Journal of Computational Physics (2020)

SAND2020-13048J

Contents lists available at ScienceDirect

Journal of Computational Physics

journal homepage: www.elsevier.com/locate/jcp

Modeling fast diffusion processes in time integration of stiff stochastic differential equations

Xiaoying Han[a,*], Habib N. Najm[b,1]

[a]*Department of Mathematics and Statistics, 221 Parker Hall, Auburn University, Auburn, AL 36849, USA*
[b]*Sandia National Laboratories, P.O.Box 969, MS 9051, Livermore, CA 94551, USA*

ARTICLE INFO

Article history:

2000 MSC: 60H10, 60H35, 60C20, 65C05, 65C30.

Keywords: Stiff stochastic differential equation, fast diffusion, linear diffusion approximation, mean-square convergence, weak convergence

ABSTRACT

Numerical algorithms for stiff stochastic differential equations are developed using linear approximations of the fast diffusion processes, under the assumption of decoupling between fast and slow processes. Three numerical schemes are proposed, all of which are based on the linearized formulation albeit with different degrees of approximation. The schemes are of comparable complexity to the classical explicit Euler-Maruyama scheme, but can achieve better accuracy at larger time steps in stiff systems. Convergence analysis is conducted for one of the schemes, that shows it to have a strong convergence order of 1/2 and a weak convergence order of 1. Approximations arriving at the other two schemes are discussed. Numerical experiments are carried out to examine the convergence of the schemes proposed on model problems.

1. Introduction

Stiff stochastic differential equations (SDEs) are relevant in the mathematical modeling of a wide range of physical systems, where multiscale stochastic processes, involving a large range of time-scales, enter in the formulation of the system governing equations. Our work is motivated by the challenges in time integration of stiff stochastic chemical systems. These systems can in principle be modeled in the discrete stochastic context by simulating pathways of the chemical master equation (CME) using stochastic methods [1], including the stochastic simulation algorithm (SSA) and its variants [2–5]. The SSA has been used extensively in this regard, with various developments for dealing with stiffness [6–11]. These methods employ a range of strategies often involving identification of fast and slow reaction

*Corresponding author: Tel.: +1-334-844-5884; fax: +1-334-844-6555;
e-mail: xzh0003@auburn.edu (Xiaoying Han)
[1]hnnajm@sandia.gov

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

subsets [11–16] with targeted approximations for handling each subset, including the identification of state-vector subsets dominated by fast/slow processes [15, 17]. These approaches rely on the observation that the relaxation time for fast processes is small relative to the slow time scales that dominate the time rate of change of the state, thereby invoking a stochastic quasi-steady state approximation (QSSA) [17–19], or allowing some other approximate means of approximation of the fast processes [12, 14, 16, 20]. Often, approximate modeling of fast reaction processes [12, 13] is done by taking the CME to the continuous stochastic limit, namely the chemical Langevin equation (CLE) [21–24], an SDE.

Our focus here is specifically in this continuous stochastic context. The CLE typically retains a significant degree of stiffness, resulting in associated computational challenges which have been the subject of ongoing research. Sotiropoulos *et al.* [25] outline a semianalytical method that relies on transformation of the SDE into two subsystems describing the evolution of distinct slow and fast varying variables, following [26]. This is followed by a decoupling step, relying on the fast-slow structure of the dynamics, which allows separate approximation of the distributions of the fast and slow variables. This enables the construction of a system of slow CLEs that can be integrated with large time steps, before reconstructing the full solution. Similarly, Thomas *et al.* [27] present a method that relies on the identification of fast and slow reactions, using forward and inverse Fourier transforms to apply approximations of the fast processes, and employ separation of the state vector into slow and fast variables. In our own earlier work with the CLE [28], we outlined an automated procedure that does not presume an *a priori* distinction between fast/slow reactions or variables, relying rather on dynamical analysis using computational singular perturbation (CSP) [29–51] to decouple fast/slow processes and provide stable large time-step time integration of the full state vector as driven by the drift terms. This is possible when appropriate statistics can be employed to judge the exhaustion of the fast dynamics of the mean-state towards an underlying slow manifold. This method, however, does not deal with the stiffness of the diffusion source terms, and their associated slow/fast structure. That is the focus of the present work.

While we have motivated the work based on stochastic stiff chemical models, highlighting associated challenges and progress, we are concerned here with stiff SDEs in general, and particularly with effective modeling, through local linearization, of fast diffusion processes, allowing efficient explicit time integration of SDEs with stiff diffusion. The paper is organized as follows. First, in Section 2, we provide a brief background on general numerical schemes for stiff SDEs and formulate the specific objectives of this work. Then, in Section 3, we propose a linearized model for handling the fast diffusion processes in stiff SDEs, and develop numerical algorithms for stiff SDEs based on this construction. In Section 4, we present a convergence analysis for one of the schemes proposed in Section 3. Illustrative, numerical experiments are then presented in Section 5, and closing remarks are given in Section 6.

2. Background and objective

There is a broad landscape of theoretical and numerical work with SDEs, covering their mathematical properties and both analytical and numerical solution methods (see, e.g., [52–55]). Numerous numerical methods are available for time integration of SDEs, and their accuracy naturally improves with decreasing time step size, with the rate of convergence of the error depending on the choice of error metric as well as the integration scheme. One of the most

widely used methods is the explicit Euler-Maruyama (EM) scheme, due to its ease of implementation. The classical explicit EM scheme has a strong convergence rate , *i.e.* the rate of convergence of the mean square error in the solution at some time instant T, of $O(\Delta t^{1/2})$ (see, e.g., [56–58]) and a weak convergence rate, *i.e.*, the rate of convergence of the error in moments, such as mean or variance/covariance, of the solution at time T, of $O(\Delta t)$. The EM formulation is briefly outlined below.

Consider a d-dimensional SDE with an m-dimensional Brownian motion

$$\mathrm{d}\boldsymbol{X}(t) = \boldsymbol{f}(\boldsymbol{X}(t))\mathrm{d}t + \sum_{k=1}^{m} \boldsymbol{g}_k(\boldsymbol{X}(t))\mathrm{d}W_k(t), \quad 0 \le t \le T, \quad \boldsymbol{X}(0) = \boldsymbol{x}_0, \tag{2.1}$$

where $\boldsymbol{X}(t) = (X^1(t), \dots, X^d(t))^T \in \mathbb{R}^d$ for every $t \ge 0$, $W_k(t)$ is a scalar Brownian motion for each $k = 1, \cdots, m$, and the functions $\boldsymbol{f} : \mathbb{R}^d \to \mathbb{R}^d$, $\boldsymbol{g} : \mathbb{R}^d \to \mathbb{R}^d$ satisfy standard assumptions for the existence and uniqueness of solutions to the system (2.1) (see, e.g., [52, 53]). Given a stepsize Δt, set $N = T/\Delta t$ and $t_n = n\Delta t$. Then for $n = 0, \cdots, N$ the solution to (2.1) on $[t_n, t_{n+1})$ given $\boldsymbol{X}(t_n) = \boldsymbol{x}_n$ is

$$\boldsymbol{X}(t) = \boldsymbol{x}_n + \int_{t_n}^{t} \boldsymbol{f}(\boldsymbol{X}(s))\mathrm{d}s + \sum_{k=1}^{m} \int_{t_n}^{t} \boldsymbol{g}_k(\boldsymbol{X}(s))\mathrm{d}W_k(s), \quad t \in [t_n, t_{n+1}). \tag{2.2}$$

The simplest one-step EM scheme applied to (2.2) gives an approximation for $\boldsymbol{x}_{n+1} \approx \boldsymbol{X}(t_{n+1})$ as

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \Delta t \boldsymbol{f}(\boldsymbol{x}_n) + \sum_{k=1}^{m} \boldsymbol{g}_k(\boldsymbol{x}_n)\Delta W_{k,n}, \quad \Delta W_{k,n} = W_k(t_{n+1}) - W_k(t_n), \quad n = 1, \cdots, N. \tag{2.3}$$

The focus of this work is computing solutions of the SDE (2.1) with stiffness in the diffusion coefficients. In particular, we are concerned with SDEs of the form (2.1) where the diffusion coefficients $\boldsymbol{g}_k(\boldsymbol{x}_n), k = 1, \dots, m$ exhibit a large range magnitudes with k. Moreover, presuming an ordering where $|\boldsymbol{g}_1| > |\boldsymbol{g}_2| > \cdots > |\boldsymbol{g}_m|$, the ideal context for the present construction is one where there exists $M \in [1, m]$ such that the magnitudes of the first M diffusion coefficients, $\boldsymbol{g}_1, \dots, \boldsymbol{g}_M$ (referred to as the fast diffusion processes), are significantly larger than those of the other diffusion coefficients $\boldsymbol{g}_{M+1}, \dots, \boldsymbol{g}_m$ (referred to as the slow diffusion processes). The EM scheme is still feasible, but requires Δt to be chosen sufficiently small to accommodate the fast variations of the states due to $\boldsymbol{g}_1, \dots, \boldsymbol{g}_M$, because of its slow rate of convergence. Higher order explicit schemes allowing larger stepsizes can be derived iteratively using Itô-Taylor approximations for SDEs (see, e.g., [54]), but most of the time are difficult to implement due to the presence of multiple stochastic integrals. Extensive work has been done towards improving the accuracy or stability of numerical schemes for SDEs with stiffness (see, e.g. [59–70]), among which numerical schemes for stiff SDEs based on the EM scheme were developed in [59–61].

One major drawback of EM, even when it is stable, is the large error in approximating diffusion at larger stepsizes when the magnitude of diffusion term is large. This is due to lower order of discretization error of the diffusion terms $\boldsymbol{g}_k(\boldsymbol{x}_n)\Delta W_{k,n}$. The aim of this work is to develop numerical schemes for the stiff SDE (2.1) that employ linearized modeling of fast diffusion processes, thereby giving more accurate approximations for the diffusion terms of larger magnitudes at larger time steps, without having to use higher order explicit schemes. More precisely, we will develop

schemes by modifying the EM time integration (2.3) to

$$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \Delta t \boldsymbol{f}(\boldsymbol{x}_n) + \sum_{k=M+1}^{m} \boldsymbol{g}_k(\boldsymbol{x}_n)\Delta W_{k,n} + \boldsymbol{\eta}_{n+1}, \quad \Delta W_{k,n} := W_k(t_{n+1}) - W_k(t_n), \tag{2.4}$$

in which $\boldsymbol{\eta}_{n+1}$ is an $\mathbb{R}^d$-valued random variable that approximates $\sum_{k=1}^{M}\int_{t_n}^{t_{n+1}} \boldsymbol{g}_k(\boldsymbol{X}(s))\mathrm{d}W_k(s)$ more accurately than $\sum_{k=1}^{M} \boldsymbol{g}_k(\boldsymbol{x}_n)\Delta W_{k,n}$ as in the EM scheme at larger step size Δt.

Given an $\mathbb{R}^d$-valued stochastic process $\boldsymbol{X}(t)$, for $n = 0, \cdots, N$ define the diffusion process $\mathscr{D}_n^{\boldsymbol{X}}(t)$ by

$$\mathscr{D}_n^{\boldsymbol{X}}(t) = \sum_{k=1}^{M}\int_{t_n}^{t} \boldsymbol{g}_k(\boldsymbol{X}(s))\mathrm{d}W_k(s), \quad t \in [t_n, t_{n+1}).$$

In particular, when $\boldsymbol{X}(t)$ satisfies the equation (2.2), $\mathscr{D}_n^{\boldsymbol{X}}(t)$ represents the fast diffusion processes in (2.2). Then the random variable $\boldsymbol{\eta}_{n+1}$ in (2.4) that we seek, is an approximation of $\mathscr{D}_n^{\boldsymbol{X}}(t_{n+1})$. The idea here is to approximate $\mathscr{D}_n^{\boldsymbol{X}}(t)$ by $\mathscr{D}_n^{\hat{\boldsymbol{X}}}(t)$, in which the $\mathbb{R}^d$-valued process $\hat{\boldsymbol{X}}$ satisfies the diffusion only SDE

$$\mathrm{d}\hat{\boldsymbol{X}}(t) = \sum_{k=1}^{M} \boldsymbol{g}_k(\hat{\boldsymbol{X}}(t))\mathrm{d}W_k(t), \quad \hat{\boldsymbol{X}}(t_n) = \boldsymbol{x}_n, \quad t \in [t_n, t_{n+1}). \tag{2.5}$$

More precisely, we first derive approximations of the solution $\hat{\boldsymbol{X}}(t)$ to the SDE (2.5) by employing local linear approximations of $\boldsymbol{g}_k$. Then we apply the approximations of $\hat{\boldsymbol{X}}(t)$ to develop various models for $\boldsymbol{\eta}_{n+1}$, which are eventually employed in the construction of numerical schemes for the original SDE (2.1).

