Easy-to-Implement One-Step Schemes for Stochastic Integration

J. Woodfield[a], A. Lobbe[a]

[a]*Department of Mathematics, Imperial College London, South Kensington Campus, London, SW7 2AZ, United Kingdom*

Abstract

Convenient, easy to implement stochastic integration methods are developed on the basis of abstract one-step deterministic order p integration techniques. The abstraction as an arbitrary one step map allows the inspection of easy to implement stochastic exponential time differencing Runge-Kutta (SETDRK), stochastic integrating factor Runge-Kutta (SIFRK) and stochastic RK (SRK) schemes. Such schemes require minimal modifications to existing deterministic schemes and converging to the Stratonovich SDE, whilst inheriting many of their desirable properties.

These schemes capture all symmetric terms in the Stratonovich-Taylor expansion, are order p in the limit of vanishing noise, can attain at least strong order $p/2$ or $p/2 - 1/2$ (parity dependent) for drift commutative noise, strong order 1 for commutative noise, and strong order 1/2 for multidimensional non-commutative noise. Numerical convergence is demonstrated using different bases of noise for 2nd, 3rd and 4th order SETDRK, SIFRK and SRK schemes for a stochastic KdV equation.

Keywords: Stochastic; Integrating Factor; Exponential time differencing; Runge-Kutta

---

Contents

1. Introduction

1.1. Motivation

Stochastic differential equations (SDEs) and stochastic partial differential equations (SPDEs) are central to ensemble forecasting, uncertainty quantification, and data assimilation. Developing dedicated solvers for SDEs and SPDEs is often impractical in large-scale applications. This work studies convergence properties of deterministic one-step integrators, which converge to an SDE under an easily implementable change of variables. Abstracting to arbitrary one-step schemes, allows convergence properties for Runge-Kutta (RK), Integrating Factor Runge-Kutta (IFRK) and Exponential Time Differencing Runge-Kutta (ETDRK) methods to be established in the stochastic context.

Whilst Integrating Factor Runge-Kutta (IFRK) methods and Runge-Kutta (RK) methods can be naturally derived for Stratonovich SDEs analogously to the deterministic case, stochastic exponential time differencing Runge-Kutta (SETDRK) schemes do not derive analogously. Here, abstract one-step map convergence theory is particularly useful for constructing SETDRK methods from existing deterministic schemes.

1.2. Deterministic Integration Schemes

Runge–Kutta (RK) methods are widely used one-step integrators for ODEs, offering arbitrary order accuracy and, in certain formulations, desirable properties such as strong stability preservation or symplecticity. Integrating Factor Runge–Kutta (IFRK) methods extend RK schemes by removing stiff linear terms via an integrating factor allowing larger timesteps. Exponential Time Differencing Runge–Kutta (ETDRK) methods achieve the same goal using exponential propagators acting on the linear stiff term and have shown to be efficient for solving stiff coupled systems of ordinary differential equations (ODEs) arising in diverse contexts ranging from American options under stochastic volatility to shallow water models of the atmosphere [51, 20, 54, 57, 8, 58, 1, 42, 6, 15]. For background on deterministic ETDRK see Cox and Matthews (2002), Kassam and Trefethen (2005) and Hochbruck and Ostermann (2010) [16, 30, 24].

1.3. Related work and literature

Whilst our work focuses on minimal-change Stratonovich-targeted deterministic ETDRK, IFRK, RK schemes there exists a range of related and extended approaches in the SDE literature, which we briefly review below.

Rüemelin in 1982 proposed a class of explicit Stochastic Runge-Kutta (SRK) schemes dependent on random variables approximating the first Stratonovich integral J_i, and gave algebraic conditions on the Double-Butcher-Tableau sufficient for the SRK scheme to converge with strong order 1/2 and commutative order 1. It was later shown in [9], that for arbitrary dimensional SDEs with a non-commuting basis of noise, the maximum strong order of a SRK method dependent on random variables representing only the first Stratonovich integral cannot exceed 1/2, due to the missing Lévy area correction.

In 1996 Burrage and Burrage [10] proposed a more general class of SRK schemes by allowing dependence on random variables approximating nested integration of Brownian motion. In the one-dimensional case ($d = 1$), including approximations of higher order nested integration such as J_{0i} further increased the possible order of such SRK methods to 1.5 see [10, 9]. However, as shown by the same authors in [9], in the multidimensional-noncommutative case the inclusion of higher order nested integration still remained maximum strong order 1/2 due to the particular structure of the SRK integrator.

In 2019 Erdogan and Lord [19] propose stochastic exponential time differencing methods that converge to an Itô system with strong order 1/2, commutative order 1. In section Appendix .4 we give additional literature review relating to stochastic exponential time differencing schemes for Itô-SDE's which are of a different derivation to those in this paper, some use Itô's isometry for the drift term whilst others can be derived as integrating factor

Runge-Kutta methods. In [33] linear A-stability is established for explicit exponential Runge-Kutta methods for weak solutions of Itô SDE's and have been successfully implemented for an Itô driven non-linear stochastic Schrödinger equation in [2].

In [4] an exponential Wagner-Platen scheme is proposed for Itô-SPDE's these schemes are fundamentally different to SRK, SIFRK, and SETDRK schemes in this paper, and rely on iterative substitutions similar in spirit to a Wagner-Platen expansion [31], rather than a RK approach. In [53], an exponential stochastic Runge-Kutta scheme based on the exponential Wagner-Platen type scheme is proposed for SPDE's which, under multiple operator-valued commutator relationships, could achieve order $3/2$.

In [41] deterministic ETD schemes are introduced which treat the nonlinear term semi-implicitly, and in [52] some of these semi-implicit schemes are extended to the stochastic context with strong order $1/2$, commutative order 1. Integrating factor Runge-Kutta methods (sometimes referred to as Runge-Kutta Lawson methods after the seminal work by Lawson [36]), were extended to the stochastic setting in [18] where a general class of stochastic Runge-Kutta-Lawson schemes are presented for SDE's with multiple linear commutative drift terms. In [7, 5, 24] deterministic IFRK and ETDRK schemes are considered as generalised deterministic ERK methods, and this is used as a basis for even more generalised stochastic integration in [56]. In [56], a general class of s-stage stochastic exponential integrators are considered following the approach [10], where preserving conformal quadratic invariants, the generalisation to multiple linear terms, and explicit B-series expansions are discussed. This paper considers abstract one-step integrators including SETDRK, SIFRK, and SRK methods, but only those which are related to deterministic integrators under a change of variables and those dependent on random variables representing the first Stratonovich integral, for ease of implementation, and because of the theoretical order barriers in [9].

1.4. Contributions

Theorem 2.2, presents the main theoretical result. For any deterministic one-step method of order p, we show that an easily implementable change of variables yields a scheme converging to a Stratonovich SDE with the following properties: strong order $1/2$ (item 1); exact representation of all symmetric terms in the Stratonovich–Taylor expansion up to order $p/2$ (item 2); drift-commutative order $\text{int}(p/2)$ for B-series methods (items 3 and 4); commutative order 1 (item 6); deterministic order P item 5.

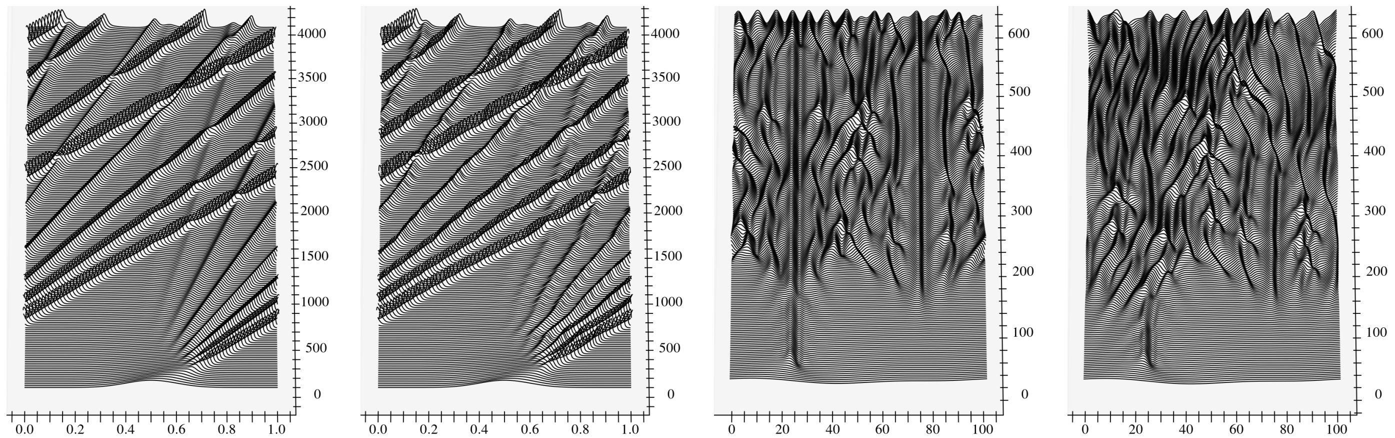
Whilst results (items 1 and 6) are known for broader SRK SIFRK SETDRK methods, [46, 56] the proof differs. Convergence follows directly from deterministic Lie-series expansions rather than the algebraic structure and analysis of the numerical scheme. Consequently, the result applies to any deterministic one-step method under a change of variables. Other properties (items 2 to 4) are specific to the particular class of methods we consider. The main numerical contribution is numerical comparison of SRK, SIFRK, and SETDRK methods of different orders, with different basis of noise. The theoretical properties in theorem 2.2 are all observed in practice. We investigate a convenient class of stochastic exponential time differencing methods (based on the deterministic integration schemes in [16, 30]) for instance SETDRK3, SETDRK4, and show they permit larger timesteps than standard SRK or SIFRK approaches, while preserving higher-order deterministic accuracy in the small-noise regime. We translate some successful IFRK schemes into the stochastic context, for instance the eSSPIFRK$^+(3,3)$ from [27].

1.5. Paper Organization

The paper is organised as follows. Section 2.1 defines the stochastically modified one-step map and the associated SDE, section 2.2 provides the convergence framework. Sections 3.1 to 3.3 introduces stochastic Runge–Kutta (SRK) methods, stochastic integrating factor Runge–Kutta (SIFRK) methods, and stochastic exponential time differencing Runge–Kutta (SETDRK) methods. Section 3.3.1 then outlines numerical implementation details for ETDRK methods.

Section 4 presents numerical results. Section 4.1 investigates multidimensional non-commutative noise, in terms of convergence and efficiency. Section 4.2 investigates commutative noise. Sections 4.3 and 4.4 investigates drift commutative noise. Section 4.5 explores decreasing noise magnitude and errors at large timesteps, whilst section 4.6 applies ETDRK4 to the two-dimensional incompressible Navier–Stokes equation.

The paper concludes in section 2. The appendices provide supporting material: section Appendix .1 derives the symmetric part of nested Stratonovich integrals, section Appendix .2 derives stochastic travelling wave solutions to the KdV equation, section Appendix .3 records commutator relationships, and section Appendix .4 gives further background and literature on SETDRK methods for Itô systems.



(a) KdV (b) KdV-Stochastic (c) KS (d) KS-Stochastic

Figure 1.1: Waterfall Plots of deterministic and stochastic KdV and KS equations under nonlinear transport noise.

2. Convenient Stochastic Integration

2.1. Setup

In this work, we study the numerical integration of stochastic differential equations (SDEs) using modified deterministic solvers. We describe the setting below. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a complete probability space with filtration $\{\mathcal{F}_t\}_{t\geq 0}$ satisfying the usual conditions. We consider the d-dimensional Stratonovich SDE

$$u_t = u_0 + \int_0^t f(s, u_s)\, ds + \sum_{m=1}^{M} \int_0^t g_m(s, u_s) \circ dW_s^m, \tag{2.1}$$

where $u_0 \colon \Omega \to \mathbb{R}^d$ denotes the initial condition, assumed $\mathcal{F}_0$-measurable and with finite second moment $\mathbb{E}||u_0||_2^2 \leq \infty$. The drift and diffusion coefficients f and $\{g_m\}_{m=1}^M$ are d-dimensional continuous functions satisfying global Lipschitz condition and linear growth bounds on the Itô-Stratonovich corrected drift and diffusion, and $\{W^i\}_{m=1}^M$ are independent standard Brownian motions adapted to $\{\mathcal{F}_t\}_{t\geq 0}$ [38, 11].

Given a deterministic one-step integration scheme

$$y^{n+1} = \Psi(y^n, \Delta t, f), \quad n = 0, 1, \ldots, N, \tag{2.2}$$

designed to approximate solutions of the d-dimensional ordinary differential equation (ODE)

$$\frac{dy}{dt} = f(t, y), \quad y(0) = y_0. \tag{2.3}$$

We define a convenient stochastic integration scheme as an abstract one-step stochastic method:

$$u^{n+1} = \widehat{\Psi}(u^n) := \Psi\left(u^n, \Delta t, f + \sum_{m=1}^{M} g_m \frac{\Delta W^m}{\Delta t}\right), \quad \text{where} \quad \Delta W^m := W^m(t^{n+1}) - W^m(t^n) \sim \sqrt{\Delta t}\,\mathcal{N}(0,1). \tag{2.4}$$

In what follows, we prove convergence results for the method (2.4) to solutions of the SDE (2.1).

2.2. Convergence

We shall study convergence in the following mean-square sense:

Definition 2.1 (Mean-square convergence)**.** Let u^N be the numerical approximation to $u(t^N)$ after N steps with constant stepsize Δt; then u_N is said to converge strongly (mean square) to u with strong global order p if $\exists C > 0$ (independent of Δt) and $\tau > 0$ such that

$$\mathbb{E}\left[||u_N - u(t_N)||_2^2\right]^{1/2} \leq C\Delta t^p, \quad \forall \Delta t \leq \tau. \tag{2.5}$$

Mean-square converge implies strong convergence under Lyapunov's inequality, and implies pathwise/strong convergence as defined in [31]. In order to prove convergence properties about one-step approximations, we use Milstein and Tretyakov's Fundamental Theorem of Mean Square Convergence (FTMSC) [40] stated below:

Theorem 2.1 (FTMSC [40])**.** Suppose the one-step approximation $\bar{u}_{t,x}(t+\Delta t) := \widehat{\Psi}(x, \Delta t, f+\frac{1}{h}\sum_{m=1}^{M} g_m \Delta W^m)$ starting from $x = u(t)$ and the exact solution $u(t+\Delta t)$ satisfies the local error estimates:

$$\|\mathbb{E}\left[\bar{u}_{t,x}(t+\Delta t) - u(t+\Delta t)\right]\|_2 \leq K\left(1 + ||x||_2^2\right)^{1/2} (\Delta t)^{p_1}, \tag{2.6}$$

$$\left[\mathbb{E}\,\|\bar{u}_{t,x}(t+\Delta t) - u(t+\Delta t)\|_2^2\right]^{1/2} \leq K\left(1 + ||x||_2^2\right)^{1/2} (\Delta t)^{p_2}, \tag{2.7}$$

for

$$p_2 \geq \frac{1}{2}, \quad p_1 \geq p_2 + \frac{1}{2}, \quad t_0 \leq t \leq T - \Delta t, \quad x \in \mathbb{R}^d. \tag{2.8}$$

Then $\bar{u}_{t,x}(t+\Delta t)$ is global mean square (strong) order $p = p_2 - 1/2$, i.e. the global approximation $\bar{u}_{t_0,u_0}(t_k)$, obtained by iterating the one-step map k times from initial condition u_0 satisfies:

$$\left[\mathbb{E}\,\|\bar{u}_{t_0,u_0}(t_k) - u(t_k)\|_2^2\right]^{1/2} \leq K\left(1 + E\,\|u_0\|_2^2\right)^{1/2} (\Delta t)^{p_2 - 1/2}, \quad k = 0, 1, \ldots, N, \quad \forall N \tag{2.9}$$

Definition 2.2 (L^0, L^i operators)**.** Given the functions f, $\{g_i\}_{i=1}^M : [0,T] \times \mathbb{R}^d \to \mathbb{R}^d$, we shall define the corresponding operators L^0, L^i, as follows

$$L^0 := \frac{\partial}{\partial t} + f^k \frac{\partial}{\partial u^k}, \quad L^i := g_i^k \frac{\partial}{\partial u^k}, \quad i = 1, ..., M, \tag{2.10}$$

where the repeated index k is summed over its range $1, ..., d$.

