$ and $\mu = \sqrt{2}$. In Figure 7, we consider the SK-ROCK and PSK-ROCK methods with $s = 1, 5, 10, 100$ stages, respectively. For a short time $T = 0.5$ (Fig. 7(a)(b)), we observe weak convergence slopes (second moment $\mathbb{E}(X(T)^2)$) one as predicted by Theorem 4.1, and the postprocessor has nearly no effect of the errors. For a long time $T = 10$ where the solution of this ergodic SDE is close to equilibrium, we observe that the weak order one of SROCK2 (Fig. 7(c)) is improved to order two using the postprocessor (Fig. 7(d)), which confirms the statement of Theorem 5.4 that the postprocessed scheme has order two of accuracy for the invariant measure. For comparison, in Figure 8, we also include the results of the PSK-ROCK and without damping

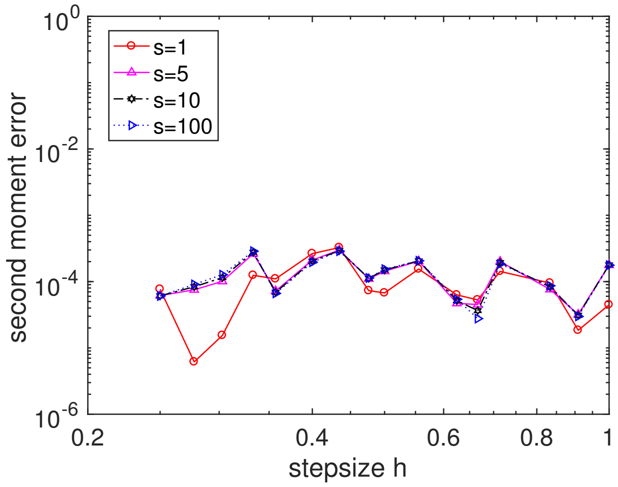



Figure 8: PSK-ROCK without damping ($\eta = 0$). Second moment error of problem (65), $T = 10, s = 1, 5, 10, 100$, using $M = 10^8$ samples (the Monte-Carlo error has size $1/\sqrt{M} = 10^{-4}$).

$\eta = 0$ using $10^8$ samples. We observe only Monte-Carlo errors with size $\simeq 10^{-4}$, which confirms that the SROCK method has no bias for the invariant measure in the absence of damping, as explained in Remark TO BE COMPLETED. Note however that this exactness results holds only for linear problems, and a positive damping parameter $\eta$ should be used for nonlinear SDEs for stabilization, as shown in Section 3.

### 6.5 Stochastic heat equation with multiplicative space-time noise

Although our analysis applies only to finite dimensional systems of SDEs, we consider the following stochastic partial differential equation (SPDE) obtained by adding multiplicative noise to the heat equation,

$$\begin{aligned}
\frac{\partial u(t,x)}{\partial t} &= \frac{\partial^2 u(t,x)}{\partial x^2} + \sigma u(t,x)\dot{W}(t,x), \quad (t,x) \in [0,T] \times [0,1] \\
u(0,x) &= 5\cos(\pi x), \;\; x \in [0,1], \\
u(t,0) &= 5, \;\; t \in [0,T], \;\; \frac{\partial u(t,1)}{\partial x} = 0, \;\; t \in [0,T]
\end{aligned} \tag{65}$$

where we descritize the Laplace operator with a standard finite difference formula. In Figure 9, we plot one realization of the SPDE using space stepsize $\Delta x = 1/100$ and timestepsize $\Delta t = 1/50$. Note that the Lipchitz constant associated to the space-discretization of (66) has size $\rho = 4\Delta x^{-2}$, and the stability condition is fulfilled for $s = 22$ stages. For comparison, the standard S-ROCK method would require $s = 46$ stages, while applying the standard Euler-Maruyama with a stable timestep $\Delta/s$ would require $s \geq \Delta t\rho/2 = 400$ intermediate steps. Notice that the initial condition in (66) satisfies the boundary conditions, which permits a smooth solution close to time $t = 0$. Taking alternatively the initial condition $u(x,0) = 1$ that does not satisfy the boundary conditions yields an inaccurate numerical solution with large oscillations close to the boundary $x = 0$. A possible remedy in such a case is to consider a larger damping parameter $\eta$.

In Figure 10a(a), we compare the number of function evaluations of the standard S-ROCK and new SK-ROCK methods when applied to the SPDE (66) with finite difference discretization with parameter $\Delta x = 1/100$. The better performance of SK-ROCK with damping $\eta = 0.05$ is due to its larger stability domain with size $\simeq 1.94s^2$ compared to the size $\simeq 0.33 \cdot s^2$ for S-ROCK.