Note that if $\{\boldsymbol{g}_k(\boldsymbol{X}(t))\}_{k=1,\cdots,M}$ are decoupled from $\{\boldsymbol{g}_k(\boldsymbol{X}(t))\}_{k=M+1,\cdots,m}$ and $\boldsymbol{f}(\boldsymbol{X}(t))$, then approximating $\mathscr{D}_n^{\boldsymbol{X}}(t)$ by $\mathscr{D}_n^{\hat{\boldsymbol{X}}}(t)$ in (2.2) does not introduce an extra error. On the other side, when $\{\boldsymbol{g}_k(\boldsymbol{X}(t))\}_{k=1,\cdots,m}$ and $\{\boldsymbol{g}_k(\boldsymbol{X}(t))\}_{k=M+1,\cdots,m}$ and/or $\boldsymbol{f}(\boldsymbol{X}(t))$ are coupled, such an approximation is associated with an error term $|\mathscr{D}_n^{\boldsymbol{X}}(t) - \mathscr{D}_n^{\hat{\boldsymbol{X}}}(t)|$, which needs to be taken into account when constructing error estimates for any numerical scheme arising from (2.5). In the context of chemical reaction systems, the magnitude of $\boldsymbol{g}_k$ is proportional to the magnitude of the propensity function of the kth reaction. Therefore, the magnitudes of $\{\boldsymbol{g}_k(\boldsymbol{X}(t))\}_{k=1,\cdots,M}$ being much larger than those of $\{\boldsymbol{g}_k(\boldsymbol{X}(t))\}_{k=M+1,\cdots,m}$ implies that at the time instant t, the reactions $1, \cdots, M$ are much faster and thus more likely to occur/"fire" than the reactions $M+1, \cdots, m$ during the next infinitesimal time period. In the well known formulation of chemical Langevin equations via the tau-leaping method [71], this can be interpreted as the reactions $1, \cdots, M$ fires much more frequently than the reactions $M+1, \cdots, m$ during any fixed period of time that fulfills the assumptions for tau-leaping [71]. Therefore, in this time period, reactions $M+1, \cdots, m$ can be regarded as "nearly frozen", thus not contributing significantly to the SDE diffusion term, while the reactions $1, \cdots, M$ are active. Accordingly, the evolution of $\hat{\boldsymbol{X}}(t)$ is only weakly coupled with the evolution of $\boldsymbol{X}(t)$. This assumed decoupling among fast and slow processes on appropriate time-scales is used in various other constructions already cited above, e.g. [27]. In the present construction, in the convergence analysis conducted in Section 4, we show that the contribution of the error term $|\mathscr{D}_n^{\boldsymbol{X}}(t) - \mathscr{D}_n^{\hat{\boldsymbol{X}}}(t)|$ is relatively small compared to the error of the EM scheme applied to $\{\boldsymbol{g}_k(\boldsymbol{X}(t))\}_{k=M+1,\cdots,m}$, when the magnitudes of $\{\boldsymbol{g}_k(\boldsymbol{X}(t))\}_{k=M+1,\cdots,m}$ are much smaller than the magnitudes of $\{\boldsymbol{g}_k(\boldsymbol{X}(t))\}_{k=1,\cdots,M}$.

In the next section , we will solve the diffusion only SDE (2.5) when each of the fast diffusion processes $\{\boldsymbol{g}_k\}_{k=1,\cdots,M}$ is approximated by a linear function of the state. The solutions are then used to model $\boldsymbol{\eta}_{n+1}$ in the modified EM scheme (2.4).

3. Linear models for fast diffusion processes

In this section we first construct approximated solutions to the SDE (2.5) using a local linear approximation of $\boldsymbol{g}_k$ for $k = 1, \cdots, M$ on each interval $[t_n, t_{n+1})$. Then we develop various models for $\boldsymbol{\eta}_{n+1}$, which we refer to as fast processes models (FPMs) based on the approximated solutions of (2.5). In particular, we develop three models for vector-valued SDEs in Subsection 3.1, and we discuss two of them in the context of scalar-valued SDEs in Subection 3.2. Numerical schemes resulting from each of the models, for integrating the original SDE (2.1), are presented at the end of each subsection.

3.1. Vector-valued diffusion only SDE

For $t \in [t_n, t_{n+1})$, set $\boldsymbol{Y}_n(t) = \hat{\boldsymbol{X}}(t) - \boldsymbol{x}_n$, where $\hat{\boldsymbol{X}}(t)$ is the solution to the SDE (2.5). Note that $\boldsymbol{Y}_n(t_{n+1}) = \mathscr{D}_n^{\hat{\boldsymbol{X}}}(t_{n+1})$, and that it will be used to approximate $\mathscr{D}_n^{\boldsymbol{X}}(t_{n+1})$ and model the random variable $\boldsymbol{\eta}_{n+1}$.

3.1.1. Linear approximation at pathwise initial state

Given $\hat{\boldsymbol{X}}(t_n) = \boldsymbol{x}_n$, approximate $\boldsymbol{g}_k(\hat{\boldsymbol{X}}(t))$ at $\boldsymbol{x}_n$ by

$$\boldsymbol{g}_k(\hat{\boldsymbol{X}}(t)) \approx \boldsymbol{b}_{k,n} + J_{k,n}(\hat{\boldsymbol{X}}(t) - \boldsymbol{x}_n), \quad t \in [t_n, t_{n+1}),$$

where $J_{k,n}$ is the Jacobian of $\boldsymbol{g}_k$ evaluated at $\boldsymbol{x}_n$ and $\boldsymbol{b}_{k,n} = \boldsymbol{g}_k(\boldsymbol{x}_n)$. Then $\boldsymbol{Y}_n(t)$ satisfies the linear SDE

$$\mathrm{d}\boldsymbol{Y}_n(t) = \sum_{k=1}^{M} (J_{k,n}\boldsymbol{Y}_n(t) + \boldsymbol{b}_{k,n})\mathrm{d}W_k(t) \text{ for } t \in [t_n, t_{n+1}), \quad \boldsymbol{Y}_n(t_n) = 0. \tag{3.1}$$

The solution to (3.1) reads (see, e.g., [52])

$$\boldsymbol{Y}_n(t) = \Phi_n(t)\left(- \int_{t_n}^{t} \Phi_n^{-1}(s) \sum_{k=1}^{M} J_{k,n}\boldsymbol{b}_{k,n}\mathrm{d}s + \int_{t_n}^{t} \Phi_n^{-1}(s) \sum_{k=1}^{M} \boldsymbol{b}_{k,n}\mathrm{d}W_k(s) \right), \tag{3.2}$$

where $\Phi_n(t) \in \mathbb{R}^{d\times d}$ is the fundamental matrix of the corresponding homogeneous equation, i.e., the solution of the homogeneous SDE

$$\mathrm{d}\Phi_n(t) = \sum_{k=1}^{M} J_{k,n}\Phi_n(t)\mathrm{d}W_k(t), \quad t \in [t_n, t_{n+1}), \quad \Phi_n(t_n) = I_d. \tag{3.3}$$

For this special SDE (3.3) with no drift and $J_{k,n}$ being autonomous on the interval $[t_n, t_{n+1})$, the fundamental matrix can be given explicitly as

$$\Phi_n(t) = \exp\left\{ -\frac{1}{2}\sum_{k=1}^{M} J_{k,n}^2(t - t_n) + \sum_{k=1}^{M} J_{k,n}(W_k(t) - W_k(t_n)) \right\}. \tag{3.4}$$

However, since equation (3.4) involves an exponential of stochastic matrices, computing $\Phi_n(t)$ directly is of high complexity and cost. A simplified formulation for Φ_n can be obtained from a one-step EM applied to the homogeneous SDE (3.3), resulting in

$$\Phi_n(t) \approx I_d + \sum_{k=1}^{M} J_{k,n}(W_k(t) - W_k(t_n)). \tag{3.5}$$

Now applying (3.5) in (3.2), setting $\Phi_n^{-1}(s) \equiv \Phi_n^{-1}(t_n)$ for $s \in [t_n, t)$ and evaluating the integrals from t_n to t_{n+1} gives

$$\boldsymbol{Y}_n(t_{n+1}) \approx \left(I_d + \sum_{k=1}^{M} J_{k,n}\Delta W_{k,n}\right) \sum_{k=1}^{M} \left(-J_{k,n}\boldsymbol{b}_{k,n}\Delta t + \boldsymbol{b}_{k,n}\Delta W_{k,n}\right), \tag{3.6}$$

where $\Delta W_{k,n}$ is defined as in (2.4). The algorithm using $\boldsymbol{\eta}_{n+1} = \boldsymbol{Y}_n(t_{n+1})$ in (3.6), which we refer to as FPM-LP (Linear approximation at Pathwise state), is summarized in Subsection 3.1.4. An upper bound of the error from approximating $\boldsymbol{Y}_n(t)$ in (3.2) at t_{n+1} by (3.6) will be provided in Section 4.

3.1.2. Linear approximation at the mean initial state

Note that $J_{k,n}$ is the Jacobian of $\boldsymbol{g}_k$ evaluated at the "pathwise" initial state $\boldsymbol{x}_n = \boldsymbol{x}_n(\omega)$, and thus is also path dependent, i.e. $J_{k,n} = J_{k,n}(\omega)$ needs to be computed for every sample path $\omega \in \Omega$. The associated computational cost can be reduced by anchoring the linear approximation of $\boldsymbol{g}_k(\hat{\boldsymbol{X}}(t))$ at the path independent mean state $\bar{\boldsymbol{x}}_n := \mathbb{E}[\boldsymbol{x}_n(\omega)]$, instead of the path dependent state $\boldsymbol{x}_n$. More precisely, approximate $\boldsymbol{g}_k(\hat{\boldsymbol{X}}(t))$ at $\bar{\boldsymbol{x}}_n$ by

$$\boldsymbol{g}_k(\hat{\boldsymbol{X}}(t)) \approx \tilde{\boldsymbol{b}}_{k,n} + \tilde{J}_{k,n}(\hat{\boldsymbol{X}}(t) - \bar{\boldsymbol{x}}_n) \text{ with } \bar{\boldsymbol{x}}_n := \mathbb{E}[\boldsymbol{x}_n(\omega)], \quad t \in [t_n, t_{n+1}),$$

where $\tilde{J}_{k,n}$ is the Jacobian of $\boldsymbol{g}_k$ evaluated at $\bar{\boldsymbol{x}}_n$ and $\tilde{\boldsymbol{b}}_{k,n} = \boldsymbol{g}_k(\bar{\boldsymbol{x}}_n)$. If we still define $\boldsymbol{Y}_n(t) = \hat{\boldsymbol{X}}(t) - \boldsymbol{x}_n$ for $t \in [t_n, t_{n+1})$, then $\hat{\boldsymbol{X}}(t) - \bar{\boldsymbol{x}}_n = \boldsymbol{Y}_n(t) + \boldsymbol{x}_n - \bar{\boldsymbol{x}}_n$ for $t \in [t_n, t_{n+1})$, and $\boldsymbol{Y}_n(t)$ satisfies the SDE

$$\mathrm{d}\boldsymbol{Y}_n(t) = \sum_{k=1}^{M} (\tilde{J}_{k,n}\boldsymbol{Y}_n(t) + \boldsymbol{c}_{k,n})\mathrm{d}W_k(t), \quad t \in [t_n, t_{n+1}), \tag{3.7}$$

where $\boldsymbol{c}_{k,n} = \tilde{\boldsymbol{b}}_{k,n} + \tilde{J}_{k,n}(\boldsymbol{x}_n - \bar{\boldsymbol{x}}_n)$. The solution to (3.7) is given by

$$\boldsymbol{Y}_n(t) = \tilde{\Phi}_n(t)\left(-\int_{t_n}^{t} \tilde{\Phi}_n^{-1}(s) \sum_{k=1}^{M} \tilde{J}_{k,n}\boldsymbol{c}_{k,n}\mathrm{d}s + \int_{t_n}^{t} \tilde{\Phi}_n^{-1}(s) \sum_{k=1}^{M} \boldsymbol{c}_{k,n}\mathrm{d}W_k(s)\right), \tag{3.8}$$

where $\tilde{\Phi}_n(t) = \exp\left\{-\frac{1}{2}\sum_{k=1}^{M} \tilde{J}_{k,n}^2(t - t_n) + \sum_{k=1}^{M} \tilde{J}_{k,n}(W_k(t) - W_k(t_n))\right\}$.