Definition 2.3 (Bracket)**.** We shall define the commutator of d-dimensional vector-valued functions $f, \{g_i\}_{i=1}^M$, through the action of L^i in the following manner

$$[f, g_i] := L^0 g_i - L^i f, \quad [g_i, g_j] := L^i g_j - L^j g_i. \tag{2.11}$$

The l-th component of this bracket agrees componentwise with the vector field generated by the commutator of the vector field L-operators, namely:

$$[L^i, L^j] := L^i L^j - L^j L^i = g_i^k \frac{\partial g_j^l}{\partial u^k}\frac{\partial}{\partial u^l} - g_j^k \frac{\partial g_i^l}{\partial u^k}\frac{\partial}{\partial u^l} = [g_i, g_j]^l \frac{\partial}{\partial u^l}, \quad i, j \in \{1, ..., M\}. \tag{2.12}$$

$$[L^0, L^i] := L^0 L^i - L^i L^0 = \left(\frac{\partial g^l}{\partial t} + f^k \frac{\partial g^l}{\partial u^k}\right)\frac{\partial}{\partial u^l} - g^k \frac{\partial f^l}{\partial u^k}\frac{\partial}{\partial u^l} = [f, g_i]^l \frac{\partial}{\partial u^l} = 0, \quad i \in \{1, ..., M\} \tag{2.13}$$

Therefore, the commutativity of f, g_i under the definition of commutator defined in eq. (2.11) is equivalent to having commutativity of the corresponding Lie operators, since componentwise:

$$[f, g_i]^k := [L^0, L^i]^k, \quad [g_i, g_j]^k := [L^i, L^j]^k. \tag{2.14}$$

Definition 2.4 (Commutative)**.** We shall say the basis of noise is commutative if $[g_i, g_j] = 0$, or equivalently if $[L^i, L^j] = 0$, in summation notation:

$$[g_i, g_j] := g_i^k(t,u)\frac{\partial g_j}{\partial u^k} - g_j^k(t,u)\frac{\partial g_i}{\partial u^k} = 0, \quad i,j \in \{1,...,M\}. \tag{2.15}$$

Definition 2.5 (Drift Commutative)**.** We shall say the basis of noise is drift commutative if $[f, g_j] = 0$, or equivalently if $[L^0, L^j] = 0$, in summation notation:

$$[f, g_i] := L^0 g_i - L^i f = \frac{\partial}{\partial t} g_i + f^k \frac{\partial g_i}{\partial u^k} - g_i \frac{\partial f}{\partial u^k} = 0, \quad j \in \{1,...,M\}. \tag{2.16}$$

We shall refer to noise as being drift commutative if both $[f, g_i] = 0$ and $[g_i, g_j] = 0$, $\forall i, j \in \{1,..,M\}$. This can be used for the following theorem:

Theorem 2.2. Let $f, g_i \in C_b^\infty\left([0,T] \times \mathbb{R}^d; \mathbb{R}^d\right)$, for all $i = 1,...,M$, such that global Lipschitz and linear growth bounds hold in the Euclidean norm and sufficient derivatives exist and are bounded. Let eq. (2.2) denote an abstract one-step deterministic global order P numerical integrator of eq. (2.3) with fixed timestep $\Delta t = t^{n+1} - t^n$ for all n. Defining $u(t)$ to be the strong solution to the Stratonovich SDE eq. (2.1). Then defining $\widehat{\Psi}(u^n)$ as a convenient stochastic one-step integrator as in eq. (2.4). Then the following properties hold:

1. If $P \geq 2$, then the sequence $\{u_n\}_{n=0}^N$ generated by eq. (2.4) converges to the solution $u(t)$ of eq. (2.1) with global (mean square) strong order 1/2. Meaning, the following estimate holds

$$\left(\mathbb{E}\|u_n - u(t_n)\|_2^2\right)^{1/2} \leq C(\Delta t)^{1/2}, \quad \forall n = 0,1,...,N, \quad \forall N. \tag{2.17}$$

2. Method eq. (2.4) captures all symmetric terms in the Stratonovich-Taylor expansion up to (strong) order $P/2$, independent of whether P is even or odd.
3. If $P \geq 2$, and P is even and the vector fields drift commute $[f, g_i] = L^0 g_i - L^i f = 0$, $[g_i, g_j] = L^i g_j - L^j g_i = 0$, $\forall i, j \in \{1,...,M\}$, then $u_{n+1} = \widehat{\Psi}(u_n)$ generated from the onestep-map eq. (2.4) from the previous approximation u_n satisfies the following *local* error estimates

$$\mathbb{E}[u_{n+1} - u(t_{n+1})] = \mathcal{O}(\Delta t^{P/2+1}), \tag{2.18}$$

$$\left(\mathbb{E}[||u_{n+1} - u(t_{n+1})||_2^2]\right)^{1/2} = \mathcal{O}(\Delta t^{P/2+1/2}), \tag{2.19}$$

   sufficient by the FTMSC theorem 2.1-[40] to prove the sequence $\{u_n\}_{n=0}^N$ generated by eq. (2.4) converges to the solution $u(t)$ of the Stratonovich SDE eq. (2.1) with strong global (mean square) order $P/2$, with the strong global error bound

$$\left(\mathbb{E}[\|u_n - u(t_n)\|_2^2]\right)^{1/2} \leq C(\Delta t)^{P/2}, \quad \forall n = 0,1,...,N, \quad \forall N. \tag{2.20}$$

4. If $P \geq 2$, and P is odd and the vector fields drift commute $[f, g_i] = 0$, $[g_i, g_j] = L^i g_j - L^j g_i = 0$, $\forall i, j \in \{1,...,M\}$, then u_{n+1} generated by eq. (2.4) from initial condition u_n satisfies the following local error estimates

$$\mathbb{E}[u_{n+1} - u(t_{n+1})] = \mathcal{O}(\Delta t^{P/2}) \tag{2.21}$$

$$\left(\mathbb{E}[||u_{n+1} - u(t_{n+1})||_2^2]\right)^{1/2} = \mathcal{O}(\Delta t^{P/2+1/2}) \tag{2.22}$$

   sufficient by the FTMSC [40] to converge to the solution $u(t)$ of the Stratonovich SDE eq. (2.1) with strong order $P/2 - 1/2$, with the strong global error bound

$$\left(\mathbb{E}[\|u_n - u(t_n)\|_2^2]\right)^{1/2} \leq C(\Delta t)^{P/2-1/2}. \tag{2.23}$$

5. If $g_i = 0$, for $i = 1,...,M$ we recover deterministic order P convergence.

6. If $P \geq 2$, and $[g_i, g_j] = L^i g_j - L^j g_i = 0, \quad \forall i, j \in \{1, ..., M\}$, then $u_{n+1} = \widehat{\Psi}(u_n)$ generated from the onestep-map eq. (2.4) from the previous approximation u_n satisfies the following *local* error estimates

$$\mathbb{E}[u_{n+1} - u(t_{n+1})] = \mathcal{O}(\Delta t^2), \tag{2.24}$$

$$\left(\mathbb{E}[||u_{n+1} - u(t_{n+1})||_2^2]\right)^{1/2} = \mathcal{O}(\Delta t^{3/2}), \tag{2.25}$$

sufficient by the FTMSC [40] to prove the sequence $\{u_n\}_{n=0}^N$ generated by eq. (2.4) converges to the solution $u(t)$ of the Stratonovich SDE eq. (2.1) with strong global (mean square) order 1, with the strong global error bound

$$\left(\mathbb{E}[\|u_n - u(t_n)\|_2^2]\right)^{1/2} \leq C(\Delta t)^1, \quad \forall n = 0, 1, ..., N, \quad \forall N. \tag{2.26}$$

Proof. To simplify notation let $g_0 = f$, $t = W^0$, and for a multi-index $\boldsymbol{j} = \{j_1, ..., j_n\} \in \{0, 1, ..., M\}^n$, denote nested Stratonovich integration of Brownian motion in $[t_n, t_{n+1})$ as

$$J_{j_1,...,j_n} := \int_{t_n}^{t_{n+1}} \left[\int_{t_n}^{s_n} ... \left[\int_{t_n}^{s_2} \circ dW^{j_n}(s_1)\right] ... \circ dW^{j_{n-1}}(s_{n-1})\right] \circ dW^{j_n}(s_n). \tag{2.27}$$

Then the k-th component of a Stratonovich-Taylor expansion of the analytic solution $u(t+\Delta t)$ of eq. (2.1) about local initial condition u_n at time t_n with timestep Δt up to strong order $(P+1)/2$ can be written compactly as

$$u^k(t_{n+1}) = u_n^k + \sum_{j=0}^{M} g_j^k J_j + \sum_{0 \leq j_1, j_2 \leq m} L^{j_1} g_{j_2}^k J_{j_1,j_2} + \sum_{0 \leq j_1, j_2, j_3 \leq M} L^{j_1} L^{j_2} g_{j_3}^k J_{j_1,j_2,j_3} + \dots \tag{2.28}$$

$$+ \sum_{0 \leq j_1, j_2, ... j_P \leq M} L^{j_1} L^{j_2} ... L^{j_{P-1}} g_{j_P}^k J_{j_1,j_2,...j_P} + O(\Delta t^{\frac{P+1}{2}}). \tag{2.29}$$

Whereas the k-th component of the Taylor expansion of the deterministic flow map up to order P from local initial condition u_n at time t_n with timestep Δt with the subsequent substitution $\Delta t f \mapsto \Delta t f + \sum_{i=1}^M g_i \Delta W^i$, takes the following local form

$$\hat{u}_{n+1}^k = u_n^k + \sum_{j=0}^{M} g_j^k J_j + \frac{1}{2} \sum_{0 \leq j_1, j_2 \leq M} L^{j_1} g_{j_2}^k J_{j_1} J_{j_2} + \frac{1}{3!} \sum_{0 \leq j_1, j_2, j_3 \leq M} L^{j_1} L^{j_2} g_{j_3}^k J_{j_1} J_{j_2} J_{j_3} + ... \tag{2.30}$$

$$+ \frac{1}{P!} \sum_{0 \leq j_1, j_2, ... j_P \leq M} L^{j_1} L^{j_2} ... L^{j_{P-1}} g_{j_P}^k J_{j_1} J_{j_2} ... J_{j_P} + O(\Delta t^{\frac{P+1}{2}}). \tag{2.31}$$

Where all terms are evaluated at (t^n, u_n) unless otherwise specified. Let $\boldsymbol{j}$ denote a multi-index length n, then define the symmetric parts of $J_{\boldsymbol{j}}$ by summation over all permutations in the symmetric group S_n, then the following identity

$$\mathrm{Sym}\left(J_{\boldsymbol{j}}\right) = \frac{1}{n!} \sum_{\sigma \in S_n} J_{j_{\sigma(1)},...,j_{\sigma(n)}} = \frac{1}{n!} J_{j_1} J_{j_2} ... J_{j_n}, \quad n = 1, 2..., \tag{2.32}$$

(proven inductively using the shuffle property of Stratonovich integrals section Appendix .1), clarifies all symmetric terms, up to strong order $P/2$ in the stochastic expansion eq. (2.29) are captured by the stochastically modified deterministic scheme in eq. (2.31) proving item 2.

The remaining non-symmetric terms can be expressed in terms of nested commutator relationships, this is a consequence of the Poincaré-Birchoff-Witt theorem but also arises in a number of different manners and in a variety of different contexts such as the Magnus-Expansion, some selected references are [48, 45, 3, 13, 50, 37, 21, 44, 12, 35]. For example, at weight order 2 we have the decomposition into symmetric and non-symmetric components expressible as commutators

$$\sum_{0 \leq j_1, j_2 \leq M} L^{j_1} g_{j_2}^k J_{j_1,j_2} = \sum_{0 \leq j_1, j_2 \leq M} L^{j_1} g_{j_2}^k \,\mathrm{Sym}(J_{j_1,j_2}) + \sum_{0 \leq j_1, j_2 \leq M} L^{j_1} g_{j_2}^k \,\mathrm{Alt}(J_{j_1,j_2}) \tag{2.33}$$

$$= \frac{1}{2} \sum_{0 \leq j_1, j_2 \leq M} L^{j_1} g^{k,j_2} J_{j_1} J_{j_2} + \frac{1}{2} \sum_{0 \leq j_1, j_2 \leq M} \left(L^{j_1} g^{k,j_2} - L^{j_2} g^{k,j_1}\right) J_{j_1,j_2}. \tag{2.34}$$

Similarly at weight order 3, decomposing J_{j_1,j_2,j_3} as a rank-3 tensor into symmetric, anti-symmetric and two remainder tensors of mixed type, then upon re-indexing and using the anti-symmetric properties of the mixed tensors in section Appendix .3 one attain the following representation

$$\sum_{0\leq j_1,j_2,j_3\leq M} L^{j_1}L^{j_2}g_{j_3}J_{j_1,j_2,j_3} = \sum_{0\leq j_1,j_2,j_3\leq M} L^{j_1}L^{j_2}g_{j_3}J_{j_1}J_{j_2}J_{j_3} + \frac{1}{3!}\sum_{0\leq j_1,j_2,j_3\leq M} \left([g_{j_1},[g_{j_2},g_{j_3}]]\right) J_{j_1,j_2,j_3} \tag{2.35}$$

$$+ \frac{1}{2}\sum_{0\leq j_1,j_2,j_3\leq M} L^{j_1}\left[g_{j_2},g_{j_3}\right] J_{j_1,j_2,j_3} + \frac{1}{2}\sum_{0\leq j_1,j_2,j_3\leq M} [L^{j_1},L^{j_2}]g_{j_3}J_{j_1,j_2,j_3}. \tag{2.36}$$

In order to prove item 1, suppose that $P \geq 2$, and f, g_i do not commute or satisfy any other higher order commutative identities, consider $u(t+\Delta t) - \hat{u}_{n+1}$, using eq. (2.29) and eq. (2.31), one attains that the leading order term is $\mathcal{O}(\Delta t)$, and given by the Lévy area correction

$$u^k(t+\Delta t) - \hat{u}^k_{n+1} = \frac{1}{2}\sum_{1\leq j_1,j_2\leq M} L^{j_1}g^k_{j_2}\,\mathrm{Alt}(J_{j_1,j_2}) + \mathcal{O}(\Delta t^{3/2}). \tag{2.37}$$

The expectation of the Lévy area is 0 and the expectation of the square of the Lévy area using Itô's isometry is Δt^2, and is sufficient to give local error estimates of at least

$$\mathbb{E}[u(t_{n+1}) - \hat{u}_{n+1}] = \mathcal{O}((\Delta t)^{3/2}), \quad \left(\mathbb{E}[||u(t_{n+1}) - \hat{u}_{n+1}||_2^2]\right)^{1/2} = \mathcal{O}((\Delta t)^1). \tag{2.38}$$

Which is sufficient by FTMSC [40] for strong order 1/2, with the *global* error estimate

$$\left(\mathbb{E}[||u(t_{n+1}) - \hat{u}_{n+1}||_2^2]\right)^{1/2} = \mathcal{O}((\Delta t)^{1/2}), \quad \forall n = 0,1,...,N. \tag{2.39}$$

In order to prove item 6, suppose $[g_i, g_j] = 0$ for $i,j = 1,...,M$, but $[g_i, g_0] \neq 0$, in this case the $\mathcal{O}(\Delta t^{3/2})$ contribution of

$$\frac{1}{3!}\sum_{0\leq j_1,j_2,j_3\leq M} \left([g_{j_1},[g_{j_2},g_{j_3}]]\right) J_{j_1,j_2,j_3},$$

is zero, consequently the leading order $\mathcal{O}(\Delta t^{3/2})$ term arises from the following term

$$\frac{1}{2}\sum_{0\leq j_1,j_2\leq M} \left(L^{j_1}g^{k,j_2} - L^{j_2}g^{k,j_1}\right) J_{j_1,j_2}. \tag{2.40}$$

Which upon re-indexing, using the commutativity condition, the shuffle property $J_{0,j_1} = -J_{j_1,0} + J_0 J_{j_1}$ gives

$$u^k(t+\Delta t) - \hat{u}^k_{n+1} = \frac{1}{2}\sum_{1\leq j_1\leq M} \left(L^0 g^{k,j_1} - L^{j_1}g^{k,0}\right) J_0 J_{j_1} + \sum_{1\leq j_1\leq M} \left(L^{j_1}g^{k,0} - L^0 g^{k,j_1}\right) J_{j_1,0} + \mathcal{O}(\Delta t^2). \tag{2.41}$$

In expectation, the leading order term in eq. (2.41) vanishes, but the square is $\mathcal{O}(\Delta t^3)$, such that

$$\mathbb{E}[u(t_{n+1}) - \hat{u}_{n+1}] = \mathcal{O}((\Delta t)^2), \quad \left(\mathbb{E}[||u(t_{n+1}) - \hat{u}_{n+1}||_2^2]\right)^{1/2} = \mathcal{O}((\Delta t)^{3/2}). \tag{2.42}$$

Sufficient by the FTMSC theorem 2.1 to establish the following global error estimate

$$\left(\mathbb{E}[||u(t_{n+1}) - \hat{u}_{n+1}||_2^2]\right)^{1/2} = \mathcal{O}((\Delta t)^1), \quad \forall n = 0,1,...,N, \tag{2.43}$$

implying global mean square order 1 for a commutative basis of noise, proving item 6.

In order to prove item 3, suppose that $P \geq 2$ is even, and $[g_i, g_j] = 0$ for $i, j = 0, ..., P$, then invoking that the non-symmetric component of a Stratonovich-Taylor expansion term is expressible in terms of commutator relationships the leading error term is

$$u^k(t_{n+1}) - \hat{u}^k_{n+1} = \sum_{0 \leq j_1, j_2, ... j_{P+1} \leq M} L^{j_1} L^{j_2} ... L^{j_P} g^k_{j_{P+1}} J_{j_1, j_2, ... j_{P+1}} + \mathcal{O}((\Delta t)^{P/2+1}) \tag{2.44}$$

$$= \frac{1}{(P+1)!} \sum_{0 \leq j_1, j_2, ... j_{P+1} \leq M} L^{j_1} L^{j_2} ... L^{j_P} g^k_{j_{P+1}} J_{j_1} J_{j_2} ... J_{j_{P+1}} + \mathcal{O}((\Delta t)^{P/2+1}). \tag{2.45}$$

Taking expectation, and noting that the expectation of the product of an odd number of symmetric random variables gives zero, gives the following local error estimate of the mean error and mean square error respectively

$$\mathbb{E}[u(t_{n+1}) - \hat{u}_{n+1}] = \mathcal{O}((\Delta t)^{P/2+1}), \quad \left(\mathbb{E}[||u(t_{n+1}) - \hat{u}_{n+1}||_2^2]\right)^{1/2} = \mathcal{O}((\Delta t)^{P/2+1/2}). \tag{2.46}$$

Therefore, by the FTMSC [40] one establishes the following global error estimate

$$\left(\mathbb{E}[||u(t_{n+1}) - \hat{u}_{n+1}||_2^2]\right)^{1/2} = \mathcal{O}((\Delta t)^{P/2}), \quad \forall n = 0, 1, ..., N, \tag{2.47}$$

implying global mean square order $P/2$ (strong), when P is even.

For item 4, if instead P is odd, one instead attains $\mathbb{E}[J_{j_1} ... J_{j_{P+1}}] = \mathcal{O}((\Delta t)^{P/2+1/2})$ sufficient for the following local error estimates

$$\mathbb{E}[u(t_{n+1}) - \hat{u}_{n+1}] = \mathcal{O}((\Delta t)^{P/2+1/2}), \quad \left(\mathbb{E}[||u(t_{n+1}) - \hat{u}_{n+1}||_2^2]\right)^{1/2} = \mathcal{O}((\Delta t)^{P/2+1/2}). \tag{2.48}$$

Therefore, using the FTMSC [40] with the previous local estimates is only sufficient to prove (strong) mean square convergence order $P/2 - 1/2$ with the following global error estimate

$$\left(\mathbb{E}[||u(t_{n+1}) - \hat{u}_{n+1}||_2^2]\right)^{1/2} = \mathcal{O}((\Delta t)^{P/2-1/2}). \tag{2.49}$$

Item 5 is true by construction. □

Remark. In the above theorem:

1. We do not assume any structure of the one step map[1]. Meaning that the theoretical convergence properties of theorem 2.2 are applicable to existing deterministic methods including RK, IFRK and ETDRK schemes.
2. We found that when using a odd order deterministic integrator the expectation of the deviation (mean of the error), was not sufficient to prove order $P/2$ by the FTMSC under drift commuting noise, and lowered the order of provable convergence by $1/2$.