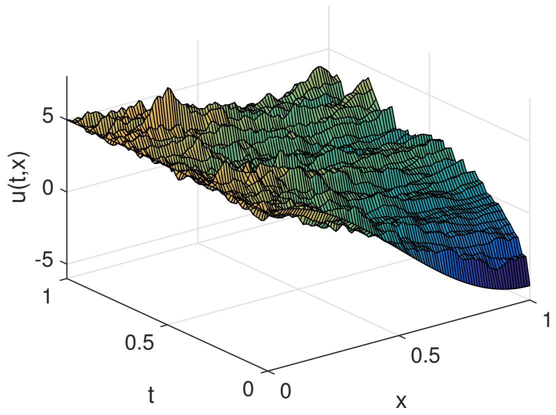



Figure 9: One realization of the SPDE (66) using SK-ROCK with discretization parameters $\Delta x = 1/100$, and $\Delta t = 1/50$, using $s = 22$ stages.

Observing the ratio of the two costs in Figure 10a(b), we see that the new SK-ROCK methods has a reduced cost for stabilization by an asymptotic factor of about $\sqrt{1.94/0.33} \simeq 2.4$ for large $s$ and large stepsizes, which confirms the stability analysis of Section 3.

# 7 Appendix

*Proof of Lemma 3.5.* We will prove the uniform convergence of the first limit only, since it will be useful in the proof of the next theorem. The others can be proved in a similar way.

First, let us write the two functions of $\eta$ in Taylor series:

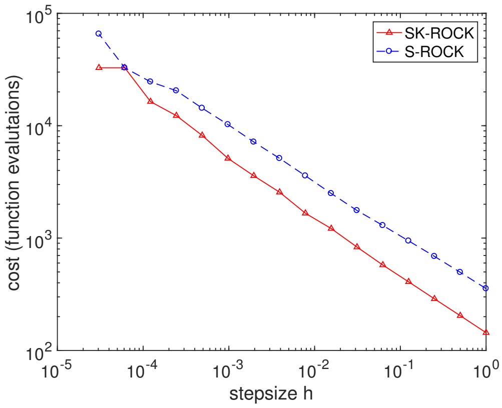
$$\begin{aligned} s^{-1}U_{s-1}(\omega_0) &= s^{-1}U_{s-1}(1+\frac{\eta}{s^2}) \\ &= s^{-1}\sum_{n=0}^{s-1}\frac{U_{s-1}^{(n)}(1)}{n!}(\frac{\eta}{s^2})^n \\ &= s^{-2}\sum_{n=0}^{s-1}\frac{T_s^{(n+1)}(1)}{n!}(\frac{\eta}{s^2})^n \\ &= \sum_{n=0}^{s-1}\left(\frac{1}{n!}\prod_{k=0}^{n}(1-\frac{k^2}{s^2})\prod_{k=0}^{n}\frac{1}{2k+1}\right)\eta^n \end{aligned}$$

Here we used the formula:

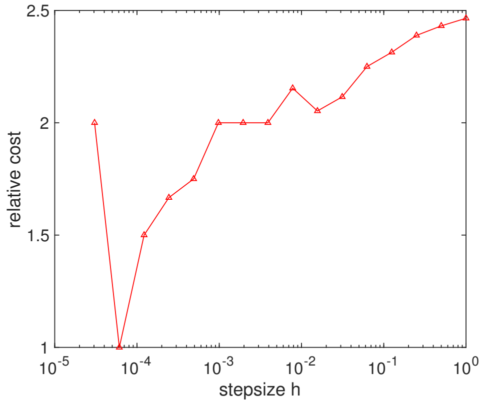
$$T_s^{(n)}(1) = \prod_{k=0}^{n-1}\frac{s^2-k^2}{2k+1}. \tag{66}$$

And,

$$\begin{aligned} \alpha(\eta) &= \sum_{n=0}^{\infty}\frac{2^n\eta^n}{(2n+1)!} \\ &= \sum_{n=0}^{\infty}\left(\frac{1}{n!}\prod_{k=0}^{n}\frac{1}{2k+1}\right)\eta^n \end{aligned}$$




(a) Cost with respect to stepsize $h$.




(b) ratio of the cost of S-ROCK method versus the cost of SK-ROCK method cost(S-ROCK)/cost(SK-ROCK).

Figure 10: Comparison of the cost (number of function evaluations) of the S-ROCK and SK-ROCK methods applied to the SPDE (66) with $\Delta x = 1/100$ as a function of the stepsize $h = 2^{-i}, i = 0, \dots, 15$.

Now let us prove the uniform convergence:

$$\begin{aligned}
\sup_{\eta\in[-\eta_0,\eta_0]} \left|s^{-1}U_{s-1}(\omega_0) - \alpha(\eta)\right| &= \sup_{\eta\in[-\eta_0,\eta_0]} \left|\sum_{n=0}^{s-1}\left(\frac{1}{n!}\prod_{k=0}^{n}(1-\frac{k^2}{s^2})\prod_{k=0}^{n}\frac{1}{2k+1}\right)\eta^n - \sum_{n=0}^{\infty}\left(\frac{1}{n!}\prod_{k=0}^{n}\frac{1}{2k+1}\right)\eta^n\right| \\
&\leq \sum_{n=0}^{s-1}\frac{\eta_0^n}{n!}\left(1-\prod_{k=0}^{n}(1-\frac{k^2}{s^2})\right)\prod_{k=0}^{n}\frac{1}{2k+1} + \sum_{n=s}^{\infty}\frac{\eta_0^n}{n!} \\
&\leq \sum_{n=0}^{s-1}\frac{\eta_0^n}{n!}\left(1-\prod_{k=0}^{n}(1-\frac{k^2}{s^2})\right)\frac{1}{2s-1} + \sum_{n=s}^{\infty}\frac{\eta_0^n}{n!}
\end{aligned}$$