Similar to (3.5)–(3.6), approximating $\tilde{\Phi}_n(t)$ by $\tilde{\Phi}_n(t) \approx I_d + \sum_{k=1}^{M} \tilde{J}_{k,n}(W_k(t) - W_k(t_n))$ in (3.8), setting $\tilde{\Phi}_n^{-1}(s) \equiv \tilde{\Phi}_n^{-1}(t_n)$ for $s \in [t_n, t)$, and evaluating the integrals from t_n to t_{n+1} gives

$$\boldsymbol{Y}_n(t_{n+1}) \approx \left(I_d + \sum_{k=1}^{M} \tilde{J}_{k,n}\Delta W_{k,n}\right) \sum_{k=1}^{M} \left(-\tilde{J}_{k,n}\boldsymbol{c}_{k,n}\Delta t + \boldsymbol{c}_{k,n}\Delta W_{k,n}\right), \tag{3.9}$$

where $\tilde{J}_{k,n}$ is the Jacobian of $\boldsymbol{g}_k$ evaluated at $\bar{\boldsymbol{x}}_n$, $\boldsymbol{c}_{k,n} = \boldsymbol{g}_k(\bar{\boldsymbol{x}}_n) + \tilde{J}_{k,n}(\boldsymbol{x}_n - \bar{\boldsymbol{x}}_n)$ and $\Delta W_{k,n}$ is defined as in (2.4). Note that since $\tilde{J}_{k,n}$ depends on the mean state at t_n, it needs to be computed only once for all samples at t_n and thus reduces the computational cost. The algorithm using $\boldsymbol{\eta}_{n+1} = \boldsymbol{Y}_n(t_{n+1})$ in (3.9), referred to as FPM-LM (Linear approximation at Mean state), is summarized in subsection 3.1.4.

3.1.3. Moment approximations

Both (3.6) and (3.9) provide pathwise approximations to $\boldsymbol{Y}_n(t_{n+1})$, i.e., for each $\omega \in \Omega$ and $\boldsymbol{x}_n(\omega) \in \mathbb{R}^d$, $\boldsymbol{Y}_n(t_{n+1}) = \boldsymbol{Y}_n(t_{n+1}, \omega)$ is approximated based on the pathwise solution of the SDE (3.1) or (3.7). An alternative is to model $\boldsymbol{\eta}_{n+1}$ in distribution by computing the moments for $\boldsymbol{Y}_n(t_{n+1})$. To this end, we revisit the SDE (3.1) and consider the first and

second moments of its solution. More precisely, let $\boldsymbol{\mu}_n(t) = \mathbb{E}[\boldsymbol{Y}_n(t)]$ be the mean and $P_n(t) = \mathbb{E}[\boldsymbol{Y}_n(t)\boldsymbol{Y}_n^T(t)]$ be the second moment of the solution $\boldsymbol{Y}_n(t)$ to (3.1), respectively. These moments satisfy the system of ordinary differential equations on $t \in [t_n, t_{n+1})$

$$\frac{\mathrm{d}\boldsymbol{\mu}_n(t)}{\mathrm{d}t} = 0, \tag{3.10}$$

$$\frac{\mathrm{d}P_n(t)}{\mathrm{d}t} = \sum_{k=1}^{M} \left(J_{k,n} P_n(t) J_{k,n}^T + J_{k,n}\boldsymbol{\mu}_n(t)\boldsymbol{b}_{k,n}^T + \boldsymbol{b}_{k,n}\boldsymbol{\mu}_n^T(t) J_{k,n}^T + \boldsymbol{b}_{k,n}\boldsymbol{b}_{k,n}^T \right). \tag{3.11}$$

Note that since $\boldsymbol{Y}(t_n) = 0$, $\boldsymbol{\mu}_n(t_n) = 0$ and $P_n(t_n) = 0$. Thus the solution to equation (3.10) is $\boldsymbol{\mu}_n(t) \equiv 0$ on $[t_n, t_{n+1})$, and the resulting covariance matrix of $\boldsymbol{Y}_n(t)$ on $[t_n, t_{n+1})$ is $C_n(t) = \mathbb{E}[\boldsymbol{Y}_n(t)\boldsymbol{Y}_n^T(t)] - 0 = P_n(t)$. Plugging into (3.11) results in the following linear matrix-valued non-homogeneous ODE for the covariance matrix

$$\frac{\mathrm{d}C_n(t)}{\mathrm{d}t} = \sum_{k=1}^{M} J_{k,n} C_n(t) J_{k,n}^T + \sum_{k=1}^{M} \boldsymbol{b}_{k,n}\boldsymbol{b}_{k,n}^T, \quad t \in [t_n, t_{n+1}). \tag{3.12}$$

The solution of (3.12) can only be derived explicitly for special cases. Here a one-step Euler approximation with initial value $C_n(t_n) = 0_n$ is applied to give $C_n(t) = (t - t_n)\sum_{k=1}^{M} \boldsymbol{b}_{k,n}\boldsymbol{b}_{k,n}^T$ for $t \in [t_n, t_{n+1})$, and in particular

$$C_n(t_{n+1}) \approx \Delta t \sum_{k=1}^{M} \boldsymbol{b}_{k,n}\boldsymbol{b}_{k,n}^T, \quad \text{with } \boldsymbol{b}_{k,n} = \boldsymbol{g}_k(\boldsymbol{x}_n).$$

The random variable $\boldsymbol{\eta}_{n+1}$ is then modeled to have the mean 0 and covariance $C_n(t_{n+1})$. In particular, set $\boldsymbol{\eta}_{n+1}$ to follow the d-variate normal distribution

$$\boldsymbol{\eta}_{n+1} \sim \mathcal{N}\left(0, \Delta t \sum_{k=1}^{M} \boldsymbol{b}_{k,n}\boldsymbol{b}_{k,n}^T\right). \tag{3.13}$$

The algorithm using $\boldsymbol{\eta}_{n+1}$ modeled by (3.13), which we refer to as FPM-MM (MoMent approximation), is summarized in subsection 3.1.4.

3.1.4. Numerical algorithms

In this subsection, we present three variations of the numerical scheme (2.3) based on the FPM-LP resulted from the model (3.6), FPM-LM resulted from (3.9), and FPM-MM resulted from (3.13), respectively, in Table 1.

For $n = 0, \cdots, N$:	
$\boldsymbol{x}_{n+1} = \boldsymbol{x}_n + \Delta t \boldsymbol{f}(\boldsymbol{x}_n) + \sum_{k=M+1}^{m} \boldsymbol{g}_k(\boldsymbol{x}_n)\Delta W_{k,n} + \boldsymbol{\eta}_{n+1}$ with	
[FPM-LP]	$\boldsymbol{\eta}_{n+1} = \left(I_d + \sum_{k=1}^{M} J_{k,n}\Delta W_{k,n}\right)\sum_{k=1}^{M}\left(-J_{k,n}\boldsymbol{b}_{k,n}\Delta t + \boldsymbol{b}_{k,n}\Delta W_{k,n}\right)$
[FPM-LM]	$\boldsymbol{\eta}_{n+1} = \left(I_d + \sum_{k=1}^{M} \tilde{J}_{k,n}\Delta W_{k,n}\right)\sum_{k=1}^{M}\left(-\tilde{J}_{k,n}\boldsymbol{c}_{k,n}\Delta t + \boldsymbol{c}_{k,n}\Delta W_{k,n}\right)$
[FPM-MM]	$\boldsymbol{\eta}_{n+1} \sim \mathcal{N}\left(0, \Delta t \sum_{k=1}^{M} \boldsymbol{b}_{k,n}\boldsymbol{b}_{k,n}^T\right)$
where	$\bar{\boldsymbol{x}}_n = \mathbb{E}[\boldsymbol{x}_n]$, $\boldsymbol{b}_{k,n} = \boldsymbol{g}_k(\boldsymbol{x}_n)$, $\boldsymbol{c}_{k,n} = \boldsymbol{g}_k(\bar{\boldsymbol{x}}_n) + \tilde{J}_{k,n}(\boldsymbol{x}_n - \bar{\boldsymbol{x}}_n)$, $J_{k,n} = \left(\frac{\partial \boldsymbol{g}_k}{\partial X_1}, \cdots, \frac{\partial \boldsymbol{g}_k}{\partial X_d}\right)\Big\vert_{\boldsymbol{X}=\boldsymbol{x}_n}$, $\tilde{J}_{k,n} = \left(\frac{\partial \boldsymbol{g}_k}{\partial X_1}, \cdots, \frac{\partial \boldsymbol{g}_k}{\partial X_d}\right)\Big\vert_{\boldsymbol{X}=\bar{\boldsymbol{x}}_n}$.

Table 1. Numerical algorithms for vector-valued SDE (2.1) based on FPMs

3.2. Scalar-valued diffusion only SDE

In this section we focus on the case where $d = 1$, for which the solution to the linear SDEs (3.1) and (3.7), as well as the solution to the linear ODE system (3.10)–(3.11) can be derived explicitly. For $x_n \in \mathbb{R}$, set $Y_n(t) = \hat{X}(t) - x_n$ where $\hat{X}(t)$ is the solution to the scalar SDE

$$\mathrm{d}\hat{X}(t) = \sum_{k=1}^{M} g_k(\hat{X}(t))\mathrm{d}W_k(t), \quad \hat{X}(t_n) = x_n, \quad t \in [t_n, t_{n+1}). \tag{3.14}$$

As in the vector case, approximations of $Y_n(t_{n+1})$ will be developed and used to model the random variable $\boldsymbol{\eta}_{n+1}$, denoted as η_{n+1} for $d = 1$, in (2.4).

3.2.1. Linear approximation at the pathwise initial state

For $n = 1, \cdots, N$, given $\hat{X}(t_n) = x_n$, approximate $g_k(\hat{X}(t))$ on $[t_n, t_{n+1})$ by

$$g_k(\hat{X}(t)) \approx b_{k,n} + a_{k,n}(\hat{X}(t) - x_n) \quad \text{with} \quad a_{k,n} = \left.\frac{\mathrm{d}g_k(\hat{X})}{\mathrm{d}\hat{X}}\right|_{\hat{X}=x_n}, \quad b_{k,n} = g_k(x_n), \quad k = 1, \cdots, M.$$

Then $Y_n(t) = \hat{X}(t) - x_n$ satisfies the scalar linear SDE

$$\mathrm{d}Y_n(t) = \sum_{k=1}^{M} (b_{k,n} + a_{k,n}Y_n(t))\mathrm{d}W_k(t), \quad Y_n(t_n) = 0, \ t \in [t_n, t_{n+1}), \tag{3.15}$$

which has a unique solution

$$Y_n(t) = \phi_n(t) \sum_{k=1}^{M} \left(b_{k,n}\Big(- a_{k,n} \int_{t_n}^{t} \phi_n^{-1}(s)\mathrm{d}s + \int_{t_n}^{t} \phi_n^{-1}(s)\mathrm{d}W_k(s)\Big)\right), \quad t \in [t_n, t_{n+1}), \tag{3.16}$$

where

$$\phi_n(t) = \exp\left\{ -\frac{1}{2}\sum_{k=1}^{M} a_{k,n}^2(t - t_n) + \sum_{k=1}^{M} a_{k,n}(W_k(t) - W_k(t_n))\right\}, \quad t \in [t_n, t_{n+1}). \tag{3.17}$$

Note that, here, $\phi_n(t)$ is the exact solution of the homogeneous SDE corresponding to (3.15), i.e., the scalar version of (3.4), but will now be computed directly instead of using the approximation (3.5). Although the integrals $\int_{t_n}^{t} \phi_n^{-1}(s)\mathrm{d}s$ and $\int_{t_n}^{t} \phi_n^{-1}(s)\mathrm{d}W_k(s)$ are still not analytically tractable, since the integrands are state independent we can approximate the integrals by their Riemann sums. To that end, divide $[t_n, t_{n+1})$ into l subintervals $t_n = t_{n,0} < \ldots < t_{n,l-1} = t_{n+1}$ with $t_{n,i+1} - t_{n,i} = \frac{\Delta t}{l} := h$. Then $t_{n,i} = t_n + ih$ and the integrals $\int_{t_{n,i}}^{t_{n,i+1}} \phi_n^{-1}(s)\mathrm{d}s$ and $\int_{t_{n,i}}^{t_{n,i+1}} \phi_n^{-1}(s)\mathrm{d}W_k(s)$ can be approximated by, respectively,

$$\begin{aligned}
\int_{t_{n,i}}^{t_{n,i+1}} \phi_n^{-1}(s)\mathrm{d}s &\approx \phi_n^{-1}(t_{n,i})h = \exp\left\{ \frac{1}{2}ih\sum_{k=1}^{M} a_{k,n}^2 - \sum_{k=1}^{M} a_{k,n}\Delta W_{k,n}^{ih}\right\} h, \\
\int_{t_{n,i}}^{t_{n,i+1}} \phi_n^{-1}(s)\mathrm{d}W_k(s) &\approx \phi_n^{-1}(t_{n,i})\mathcal{N}_k(0, h) = \exp\left\{ \frac{1}{2}ih\sum_{k=1}^{M} a_{k,n}^2 - \sum_{k=1}^{M} a_{k,n}\Delta W_{k,n}^{ih}\right\} \Delta W_{k,n}^{h},
\end{aligned}$$

where $\Delta W_{k,n}^{ih} = W_k(t_n + ih) - W_k(t_n)$ for $i = 0, \cdots, l-1$. As a direct consequence,

$$\int_{t_n}^{t_{n+1}} \phi_n^{-1}(s)\mathrm{d}s = h\sum_{i=0}^{l-1} e^{\frac{1}{2}ih\sum_{k=1}^{M} a_{k,n}^2 - \sum_{k=1}^{M} a_{k,n}\Delta W_{k,n}^{ih}}, \quad \int_{t_n}^{t_{n+1}} \phi_n^{-1}(s)\mathrm{d}W_k(s) = \Delta W_{k,n}^{h}\sum_{i=0}^{l-1} e^{\frac{1}{2}ih\sum_{k=1}^{M} a_{k,n}^2 - \sum_{k=1}^{M} a_{k,n}\Delta W_{k,n}^{ih}}. \tag{3.18}$$