Definition 2.6. We shall refer to a scheme with quadruple order (P_d, P_{dc}, P_c, P_s). P_d denotes the deterministic global order of convergence. P_{dc} denotes the mean square strong stochastic global order convergence for a drift commutative basis of noise. P_c denotes mean square strong stochastic global order convergence for a commutative basis of noise. P_s refers to the strong stochastic order for general multidimensional noise.

Remark. Informally, theorem 2.2 states any deterministic one-step map of order $p \geq 2$ modified to solve a SDE under the transform eq. (2.4) has deterministic, drift commutative, commutative, stochastic order of at least

$$(P_d, P_{dc}, P_c, P_s) = (p, \text{int}(p/2), 1, 1/2), \tag{2.50}$$

understood in the global mean square sense. There are typical regularity caveats, and the P_{dc} condition requires the method to have a B-series expansion such as in RK-IFRK-ETDRK methods.

[1] with the exception of items 3 and 4 requiring a B-series expansion of the numerical method, true for RK,IFRK,ETDRK schemes section Appendix .5.

3. Methods

3.1. Stochastic Runge-Kutta (SRK) methods

The simplest example of a convenient integration technique is to take any deterministic order p Runge-Kutta scheme uniquely specified by the Butcher-Tableau (A, b, c) under the following mapping

$$f \mapsto f + \sum_{m=1}^{M} g_m \frac{\Delta W_m}{\Delta t}, \tag{3.1}$$

to give the following SRK scheme:

Method 3.1 (SRK)**.**

$$u^{(i)} = u^n + \Delta t \sum_{j=1}^{s} a_{ij} \left(f(t^n + c_j \Delta t, u^j) + g_m \left(t^n + c_j \Delta t, u^{(j)} \right) \frac{\Delta W^m}{\Delta t} \right), \quad i = 1, ..., s, \tag{3.2}$$

$$u^{n+1} = u^n + \Delta t \sum_{i=1}^{s} b_i \left(f(t^n + c_i \Delta t, u^{(i)}) + g_m \left(t^n + c_i \Delta t, u^{(i)} \right) \frac{\Delta W^m}{\Delta t} \right). \tag{3.3}$$

Theorem 2.2 applies to method 3.1, so that order $P \geq 2$ Butcher-Tableau conditions are sufficient for convergence properties: $(P_d, P_{dc}, P_c, P_s) = (p, [p/2], 1, 1/2)$, capturing all symmetric terms in the Stratonovich-Taylor expansion. Second order Butcher-tableau conditions for $P_c, P_s = 1, 1/2$ are consistent with the convergence results derived by Rüemelin (compare eq. (2.42) to Theorem 4 in [46]).
We shall now define three practical SRK schemes tested in this paper:

Method 3.2 (SSP22)**.**

$$u^{(1)} = u^n + \Delta t \left(f(t^n, u^n) + g_m(t^n, u^n) \frac{\Delta W^m}{\Delta t} \right) \tag{3.4}$$

$$u^{n+1} = \frac{1}{2} u^n + \frac{1}{2} \left(u^{(1)} + \Delta t \left(f(t^n + \Delta t, u^{(1)}) + g_m(t^n + \Delta t, u^{(1)}) \frac{\Delta W^m}{\Delta t} \right) \right) \tag{3.5}$$

Theorem 2.2 implies convergence properties $(P_d, P_{dc}, P_c, P_s) = (2, 1, 1, 1/2)$. Strong stability in the stochastic context is discussed in [55].

Remark (Notation)*.* We shall present schemes in the time independent case where $f(t, u) = f(u)$, $g_m(t, u) = g_m(u)$, and adopt the shorthand $G(u)\Delta S = g_m(u)\Delta W^m$ for presentation purposes, this is wlog and the time dependent scheme can be inferred from method 3.1-eqs. (3.2) and (3.3). More specifically: recover c from the consistency condition $c = Ae$ and evaluate temporally dependent functions at $t^n + c_i \Delta t$ in stage i.

Method 3.3 (SSP33)**.**

$$u^{(1)} = u^n + \Delta t \left(f(u^n) + G(u^n) \frac{\Delta S}{\Delta t} \right) \tag{3.6}$$

$$u^{(2)} = \frac{3}{4} u^n + \frac{1}{4} \left(u^{(1)} + \Delta t \left(f(u^{(1)}) + G(u^{(1)}) \frac{\Delta S}{\Delta t} \right) \right) \tag{3.7}$$

$$u^{(3)} = \frac{1}{2} u^n + \frac{2}{3} \left(u^{(2)} + \Delta t \left(f(u^{(2)}) + G(u^{(2)}) \frac{\Delta S}{\Delta t} \right) \right) \tag{3.8}$$

Theorem 2.2 implies convergence properties $(P_d, P_{dc}, P_c, P_s) = (3, 1, 1, 1/2)$ and symmetric terms of order 3/2 are captured. For the temporally dependent version use $c = (0, 1, 1/2)$, method 3.1. Strong stability in the stochastic context is discussed in [55].

Method 3.4 (SRK4)**.** The SRK4 method is given by

$$u^{(1)} = u_n \tag{3.9}$$

$$u^{(2)} = u^n + \frac{\Delta t}{2}\left(f\left(u^{(1)}\right) + G\left(u^{(1)}\right)\frac{\Delta S}{\Delta t}\right), \tag{3.10}$$

$$u^{(3)} = u^n + \frac{\Delta t}{2}\left(f\left(u^{(2)}\right) + G\left(u^{(2)}\right)\frac{\Delta S}{\Delta t}\right), \tag{3.11}$$

$$u^{(4)} = u^n + \Delta t\left(f\left(u^{(3)}\right) + G\left(u^{(3)}\right)\frac{\Delta S}{\Delta t}\right), \tag{3.12}$$

$$u^{n+1} = u^n + \Delta t\left(\frac{1}{6}\mathcal{F}\left(u^{(1)}\right) + \frac{1}{3}\mathcal{F}\left(u^{(1)}\right) + \frac{1}{3}\mathcal{F}\left(u^{(2)}\right) + \frac{1}{6}\mathcal{F}\left(u^{(3)}\right)\right). \tag{3.13}$$

$$\mathcal{F}(a) := f(a) + \frac{G(a)\Delta S}{\Delta t}. \tag{3.14}$$

For the temporally varying version $c = (0, 1/2, 1/2, 1)$. Theorem 2.2 implies convergence properties $(P_d, P_{dc}, P_c, P_s) = (4, 2, 1, 1/2)$ and symmetric terms up to order 2 are captured.

Remark (Implementation convenience). A single optimised function $\mathcal{F}$, can be called at each stage in the convenient SRK approach (see method 3.4).

3.2. Stochastic Integrating Factor Runge Kutta (SIFRK) methods

We shall recall the derivation of SIFRK methods, before subsequently showing they can be alternatively derived as deterministic IFRK methods under a change of variables. Consider the Stratonovich SDE eq. (2.1), admitting the following decomposition of drift f into linear L and nonlinear N terms:

$$d_t u = \underbrace{[Lu + N(t,u)]}_{f}\, dt + g_m(t,u) \circ dW^m(t). \tag{3.15}$$

Where the vector-valued functions $g_m(t,u)$ are integrated against the increments of the m-th i.i.d Brownian motion W^m in a Stratonovich sense, and summation over the repeated index $m = 1, ..., M$ is assumed. It is assumed L is a linear matrix, and N is a nonlinear vector valued function. SIFRK methods can be derived, by first multiplying eq. (3.15) by the integrating factor e^{-Lt} to rearrange into the SDE

$$d_t(e^{-Lt}u) = e^{-Lt}N(t,u)dt + e^{-Lt}g_m(t,u) \circ dW^m. \tag{3.16}$$

By transforming variables, one can write eq. (3.16) as the non-autonomous SDE

$$d_t w = H(t,w)dt + K_m(t,w) \circ dW^m, \tag{3.17}$$

where $w = e^{-Lt}u$, $H(t,w) := e^{-Lt}N(t, e^{Lt}w)$, $K_m(t,w) := e^{Lt}g_m(t,u)$. A one-step RK scheme (specified by the Butcher tableau (A,b) and the temporal consistency conditions $c = Ae$) applied to eq. (3.17) gives

$$k^{(i)} = w_n + \Delta t\sum_{j=1}^{s} a_{ij}H\left(t_n + c_j\Delta t, k^{(j)}\right) + \sum_{j=1}^{s} a_{ij}K_m\left(t_n + c_j\Delta t, k^{(j)}\right)\Delta W^m, \quad i = 1,\dots,s, \tag{3.18}$$

$$w_{n+1} = w_n + \Delta t\sum_{i=1}^{s} b_i H\left(t_n + c_i\Delta t, k^{(i)}\right) + \sum_{i=1}^{s} b_i K_m\left(t_n + c_i\Delta t, k^{(i)}\right)\Delta W^m. \tag{3.19}$$

Transforming back into the variable u and multiplying through by $e^{L(t^n + c_i\Delta t)}$ gives the SIFRK method:

Method 3.5 (SIFRK)**.**

$$u^{(i)} = e^{Lc_i\Delta t}u^n + \Delta t\sum_{j=1}^{s} a_{ij}e^{L(c_i - c_j)\Delta t}\left(N(t^n + c_j\Delta t, u^j) + g_m(t^n + c_j\Delta t, u^{(j)})\frac{\Delta W^m}{\Delta t}\right), \quad i = 1, ..., s, \tag{3.20}$$

$$u^{n+1} = e^{L\Delta t}u^n + \Delta t\sum_{i=1}^{s} b_i e^{L(1-c_i)\Delta t}\left(N(t^n + c_i\Delta t, u^{(i)}) + g_m(t^n + c_i\Delta t, u^{(i)})\frac{\Delta W^m}{\Delta t}\right). \tag{3.21}$$

It is apparent from eqs. (3.20) and (3.21) that this class of SIFRK schemes are deterministic IFRK method under the substitution

$$N \mapsto N + \sum_{m=1}^{M} g_m \frac{\Delta W_m}{\Delta t}, \tag{3.22}$$

and convergence follows directly from theorem 2.2 since eq. (3.22) is equivalent to

$$\Delta t f \mapsto \Delta t f + \sum_{m=1}^{M} g_m \Delta W^m. \tag{3.23}$$

The Butcher tableau corresponding to the RK4 scheme

$$(A, b) = \begin{array}{c|cccc} 0 & & & & \\ 1/2 & 1/2 & & & \\ 1/2 & 0 & 1/2 & & \\ 1 & 0 & 0 & 1 & \\ \hline & 1/6 & 1/3 & 1/3 & 1/6 \end{array}, \tag{3.24}$$

gives rise to the following Stochastic integrating factor Runge-Kutta scheme:

Method 3.6 (IFSRK4)**.** The IFSRK4 method is given by

$$u^{(1)} = u_n \tag{3.25}$$

$$u^{(2)} = \mathrm{e}^{\frac{\Delta t}{2}L} u^n + \frac{\Delta t}{2} \mathrm{e}^{\frac{\Delta t}{2}L} N\left(u^{(1)}\right) + \frac{1}{2} \mathrm{e}^{\frac{\Delta t}{2}L} G\left(u^{(1)}\right) \Delta S, \tag{3.26}$$

$$u^{(3)} = \mathrm{e}^{\frac{\Delta t}{2}L} u^n + \frac{\Delta t}{2} N\left(u^{(2)}\right) + \frac{1}{2} G\left(u^{(2)}\right) \Delta S, \tag{3.27}$$

$$u^{(4)} = \mathrm{e}^{\Delta t L} u^n + \Delta t \mathrm{e}^{\frac{\Delta t}{2}L} N\left(u^{(3)}\right) + \mathrm{e}^{\frac{\Delta t}{2}L} G\left(u^{(3)}\right) \Delta S, \tag{3.28}$$

$$u^{n+1} = \mathrm{e}^{\Delta t L} u^n + \Delta t \left(\frac{1}{6} \mathrm{e}^{\Delta t L} \mathcal{N}\left(u^{(1)}\right) + \frac{1}{3} \mathrm{e}^{\frac{\Delta t}{2}L} \mathcal{N}\left(u^{(2)}\right) + \frac{1}{3} \mathrm{e}^{\frac{\Delta t}{2}L} \mathcal{N}\left(u^{(3)}\right) + \frac{1}{6} \mathcal{N}\left(u^{(4)}\right) \right). \tag{3.29}$$

Where

$$\mathcal{N}(a) = N(a) + \frac{G(a)\Delta S}{\Delta t}. \tag{3.30}$$

Again we present schemes in the time independent case where $N(t, u) = N(u)$, $G(u)\Delta S = g_m(u)\Delta W^m$ for presentation purposes as the time dependent case can be inferred from eqs. (3.20) and (3.21)). Below we define two (stochastic Stratonovich modified schemes) from [27] below.

Method 3.7 (eSSPIFSRK$^+$(2,2))**.** The eSSPIFSRK$^+$(2,2) method is given by:

$$k^1 = e^{L\Delta t} \left(u^n + \Delta t N(u^n) + G(u^n)\Delta S\right) \tag{3.31}$$

$$u^{n+1} = \frac{1}{2} e^{L\Delta t} u^n + \frac{1}{2} e^{L\Delta t} \left(k^1 + \Delta t N(k^1) + G(k^1)\Delta S\right) \tag{3.32}$$

Method 3.8 (eSSPIFSRK$^+$(3,3))**.** The eSSPIFSRK$^+$(3,3) method is given by:

$$k^1 = \frac{1}{2} e^{\frac{2}{3}\Delta t L} u^n + \frac{1}{2} e^{\frac{2}{3}\Delta t L} \left(u^n + \frac{4}{3}\Delta t N(u^n) + \frac{4}{3} G(u^n)\Delta S\right) \tag{3.33}$$

$$k^2 = \frac{2}{3} e^{\frac{2}{3}\Delta t L} u^n + \frac{1}{3} \left(k^1 + \frac{4}{3}\Delta t N(k^1) + \frac{4}{3} G(k^1)\Delta S\right) \tag{3.34}$$

$$u^{n+1} = \frac{59}{128} e^{\Delta t L} u^n + \frac{15}{128} e^{\Delta t L} \left(u^n + \frac{4}{3}\Delta t N\left(u^n\right) + \frac{4}{3} G(u^n)\Delta S\right) + \frac{27}{64} e^{\frac{1}{3}\Delta t L} \left(k^2 + \frac{4}{3}\Delta t N\left(k^2\right) + \frac{4}{3} G(k^2)\Delta S\right). \tag{3.35}$$

3.3. Stochastic Exponential Time Differencing Runge Kutta (SETDRK) methods

Consider the following SDE,

$$d_t u = [N(t,u) + Lu]\, dt + [g_m(t,u)] \circ dW_t^m, \tag{3.36}$$

Using an integrating factor technique, one can write this in integral form between t^n and t^{n+1} as follows

$$u^{n+1} = e^{L\Delta t}u^n + e^{L\Delta t}\int_0^{\Delta t} e^{-Ls}N(t_n+s, u(t_n+s))ds + e^{L(\Delta t + t^n)}\int_{t^n}^{t^{n+1}} e^{-Ls}g_m(s,u(s)) \circ dW_s^m. \tag{3.37}$$

Unlike SRK or SIFRK methods, SETDRK, do not derive analogously to the deterministic ETDRK scheme, and to approximate eq. (3.37), we explicitly require the use of the transform in eq. (3.22) and theorem 2.2 to use any deterministic order 2 or higher ETDRK integration method. Following [24] the general case with eq. (3.22) is given by:

Method 3.9 (CSETDRK)**.**

$$k_i = \chi_i(\Delta t L)\, u_n + \Delta t \sum_{j=1}^{s} a_{ij}(\Delta t L)\left(N(t_n + c_j\Delta t, k_j) + \frac{g_m(t_n + c_j\Delta t, k_j)\,\Delta W^m}{\Delta t}\right) \tag{3.38}$$

$$u_{n+1} = \chi(\Delta t L)\, u_n + \Delta t \sum_{i=1}^{s} b_i(\Delta t L)\left(N(t_n + c_j\Delta t, k_i) + \frac{g_m(t_n + c_j\Delta t)\Delta W^m}{\Delta t}\right) \tag{3.39}$$

Here, χ, χ_i, a_{ij}, and b_i are not just coefficients, but are constructed from rational expressions involving the exponential evaluated at $(\Delta t L)$. For consistency it is assumed $\chi(0) = \chi_i(0) = 1$, and the underlying Runge-Kutta method is denoted by the coefficients $(a_{i,j}(0), b_j(0))$, typically assumed to satisfy $c_i = A(0)e$, $b^T(0)e = 1$. It is normally considered $\chi(z) = \mathrm{e}^z$ and $\chi_i(z) = \mathrm{e}^{c_i z}, 1 \le i \le s$. We refer to [24] for the general construction of deterministic order 2 and higher ETDRK methods, which require specific construction of χ, χ_i, a_{ij}, and b_i.

Remark. Fortunately, in light of theorem 2.2, one need not know the general construction of ETDRK methods, but simply adopt any deterministic order $p \ge 2$ method, under the above change of variables:

$$\mathcal{N}(t^n, u^n) := N(t_n, u_n) + g_m(t_n, u_n)\frac{\Delta W^m}{\Delta t}. \tag{3.40}$$

One can inspect a convenient stochastic ETDRK2 (modified from [16]) as follows.