The second term of the right hand side clearly goes to zero as $s$ goes to infinity since it is the remainder of a convergent series. For the first term we have:

$$\begin{aligned}
h_s(n) &:= \frac{\eta_0^n}{n!}\left(1-\prod_{k=0}^{n}(1-\frac{k^2}{s^2})\right)\frac{1}{2s-1} \to 0 \textit{ as } s\to\infty \textit{ for all fixed } n, \\
\textit{and} \quad & h_s(n) \leq k(n) := \frac{\eta_0^n}{n!} \textit{ with } \sum_{n=0}^{\infty}\frac{\eta_0^n}{n!} < \infty
\end{aligned}$$

So the dominated convergence theorem implies that:

$$\sum_{n=0}^{s-1}\frac{\eta_0^n}{n!}\left(1-\prod_{k=0}^{n}(1-\frac{k^2}{s^2})\right)\frac{1}{2s-1} \to 0 \textit{ as } s\to\infty,$$

and then the uniform convergence is proved. □

## References


[1] A. Abdulle. Fourth order Chebyshev methods with recurrence relation. *SIAM J. Sci. Comput.*, 23(6):2041–2054, 2002.

[2] A. Abdulle. *Explicit Stabilized Runge–Kutta Methods*, pages 460–468. Encyclopedia of Applied and Computational Mathematics, Springer Berlin Heidelberg, 2015.

[3] A. Abdulle and S. Cirilli. S-ROCK: Chebyshev methods for stiff stochastic differential equations. *SIAM J. Sci. Comput.*, 30(2):997–1014, 2008.

[4] A. Abdulle and T. Li. S-ROCK methods for stiff Ito SDEs. *Commun. Math. Sci.*, 6(4):845–868, 2008.

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

[6] 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(4):827–840, 2013.

[7] A. Abdulle, G. Vilmart, and K. C. Zygalakis. Weak second order explicit stabilized methods for stiff stochastic differential equations. *SIAM J. Sci. Comput.*, 35(4):A1792–A1814, 2013.

[8] E. Buckwar and C. Kelly. Towards a systematic linear stability analysis of numerical methods for systems of stochastic differential equations. *SIAM Journal on Numerical Analysis*, 48(1):298–321, 2010.

[9] K. Burrage, P. Burrage, and T. Tian. Numerical methods for strong solutions of stochastic differential equations: an overview. *Proc. R. Soc. Lond. Ser. A Math. Phys. Eng. Sci.*, 460(2041):373–402, 2004.

[10] T. Gard. *Introduction to stochastic differential equations*. Marcel Dekker, New York, 1988.

[11] E. Hairer and G. Wanner. *Solving ordinary differential equations II. Stiff and differential-algebraic problems*. Springer-Verlag, Berlin and Heidelberg, 1996.

[12] D. J. Higham. Mean-square and asymptotic stability of the stochastic theta method. *SIAM J. Numer. Anal.*, 38(3):753–769, 2000.

[13] P. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*. Springer-Verlag, Berlin and New York, 1992.

[14] B. Leimkuhler and C. Matthews. Rational construction of stochastic numerical methods for molecular sampling. *Appl. Math. Res. Express.*, 2013(1):34–56, 2013.

[15] B. Leimkuhler, C. Matthews, and M. V. Tretyakov. On the long-time integration of stochastic gradient systems. *Proc. R. Soc. A*, 470(2170), 2014.

[16] G. N. Mil′shteĭn. A theorem on the order of convergence of mean-square approximations of solutions of systems of stochastic differential equations. *Teor. Veroyatnost. i Primenen.*, 32(4):809–811, 1987.

[17] G. Milstein. Weak approximation of solutions of systems of stochastic differential equations. *Theory Probab. Appl.*, 30(4):750–766, 1986.

[18] G. Milstein and M. Tretyakov. *Stochastic numerics for mathematical physics*. Scientific Computing. Springer-Verlag, Berlin and New York, 2004.

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

[20] A. Tocino. Mean-square stability of second-order Runge-Kutta methods for stochastic differential equations. *J. Comput. Appl. Math.*, 175(2):355–367, 2005.

[21] P. Van der Houwen and B. Sommeijer. On the internal stage Runge-Kutta methods for large m-values. *Z. Angew. Math. Mech.*, 60:479–485, 1980.

[22] J. Verwer. Explicit Runge-Kutta methods for parabolic partial differential equations. *Special issue of Appl. Num. Math.*, 22:359–379, 1996.

[23] G. Vilmart. Postprocessed integrators for the high order integration of ergodic SDEs. *SIAM J. Sci. Comput.*, 37(1):A201–A220, 2015.