Using (3.17) and (3.18) in (3.16) gives

$$Y_n(t_{n+1}) \approx \phi_n(t_{n+1}) \sum_{k=1}^{M} \Big(b_{k,n} \mathcal{I}^l_{k,n} \Big(- a_{k,n} h + \Delta W^h_{k,n} \Big) \Big), \tag{3.19}$$

where

$$\phi_n(t_{n+1}) \approx \exp\left\{ -\frac{1}{2} \sum_{k=1}^{M} a^2_{k,n} \Delta t + \sum_{k=1}^{M} a_{k,n} \Delta W_{k,n} \right\}, \quad \mathcal{I}^l_{k,n} = \sum_{i=0}^{l-1} e^{\frac{1}{2} ih \sum_{k=1}^{M} a^2_{k,n} - \sum_{k=1}^{M} a_{k,n} \Delta W^{ih}_{k,n}}.$$

In particular, when $l = 1$ and $h = \Delta t$, the formula (3.19) becomes

$$Y_n(t_{n+1}) \approx e^{-\frac{1}{2} \sum_{k=1}^{M} a^2_{k,n} \Delta t + \sum_{k=1}^{M} a_{k,n} \Delta W_{k,n}} \sum_{k=1}^{M} \left(b_{k,n} \left(-a_{k,n} \Delta t + \Delta W_{k,n} \right) \right). \tag{3.20}$$

The algorithm using $\eta_{n+1} \approx Y_n(t_{n+1})$ in (3.20) is summarized in subsection 3.2.3.

3.2.2. Moment approximations

As in the vector case, we can also model η_{n+1} by a scalar-valued random variable. Unlike the vector case, now the ordinary differential equations satisfied by the first and second moments for $Y_n(t)$ can be solved explicitly. We denote by $\mu_n(t) = \mathbb{E}[Y_n(t)]$ and $\nu_n(t) = \mathbb{E}[Y_n^2(t)]$ the first and second moments of $Y_n(t)$ in (3.16), respectively. These moments satisfy the following system of ODEs on $[t_n, t_{n+1})$

$$\begin{aligned} \frac{\mathrm{d}\mu_n(t)}{\mathrm{d}t} &= 0, \quad \mu_n(t_n) = 0, \\ \frac{\mathrm{d}\nu_n(t)}{\mathrm{d}t} &= \sum_{k=1}^{M} a^2_{k,n} \nu_n(t) + 2\mu_n(t) \sum_{k=1}^{M} a_{k,n} b_{k,n} + \sum_{k=1}^{M} b^2_{k,n}, \quad \nu_n(t_n) = 0, \end{aligned}$$

which can be solved analytically to give

$$\mu_n(t) = 0, \quad \nu_n(t) = \frac{\sum_{k=1}^{M} b^2_{k,n}}{\sum_{k=1}^{M} a^2_{k,n}} \left(e^{\sum_{k=1}^{M} a^2_{k,n} (t - t_n)} - 1 \right).$$

In particular at $t = t_{n+1}$ we have

$$\mathbb{E}[Y_n(t_{n+1})] = 0, \quad \mathrm{Var}[Y_n(t_{n+1})] = \frac{\sum_{k=1}^{M} b^2_{k,n}}{\sum_{k=1}^{M} a^2_{k,n}} \left(e^{\sum_{k=1}^{M} a^2_{k,n} \Delta t} - 1 \right). \tag{3.21}$$

Consequently, we can model η_{n+1} by a random variable with mean and variance given by (3.21), and in particular, we set η_{n+1} to be the normal random variable:

$$\eta_{n+1} \sim \mathcal{N}\left(0, \frac{\sum_{k=1}^{M} b^2_{k,n}}{\sum_{k=1}^{M} a^2_{k,n}} \left(e^{\sum_{k=1}^{M} a^2_{k,n} \Delta t} - 1 \right) \right). \tag{3.22}$$

The algorithm using η_{n+1} modeled by (3.22) is summarized in subsection 3.2.3.

3.2.3. Numerical algorithms

In this subsection, we present two variations of the numerical scheme (2.3) for its scalar case, based on the FPM-LP scheme resulting from the model (3.20), and the FPM-MM scheme resulting from (3.22), respectively, in Table 2.

For $n = 0, \cdots, N$:

$x_{n+1} = x_n + \Delta t f(x_n) + \sum_{k=M+1}^{m} g_k(x_n)\Delta W_{k,n} + \eta_{n+1}$ with

[FPM-LP] $\quad \eta_{n+1} = e^{-\frac{1}{2}\sum_{k=1}^{M} a_{k,n}^2 \Delta t + \sum_{k=1}^{M} a_{k,n}\Delta W_{k,n}} \sum_{k=1}^{M} \left(b_{k,n}\left(-a_{k,n}\Delta t + \Delta W_{k,n}\right)\right)$

[FPM-MM] $\quad \eta_{n+1} \sim \mathcal{N}\left(0, \frac{\sum_{k=1}^{M} b_{k,n}^2}{\sum_{k=1}^{M} a_{k,n}^2}\left(e^{\sum_{k=1}^{M} a_{k,n}^2 \Delta t} - 1\right)\right)$

where $\quad a_{k,n} = \left.\frac{\mathrm{d}g_k(X)}{\mathrm{d}X}\right|_{X=x_n}, \quad b_{k,n} = g_k(x_n)$

Table 2. Numerical algorithms for scalar-valued SDE (2.1) **FPMs**

4. Convergence Analysis

In this section we conduct convergence analysis for the FPM-LP scheme. In particular, we will show that the FPM-LP scheme converges with a strong order of 1/2, and a weak order of 1. For simplicity of exposition, we present the analysis for scalar-valued SDEs. The same convergence order can be obtained for vector-valued SDEs following the same procedure with more complicated expressions [54]. For the reader's convenience we restate the one-step scheme on $[t_n, t_{n+1})$ summarized in Table 2

$$x_{n+1} = x_n + \Delta t f(x_n) + \sum_{k=M+1}^{m} g_k(x_n)\Delta W_{k,n} + \eta_{n+1}, \tag{4.1}$$

where η_{n+1} is an approximation of the unique solution (3.16) to the diffusion only linear SDE (3.15) which is obtained from the linear expansion of g_k at each state x_n in the diffusion only SDE (2.5). The goal is to estimate $\mathcal{E} := \max_{n=1,\cdots,N} \mathbb{E}[|X(t_n) - x_n|]$ for the strong convergence, and $\mathfrak{E} := \max_{n=1,\cdots,N} |\mathbb{E}[\psi(X(t_n))] - \mathbb{E}[\psi(x_n)]|$ for an appropriate class of test functions ψ for the weak convergence, respectively.

4.1. Strong convergence

We estimate here the error $\mathcal{E} := \max_{n=1,\cdots,N} \mathbb{E}[|X(t_n) - x_n|]$, in which x_n is computed using FPM-LP, as in Table 2. To that end, consider the the piecewise interpolation process of (4.1) using FPM-LP:

$$x(t) = x_n + (t - t_n) f(x_n) + \sum_{k=M+1}^{m} g_k(x_n)(W_k(t) - W_k(t_n)) + \eta(t), \quad t \in [t_n, t_{n+1}), \tag{4.2}$$

where

$$\eta(t) = \phi_n(t) \sum_{k=1}^{M} \Big(b_{k,n}\left(-a_{k,n}(t - t_n) + (W_k(t) - W_k(t_n))\right)\Big), \quad t \in [t_n, t_{n+1}). \tag{4.3}$$

Note that $\eta(t)$ is an approximation for $Y_n(t)$ on $t \in [t_n, t_{n+1})$, where $Y_n(t)$ is the solution to the linear SDE (3.15), as expressed in (3.16). We first consider the piecewise interpolation process of (4.1) using (3.16)

$$y(t) = x_n + (t - t_n) f(x_n) + \sum_{k=M+1}^{m} g_k(x_n)(W_k(t) - W_k(t_n)) + Y_n(t), \quad t \in [t_n, t_{n+1}),$$

and estimate $\mathbb{E}\left[\sup_{0\le t\le T} |X(t) - y(t)|^2\right]$.

For convenience of the analysis in the sequel, it is natural to write the above equation in its integral form

$$y(t) = x_0 + \int_0^t f(\tilde{x}(s))\mathrm{d}s + \sum_{k=M+1}^{m} \int_0^t g_k(\tilde{x}(s))\mathrm{d}W_k(s) + \int_0^t \mathrm{d}Y(s), \tag{4.4}$$

where $Y(t) = Y_n(t)$ for $t \in [t_n, t_{n+1})$ and $\tilde{x}(t) \equiv x_n$ for $t \in [t_n, t_{n+1})$. Note that $\tilde{x}(t)$ is the piecewise constant process of the FPM-LP discrete solution, $x(t)$ is the piecewise linear interpolation process of the FPM-LP discrete solution, and $y(t)$ is the piecewise solution of the SDE (2.1) in which the drift and slow diffusion are computed using the EM scheme, and the fast diffusion is replaced by the solution to the linear SDE (3.15). They all coincide with the discrete numerical solution x_n at each grid point, i.e., $x(t_n) = \tilde{x}(t_n) = y(t_n) = x_n$ for $n = 1, \cdots, N$.

Set $\mathcal{E}_0(t) := |X(t) - y(t)|$, then the strong discretization error $\mathcal{E}$ satisfies

$$\begin{aligned} \mathcal{E} = \max_{n=1,\cdots,N} \mathbb{E}\left[|X(t_n) - x_n|\right] &\leq \max_{n=1,\cdots,N} \mathbb{E}\left[|X(t_n) - y(t_n)|\right] + \max_{n=1,\cdots,N} \mathbb{E}\left[|y(t_n) - x_n|\right] \\ &\leq \left(\mathbb{E}\Big[\sup_{0\leq t\leq T} \mathcal{E}_0^2(t)\Big]\right)^{1/2} + \max_{n=1,\cdots,N} \mathbb{E}\left[|y(t_n) - x_n|\right]. \end{aligned} \tag{4.5}$$

The first step is to estimate $\mathbb{E}[\sup_{0\leq t\leq T} \mathcal{E}_0^2(t)]$. Recalling that $\hat{X}(t)$ is the piecewise solution to the scalar-valued diffusion-only SDE (2.5), using $Y_n(t) = \hat{X}(t) - x_n$ for $t \in [t_n, t_{n+1})$, and inserting the equation (3.15) in (4.4), we obtain the following expression for $y(t)$ equivalent to (4.4) but more convenient for the analysis in the sequel,

$$y(t) = x_0 + \int_0^t f(\tilde{x}(s))\mathrm{d}s + \sum_{k=M+1}^{m} \int_0^t g_k(\tilde{x}(s))\mathrm{d}W_k(s) + \sum_{k=1}^{M} \int_0^t \Big(g_k(\tilde{x}(s)) + g_k'(\tilde{x}(s))(\hat{X}(s) - \tilde{x}(s))\Big)\mathrm{d}W_k(s). \tag{4.6}$$