Method 3.10 (SETDRK2)**.** The Stochastic-ETDRK2 scheme is given by:

$$k_1 = e^{L\Delta t}u^n + A_1\mathcal{N}(t^n, u^n), \tag{3.41}$$

$$u_{n+1} = k_1 + A_2\left(\mathcal{N}(t_n + \Delta t, k_1) - \mathcal{N}(t_n, u_n)\right) \tag{3.42}$$

where

$$A_1 = \frac{e^{L\Delta t} - 1}{L}, \quad A_2 = \frac{\left(1 + \Delta t L - e^{L\Delta t}\right)}{(\Delta t)L^2}. \tag{3.43}$$

Theorem 2.2 implies this scheme has deterministic, drift commutative, commutative, and general mean square order $(P_d, P_{dc}, P_c, P_s) = (2, 1, 1, 0.5)$, captures symetric terms in the Stratonovich-Taylor expansion to order 1. The approximation of A_i is discussed in section 3.3.1

Similarly, one can inspect a convenient stochastic ETDRK3 (modified from [26, 16])

Method 3.11 (SETDRK3)**.** The Convenient-Stochastic-ETDRK3 scheme is given by:

$$k^1 = u_n e^{L\Delta t/2} + B_1 \mathcal{N}(t_n, u_n), \tag{3.44}$$
$$k^2 = u_n e^{L\Delta t} + B_2 \left(2\mathcal{N}\left(t_n + \Delta t/2, k^1\right) - \mathcal{N}(t_n, u_n)\right), \tag{3.45}$$
$$u_{n+1} = u_n e^{L\Delta t} + B_3 \mathcal{N}(t_n, u_n) + B_4 \mathcal{N}\left(t_n + \Delta t/2, k^1\right) + B_5 \mathcal{N}\left(t_n + \Delta t, k^2\right). \tag{3.46}$$

Where

$$B_1 = \frac{\left(e^{L\Delta t/2} - 1\right)}{L}, \quad B_2 = \frac{\left(e^{L\Delta t} - 1\right)}{L}, \quad B_3 = \frac{-4 - \Delta t L + e^{L\Delta t}\left(4 - 3\Delta t L + \Delta t^2 L^2\right)}{\Delta t^2 L^3}, \tag{3.47}$$
$$B_4 = 4\frac{2 + \Delta t L + e^{L\Delta t}(-2 + \Delta t L)}{\Delta t^2 L^3}, \quad B_5 = \frac{-4 - 3\Delta t L - \Delta t^2 L^2 + e^{L\Delta t}(4 - \Delta t L)}{\Delta t^2 L^3}. \tag{3.48}$$

Theorem 2.2 gives deterministic, drift commutative, commutative, and general mean square order $(P_d, P_{dc}, P_c, P_s) = (3, 1, 1, 0.5)$. The approximation of B_i is discussed in section 3.3.1

Similarly one can inspect, the convenient Stratonovich extension of Cox-Mathews fourth order integration scheme originally derived in the deterministic setting by symbolic computation in [16].

Method 3.12 (SETDRK4)**.** The Stochastic ETDRK4 scheme is given by:

$$a_n = e^{L\Delta t/2} u_n + E_0 \mathcal{N}(t_n, u_n), \tag{3.49}$$
$$b_n = e^{L\Delta t/2} u_n + E_0 \left[\mathcal{N}(t_n + \Delta t/2, a_n)\right], \tag{3.50}$$
$$c_n = e^{L\Delta t/2} a_n + E_0 \left[2\mathcal{N}(t_n + \Delta t/2, b_n) - \mathcal{N}(t_n, u_n)\right], \tag{3.51}$$
$$u_{n+1} = e^{L\Delta t} u_n + E_1 \mathcal{N}(t_n, u_n) + E_2 \left[N(t_n + \Delta t/2, a_n) + \mathcal{N}(t_n + \Delta t/2, b_n)\right] + E_3 \mathcal{N}(t_n + \Delta t, c_n), \tag{3.52}$$

where

$$E_0 := L^{-1}\left(e^{L\Delta t/2} - I\right), \tag{3.53}$$
$$E_1 := \Delta t^{-2} L^{-3}\left[-4 - L\Delta t + e^{L\Delta t}\left(4 - 3L\Delta t + (L\Delta t)^2\right)\right], \tag{3.54}$$
$$E_2 := \Delta t^{-2} L^{-3}\left[4 + 2L\Delta t + e^{L\Delta t}(-4 + 2L\Delta t)\right], \tag{3.55}$$
$$E_3 := \Delta t^{-2} L^{-3}\left[-4 - 3L\Delta t - (L\Delta t)^2 + e^{L\Delta t}(4 - L\Delta t)\right]. \tag{3.56}$$

Theorem 2.2 gives deterministic, drift commutative, commutative, and general mean square order $(P_d, P_{dc}, P_c, P_s) = (4, 2, 1, 0.5)$. The approximation of E_i is discussed in section 3.3.1

3.3.1. Numerical implementation details

We consider the following motivating SPDE, as a prototypical nonlinear SPDE with quadratic nonlinearity, linear stiff term and nonlinear noise,

$$d_t u + \left(\underbrace{\frac{c_1}{2}(u^2)_x}_{Nonlinear} + \underbrace{c_0 u_x + c_2 u_{xx} + c_3 u_{xxx} + c_4 u_{xxxx}}_{Linear}\right) dt + \underbrace{\sum_{m=1}^{M} (\xi_m(x) u)_x \circ dW^m}_{Nonlinear} = 0. \tag{3.57}$$

Subcases include the Heat equation, Linear advection equation, Burger's equation, KdV equation, KS equation, under a nonlinear flux form transport noise interpreted in the Stratonovich sense. Taking the Fourier transform gives,

$$d\hat{u} + \underbrace{\frac{ikc_1}{2}\widehat{(u^2)}}_{-N(u)} + \underbrace{(c_0 ik\hat{u} - k^2 c_2 \hat{u} - c_3 ik^3 \hat{u} + c_4 k^4 \hat{u})}_{-Lu} + \sum_{m=1}^{M} \underbrace{ik\widehat{(\xi_p u)}}_{g_m(u)} \circ dW^m = 0 \tag{3.58}$$

where k denotes the wave-number (scaled by $2\pi/L_x$ where L_x is the length of the domain). By taking a finite dimensional approximation of eq. (3.58) one has a discrete analogue to eq. (3.58), a finite-dimensional system of SDE's of the following form

$$d_t u_h = [N(\boldsymbol{u}_h, t) + L u_h]\, dt + \sum_{m=1}^{M} g_m(u_h, t) \circ dW_t^m. \tag{3.59}$$

In our work we use a pseudo-spectral discretisation where the nonlinear terms N, g_m of quadratic type are computed by multiplication in real space, then Fourier transformed, multiplied by the wave numbers and dealiased through a 2/3-rds rule before inverse Fourier transformed.

Due to numerical instability highlighted in [23] the rational expressions involving the exponential $\{A_i\}_{i=1,2}$, $\{B_i\}_{i=1,2,3,4,5}$, $\{E_i\}_{i=1,2,3,4}$, in method 3.10, method 3.11, method 3.12 respectively, have removable singularities and require approximation, and in this work are computed before runtime following Kassam and Trefethen's, numerical quadrature of contour integration in the complex plane [30]. As pointed out in [30] the practical computation of E_i, using the following contour integration

$$f(\boldsymbol{L}) = \frac{1}{2\pi i} \oint_{\Gamma} f(t)(t\boldsymbol{I} - \boldsymbol{L})^{-1} dt, \tag{3.60}$$

has significant simplification when the operator diagonalises and the trapezium rule can be employed around roots of a circle in the complex plane centred at specific points relating to eigenvalues of $\boldsymbol{L}$ [30]. The trapezium rule (in this complex plane context) converges exponentially fast, and is done before runtime. Eigenvalues of $\boldsymbol{L}$ depend on the spatial discretisation, and number of spatial points used, for even diffusive operators such as the KS equation $-k^2 + k^4$ eigenvalues are near the negative real axis, and for dispersive odd-order operators like the KdV equation eigenvalues are close to the imaginary axis $-ik^3$, such differences can be accounted for when defining the contour integration for increased performance [30]. The contour integration method applies more generally and includes operators that do not diagonalise [30]. The numerical computation of matrix exponentials and rational expressions involving exponential operators can be attained through Chebyshev approximation, diagonalisation, Padé approximation [17], Leja approximation, contour integration, Krylov subspace methods [22], scaling and squaring [14], see [24] for an overview.

4. Numerical results

In this section we show theorem 2.2 has practical consequences for numerical integration.

1. For general non-commutative noise we show mean square order $\mathcal{O}(\Delta t^{1/2})$, theorem 2.2-item 1.
2. For commutative noise we show mean square order $\mathcal{O}(\Delta t^{1})$ theorem 2.2-item 6.
3. For drift commutative noise we show order $\mathcal{O}(\Delta t^{\mathrm{int}(p/2)})$ theorem 2.2-item 3+item 4.
4. In the limit of small noise, large timestep Δt and no noise we show $\mathcal{O}(\Delta t^{p})$ theorem 2.2-item 5.

We show each of these convergence behaviors in the context of SIFRK, SETDRK, and SRK methods of orders $p = 2, 3, 4$, and facilitate relative comparisons between all 9 methods of different types and orders for different examples.

4.1. Experiment 1: General Multidimensional noise, convergence and efficiency.

Theorem 2.2, item 1 indicates all numerical schemes presented in this paper converge to the Stratonovich system with strong order 1/2 (for general multidimensional non-commutative noise). However, theorem 2.2 item 2 also indicates that all higher order symmetric terms up to $p/2$ are captured when using a deterministic scheme of order p. This experiment determines to what extent both items 1 and 2 are observed in practice for general multidimensional non-commutative noise, by considering the following initial value problem

$$d_t u + u u_x + u_{xxx} + \sum_{m=1}^{M} (\xi_m u)_x \circ dW_t = 0, \quad \xi_m(x) = \frac{\sin(2\pi x m)}{100 m}, \quad u(x, 0) = \exp(-50(x - 0.5)^2). \tag{4.1}$$

This setup does not admit an analytic pathwise solution and the noise (plotted in fig. 4.1) is non-commutative in the infinite dimensional Fréchet operator sense $[g_i, g_j] = (\xi_i(\xi_j)_{xx} - \xi_j(\xi_i)_{xx})\, u + ((\xi_i)_x\xi_j - (\xi_j)_x\xi_i)u_x \neq 0$, nor drift commutes $[f_i, g_i] \neq 0$. A high resolution solution u_e is used as the reference solution and plotted in fig. 1.1b.

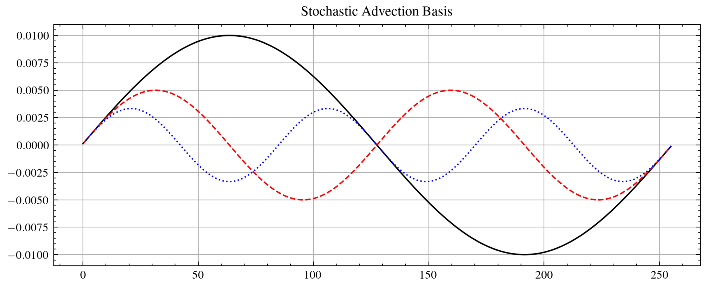
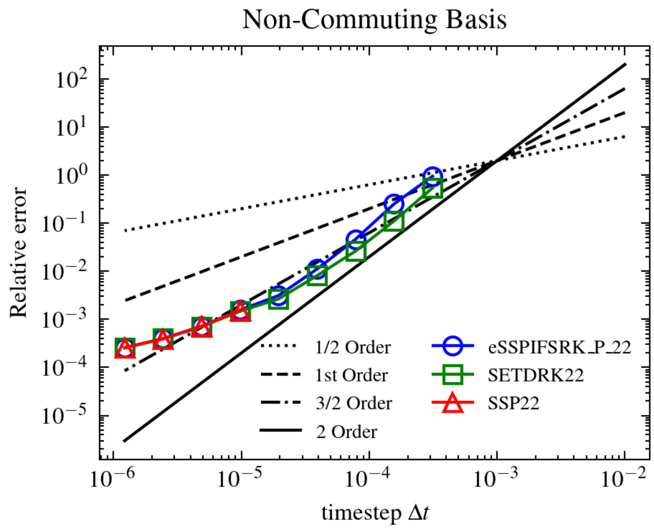


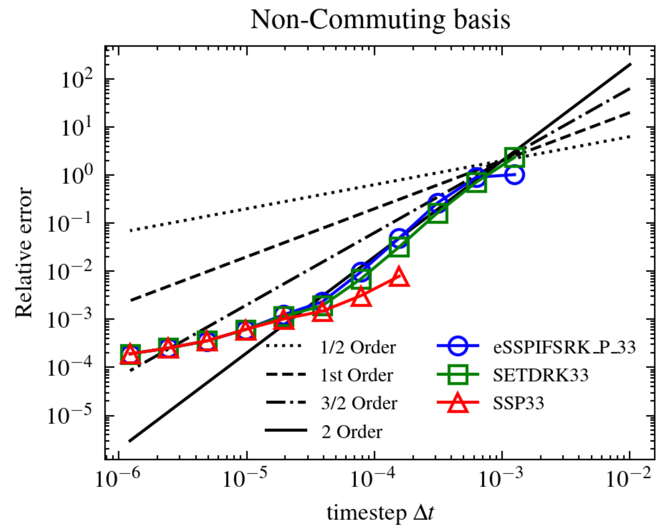
Figure 4.1: Non-commutative basis $\{\xi_m\}_{m=1,2,3}$

In fig. 4.2 we plot the results, in fig. 4.2a, fig. 4.2b, fig. 4.2c, we respectively plot the pathwise relative L2 errors of the second order, third order, and fourth order schemes employed in this paper as compared with the high resolution reference. This is to facilitate a comparison between SRK, SIFRK and SETDRK schemes. If a figure omits a data point, the numerical method failed to converge at that temporal resolution, and this is adopted for all figures. In fig. 4.2d all convergence results are overlayed on the same graph as to facilitate comparison between methods of different orders.

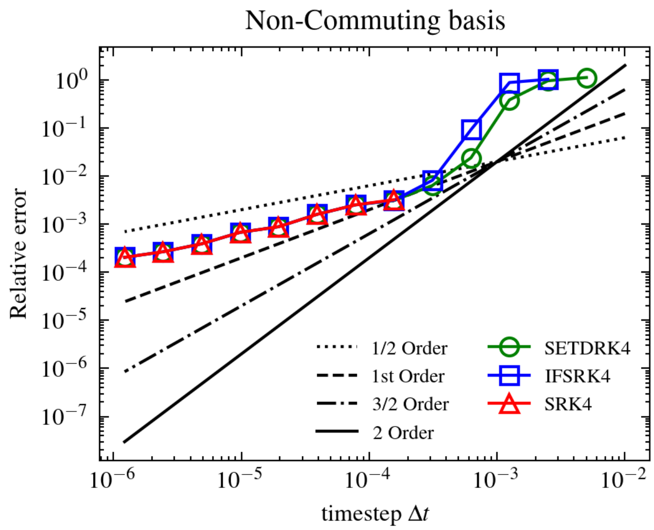
Remark. Order lines are not placed in the same position in each plot, and the y axis are different amongst figures, so fig. 4.2d is most appropriate to compare schemes of different orders. The other figures in fig. 4.2a, fig. 4.2b, fig. 4.2c, are for comparing SETDRK, SIFRK and SRK methods of equal order.



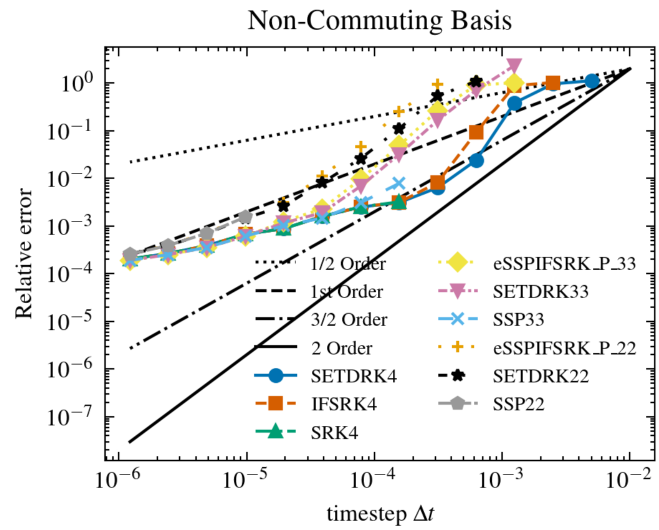
(a) Comparison of 2nd order methods non-commuting basis.



(b) Comparison of 3-rd order methods non-commuting basis.



(c) Comparison of 4th order methods non-commutative basis.



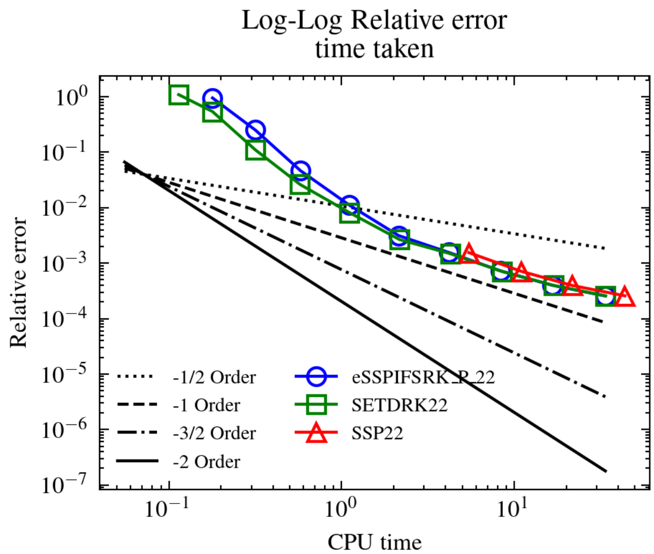
(d) Comparison of all methods non-commutative basis.

Figure 4.2: Temporal convergence for non-commutative noise. In all schemes we see that for non-commutative noise, we attain strong order 1/2 in the limit of $\Delta t \to 0$ in fig. 4.2d. SETDTK and SIFRK methods could be computed at timesteps larger than RK methods. SETDTK had errors slightly lower than SIFRK methods. In fig. 4.2d when using larger timesteps there was significant practical merit in using higher order schemes.

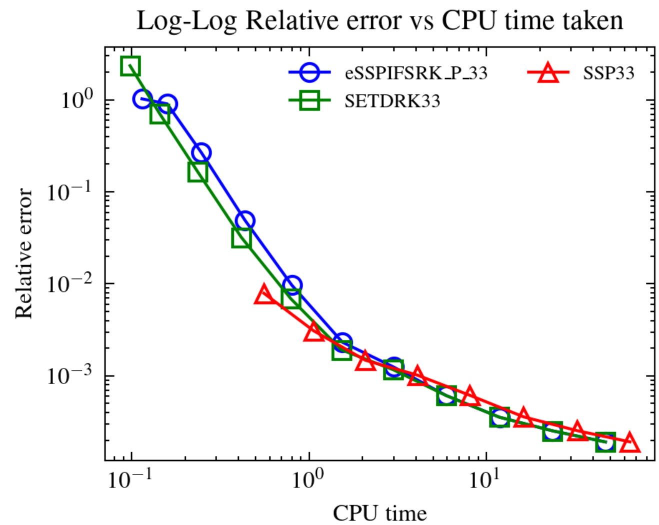
As shown in fig. 4.2a SETDRK22 and IFSRK22 can take timesteps 32 times larger than SSPRK22. Errors are slightly lower in SETDRK22 than IFSRK22. As shown in fig. 4.2b SETDRK33 and SIFRK33 can take timesteps 8 times larger than SSP33. As shown in fig. 4.2b SETDRK4 and IFRK4 can take timesteps 32, 16 times larger than RK4 respectively. SETDRK4 has smaller error than IFRK4 at large timesteps.

In fig. 4.2d all schemes are $\mathcal{O}(\Delta t^{1/2})$ in the limit of small Δt in accordance with theorem 2.2 item 1. Indicating decreasing benefit in using increasingly higher order Runge-Kutta, SIFRK and SETDRK approaches in the limit of small noise due to the dominant missing Lévy area approximation. However, as simultaneously shown in fig. 4.2d when using larger and intermediate timesteps, there was significant practical merit in using a higher-order SRK, SETDRK or SIFRK approach in terms of absolute error.
We run the same convergence experiment and at each temporal resolution we recompute the entire simulation 5 times and take an average for a CPU-time (per resolution per method). We then plot a relative error vs CPU time graphs in fig. 4.3a, fig. 4.3b, fig. 4.3b and fig. 4.8d. Figures fig. 4.3a, fig. 4.3b, fig. 4.3b present CPU-time versus relative error using second, third, and fourth-order schemes, respectively. Figure 4.8d collates the performance of all methods. Ideally, a method has small error for a given computational cost, if one scheme

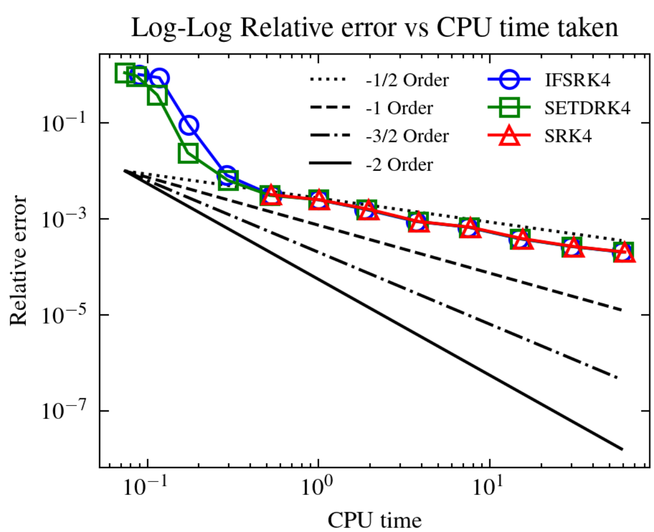
is below another scheme it is deemed more computationally efficient than another for a given CPU cost. This experiment takes into account the exponential operator cost and the cost associated with using more stages.



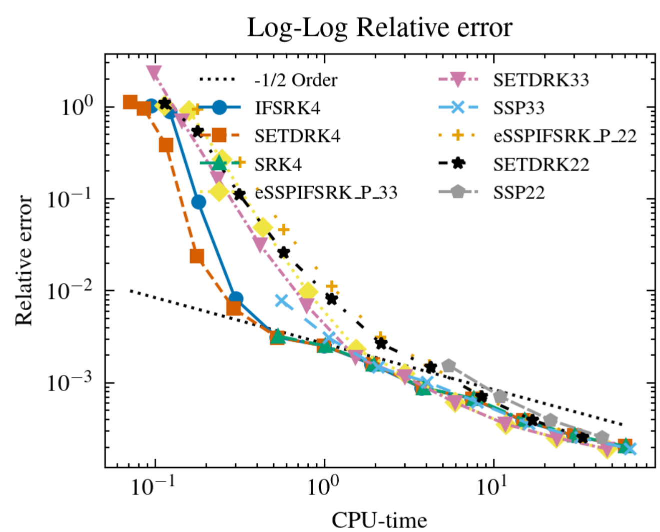
(a) Non-commuting comparison of 2nd order methods CPU-time vs Relative error.



(b) Non-commuting comparison of 3-rd order methods CPU-time vs Relative error.



(c) Non-commuting comparison of 4th order methods CPU-time vs Relative error.



(d) Non-Commuting comparison of all methods CPU-time vs Relative error.

Figure 4.3: CPU-time vs Relative error

Figure 4.3a demonstrates that among the three second-order schemes tested, SETDRK22 provides the best accuracy for a given CPU time. Its curve lies below that of SSP22 and eSSPIFSRK^{+}22 across the entire range. Figure 4.3b, demonstrates that amongst the three third-order schemes tested, SETDRK33 provides the best accuracy for a given CPU time. Its curve lies below that of eSSPIFSRK^{+}33 and SIFRK33 for the majority of CPU times; however SSP33 showed slight computational advantage at intermediate CPU times. Figure 4.3c, demonstrates that amongst the three fourth-order schemes tested (IFSRK4, SETDRK4, and SRK4) the CPU-ERROR plot collapse almost exactly onto a single curve (when using small timesteps). Indicating similar efficiency of all schemes when running at small timesteps. In the same figure fig. 4.3c SETDRK4 was more cost efficient than IFSRK4 and SRK4 in some small CPU time regimes (corresponding to using large timesteps). Overall figs. 4.3a to 4.3c demonstrate that the computational overhead of exponential-type operators was justifiable, and the ETDRK schemes were overall more accurate per CPU time.
Figure 4.3d places all 9 methods on the same plot and scale. SETDRK4 is more accurate than SSP22 at all

CPU times investigated. This takes into account the cost of using exponentials and more stages. In regimes corresponding to taking increasingly small timesteps (corresponding to the bottom right corner in Figure 4.3d) the advantage to using a higher order scheme is increasingly diminished for non-commutative noise when the stage cost is taken into account. Here eSSPIFSRK$^+$33 emerges as being more computationally efficient than SETDRK4. One could speculate that when using extremely small timesteps (smaller than 10^{-6}) the trend observed in fig. 4.3d would continue and second order methods such as SSP22 would eventually be more cost effective than ETDRK4 in the stochastic context since they have the same strong order but require fewer stages. However, in practice many applications require modest timesteps and for our examples the higher methods were overall the most cost effective, even when stage cost and exponentials were taken into account and the noise is noncommutative. In some regimes, fourth order ETDRK schemes deliver an order of magnitude smaller error than second-order and third order methods for the same CPU time. It is also clear that the CPU overhead of using exponential operators can be mitigated enough to be as efficient or more efficient than a SRK method, whilst giving the possibility of larger timesteps. CPU-time depends on implementation, software, hardware, we ran all experiments on a M2-mac using python with @jax.jit in JAX.

4.2. Experiment 2: Commutative noise

We consider the same IVP as eq. (4.1), but define a $C^\infty(\mathbb{T})$ compactly supported basis of noise as follows:

$$\xi_j(x) = \mathbb{I}_{\{|x-c_j|<w/2\}} \exp\left[-\frac{1}{1-\left(\frac{2(x-c_j)}{w}\right)^2}\right], \quad c_j = x_{\min} + jw, \quad j = 1,\dots,M, \quad w = \frac{x_{\max}-x_{\min}}{M+1}, \quad M = 3. \tag{4.2}$$

plotted in fig. 4.1

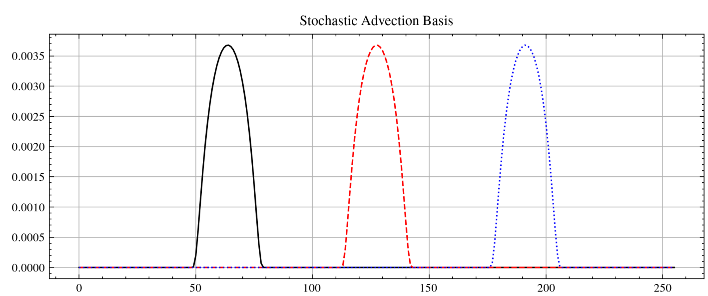
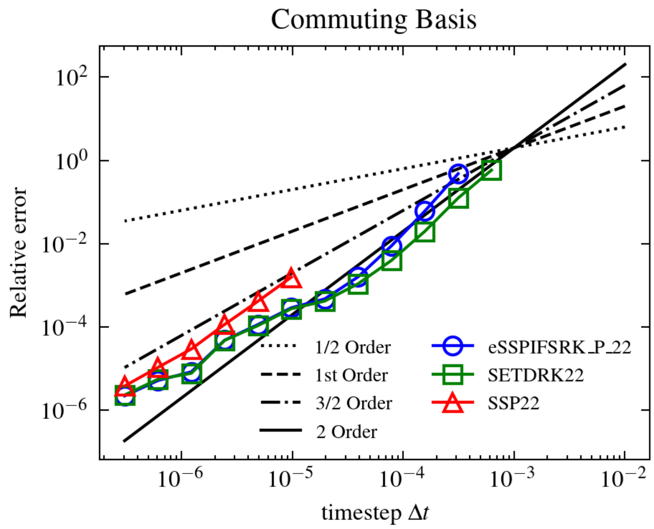
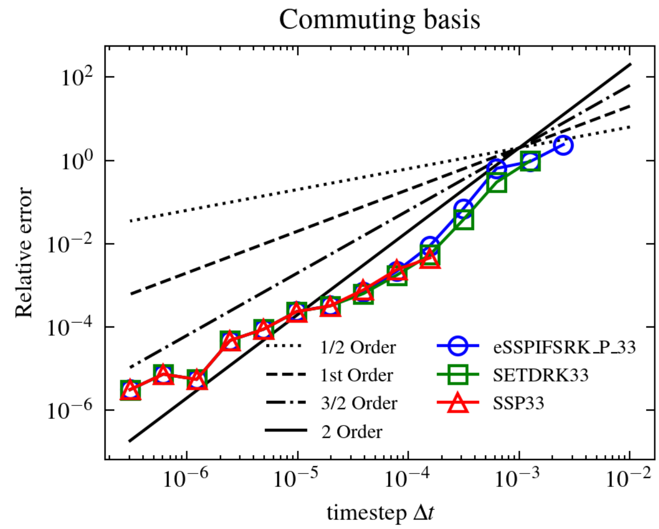


Figure 4.4: Commutative basis $\{\xi_m\}_{m=1,2,3}$.

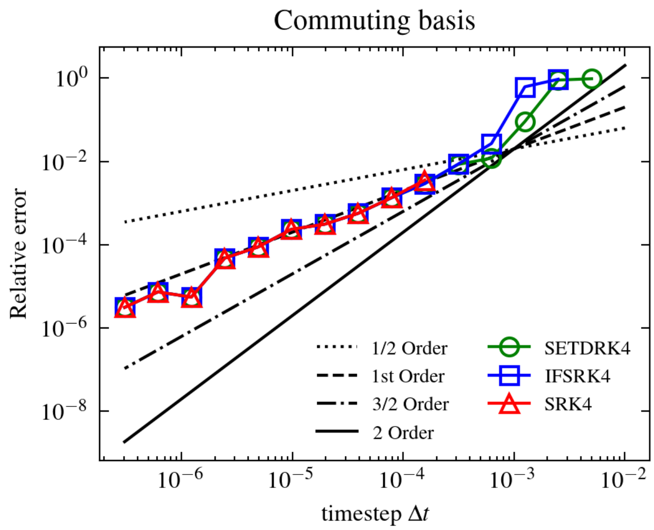
This setup does not admit an analytic path-wise solution. The noise commutes by compactness in the Fréchet operator sense $[g_i, g_j] = [(\xi_i u)_x, (\xi_j u)_x] = 0$ when $i \neq j$ but does not drift commute $[f, g_j] \neq 0$, therefore theorem 2.2-item 6 implies strong order 1 for all schemes. A high realization solution u_e is used as the reference solution. In figures figs. 4.5a to 4.5c, we plot the convergence for the second order, third order, fourth order schemes, in fig. 4.5d all 9 methods relative error are plotted on the same graph.



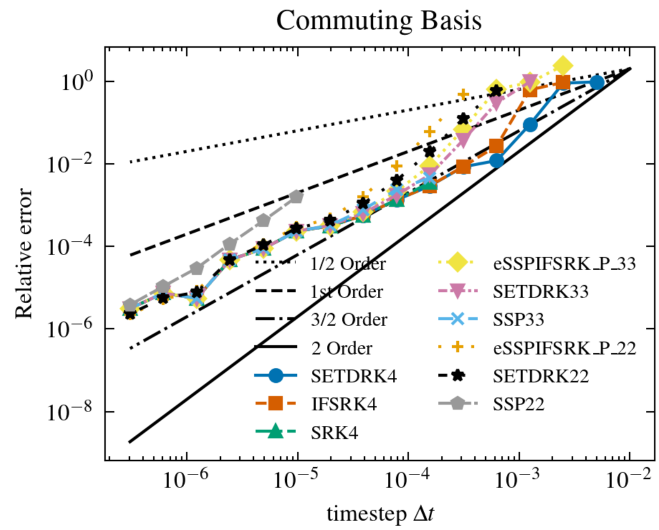
(a) Comparison of 2nd order methods commutative noise.



(b) Comparison of 3rd order methods commutative noise.



(c) Comparison of 4th order methods commutative noise.



(d) Comparison of all methods commutative noise.

Figure 4.5: Commutative noise convergence order 1 is verified theorem 2.2-item 6.

Although figs. 4.5a and 4.5b appear to have better convergence than $\mathcal{O}(\Delta t^1)$ observed in fig. 4.5c, this is because they have not reached the true order $\mathcal{O}(\Delta t^1)$ of convergence yet as demonstrated in fig. 4.5d.
ETDRK typically outperform IFRK methods and these outperform SRK methods in terms of not only accuracy but ability to run at larger timesteps, this was true independent of the order, but most apparent for the second order methods in fig. 4.5a. In terms of accuracy, 4th outperforms 3rd, which outperforms 2nd, as is clear in fig. 4.5d, but the increase in accuracy becomes increasingly diminished in the limit of small timesteps. We verify theorem 2.2-item 6, $\mathcal{O}(\Delta t^1)$, but simultaneously that higher-order schemes are more accurate, as a consequence of items 2 and 5

4.3. Experiment 3: Drift Commutative noise.

Theorem 2.2 item 3 and item 4 indicates one may attain higher-order convergence for drift commutative noise, but due to the abstract form of the integration method, when P is odd, the FTMSC was not sufficient to prove order $P/2$, but $P/2-1/2$ instead. The practical implications of such a theorem will be investigated numerically. To test the theoretical order of the scheme, we use stochastic travelling wave solutions to the following stochastic KdV equation under constant advection noise.

$$d_t u + uu_x + u_{xxx} + au_x \circ dW_t = 0, \quad a \in \mathbb{R}. \tag{4.3}$$

By transforming into the stochastic moving frame $X(t,x) := x - aW(t)$. One (section Appendix .2) attains the analytic pathwise stochastic travelling wave solutions to the KdV equation

$$u(x,t) = 3\beta \operatorname{sech}^2\left(\frac{\sqrt{\beta}}{2}(x - \beta t - aW(t))\right). \tag{4.4}$$

This particular setup not only admits analytic pathwise solutions, but also satisfies an operator version of the drift commutativity condition. Denote the drift and diffusion operators by

$$F(u) = (u^2/2)_x + u_{xxx}, \quad G(u) = au_x. \tag{4.5}$$

The Fréchet derivatives with respect to u in direction v are

$$DG(u)[v] = av_x, \quad DF(u)[v] = (uv)_x + v_{xxx}. \tag{4.6}$$

These can be used to verify the infinite dimensional generalisation of the drift commutativity condition considered in theorem 2.2-items 3 and 4 as follows:

$$[F, G] = DG(u)[F(u)] - DF(u)[G(u)] := a((u^2/2)_{xx} + u_{xxxx}) - (uau_x)_x - au_{xxxx} = 0. \tag{4.7}$$

In order to verify the pathwise (strong) convergence of the scheme we compute a sequence of Brownian paths of decreasing resolution [38], and compare the relative L2 error of the numerical method in comparison to the analytic solution.
We use the following parameters: $\Delta t = 10^{-4}(1/2^i)$, $i \in \{1, ..., 10\}$, $t_{max} = 0.1$, $a = 1$, $n_x = 256$, $\xi = 1$, $u_0 = 3\beta \operatorname{sech}^2(\frac{\sqrt{\beta}}{2}x)$ where $\beta = 8^2$ and compare the relative spacetime L^2 error as between the numerical solution and the analytic solution generated using eq. (4.4). In fig. 4.6 the analytic and numerical solution are plotted at the highest resolution, in addition to the convergence plot on the right in the same figure.