Then it follows from (4.6), the scalar version of (2.2), and Cauchy-Schwarz that

$$\begin{aligned} \mathbb{E}\left[\sup_{0\leq t\leq T} \mathcal{E}_0^2(t)\right] &= \mathbb{E}\left[\sup_{0\leq t\leq T} \Big| \int_0^t (f(X(s)) - f(\tilde{x}(s)))\mathrm{d}s + \sum_{k=M+1}^{m} \int_0^t (g_k(X(s) - g_k(\tilde{x}(s)))\mathrm{d}W_k(s) \right. \\ &\qquad \left. + \sum_{k=1}^{M} \int_0^t \Big(g_k(X(s)) - g_k(\tilde{x}(s)) - g_k'(\tilde{x}(s))(\hat{X}(s) - \tilde{x}(s))\Big)\mathrm{d}W_k(s)\Big|^2\right] \\ &\leq 3\left(\mathbb{E}\left[\sup_{0\leq t\leq T} \mathcal{E}_1^2(t)\right] + \mathbb{E}\left[\sup_{0\leq t\leq T} \mathcal{E}_2^2(t)\right] + \mathbb{E}\left[\sup_{0\leq t\leq T} \mathcal{E}_3^2(t)\right]\right), \end{aligned} \tag{4.7}$$

where

$$\begin{aligned} \mathcal{E}_1^2(t) &= \left|\int_0^t (f(X(s)) - f(\tilde{x}(s)))\mathrm{d}s\right|^2, \quad \mathcal{E}_2^2(t) = \left|\sum_{k=M+1}^{m} \int_0^t (g_k(X(s) - g_k(\tilde{x}(s)))\mathrm{d}W_k(s)\right|^2, \\ \mathcal{E}_3^2(t) &= \left|\sum_{k=1}^{M} \int_0^t \Big(g_k(X(s)) - g_k(\tilde{x}(s)) - g_k'(\tilde{x}(s))(\hat{X}(s) - \tilde{x}(s))\Big)\mathrm{d}W_k(s)\right|^2. \end{aligned}$$

Throughout this section it is assumed that

(**A1**) the functions $f : \mathbb{R} \to \mathbb{R}$ and $g_k : \mathbb{R} \to \mathbb{R}$ for $k = 1, \cdots, m$ are continuously differentiable and there exist positive constants L_f, L_k such that

$$|f(x) - f(y)| \leq L_f|x - y|, \quad |g_k(x) - g_k(y)| \leq L_k|x - y|, \quad k = 1, \cdots, m, \quad \forall\ x, y \in \mathbb{R}.$$

(**A2**) there exists $\Lambda_T > 0$ such that

$$\mathbb{E}\left[\sup_{0\leq t\leq T} |x(t)|^p\right] \vee \mathbb{E}\left[\sup_{0\leq t\leq T} |X(t)|^p\right] \vee \mathbb{E}\left[\sup_{0\leq t\leq T} |\hat{X}(t)|^p\right] \leq \Lambda_T, \quad \forall\ p \geq 1.$$

(**A3**) There exists $\gamma \geq 1$ and a positive constant λ such that $|g_k'(x)| \leq \lambda(1 + |x|^\gamma)$.

It follows immediately from the Assumption (**A1**) that

$$f^2(x) \le 2\left(f^2(0) + L_f^2 x^2\right), \quad g_k^2(x) \le 2\left(g_k^2(0) + L_k^2 x^2\right), \quad k = 1, \cdots, m \quad \forall\, x \in \mathbb{R}. \tag{4.8}$$

Given any $s \in [0, T)$, let n_s be the integer for which $s \in [t_{n_s}, t_{n_s+1})$. Then by using (4.6) and the scalar version of (2.5) we obtain

$$y(s) - x_{n_s} = \int_{t_{n_s}}^{s} f(x_{n_s})\mathrm{d}\tau + \sum_{k=M+1}^{m} \int_{t_{n_s}}^{s} g_k(x_{n_s})\mathrm{d}W_k(\tau) + \sum_{k=1}^{M} \int_{t_{n_s}}^{s} \left(g_k(x_{n_s}) + g_k'(x_{n_s})(\hat{X}(\tau) - x_{n_s})\right)\mathrm{d}W_k(\tau) \quad \forall\ s \in [0, T). \tag{4.9}$$

We next estimate the error $\mathbb{E}[\sup_{0\le t\le T} \mathcal{E}_0^2(t)]$ in four steps. First, in Lemma 1 below we provide an estimate of $\mathbb{E}\left[\left|\hat{X}(s) - x_{n_s}\right|^p\right]$, which is crucial for subsequent analysis. Then in Lemmas 2–4 below we provide upper bounds for $\mathbb{E}\left[\sup_{0\le t\le T} \mathcal{E}_1^2(t)\right]$, $\mathbb{E}\left[\sup_{0\le t\le T} \mathcal{E}_2^2(t)\right]$, and $\mathbb{E}\left[\sup_{0\le t\le T} \mathcal{E}_3^2(t)\right]$, respectively. Detailed proof of each Lemma is presented in the Appendix.

Lemma 1. *Let Assumptions (**A1**) and (**A2**) hold. Then for every even number $p \ge 2$ there exists $C_{p,T} > 0$ independent of Δt such that*

$$\mathbb{E}\left[\left|\hat{X}(s) - x_{n_s}\right|^p\right] \le C_{p,T}(\Delta t)^{p/2}, \quad \forall\, s \in [0, T].$$

Proof. See Appendix A.1. □

Lemma 2. *Let Assumptions (**A1**) – (**A3**) hold. Then there exists $C_T > 0$ independent of Δt such that*

$$\mathbb{E}\left[\sup_{0\le t\le T} \mathcal{E}_1^2(t)\right] \le 4TL_f^2 \int_0^T \mathbb{E}\Big[\sup_{0\le t\le s} \mathcal{E}_0^2(t)\Big]\mathrm{d}s + C_T(\Delta t)(\Delta t + 1).$$

Proof. See Appendix A.2. □

Lemma 3. *Let Assumptions (**A1**) – (**A3**) hold. Then there exists $C_T > 0$ independent of Δt such that*

$$\mathbb{E}\left[\sup_{0\le t\le T} \mathcal{E}_2^2(t)\right] \le 4(m - M)\sum_{k=M+1}^{m} L_k^2 \int_0^T \mathbb{E}\Big[\sup_{0\le t\le s} \mathcal{E}_0^2(t)\Big]\mathrm{d}s + C_T\Delta t(\Delta t + 1).$$

Proof. See Appendix A.3. □

Lemma 4. *Let Assumptions (**A1**) - (**A3**) hold, and in addition assume that*

*(**A4**) g_k is twice continuously differentiable for $k = 1, \cdots, M$, and there exists $D_k > 0$ such that*

$$|g_k'(x) - g_k'(y)| \le D_k|x - y| \quad \textit{for all } x, y \in \mathbb{R}.$$

Then the error term $\mathcal{E}_3^2(t)$ satisfies

$$\mathbb{E}\left[\sup_{0\le t\le T} \mathcal{E}_3^2(t)\right] \le C\Delta t \int_0^T \mathbb{E}\left[\sup_{0\le t\le s} \mathcal{E}_0^2(t)\right]\mathrm{d}s + 8MTC_T\Delta t \sum_{k=1}^{M}\left(L_k^2 + D_k^2\Delta t\right).$$

Proof. See Appendix A.4. □

With the preparation above, we can obtain an upper bound for $\mathbb{E}\left[\sup_{0\le t\le T}\mathcal{E}_0^2(t)\right]$, stated in the following Lemma.

Lemma 5. *Let Assumptions (**A1**) – (**A4**) hold. Then there exists C_T independent of Δt such that*

$$\mathbb{E}\left[\sup_{0\le t\le T}\mathcal{E}_0^2(t)\right]\le C_T\Delta t.$$

Proof. Collecting estimates of $\mathbb{E}\left[\sup_{0\le t\le T}\mathcal{E}_1^2(t)\right]$, $\mathbb{E}\left[\sup_{0\le t\le T}\mathcal{E}_2^2(t)\right]$, and $\mathbb{E}\left[\sup_{0\le t\le T}\mathcal{E}_3^2(t)\right]$ obtained in Lemmas 2–4, respectively, and inserting them into (4.7) gives

$$\mathbb{E}\left[\sup_{0\le t\le T}\mathcal{E}_0^2(t)\right]\le 4\left(TL_f^2+(m-M)\sum_{k=M+1}^{m}L_k^2+C\Delta t\right)\int_0^T\mathbb{E}\Big[\sup_{0\le t\le s}\mathcal{E}_0^2(t)\Big]\mathrm{d}s+\mathcal{R}(\Delta t),\tag{4.10}$$

where C is defined in (A.13), and $\mathcal{R}(\Delta t)=\mathcal{R}_1\cdot\Delta t+\mathcal{R}_2\cdot(\Delta t)^2$ with

$$\begin{aligned}\mathcal{R}_1 &= 2(m+1)T\left(TL_f^2+4(m-M)\sum_{k=M+1}^{m}L_k^2\right)\sum_{k=1}^{m}\left(g_k^2(0)+L_k^2\Lambda_T\right)+8MTC_T\sum_{k=1}^{M}L_k^2,\\ \mathcal{R}_2 &= 2(m+1)T\left(TL_f^2+4(m-M)\sum_{k=M+1}^{m}L_k^2\right)\left(f^2(0)+L_f^2\Lambda_T+c_T\right)+8MTC_T\sum_{k=1}^{M}D_k^2,\end{aligned}$$

where c_T is the constant in (A.4) and C_T is the constant in Lemma 1. Applying Gronwall's Lemma to (4.10) results in

$$\mathbb{E}\left[\sup_{0\le t\le T}\mathcal{E}_0^2(t)\right]\le\Delta t\left(\mathcal{R}_1+\mathcal{R}_2\Delta t\right)e^{4T\left(TL_f^2+(m-M)\sum_{k=M+1}^{m}L_k^2+4M\sum_{k=1}^{M}L_k^2\right)}.$$

The proof is complete. □

Remark 1. *In the proof of Lemma 4, the mean-square difference between $X(t)$ and $\hat{X}(t)$ is given by* (A.12)*, which represents the error due to the approximation of $\mathscr{D}_n^X(t)$ by $\mathscr{D}_n^{\hat{X}}(t)$. Its contribution to the total error is given by the $C\Delta t$ term in* (4.10)*. Recall that the error of approximating $\mathscr{D}_n^X(t)$ by an EM scheme in its corresponding formula* (4.10) *is $M\sum_{k=1}^{M}L_k^2$. Then, after a closer look at the constant C, the error due to approximating $\mathscr{D}_n^X(t)$ by $\mathscr{D}_n^{\hat{X}}(t)$ is of order $\Delta t(\Delta tL_f^2+\sum_{k=M+1}^{M}L_k^2)e^{(m+1)\sum_{k=1}^{M}L_k^2\Delta t}$ in* (4.10)*, which is small compared to that of the EM scheme, when the magnitude of $\Delta tL_f^2+\sum_{k=M+1}^{M}L_k^2$ is smaller than $\sum_{k=1}^{M}L_k^2$. With the same order of magnitude of the initial values $f(0)$ and $g_k(0)$, this means that the magnitudes of the fast diffusion processes are much larger than those of the slow diffusion processes.*

Remark 2. *The analysis in Lemmas 1 –5 above can all be generalized to vector-valued SDEs, and results in the same order of convergence rate but larger coefficients. More precisely, the coefficients also depend on the dimension of the state vector.*

We are now ready to construct the strong order of convergence for the FPM-LP scheme. In fact, by (4.5) it remains to estimate $\max_{n=1,\cdots,N}\mathbb{E}\left[|y(t_n)-x_n|\right]$, which is presented in the following Lemma.

Lemma 6. *Let Assumptions (**A1**) – (**A3**) hold. Then there exists $C_T>0$ such that*

$$\max_{n=1,\cdots,N}\mathbb{E}[|y(t_n)-x_n|]\le\sup_{0\le t\le T}\mathbb{E}[|\eta(t)-Y(t)|]\le C_T(\Delta t)^{1/2}.$$

Proof. See Appendix A.5. □

Remark 3. *The order 1/2 error term* $\max_{n=1,\cdots,N}\mathbb{E}[|y(t_n)-x_n|]$ *in Lemma 6 results from the mean-square error* $\mathbb{E}\left[\left(\int_{t_n}^{t}(\phi_n^{-1}(s)-1)\mathrm{d}W_k(s)\right)^2\right]$ *in* (A.19). *It can be improved to an order 1 or higher by setting* $l\geq 2$ *in the approximation* (3.18).

Finally, applying Lemma 6 to (4.5), and using Lemma 5, immediately gives the convergence rate to the FPM-LP scheme, stated below.