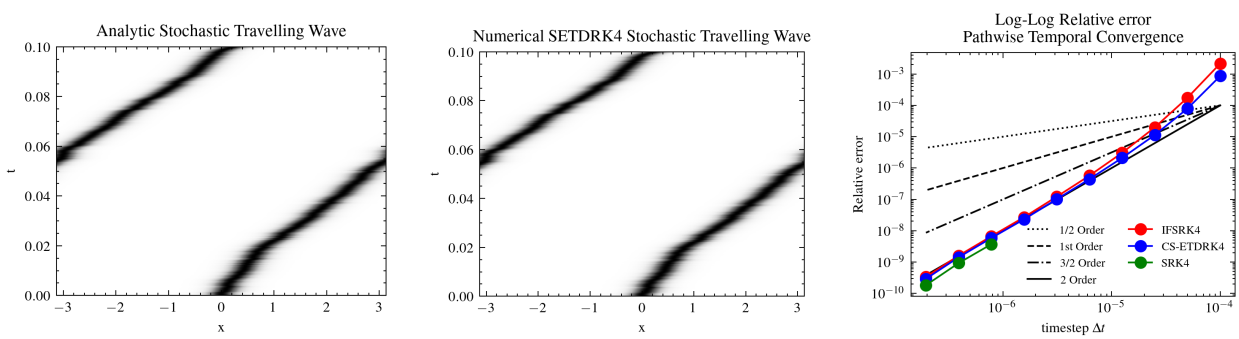
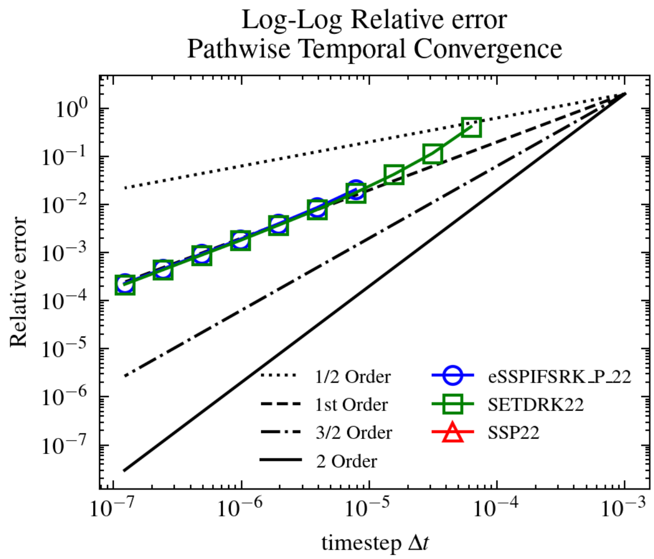


Figure 4.6: Analytic stochastic travelling wave solution to the KdV equation (left) and the numerical solution when using dealiased spectral method in space, and the SETDRK4 scheme in time (middle). Temporal convergence (relative L2 spacetime error) of some 4-th order methods, are plotted (right).

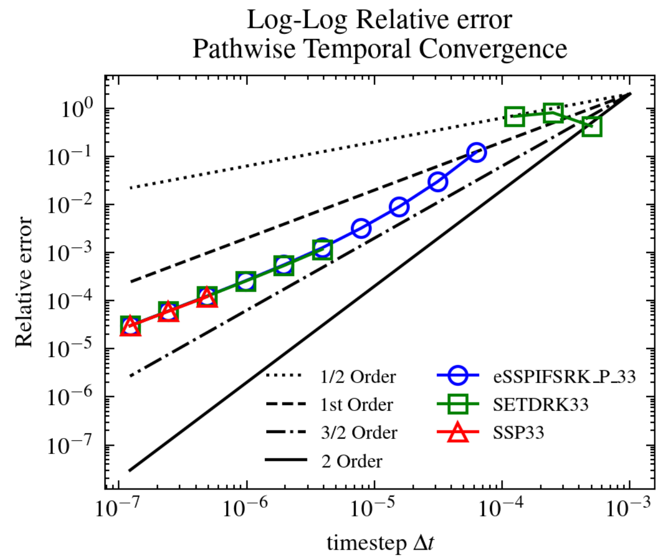
Figure 4.6 verifies, all 4-th order schemes (SRK4, IFSRK4 and ETDRK4) attain the drift commuting mean square(strong) order $\mathcal{O}(\Delta t^2)$ consistent with theorem 2.2-item 3. In fig. 4.6 the SETDRK4 scheme and the IFSRK4 scheme can run at timesteps at least 2^7 times larger than the SRK4 scheme. In fig. 4.6 the SETDRK4 scheme is more accurate than the SIFRK4 scheme at all timesteps.

4.4. Experiment 3a: Drift Commutative noise continued.

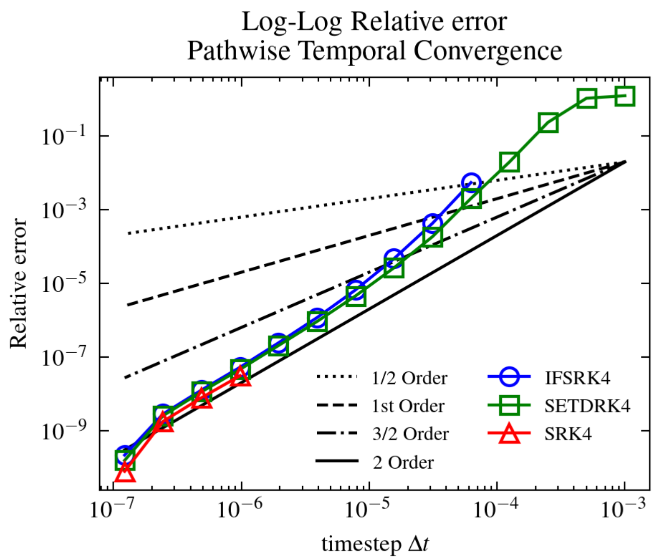
In this example we plot relative error for all 9 timestepping schemes over a wider range of possible timesteps. If a scheme is omitted from the plot at any timestep value, then the numerical scheme could not run stably at that temporal resolution. We use the following parameters: $\Delta t = 10^{-3}(1/2^i)$, $i \in \{1, ..., 14\}$, $t_{max} = 0.01$, $n_x = 256$, $\xi = 1$, $u_0 = 3\beta \operatorname{sech}^2(\frac{\sqrt{\beta}}{2}x)$ where $\beta = 9^2$, and use the final timestep to calculate the relative L^2 error from the analytic solution. In fig. 4.7 we plot the convergence of the second order methods. In fig. 4.7b we plot the convergence of the third order methods. In fig. 4.7c we plot the convergence of the fourth order methods. In fig. 4.7d we overlay convergence results for all schemes onto one graph.



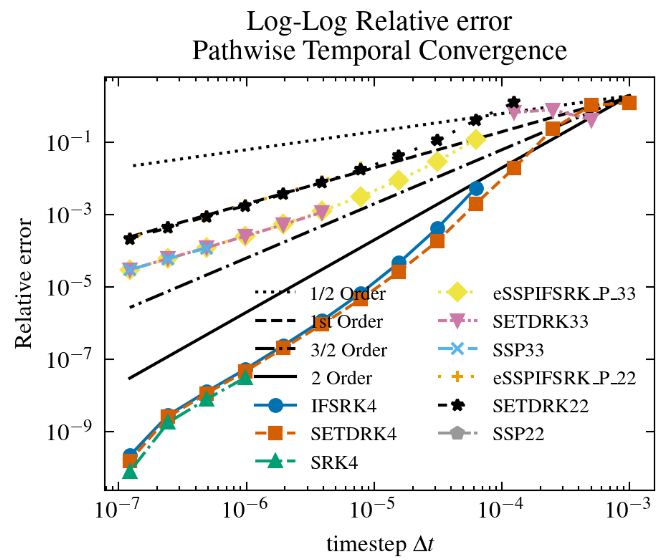
(a) Comparison of 2nd order methods, drift commutative noise.



(b) Comparison of 3-rd order methods drift commutative noise.



(c) Comparison of 4-th order methods drift commutative noise.



(d) Comparison of all methods drift commutative noise.

Figure 4.7

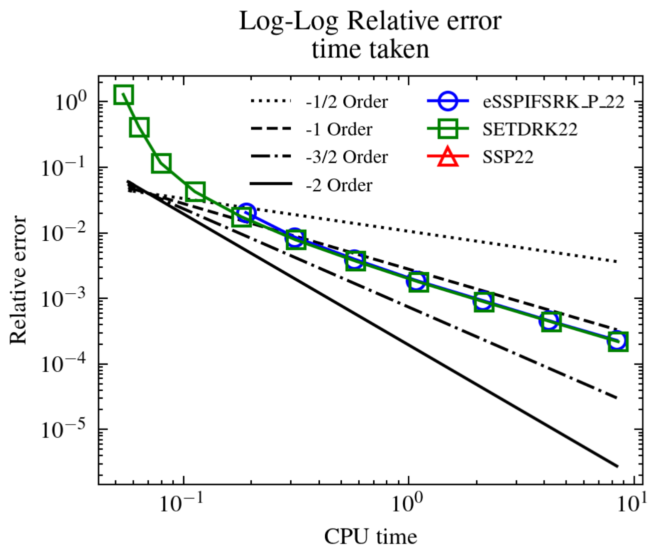
In fig. 4.7d, we see that Theorem 2.2-item 3 and Theorem 2.2-item 4 are observed in practice for SETDRK, SIFRK and SRK time integration schemes. Theorem 2.2-item 3 predicts that for the second order methods $\mathcal{O}(\Delta t^1)$ convergence and for the fourth order methods $\mathcal{O}(\Delta t^2)$ convergence. Theorem 2.2-item 4 predicts that for the third order methods $\mathcal{O}(\Delta t^1)$ convergence.

In fig. 4.7a stochastic SSP22 did not converge for any timestep, eSSPIFSRK22 converged with strong order $\mathcal{O}(\Delta t^1)$, SETDRK22 also converged with strong order $\mathcal{O}(\Delta t^1)$ and was marginally more accurate than eS-SPIFRK22 and ran at a timestep 8 times as large.

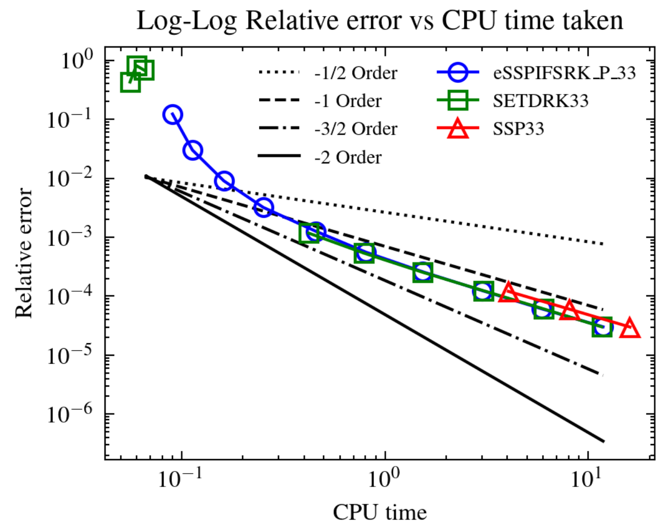
In fig. 4.7b SSP33, SETDRK33 and eSSPIFSRK33 converged with strong order $\mathcal{O}(\Delta t^1)$ as theoretically predicted in theorem 2.2-item 4. The SSP33 scheme was only stable for small timesteps. SETDRK and eSSPIFSRK methods were capable of larger timesteps than the SSP33 method. As shown in fig. 4.7b eSSPIFSRK33 was capable of running at timesteps SETDRK33 could not and vice versa.

In fig. 4.7c SRK4, SETDRK4 and SIFRK4 all converged with strong order $\mathcal{O}(\Delta t^2)$ for small timesteps, as theoretically predicted in item 3 in theorem 2.2. In fig. 4.7c at the few points SRK4 converged it was the most accurate, SETDRK4 was more accurate than the IFRK4 scheme at all timesteps. IFRK4 was capable of timesteps over 2^6 larger than SRK4. SETDRK4 was capable of timesteps 2^{10} larger than the SRK4. In fig. 4.7c the SETDRK4 schemes was calculated as having strong order 3.6 at large timesteps but order 2 at small

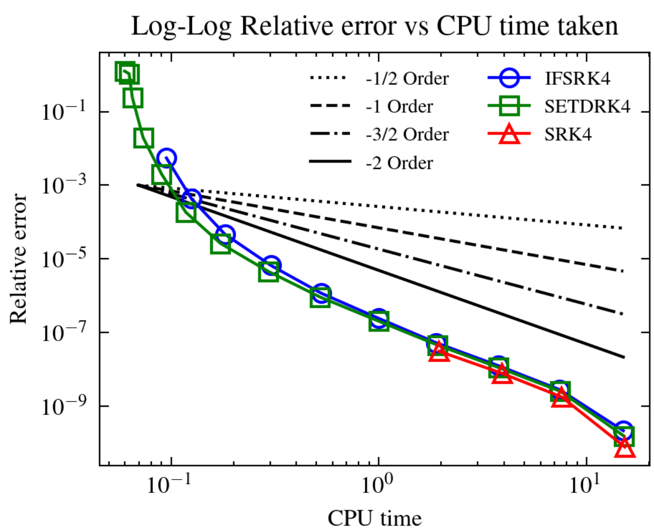
timesteps.
In fig. 4.7d we overlay all the plots on the same figure. We see significant practical improvements in accuracy by using deterministic fourth order accurate schemes as opposed to deterministic second order ones or third order ones, for instance SETDRK4 attained similar error to eSSPIFSRK33 and SETDRK3 at a timestep 2^7 larger.
We run the convergence tests at each temporal instance 5 times and average for a the CPU-time. The CPU-time is plotted on the x-axis and the relative error is plotted on the y-axis. In fig. 4.8, we observe the $\mathrm{int}(P/2)$ order behaviour even when taking into account the CPU-time. This indicates that the increase in accuracy is not lost when taking into account the stage cost and the exponential evaluation cost.



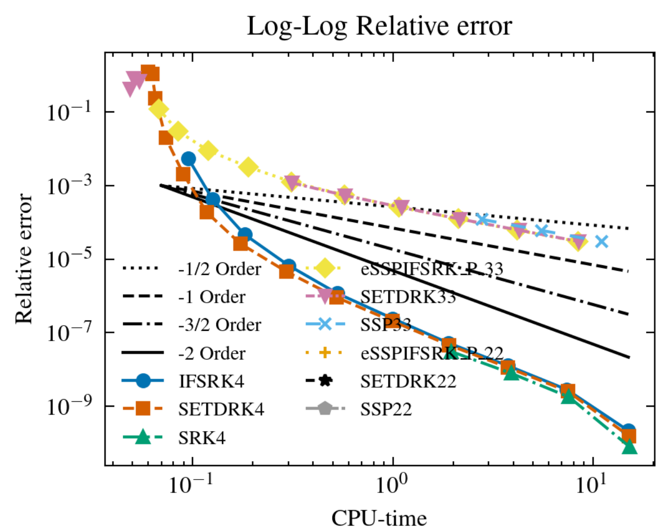
(a) Comparison of 2nd order methods CPU-time vs Relative error.



(b) Comparison of 3-rd order methods CPU-time vs Relative error.



(c) Comparison of 4th order methods CPU-time vs Relative error.



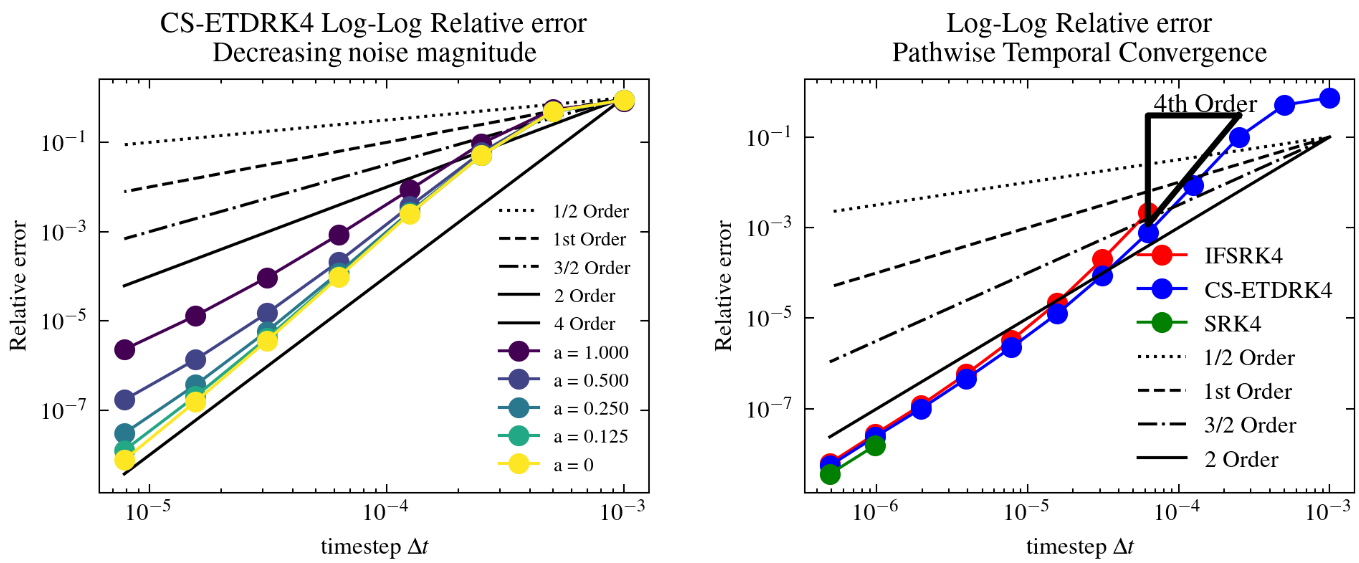
(d) Comparison of all methods CPU-time vs Relative error.

Figure 4.8: CPU-time vs Relative error, this experiment demonstrates that the "increase" in computational cost (by using more stages or exponentials) does not offset the increased performance in terms of decreasing error. By plotting CPU time instead of temporal resolution on the x-axis, cost is taken into account. We observe theoretically predicted rates of relative error decrease, consistent with assuming that the CPU-cost of stages scales linearly and the cost of the exponential operator is done before runtime, i.e. we attain $\mathcal{O}(1)$, $\mathcal{O}(1)$, $\mathcal{O}(2)$, for deterministic order $2, 3, 4$ schemes predicted by theorem 2.2.

4.5. Experiment 4: Decreasing noise magnitude and large timestep order.

We solve the SETDRK4 scheme under the drift commutative setup but with decreasing noise magnitudes $\xi = a = \{1, 1/2, 1/4, 1/8\}$, the results are plotted in fig. 4.9a where we observe deterministic order 4, emerging in the limit $\xi = a \to 0$ as expected from item 5, in theorem 2.2.

In fig. 4.9b, we replot fig. 4.7c and plot a 4-th order gradient where we observe that at large timesteps the deterministic term in the SDE dominates the stochastic term (i.e. $f\Delta t >> g\Delta W$), such that when taking large timesteps deterministic convergence order can play a role in the error. This is observed in fig. 4.9b particularly when employing ETDRK or IFRK schemes which can take larger timesteps. As shown in the previous convergence plots (fig. 4.2d) high order deterministic convergence occurs at large timesteps independent of whether the noise is commutative or drift commutative further exemplifying the importance of capturing higher order deterministic dynamical terms (item 5 from theorem 2.2).