Theorem 1. *Let Assumptions (***A1***) – (***A4***) hold. Then the FPM-LP scheme has a strong convergence order of 1/2, i.e., there exists* C_T *independent of* Δt *such that*

$$\mathcal{E}=\max_{n=1,\cdots,N}\mathbb{E}\left[|X(t_n)-x_n|\right]\leq C_T\Delta t^{1/2},$$

where x_n *is computed according to FPM-LP.*

Per Remark 2, the result in Theorem 1 still holds for vector-valued SDEs but with a larger value of C_T that also depends on d. On the other side, note that unlike the scalar-valued FPM-LP in which $\phi_n(t)$ is directly computed, for the vector-valued FPM-LP scheme $\Phi_n(t)$ is a stochastic matrix which is approximated to reduce the computational cost on matrix exponentials. Consequently, there is one extra error term in $\max_{n=1,\cdots,N}\mathbb{E}\left[|y(t_n)-x_n|\right]$ which results from approximating $\Phi_n(t)$ in (3.4) by (3.5). More precisely, instead of (A.14), the piecewise extension of η_{n+1} now becomes

$$\zeta(t)=\left(1+\sum_{k=1}^{M}a_{k,n}(W_k(t)-W_k(t_n))\right)\sum_{k=1}^{M}\left(b_{k,n}\Big(-a_{k,n}\int_{t_n}^{t}\mathrm{d}s+\int_{t_n}^{t}\mathrm{d}W_k(s)\Big)\right),\quad t\in[t_n,t_{n+1}),$$

and the error of the FPM-LP due to approximation of $Y_n(t)$ is now

$$\mathbb{E}[|\zeta(t)-Y_n(t)|]\leq\mathbb{E}[|\zeta(t)-\eta(t)|]+\mathbb{E}[|\eta(t)-Y_n(t)|]. \tag{4.11}$$

By Hölder inequality, sequential representation of $\phi_n(t)$, Itô isometry and boundedness of $\mathbb{E}[a_{k,n}^2]$ and $\mathbb{E}[b_{k,n}^2]$we have

$$\begin{aligned}(\mathbb{E}[|\zeta(t)-\eta(t)|])^2 &\leq 2\mathbb{E}\left[\left(\sum_{k=1}^{M}\Big(b_{k,n}\big(-a_{k,n}\int_{t_n}^{t}\mathrm{d}s+\int_{t_n}^{t}\mathrm{d}W_k(s)\big)\Big)\right)^2\right]\\ &\quad\cdot\mathbb{E}\left[\left|1+\sum_{k=1}^{M}a_{k,n}(W_k(t)-W_k(t_n))-\sum_{j=0}^{\infty}\frac{1}{j!}\big(-\frac{1}{2}\sum_{k=1}^{M}a_{k,n}^2(t-t_n)+\sum_{k=1}^{M}a_{k,n}(W_k(t)-W_k(t_n))\big)^j\right|^2\right]\\ &\leq 2M\sum_{k=1}^{M}\Big(\mathbb{E}[b_{k,n}^2](\mathbb{E}[a_{k,n}^2](\Delta t)^2+\Delta t)\Big)\cdot\mathbb{E}\left[\Big(\frac{1}{2}\sum_{k=1}^{M}a_{k,n}^2\Delta t+O(\Delta t)\Big)^2\right]\leq c(\Delta t)^3,\end{aligned} \tag{4.12}$$

which, along with (4.11) imply that using the approximation (3.5) of $\Phi_n(t)$ in FPM-LP does not change the strong order of convergence presented in Theorem 1.

Remark 4. *The error term* $\mathbb{E}[|\zeta(t)-\eta(t)|]$ *is introduced by the approximation of matrix exponentials. Although it is of order* $O(\Delta t^{3/2})$, *its magnitude can be comparable to* $\Delta t^{1/2}$ *when the magnitudes of* $b_{k,n}^2$, *i.e.,* $g_k^2(x_n)$, *are larger than* Δt^{-2}. *Therefore at* Δt *much larger than the reciprocal of* $|g_k(x_n)|$ *this part of the error may become dominant, and may cause the FPM-LP scheme to be less accurate than the EM scheme.*

4.2. Weak convergence

In this subsection we present an estimate of the error at the final time T

$$\mathfrak{E} := \left|\mathbb{E}[\psi(x_N)] - \mathbb{E}[\psi(X(T))]\right|,$$

for the FPM-LP scheme in the following Theorem.

Theorem 2. *Let Assumptions (**A1**) – (**A4**) hold. Then the FPM-LP scheme has a weak convergence order of 1, i.e., there exists C_T independent of Δt such that $\mathfrak{E} \leq C_T \Delta t$.*

Proof. The proof is based on the Feynman-Kac formula. See Appendix A.6. □

Remark 5. *The FPM-MM scheme uses a Gaussian process with the same moments to approximate $Y_n(t)$ and thus does not have pathwise convergence. On the other hand, by construction we have $\mathbb{E}[x(T)] = \mathbb{E}[y(T)]$ and $\mathbb{E}[x^2(T)] = \mathbb{E}[y^2(T)]$. Therefore for special class of test functions such as quadratic functions, or globally Liptchtiz continuous functions, $\mathfrak{E}_2 := \left|\mathbb{E}[\psi(x(T))] - \mathbb{E}[\psi(y(T))]\right| = 0$, and the FPM-MM scheme also has a weak convergence order of 1.*

5. Numerical Illustrations

We illustrate empirical observations of the convergence of each scheme using select model problems. We begin with a discussion of a scalar SDE, followed by vector SDEs.

5.1. Scalar-valued linear SDE

We highlight one observation regarding the advantage of the proposed integration FPM-LP scheme, relative to EM, for a scalar drift-free stiff SDE involving two Brownian motions. Given the scalar state $X(t)$, we let

$$\boldsymbol{f}(X(t)) \equiv 0; \qquad g_1(X(t)) = 4X(t), \quad g_2(X(t)) = 0.04X(t) \tag{5.1}$$

and integrate the system with an initial condition $X(0) = 1$. The two BMs are chosen to differ in magnitude by 100×, accordingly the rate of diffusive spread due to g_1 is two orders of magnitude faster than that due to g_2. We integrate this system using both EM and the above FPM-LP integrator. In this scalar case, the exponential in the ϕ_n expression is easily computed (there is no approximation of the exponential term as was proposed for the vector case by the one-step ODE integral) resulting in the formulation shown in Eq. (3.20). Resulting sample paths are illustrated in Figure 1. The exact solution of this system can be easily formulated, and is used as a reference for computing the strong convergence of each integrator, EM and FPM-LP, also shown in Fig. 1, using 4.8M samples. As can be seen, both integrators exhibit, in the limit of small time step, the expected $(\Delta t)^{1/2}$ slope. This convergence is naturally lost in the high Δt range. However, what's interesting to note is that the large-Δt error for EM is an order of magnitude higher than that for FPM-LP. This observation, albeit in a linear SDE, highlights the motivation behind the present development, namely that linearized modeling of the unresolved fast processes provides improved accuracy versus the alternative of integrating them directly using large time steps. Note that, when this system is computed with $g_1() = g_2() = 0.04X(t)$, this large difference between the errors from both schemes disappears, with the error in both

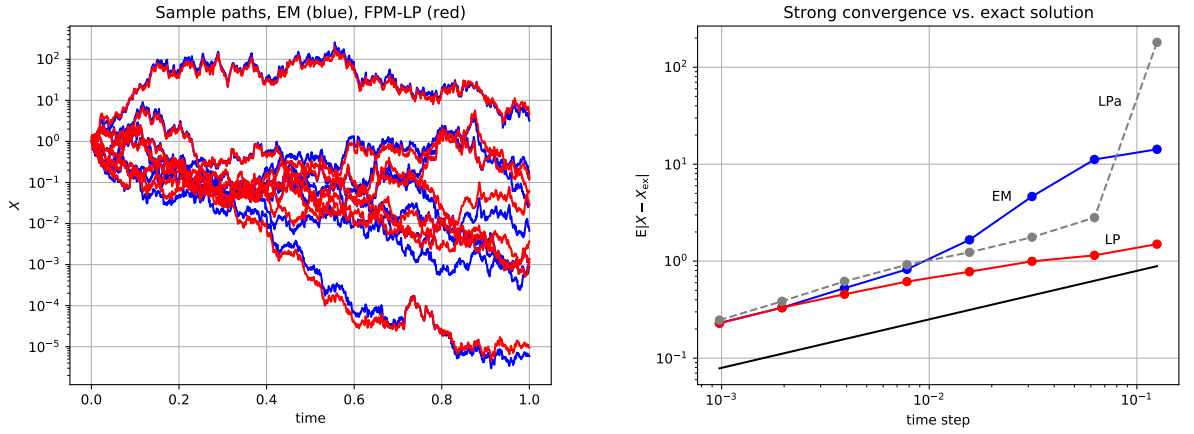

Fig. 1. Illustrated sample paths for X (left), and convergence (right), linear scalar stiff SDE with EM (blue) & FPM-LP (red) integrators. The right frame also includes the observed convergence of the error for the FPM-LP solution employing the approximate one time-step ODE computation of ϕ_n (grey), labeled "LPa".

integrators at the largest Δt being $O(10^{-4})$, and the error reduction due to FPM-LP being $\sim 1\%$. Further, it is useful to add that, if the one time-step approxmiation of the ODE (3.3) solution via (3.5), introduced in the vector case, is employed here, the associated approximate-FPM-LP error, shown in Fig. 1 as "LPa" are higher than FPM-LP, and slightly higher than EM at small Δt, but they retain a significant error advantage over EM in the large Δt region, up to and including the next to largest time step size. On the other hand, as can be seen in the figure, the approximate-FPM-LP error grows appreciably to $O(10^2)$ for the largest time step size. This is indeed a manifestation of the coarse approximation of the ODE (3.3) solution for large time steps, and an indication of the necessity to estimate and control the error in this approximation, employing a better approximation in the large Δt context as necessary, *e.g.* employing multiple internal time steps for the ODE solution.

5.2. Vector-valued model SDEs

Consider the SDE (Eq. 2.1) in $\mathbb{R}^d$ where $d = 2$, with $m = 3$ Brownian motions, and with p-order polynomial drift and diffusion terms specified as follows

$$\boldsymbol{f}(\boldsymbol{X}(t)) = (A\boldsymbol{X}(t))^{\circ p}\,; \qquad \boldsymbol{g}_k(\boldsymbol{X}(t)) = (B_k\boldsymbol{X}(t))^{\circ p}\,, \quad k = 1, \cdots, m \tag{5.2}$$

where A and B_k are $d \times d$ real matrices, and the operation $()^{\circ p}$ denotes the Hadamard power, where the quantity inside the paranthesis is raised element-wise to the power p [72]. Thus, for $\boldsymbol{Y} = (Y_1, \cdots, Y_n)$, we have $\boldsymbol{Y}^{\circ p} := (Y_1^p, \cdots, Y_n^p)$. We will consider cases with $p \in \{1, 2\}$. We use the following A, B_k matrices

$$A = \alpha \begin{bmatrix} 1 & 2 \\ 3 & -4 \end{bmatrix}, \qquad B_1 = \beta_1 \begin{bmatrix} -1 & 2 \\ 3 & -6 \end{bmatrix}, \qquad B_2 = \beta_2 \begin{bmatrix} 3 & -2 \\ -3 & 8 \end{bmatrix}, \qquad B_3 = \beta_3 \begin{bmatrix} 1 & 4 \\ 6 & -9 \end{bmatrix} \tag{5.3}$$

with the scaling coefficients $\alpha = 0.1$, while $\beta = (\beta_1, \beta_2, \beta_3)$ is chosen differently for different equation systems[2]. We further note that these (A, B_k) matrices are chosen to be commuting ($AB_k = B_kA$, $k = 1, 2, 3$) in order to facilitate

[2]We make the following choices: $\beta = (0.05, 0.05, 5 \times 10^{-7})$ for the linear ($p = 1$) case, and $(0.04, 0.04, 5 \times 10^{-4})$ for the quadratic ($p = 2$) case.

the derivation of an analytical solution for the linear case, for purposes of convergence testing. Thus, for $p = 1$, the solution to the autonomous linear SDE (2.1) with drift and diffusion terms defined by (5.2) can be given explicitly when the matrices $A, B_1, \cdots, B_m$ commute, as

$$\boldsymbol{X}(t) = \boldsymbol{x}_0 \exp\left\{\left(A - \frac{1}{2}\sum_{k=1}^{m} B_k^2\right)(t - t_0) + \sum_{k=1}^{m} B_k \left(W_k(t) - W_k(t_0)\right)\right\}. \tag{5.4}$$

The amplitudes of the diffusion source terms $\boldsymbol{g}_k()$ determine the strength of the diffusion processes in the system, and the associated speed of the diffusive spread of the random paths of the system. The chosen values of matrices (A, B_k) control the strengths of the drift and diffusion terms for any given state. By design, these choices provide, for the two cases $p = (1, 2)$, diffusion source terms $\boldsymbol{g}_k(\boldsymbol{X}(t))$ with $(\boldsymbol{g}_1, \boldsymbol{g}_2)$ having component amplitudes always larger than the corresponding amplitudes in $\boldsymbol{g}_3$, thus providing a system with two "fast" diffusion terms and one slower term. In this relative context, we define M as the number of fast diffusion source terms. Of course, M can vary during a simulation, and has to be determined for any given state, but our illustration example cases are constructed to have $M = 2$ for the full state trajectory given the initial conditions $\boldsymbol{x}_0 = (-100, 100)$ in the linear case, and $\boldsymbol{x}_0 = (1, 1)$ in the quadratic case, and the chosen integration time $T = 0.5$.

In order to examine the temporal convergence of the solution, we integrate the above model system using different time integrators, employing a set of fixed $\Delta t = T/N$ time steps, with N values related by factors of 2, specifically $N = 8, \cdots, 512$. Sample paths computed with FPM-LP are illustrated, for the finest Δt, for the quadratic case ($p = 2$), in Figure 2. Clearly, the dynamics of X_1 are dominated by drift, while X_2 is dominated by diffusion processes. Inspection of the contributions of fast-diffusion, vs drift+slow-diffusion to the RHS of X_1 and X_2 illustrate the larger magnitude of contributions of fast diffusion and lower magnitude of drift+slow-diffusion, to the RHS of X_2, vs. the corresponding contributions to the RHS of X_1, which is consistent with the observed behavior in Fig. 2. We note that this is not necessary for the present construction, which, rather, relies on distinguishing fast/slow diffusion source terms and does not require the dominance of fast diffusion in specific components of the state. For the same time step choice, sample paths with the EM, FPM-LP, and FPM-LM integrators, relying on the same random seed and thus the same Brownian motions (BMs), are essentially indistinguishable on this scale (not shown). On the other hand, FPM-MM exhibits random paths that are not derived from the same BM paths, as the BM random variables used in the fast modes are substituted by the random vector $\boldsymbol{\eta}$ in Eq. (3.13), and are not comparable path-wise (also not shown). Thus, there is no expectation of strong convergence from FPM-MM, only weak convergence.

We will rely on estimation of various error norms among the different solutions and examination of the error norm convergences. We estimate error norm statistics employing 24M samples for each case. In order to arrive at empirical estimates of temporal convergence of an SDE time integrator, the BMs employed with different choices of Δt need to use subintegrals of the same BM. A straightforward way to do this is to compute a discretized BM at the finest time step, and let the coarser time step BMs be derived based on discretized integrals over subintervals of this BM. Thus, with $h_q \equiv (\Delta t)_q$, let $h_q = h_0/2^q$ with $q = 0, 1, \cdots, Q$, where $h_0 = T/N_0$ is the coarsest time step, and $N_q = N_0 2^q$ is the number of time steps for a given resolution q, we compute the finest time step BM path $P_Q = \{\xi_1, \cdots, \xi_{N_Q}\}$, where

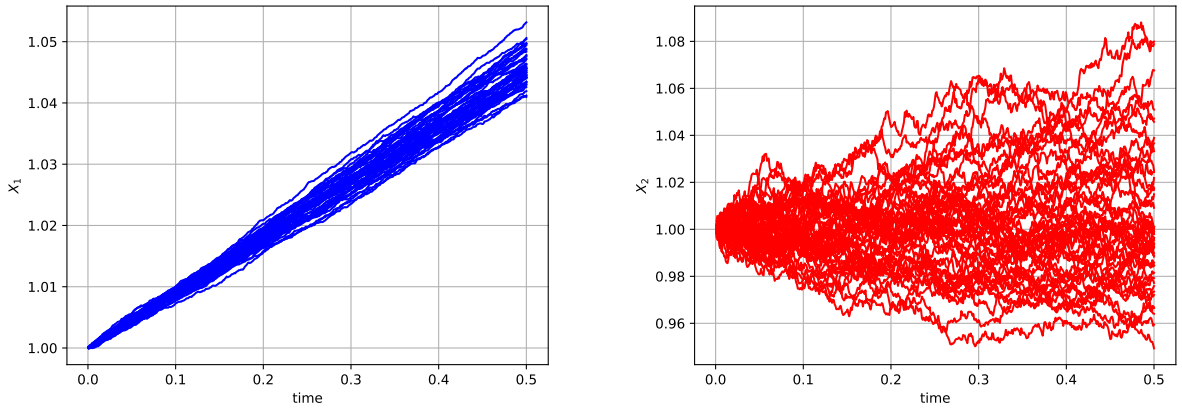

Fig. 2. Illustrated sample paths computed with FPM-LP for X_1 (left) and X_2 (right), quadratic SDE.

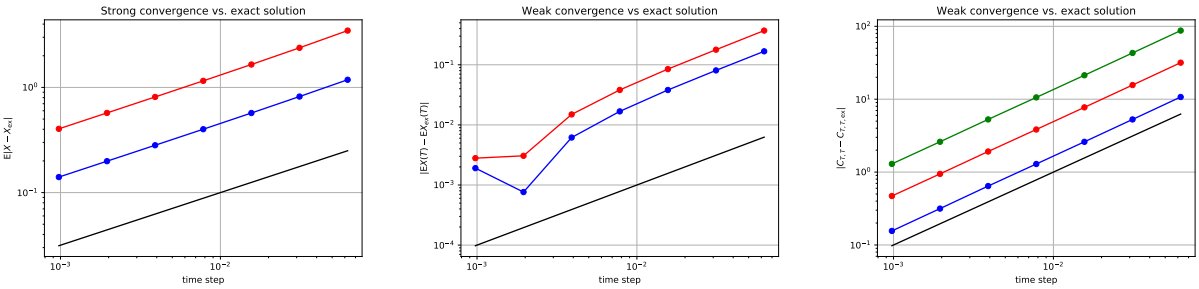

Fig. 3. Convergence of EM vs the exact solution for the linear SDE system. The left frame exhibits the strong convergence, showing the decay of the expectation of the absolute value of the solution error at time T, for X_1 (blue) and X_2 (red). The middle and right frames exhibit the weak convergence. The middle frame shows the decay of the error in the expectation of the solution for X_1 (blue) and X_2 (red), while the right frame shows the decay of the error in the covariance matrix components, all at time T. The covariance matrix components are denoted with blue (1,1), red (1,2), and green (2,2). The black lines illustrate the expected slopes in each case.

$\xi \sim \mathcal{N}(0, h_Q)$. And, for any time step $h_q = 2^{Q-q}h_Q$, we have, at time $t_n = nh_q$,

$$\Delta W_n = W_{n+1} - W_n = \int_{t_n}^{t_n+h_q} \mathrm{d}W(t) \approx \sum_{s=n\cdot 2^{Q-q}}^{(n+1)2^{Q-q}} \xi_s. \tag{5.5}$$

Considering first EM, we illustrate the observed convergence to the exact solution (5.4) for the linear SDE system. We show the $O(\Delta t^{1/2})$ strong convergence, namely the convergence of the estimate of the expectation of the absolute solution error, with respect to the exact solution, at final time T, in Fig. 3, as would be expected. We also observe the expected $O(\Delta t)$ weak convergence, at T, of the estimated error in the solution mean vector $\boldsymbol{\mu}_T = \mathbb{E}\boldsymbol{X}(T)$ and covariance matrix $\boldsymbol{C}_{T,T} = \mathbb{E}[\boldsymbol{X}(T) - \boldsymbol{\mu}_T)(\boldsymbol{X}(T) - \boldsymbol{\mu}_T)^T]$, relative to the exact solution, also shown in Fig. 3.

The FPM-LP and LM integrators show similar results, as shown e.g. in Fig. 4, including similar slopes and levels of error as for EM, as would be expected for this linear case, where the linearization and mean-based Taylor series approximations are exact. The figure also shows that FPM-MM has similar weak convergence results in terms of slope and error magnitude, albeit with more noise, but it exhibits no strong convergence vs the exact solution, as ought to be expected given the algorithm construction.

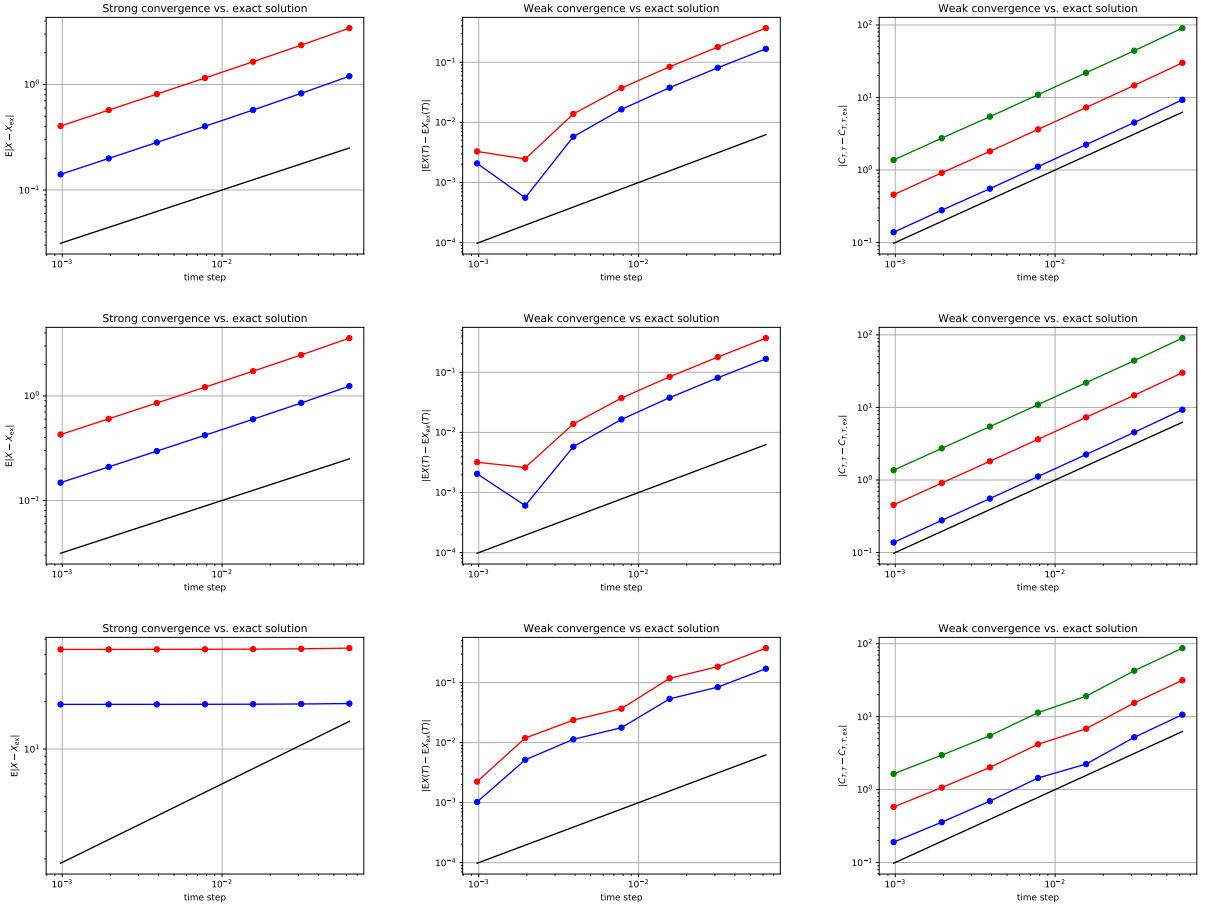

Fig. 4. Convergence of FPM-LP (top row), FPM-LM (middle row), and FPM-MM (bottom row) vs the exact solution for the linear SDE system. The format-structure of the plots follows that in Fig. 3. Results show the expected first order weak convergence in all cases, the absence of strong convergence for MM, and the 1/2-order strong convergence in the other two cases.