(a) Decreasing noise magintude

(b) Large timesteps

Figure 4.9: Higher order convergence assiciated with small noise or large timesteps.

4.6. Experiment 5: 2D Incompressible Navier Stokes Equation.

This example simply indicates SETDRK methods are adoptable to higher-dimensional fluid-dynamics applications, and may be helpful to avoid timestep restrictions associated with linear stiff operators. The 2D incompressible Navier-Stokes equation in vorticity formulation under a stochastic perturbation to the momentum in the Lie-Poisson bracket takes the following form:

$$\underbrace{dq + J\left(\psi, qdt + \sum_{m=1}^{M} \xi_m(x,y) \circ dW\right)}_{N+g_m \circ dW} = \underbrace{\mu\Delta q - \nu\Delta^2 q}_{-Lu}, \quad J(a,b) := a_x b_y - b_x a_y, \quad \psi = -\Delta^{-1} q, \quad u = -\nabla^{\perp}\psi. \tag{4.8}$$

In the absence of viscosity and hyperviscosity $(\mu, \nu) = (0, 0)$ this type of stochastic perturbation (originally appearing more generally in [25] under the acronym SFLT) preserves energy but also the first integral of vorticity since $J(a,b) = \{a,b\} = \mathrm{div}(-[\nabla^{\perp} a]b)$ and

$$d_t\left(\frac{1}{2}\int_\Omega u^2 d^2x\right) = d_t\left(\frac{1}{2}\int_\Omega -\psi q d^2x\right) = -\frac{1}{2}\int J(\psi^2, qdt + \sum_{m=1}^{M} \xi_m \circ dW) = 0, \quad d_t\left(\frac{1}{2}\int_\Omega q d^2x\right) = 0. \tag{4.9}$$

Therefore, with diffusion and hyperdiffusion the above stochastic perturbations do not change the usual global 2D deterministic energy balance of the NS equation and the stochastic system obeys:

$$d_t \frac{1}{2}||u||_2^2 = -\mu||\nabla u||_2^2 - \nu||\Delta u||_2^2 \leq 0. \tag{4.10}$$

But allows spread per ensemble member. Viscosity dissipates energy, and hyper-viscosity controls unresolved small scales; however, unless μ, ν are adjusted carefully to be small and proportional to the grid, both diffusion and hyper-diffusion can result in additional CFL restrictions for explicit schemes. Under a flux-form de-aliased Fourier spatial discretisation of eq. (4.8), one identifies non-linear contributions N, g_m and linear contributions Lu which can be treated with SETDRK4 avoiding additional CFL restrictions by treating diffusion and hyperdiffusion exponentially. We employ the Kassam-Trefethen [30] contour integration in the complex plane for E_i and the SETDRK4 method 3.12, equispaced snapshots of the solution for $t \in [0, 80]$ are plotted in fig. 4.10 where the stochastic perturbation destabilises the deterministic symmetric flow pattern, whilst keeping ensemble members with the same energy and integral of vorticity dissipation as the deterministic system. This example indicates SETDRK with contour integration is adaptable to higher dimensional problems.

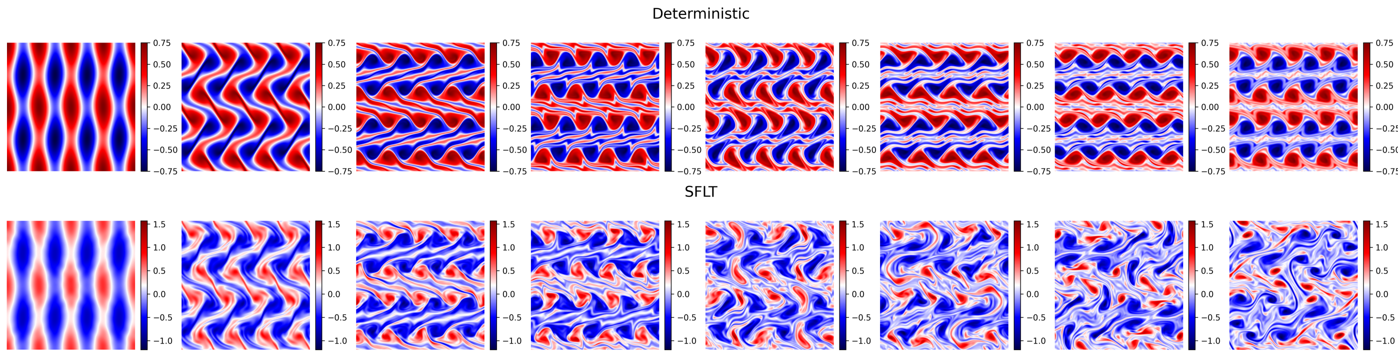


Figure 4.10: Comparison between Deterministic NS and one realisation of a SFLT perturbed NS ensemble.

5. Conclusion

Theoretically we prove and numerically demonstrate that deterministic IFRK–ETDRK–RK integrators can be modified to converge to a Stratonovich SDE. For general non-commutative noise, the methods achieve mean-square order $\mathcal{O}(\Delta t^{1/2})$ (theorem 2.2-item 1). In the commutative noise case, the order improves to $\mathcal{O}(\Delta t^1)$ (theorem 2.2-item 6), while in the presence of drift commutativity the order further increases to $\mathcal{O}(\Delta t^{\mathrm{int}(p/2)})$ (theorem 2.2-item 3–item 4). Finally, in the limit of small noise, large timestep Δt, or absence of noise, deterministic order $\mathcal{O}(\Delta t^p)$ is recovered (theorem 2.2-item 5).

In comparing SETDRK, SIFRK, and SRK schemes of the same deterministic order, we found that SETDRK and SIFRK methods were able to take timesteps typically orders of magnitude larger than those permitted by SRK methods. Among these approaches, SETDRK methods were generally more accurate than SIFRK methods. Furthermore, the additional overhead associated with exponential operator evaluations in both SETDRK and SIFRK methods was negligible, since these terms were precomputed prior to runtime, making their cost justifiable when compared to SRK methods of the same order.

In comparing SRK, SIFRK, and SETDRK methods of different deterministic orders, our experiments demonstrate that higher-order methods provide clear benefits at large timesteps, for small noise amplitudes, and under commutativity conditions. For non-commutative noise where second-order methods could be theoretically expected to be more cost-effective in the limit $\Delta t \to 0$ due to the dominant missing Lévy area term and reduced stage cost, we numerically observed higher-order methods achieving better accuracy per unit CPU cost. These results indicate that higher-order SRK, SIFRK, and SETDRK schemes remain competitive and effective in practical applications where timesteps cannot be taken arbitrarily small, or where the noise is small.

Several open directions follow from this work. The development of generalised additive SGARK, SIFGARK, and SETDGARK schemes, allowing separate treatment of drift and diffusion within the generalised additive Runge–Kutta framework [47, 46], is a promising avenue for improving efficiency. Particularly, detailed studies using splitting methods could be used to investigate the importance of capturing the symmetric terms in the Stratonovich-Taylor expansion as opposed to high order deterministic terms. Extensions of strong stability preserving (SSP) theory to stochastic integrating factor methods (SSP-SIFRK) and to stochastic exponential time-differencing Runge–Kutta methods (SSP-SETDRK) remain to be established following [27, 43, 55]. Stochastic analogues of A-stability [31] have not yet been developed for some of the methods introduced including the SETDRK4 scheme. The consequences of stage order reduction, well understood for deterministic ETDRK methods [49, 34, 24], requires investigation in the stochastic context.

While our convergence analysis focused on the stochastic Korteweg–de Vries equation, the methodology extends naturally to other stiff SPDEs, including the Kuramoto–Sivashinsky, Heat, Burgers, Navier–Stokes, Complex Ginzburg–Landau, and Quasi-geostrophic equations under stochastic transport and forcing. Providing an opportunity for easily implementable stochastic integration for ensemble forecasting and uncertainty quantification methodology.

6. Data availability

The figures, data, and code used in the creation of this document can be found in the following library https://github.com/jameswoodfield/Particle_Filter.

Acknowledgements

Grateful to R. Hu, for valuable comments leading to the improvement of this manuscript, and enjoyable discussions with D. Crisan, D. Holm, O. Street, M. Singh and E. Fausti. Both J. Woodfield and A. Lobbe have been supported during the present work by the European Research Council (ERC) Synergy grant "Stochastic Transport in Upper Ocean Dynamics" (STUOD) – DLV-856408.

References

[1] M. Ahmat and J. Qiu. Compact ETDRK scheme for nonlinear dispersive wave equations. *Computational and Applied Mathematics*, 40(8):286, 2021.

[2] R. Anton and D. Cohen. Exponential integrators for stochastic Schrödinger equations driven by Itô noise. *Journal of Computational Mathematics*, pages 276–309, 2018.

[3] F. Baudoin. *An introduction to the geometry of stochastic flows.* World Scientific, 2004.

[4] S. Becker, A. Jentzen, and P. E. Kloeden. An exponential Wagner–Platen type scheme for SPDEs. *SIAM Journal on Numerical Analysis*, 54(4):2389–2426, 2016.

[5] C. Besse, G. Dujardin, and I. Lacroix-Violet. High Order Exponential Integrators for Nonlinear Schrödinger Equations with Application to Rotating Bose–Einstein Condensates. *SIAM Journal on Numerical Analysis*, 55(3):1387–1411, 2017.

[6] G. Beylkin, J. M. Keiser, and L. Vozovoi. A new class of time discretization schemes for the solution of nonlinear PDEs. *Journal of computational physics*, 147(2):362–387, 1998.

[7] A. Bhatt and B. E. Moore. Structure-preserving exponential Runge–Kutta methods. *SIAM Journal on Scientific Computing*, 39(2):A593–A612, 2017.

[8] H. P. Bhatt and A. Khaliq. A compact fourth-order L-stable scheme for reaction–diffusion systems with nonsmooth data. *Journal of Computational and Applied Mathematics*, 299:176–193, 2016.

[9] K. Burrage and P. Burrage. General order conditions for stochastic runge-kutta methods for both commuting and non-commuting stochastic ordinary differential equation systems. *Applied Numerical Mathematics*, 28(2-4):161–177, 1998.

[10] K. Burrage and P. M. Burrage. High strong order explicit runge-kutta methods for stochastic ordinary differential equations. *Applied Numerical Mathematics*, 22(1-3):81–101, 1996.

[11] K. Burrage and P. M. Burrage. Order conditions of stochastic Runge–Kutta methods by B-series. *SIAM Journal on Numerical Analysis*, 38(5):1626–1646, 2000.

[12] F. Castell. Asymptotic expansion of stochastic flows. *Probability theory and related fields*, 96(2):225–239, 1993.

[13] K.-T. Chen. Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. *Annals of Mathematics*, 65(1):163–178, 1957.

[14] M.-H. Chen, Y. Wu, X. Feng, and S. Sun. Parallelization of the Exponential Time Differencing Method for Solving Diffuse-Interface Models. *Journal of Scientific Computing*, 104(3):1–26, 2025.

[15] C. Clancy and J. A. Pudykiewicz. On the use of exponential time integration methods in atmospheric models. *Tellus A: Dynamic Meteorology and Oceanography*, 65(1):20898, 2013.

[16] S. M. Cox and P. C. Matthews. Exponential time differencing for stiff systems. *Journal of Computational Physics*, 176(2):430–455, 2002.

[17] H. Dai, C. Wang, and Q. Huang. Exponential time differencing-Pad'e finite element method for nonlinear convection-diffusion-reaction equations with time constant delay. *Journal of computational mathematics*, 41(3), 2023.

[18] K. Debrabant, A. Kværnø, and N. C. Mattsson. Runge–Kutta Lawson schemes for stochastic differential equations. *BIT Numerical Mathematics*, 61(2):381–409, 2021.

[19] U. Erdoğan and G. J. Lord. A new class of exponential integrators for SDEs with multiplicative noise. *IMA Journal of Numerical Analysis*, 39(2):820–846, 2019.

[20] Z. Fu, J. Shen, and J. Yang. Higher-order energy-decreasing exponential time differencing Runge-Kutta methods for gradient flows. *Science China Mathematics*, pages 1–20, 2024.

[21] J. Gaines. A basis for iterated stochastic integrals. *Mathematics and Computers in Simulation*, 38(1-3):7–11, 1995.

[22] S. Geiger, G. Lord, and A. Tambue. Exponential time integrators for stochastic partial differential equations in 3D reservoir simulation. *Computational Geosciences*, 16:323–334, 2012.

[23] N. J. Higham. *Accuracy and stability of numerical algorithms*. SIAM, 2002.

[24] M. Hochbruck and A. Ostermann. Exponential integrators. *Acta Numerica*, 19:209–286, 2010.

[25] D. D. Holm and R. Hu. Stochastic effects of waves on currents in the ocean mixed layer. *Journal of Mathematical Physics*, 62(7), 2021.

[26] A. Iserles. *A first course in the numerical analysis of differential equations*. Cambridge university press, 2009.

[27] L. Isherwood, Z. J. Grant, and S. Gottlieb. Strong stability preserving integrating factor Runge–Kutta methods. *SIAM Journal on Numerical Analysis*, 56(6):3276–3307, 2018.

[28] M. Jardak, I. Navon, and M. Zupanski. Comparison of sequential data assimilation methods for the Kuramoto–Sivashinsky equation. *International journal for numerical methods in fluids*, 62(4):374–402, 2010.

[29] A. Jentzen and P. E. Kloeden. Overcoming the order barrier in the numerical approximation of stochastic partial differential equations with additive space–time noise. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 465(2102):649–667, 2009.

[30] A.-K. Kassam and L. N. Trefethen. Fourth-order time-stepping for stiff PDEs. *SIAM Journal on Scientific Computing*, 26(4):1214–1233, 2005.

[31] P. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Stochastic Modelling and Applied Probability. Springer Berlin Heidelberg, 1991.

[32] Y. Komori and K. Burrage. A stochastic exponential Euler scheme for simulation of stiff biochemical reaction systems. *BIT Numerical Mathematics*, 54:1067–1085, 2014.

[33] Y. Komori, D. Cohen, and K. Burrage. Weak second order explicit exponential Runge–Kutta methods for stochastic differential equations. *SIAM Journal on Scientific Computing*, 39(6):A2857–A2878, 2017.

[34] S. Krogstad. Generalized integrating factor methods for stiff PDEs. *Journal of Computational Physics*, 203(1):72–88, 2005.

[35] H. Kunita. On the representation of solutions of stochastic differential equations. In *Séminaire de Probabilités XIV 1978/79*, pages 282–304. Springer, 1978.

[36] J. D. Lawson. Generalized Runge-Kutta processes for stable systems with large Lipschitz constants. *SIAM Journal on Numerical Analysis*, 4(3):372–380, 1967.

[37] X. Liu and C. Li. Discretization of stochastic differential equations by the product expansion for the Chen series. *Stochastics: An International Journal of Probability and Stochastic Processes*, 60(1-2):23–40, 1997.

[38] G. J. Lord, C. E. Powell, and T. Shardlow. *An introduction to computational stochastic PDEs*, volume 50. Cambridge University Press, 2014.

[39] G. J. Lord and A. Tambue. Stochastic exponential integrators for the finite element discretization of SPDEs for multiplicative and additive noise. *IMA Journal of Numerical Analysis*, 33(2):515–543, 2013.

[40] G. N. Milstein and M. V. Tretyakov. *Stochastic numerics for mathematical physics*, volume 39. Springer, 2004.

[41] Q. Nie, Y.-T. Zhang, and R. Zhao. Efficient semi-implicit schemes for stiff systems. *Journal of Computational Physics*, 214(2):521–537, 2006.

[42] J. Niegemann, L. Tkeshelashvili, and K. Busch. Higher-order time-domain simulations of Maxwell's equations using Krylov-subspace methods. *Journal of Computational and Theoretical Nanoscience*, 4(3):627–634, 2007.

[43] M. Qin, Q. Zhai, and R. Zhang. Energy stability and maximum principle of skeletal finite element method for allen–cahn equation with exponential time differencing schemes. *Journal of Scientific Computing*, 105(1):1–29, 2025.

[44] R. Ree. Lie elements and an algebra associated with shuffles. *Annals of Mathematics*, 68(2):210–220, 1958.

[45] C. Reutenauer. Free lie algebras. In *Handbook of algebra*, volume 3, pages 887–903. Elsevier, 2003.

[46] W. Rüemelin. Numerical treatment of stochastic differential equations. *SIAM Journal on Numerical Analysis*, 19(3):604–613, 1982.

[47] A. Sandu and M. Günther. A generalized-structure approach to additive runge–kutta methods. *SIAM Journal on Numerical Analysis*, 53(1):17–42, 2015.

[48] J.-P. Serre. *Lie algebras and Lie groups: 1964 lectures given at Harvard University*. Springer, 2009.

[49] K. Strehmel and R. Weiner. B-convergence results for linearly implicit one step methods. *BIT Numerical Mathematics*, 27(2):264–281, 1987.

[50] R. S. Strichartz. The Campbell-Baker-Hausdorff-Dynkin formula and solutions of differential equations. *Journal of Functional Analysis*, 72(2):320–345, 1987.

[51] J. Sun, H. Zhang, X. Qian, and S. Song. A family of structure-preserving exponential time differencing Runge–Kutta schemes for the viscous Cahn–Hilliard equation. *Journal of Computational Physics*, 492:112414, 2023.

[52] C. Ta, D. Wang, and Q. Nie. An integration factor method for stochastic and stiff reaction–diffusion systems. *Journal of computational physics*, 295:505–522, 2015.

[53] C. von Hallern, R. Missfeldt, and A. Rössler. An exponential stochastic Runge–Kutta type method of order up to 1.5 for SPDEs of Nemytskii-type. *IMA Journal of Numerical Analysis*, page drae064, 2024.

[54] H. Wang, J. Sun, H. Zhang, and X. Qian. A novel up to fourth-order equilibria-preserving and energy-stable exponential Runge-Kutta framework for gradient flows. *CSIAM Trans. Appl. Math.*, 6(1):106–147, 2025.

[55] J. Woodfield. Strong stability preservation for stochastic partial differential equations. *arXiv preprint arXiv:2411.11172*, 2024.

[56] G. Yang, K. Burrage, Y. Komori, and X. Ding. A new class of structure-preserving stochastic exponential Runge-Kutta integrators for stochastic differential equations. *BIT Numerical Mathematics*, 62(4):1591–1623, 2022.

[57] M. Yousuf and A. Khaliq. An efficient ETD method for pricing American options under stochastic volatility with nonsmooth payoffs. *Numerical Methods for Partial Differential Equations*, 29(6):1864–1880, 2013.

[58] H. Zhang and W. Haifeng. Characterizing the stabilization size of a third-order one-parameter ETDRK scheme for the Swift–Hohenberg equation. 2024.

Appendix .1. Symmetric part of nested Stratonovich integration

Assume the following inductive hypothesis

$$\mathrm{Sym}\left(J_{\boldsymbol{j}}\right) = \frac{1}{n!}\sum_{\sigma\in S_n} J_{j_{\sigma(1)},\dots,j_{\sigma(n)}} = \frac{1}{n!} J_{j_1} J_{j_2} \dots J_{j_n}, \quad n = 1, 2, \dots N. \tag{.1}$$

We shall recall the shuffle product formula for nested integration:

$$J_{j_1,\dots j_m} J_{k_1\dots k_n} = \sum_{\sigma\in \mathrm{Sh}(m,n)} J_{\sigma(j_1,\dots,j_m;k_1,\dots,k_n)} \tag{.2}$$

where $\mathrm{Sh}(m,n)$ denotes the set of shuffles length $m+n$, preserving the order of $j_1,\dots j_m$ and $k_1,\dots,k_n$. A subcase of this formula gives the following identity:

$$J_{j_1,\dots j_n} J_{j_{n+1}} = \sum_{\sigma\in \mathrm{Sh}(n,1)} J_{\sigma(j_1,\dots,j_n;j_{n+1})} = \sum_{l=0}^{n} J_{j_1,\dots,j_l,j_{n+1},j_{l+1},\dots,j_n}. \tag{.3}$$

$n=1$ is true trivially, and eq. (.3) can be used with $n=2$ to give $J_{j_1}J_{j_2} = J_{j_1,j_2} + J_{j_2,j_1}$ then in the case $n=3$ one can use eq. (.3) recursively,

$$J_{j_1}J_{j_2}J_{j_3} = J_{j_1,j_2}J_{j_3} + J_{j_2,j_1}J_{j_3} = \sum_{\sigma\in S_3} J_{j_{\sigma(1)},j_{\sigma(2)},j_{\sigma(3)}}. \tag{.4}$$

Assume eq. (.1) up to finite N as an inductive hypothesis, using eq. (.3) we have

$$J_{j_1}\dots J_{j_N}J_{j_{N+1}} = \left(\sum_{\sigma\in S_N} J_{j_{\sigma(1)},\dots,j_{\sigma(N)}}\right) J_{j_{N+1}} = \sum_{\sigma\in S_N}\left(J_{j_{N+1},j_{\sigma(1)},\dots,j_{\sigma(N)}} + \dots + J_{j_{\sigma(1)},\dots,j_{\sigma(N)},j_{N+1}}\right) \tag{.5}$$

$$= \sum_{\sigma\in S_{n+1}} J_{j_{\sigma(1)},\dots,j_{\sigma(N+1)}} \tag{.6}$$

upon dividing by $(N+1)!$, the inductive hypothesis is proven for $n+1$.

Appendix .2. Derivation of stochastic travelling wave solutions to the KdV equation

By Itô's formula and the Chain rule, one can compute

$$d_t u = d_t u(t, X(x,t)) = \partial_t u + \partial_X u d_t X = \partial_t u - a u_x dW. \tag{.7}$$

This allows eq. (4.3) to be transformed into the deterministic KdV equation

$$\partial_t u + u u_x + u_{xxx} = 0. \tag{.8}$$

Seeking a travelling wave solution of the form

$$u = f(\xi) \quad \text{where} \quad \xi = x - \beta t. \tag{.9}$$

We can deduce that the stochastic travelling wave solution of eq. (.8) satisfies the ODE

$$-\beta f'(\xi) + f(\xi) f'(\xi) + f^{'''}(\xi) = 0. \tag{.10}$$

Assuming vanishing contribution of u, u_x, u_{xx} at positive and negative infinity, we can derive the travelling wave solutions of eq. (.10) by integrating w.r.t. ξ and then multiplying by $f'(\xi)$ and then integrating again w.r.t. ξ. using vanishing contributions at the endpoints so constants of integration disappear. Then using the separation of variables one attains

$$\int_0^F \frac{1}{f\sqrt{\beta - \frac{f}{3}}} df = \int_0^\xi d\xi \tag{.11}$$

upon letting $f = 3\beta \operatorname{sech}^2(g)$, $df = -6\beta \tanh(g) \operatorname{sech}^2(g) dg$ one can derive the travelling wave solution eq. (4.4) to the stochastic KdV equation eq. (4.3).

Appendix .3. Commutator relationships

This section is aimed at demonstrating that non-symmetric parts of a Stratonovich-Taylor expansion can be expressed in terms of nested commutator relationships of g_i explicitly. Firstly for $P = 2$,

$$\sum_{0 \le j_1, j_2 \le m} L^{j_1} g^k_{j_2} \operatorname{Alt}(J_{j_1,j_2}) = \frac{1}{2} \sum_{0 \le j_1, j_2 \le m} \left(L^{j_1} g^{k,j_2} - L^{j_2} g^{k,j_1} \right) J_{j_1,j_2} = \frac{1}{2} \sum_{0 \le j_1, j_2 \le m} \left([g_{j_1}, g_{j_2}]\right)^k J_{j_1,j_2}. \tag{.12}$$

For $P = 3$, consider the fully antisymetric term and use the alternating tensor identity and linearity of scalar multiplication gives

$$\begin{aligned}
\sum_{j_1,j_2,j_3} L^{j_1} L^{j_2} g^k_{j_3} \operatorname{Alt}(J_{j_1,j_2,j_3}) &= \frac{1}{3!} \sum_{j_1,j_2,j_3} L^{j_1} L^{j_2} g^k_{j_3} \sum_{\sigma \in S_3} \operatorname{sign}(\sigma) J_{j_{\sigma(1)}, j_{\sigma(2)}, j_{\sigma(3)}} && (.13)\\
&= \frac{1}{3!} \sum_{j_1,j_2,j_3} \sum_{\sigma \in S_3} \operatorname{sign}(\sigma) L^{j_1} L^{j_2} g^k_{j_3} J_{j_{\sigma(1)}, j_{\sigma(2)}, j_{\sigma(3)}} && (.14)\\
&= \frac{1}{3!} \sum_{\sigma \in S_3} \operatorname{sign}(\sigma) \sum_{j_1,j_2,j_3} L^{j_1} L^{j_2} g^k_{j_3} J_{j_{\sigma(1)}, j_{\sigma(2)}, j_{\sigma(3)}} && (.15)\\
&= \frac{1}{3!} \sum_{\sigma \in S_3} \operatorname{sign}(\sigma^{-1}) \sum_{i_1,i_2,i_3} L^{i_{\sigma^{-1}(1)}} L^{i_{\sigma^{-1}(2)}} g^k_{i_{\sigma^{-1}(3)}} J_{i_1,i_2,i_3} && (.16)\\
&= \sum_{i_1,i_2,i_3} \operatorname{Alt}(L^{i_1} L^{i_2} g^k_{i_3}) J_{i_1 i_2 i_3} && (.17)
\end{aligned}$$

Where in eq. (.16), we fix the permutation σ and reindex $(i_1, i_2, i_3) = (j_{\sigma(1)}, j_{\sigma(2)}, j_{\sigma(3)})$, such that $i_{\sigma^{-1}(1)}, i_{\sigma^{-1}(2)}, i_{\sigma^{-1}(3)} = j_1, j_2, j_3$, and noting $\operatorname{sign}(\sigma) = \operatorname{sign}(\sigma^{-1})$. Finally by noting that the 3! permutations of S_3 can be paired sharing the last digit $((1,2,3),(2,1,3)), ((1,3,2),(3,1,2)), ((2,3,1),(3,2,1))$ with opposite signed parity to form a sum of $3!/2 = 3$ nested commutator relationships to give

$$\sum_{j_1,j_2,j_3} L^{j_1}L^{j_2}g^k_{j_3}\,\mathrm{Alt}(J_{j_1,j_2,j_3}) = \frac{1}{3!}\sum_{j_1,j_2,j_3}\left([g_{j_1},[g_{j_2},g_{j_3}]]\right)^k J_{j_1,j_2,j_3}. \tag{.18}$$

A rank 3 tensor does not solely decompose into symetric and antisymetric parts, we rely on the particular deconstruction below:

$$J_{i,j,k} = \mathrm{Sym}(J_{i,j,k}) + \mathrm{Alt}(J_{i,j,k}) + N^1_{i,j,k} + N^2_{i,j,k} \tag{.19}$$

Where the terms in the decomposition are

$$\mathrm{Sym}(J_{i,j,k}) = \frac{1}{3!}\left(J_{i,j,k} + J_{j,k,i} + J_{k,j,i} + J_{i,k,j} + J_{j,i,k} + J_{k,i,j}\right) \tag{.20}$$

$$\mathrm{Alt}(J_{i,j,k}) = \frac{1}{3!}\left(J_{i,j,k} + J_{j,k,i} + J_{k,j,i} - J_{i,k,j} - J_{j,i,k} - J_{k,i,j}\right) \tag{.21}$$

$$N^1_{i,j,k} = \frac{1}{3}\left(J_{i,j,k} - J_{i,k,j} + J_{j,i,k} - J_{k,i,j}\right) \tag{.22}$$

$$N^2_{i,j,k} = \frac{1}{3}\left(J_{i,j,k} - J_{j,i,k} + J_{i,k,j} - J_{j,k,i}\right) \tag{.23}$$

The decomposition of a rank 3 tensor is not unique, the above representation was chosen for the following antisymmetric properties on the decomposition of the remainder

$$N^1_{i,j,k} = -N^1_{i,k,j}, \quad N^2_{i,j,k} = -N^1_{j,i,k}. \tag{.24}$$

These properties allow the following rearrangement in terms of commutators:

$$\sum_{j_1,j_2,j_3} L^{j_1}L^{j_2}g^k_{j_3}N^1_{j_1,j_2,j_3} = \frac{1}{2}\sum_{j_1,j_2,j_3} L^{j_1}\,[g_{j_2},g_{j_3}]\,J_{j_1,j_2,j_3} \tag{.25}$$

$$\sum_{j_1,j_2,j_3} L^{j_1}L^{j_2}g^k_{j_3}N^2_{j_1,j_2,j_3} = \frac{1}{2}\sum_{j_1,j_2,j_3} [L^{j_1},L^{j_2}]g^k_{j_3}J_{j_1,j_2,j_3} \tag{.26}$$

Therefore at order 3 we have

$$\sum_{j_1,j_2,j_3\in\{0,\dots,m\}^3} L^{j_1}L^{j_2}g^k_{j_3}J_{j_1,j_2,j_3} = \frac{1}{3!}\sum_{j_1,j_2,j_3}\left([g_{j_1},[g_{j_2},g_{j_3}]]\right)^k J_{j_1,j_2,j_3} + \frac{1}{2}\sum_{j_1,j_2,j_3} L^{j_1}\,[g_{j_2},g_{j_3}]^k\,J_{j_1,j_2,j_3} \tag{.27}$$

$$+\frac{1}{2}\sum_{j_1,j_2,j_3}[L^{j_1},L^{j_2}]g^k_{j_3}J_{j_1,j_2,j_3} + \sum_{j_1,j_2,j_3\in\{0,\dots,m\}^3} L^{j_1}L^{j_2}g^k_{j_3}\,\mathrm{Sym}(J_{j_1,j_2,j_3}). \tag{.28}$$

Appendix .4. SETDRK methods for Itô systems, introduction and literature

We review Itô-SETDRK schemes, in the simplified setting of one-dimensional Itô-SDEs,

$$d_t u = [N(u,t) + Lu]\,dt + [g(u,t)]\,dW_t. \tag{.29}$$

Where u is a one dimensional stochastic process, $N(u,t)$ is a nonlinear function, L is a linear function, and $g(u,t)$ is a nonlinear function, dW_t is a one dimensional Brownian motion. By multiplying eq. (.29) through by the integrating factor e^{-Lt}, and integrating between t^n and t^{n+1} one may write eq. (.29) in integral form as follows

$$u^{n+1} = e^{L\Delta t}u^n + e^{L\Delta t}\int_0^{\Delta t} e^{-Ls}N(u(t_n+s),t_n+s)ds + e^{L(\Delta t + t^n)}\int_{t^n}^{t^{n+1}} e^{-Ls}g(u(s),s)dW_s. \tag{.30}$$

where $\Delta t = t^{n+1} - t^n$. Stochastic exponential time differencing methodology arises when one considers approximating the above integral representation of the SDE.

Method Appendix .1 (SIFEM- Stochastic integrating factor Euler Maruyama)**.** Approximating both the integrands in eq. (.30) by evaluating the integrands at the leftmost temporal point t_n, leads to the following scheme

$$u^{n+1} = e^{L\Delta t}(u^n + \Delta t N(u_n, t_n) + g(u_n, t_n)\Delta W). \tag{.31}$$

this converges to the Itô SDE and treats the linear operator with an exponential operator. This scheme is refereed to as SETDM0 in [39]. Whilst this scheme does arise from approximating the integrals in eq. (.30) one can also naturally derive this scheme as a stochastic integrating factor Euler Maruyama scheme SIFEM.

Method Appendix .2 (SETDM10- Drift Exponential Time Differencing)**.** Another approach following [16], approximates N in eq. (.30) by a constant $N_n = N(u(t_n), t_n)$ and integrates $e^{-Ls}ds$ exactly to give the following scheme

$$u^{n+1} = e^{L\Delta t}u^n + N_n(e^{L\Delta t} - 1)L^{-1} + e^{L\Delta t}g_n\Delta W. \tag{.32}$$

This can be found under the name SETDM1 scheme in [39].

Method Appendix .3 (SETDM01-Drift-Diffusion Stochastic exponential time differencing)**.** In an analogous manner to the deterministic setting, if one instead approximates $g = g_n$ by a constant, one cannot directly compute the stochastic integral but instead it's distribution can be attained through Itô's Isometry

$$\int_{t^n}^{t^{n+1}} e^{-Ls}dW \sim \mathcal{N}\left(0, \int_{t^n}^{t^{n+1}} \left(e^{-Ls}\right)^2 ds\right).$$

Sampling from this distribution leads to the following first order Stochastic Exponential time differencing scheme:

$$u^{n+1} = e^{L\Delta t}u^n + N_n(e^{L\Delta t} - 1)L^{-1} + g_n\sqrt{\frac{1}{2}\left(e^{2L\Delta t} - 1\right)L^{-1}}\Delta Z, \quad \Delta Z \sim \mathcal{N}(0,1) \tag{.33}$$

Such schemes can be found for example in [28, 29].
Consider the Taylor expansion of the exponential operators in method Appendix .1, method Appendix .2, method Appendix .3, these are respectively

$$e^{L\Delta t} = (1 + L\Delta t + (L\Delta t)^2 + ...) \tag{.34}$$

$$(e^{L\Delta t} - 1)L^{-1} = \Delta t + L\Delta t^2 - ... \tag{.35}$$

$$\sqrt{(2L)^{-1}\left(e^{2L\Delta t} - 1\right)} = \sqrt{\Delta t + L\Delta t^2 + ...} \tag{.36}$$

and allows all schemes to be convergent to the Itô system and at leading order, equal to the Euler-Maruyama scheme. Therefore, with this as motivation, if one finds another perhaps more convenient exponential operator, agreeing at leading order to the desired SDE, one can derive a new method. Consider modifying method Appendix .3, with the following transformation

$$\sqrt{\frac{1}{2L}\left(e^{2L(\Delta t)} - 1\right)} \mapsto \frac{e^{L\Delta t} - 1}{L\sqrt{\Delta t}}, \tag{.37}$$

these expressions agree at leading order and this results in the following method Appendix .4.

Method Appendix .4 (Convenient stochastic exponential time differencing Itô)**.**

$$u^{n+1} = e^{L\Delta t}u^n + N_n\frac{e^{L\Delta t} - 1}{L} + g_n\frac{e^{L\Delta t} - 1}{L\Delta t}\Delta W, \tag{.38}$$

The motivation is that this can be rearranged into

$$u^{n+1} = e^{L\Delta t}u^n + \frac{e^{L\Delta t} - 1}{L}\left(N_n + g_n\frac{\Delta W}{\Delta t}\right), \tag{.39}$$

requires little change to any deterministic ETDRK1 codes, whilst converges to the Itô SPDE with strong order 1/2, whilst benefiting from the exponential treatment of the linear operator. This scheme is found in [32] as the stochastic exponential Euler (SEE) scheme, derived under different considerations. This observation also leads to the ability to generate SETDRK methods such as those in this paper.

Appendix .5. Assumption

Equation (2.44) makes an assumption on the deterministic order p one-step map, namely that the map $\Delta t f \mapsto \Delta t f + g_m \Delta W^m$ gives a stochastic term of order $\mathcal{O}(\Delta t^{\frac{p+1}{2}})$, this is true for methods associated with B-series expansions i.e. RK, IFRK, ETDRK methods.

For example, a deterministic RK scheme has a leading order error at $\mathcal{O}(\Delta t^{p+1})$ which can be written as a linear combination of $p+1$ elementary differentials of f. This allows the error to be written as a linear combination of $p+1$ elementary differentials of $\Delta t f$ such that the subsequent transformation $f\Delta t \mapsto f\Delta t + g_m \Delta W^m$, gives leading order truncation error $\mathcal{O}(\Delta t^{\frac{p+1}{2}})$. For instance the 2nd order Heun-method, takes the local truncation error

$$err = \frac{\Delta t^3}{12}\left(f''(y)[f(y), f(y)] - f'(y)^2 f(y)\right). \tag{.40}$$

Which under $f\Delta t \mapsto f\Delta t + g_p \Delta W^p$ reveals an $\mathcal{O}(\Delta t^{3/2})$ error term.