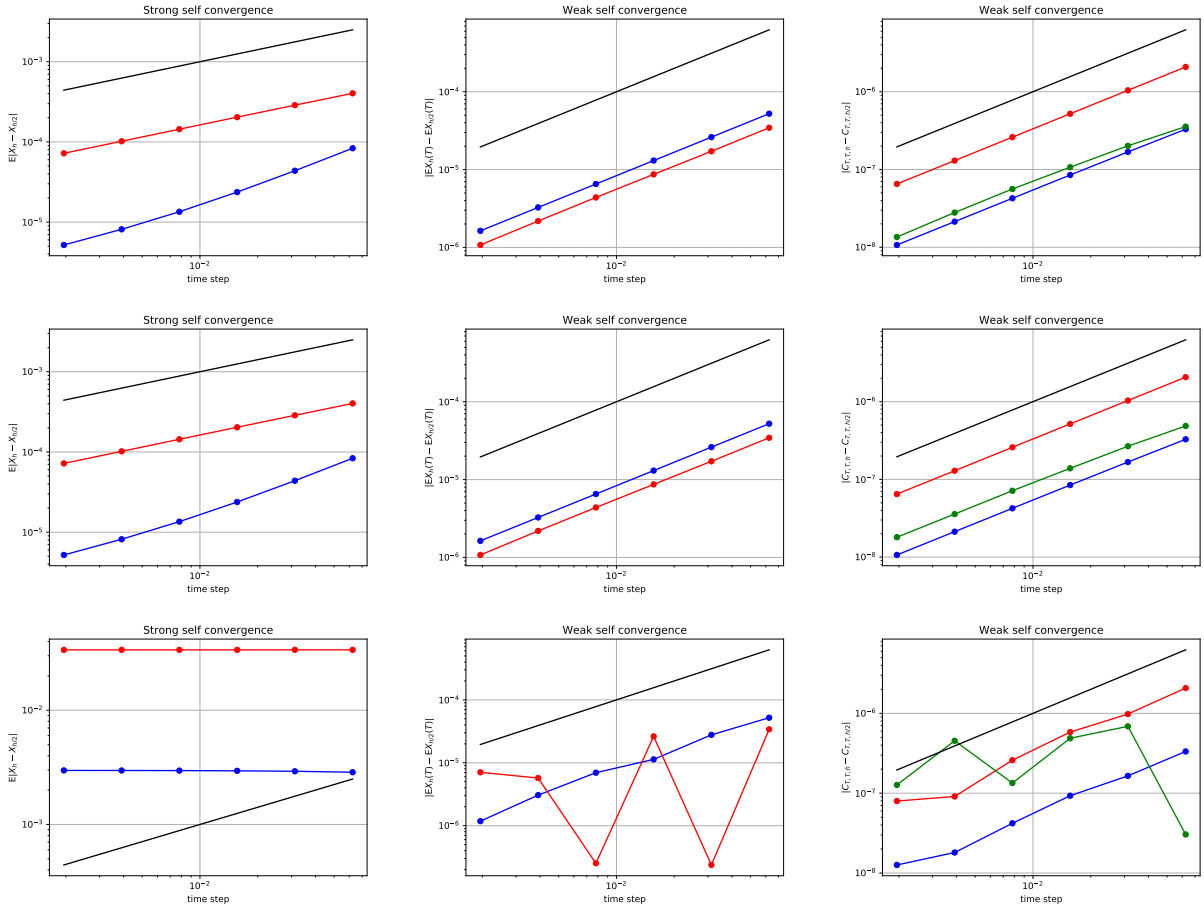

Fig. 5. Self convergence of FPM-LP (top row), FPM-LM (middle row), and FPM-MM (bottom row) for the quadratic SDE system. The format-structure of the plots follows that in Fig. 3.

Let us consider next the convergence of the quadratic SDE. We examine first self-convergence, where the error is defined as the difference in the solution (strong) or its moments (weak) between two successive time-step refinements, with (h_q, h_{q+1}). Results are shown in Fig. 5 for FPM-LP, FPM-LM, and FPM-MM. We see that LP and LM both exhibit the expected strong/weak convergence rates. On the other hand, again, as would be expected, MM has no strong self-convergence, and, while weak convergence is discernible in terms involving X_1, which had relatively minor fast-diffusion role, there is no robust observation of weak convergence for terms dominated by X_2, where the system dynamics highlight the role of strong/fast diffusion. The fact that the random vector $\boldsymbol{\eta}$ samples are wholly unrelated between different time step cases explains this challenge in observing weak convergence in terms dominated by strong diffusion. Whereas X_1 terms, dominated by drift and slow diffusion, and hence being handled using EM, with the above BM connection between time step cases, show clear convergence. We note that we also examined the weak self-convergence in terms of the components of the covariance matrix between times $(T/2, T)$ with similar conclusions.

We also examine the convergence with respect to the EM solution at the finest time step h_Q, which we refer to as

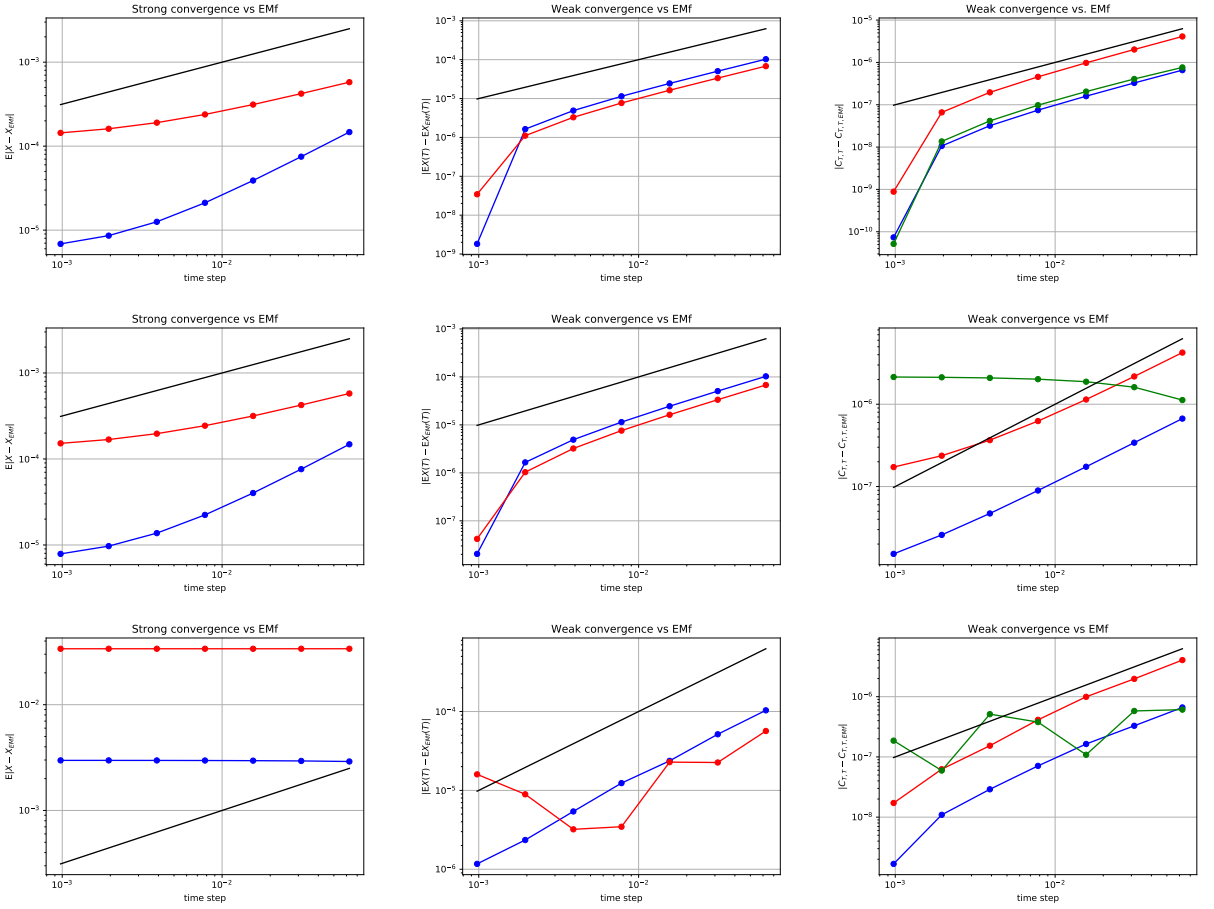

Fig. 6. Convergence of FPM-LP (top row), FPM-LM (middle row), and FPM-MM (bottom row) vs EMf (EM computed at the finest Δt) for the quadratic SDE system. The format-structure of the plots follows that in Fig. 3.

EMf. Results are shown in Fig. 6 for FPM-LP, FPM-LM, and FPM-MM. Considering first strong convergence, we see convergence at the $\Delta t^{1/2}$ slope for FPM-LP and LM, both curves flattening out at some point at small Δt. Since the EMf solution is not the true solution, this is expected, and we have observed that the floor can be lowered by taking smaller h_Q values. Note that X_2 convergence rate decays earlier and its ultimate error versus EMf is higher than X_1. This is consistent with X_2 being more impacted with faster diffusion, and thus with our associated approximation with linearization, thus deviating from EMf by a larger amount than X_1. Of course, MM exhibits no strong convergence, again as expected. As to weak convergence, we see that all components of $\boldsymbol{\mu}_T$ and $\boldsymbol{C}_{T,T}$ exhibit good $O(\Delta t)$ convergence rate of FPM-LP to EMf. On the other hand, for FPM-LM, the mean-based approximation has a discernible impact, resulting in higher error-floor in the $(1,2)$ term of the covariance matrix, and, more notably, losing convergence entirely in the $(2,2)$ diagonal term. Similar observations are evident in the $\boldsymbol{C}(T/2,T)$ components, not shown. Note that neither decreasing h_Q nor increasing the number of samples could result in a discernible convergence in this last quantity. We attribute this lack of convergence to the Taylor series mean-based approximation, given the large spread in X_2 due to diffusion. The resulting errors are not Δt dependent, and result in clear bias in the results in this case. A robust implementation would include some error-detection scheme that estimates the expected Taylor series error impact on quantities of interest, and employs the LM approximation only when desireable thresholds are met. Finally, for FPM-MM, we see the expected weak convergence in the X_1 mean, as well as the $(1,1)$ and $(1,2)$ covariance matrix terms, however, the convergence of moments of X_2 are noisy, and not reliable for observation of an empirical convergence rate.

6. Closing remarks

We have outlined the utility of a linearized approximation of fast diffusion processes in stiff SDEs, that arrived at the explicit FPM-LP scheme, and associated further approximations resulting in FPM-LM and MM. The FPM-LP scheme has the same orders of strong and weak convergence as Euler Maruyama. However, they differ in accuracy due to three approximations in FPM-LP in the time integration of the fast processes. These are: (i) local linearization of the fast diffusion source terms, (ii) decoupling fast and slow processes, and (iii) approximation of a matrix exponential by its linear components. Our analysis shows that errors due to (i) are higher order than the EM error in integrating the fast processes. The magnitude of (ii) is also small with respect to this EM error when there is a clear separation of scales. Finally, (iii) is also smaller than this EM error, except for cases where Δt is much larger than the reciprocals of the magnitudes of the fast source terms $|g_k(x_n)|$ for $k = 1, \cdots, M$ (Remark 4). Thus FPM-LP outperforms the EM scheme for stiff systems with clear separation of magnitudes of fast and slow processes, as long as Δt is not oversized. This latter condition can be controlled, allowing the elimination of this last approximation as necessary when using very large time steps, with attendant costs. Our numerical experiments highlight the convergence of the error in each of the three variants of the proposed scheme, for both linear and nonlinear model systems. We illustrated the expected strong and weak self-convergence of FPM-LP and LM. We also showed weak self-convergence of terms not-dominated by fast diffusion for MM. We also illustrated strong and weak convergence of FPM-LP relative to EM. We observed similarly strong convergence for FPM-LM, but weak convergence was found only in the mean, and in

covariance terms not dominated by fast diffusion. For MM convergence with respect to EM, we only observed weak convergence, and that only in terms not dominated by fast diffusion. We can conclude, regarding LM specifically, that, while it is certainly less expensive than LP, its use ought to be based on some estimation of the error resulting from the mean-based Taylor series approximation outlined above, particularly for terms dominated by fast diffusion. As for MM, while also efficient, it does not exhibit strong convergence, and it only exhibits weak convergence in terms not-dominated by fast diffusion.

Acknowledgments

The authors acknowledge technical discussions with Prof. Mauro Valorani, Sapienza University of Rome, that helped with the development of ideas. This work was partially supported by the Simons Foundation (Collaboration Grants for Mathematicians No. 419717), and by the US Department of Energy (DOE), Office of Basic Energy Sciences (BES) Division of Chemical Sciences, Geosciences, and Biosciences. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525. This paper describes objective technical results and analysis. Any subjective views or opinions that might be expressed in the paper do not necessarily represent the views of the U.S. Department of Energy or the United States Government.

References

[1] D. L. Bunker, B. Garrett, T. Kleindienst, G. S. Long, Discrete simulation methods in combustion kinetics, Combustion and Flame 23 (1974) 373 – 379.
[2] D. Gillespie, A general method for numerically simulating the stochastic time evolution of coupled chemical reactions, J. Comput. Phys. 22 (1976) 403–434.
[3] D. Gillespie, Exact stochastic simulation of coupled chemical reactions, Journal of Physical Chemistry 81 (1977) 2340–2361.
[4] D. T. Gillespie, Stochastic simulation of chemical kinetics, Annu. Rev. Phys. Chem. 58 (2007) 35 – 55.
[5] T. Székely, K. Burrage, Stochastic simulation in systems biology, Computational and Structural Biotechnology Journal 12 (2014) 14 – 25.
[6] D. T. Gillespie, Approximate accelerated stochastic simulation of chemically reacting systems., Journal of Chemical Physics 115 (2001) 1716 – 33.
[7] M. Rathinam, L. R. Petzold, Y. Cao, D. T. Gillespie, Stiffness in stochastic chemically reacting systems: the implicit tau-leaping method., Journal of Chemical Physics 119 (2003) 12784 – 94.
[8] Y. Cao, L. R. Petzold, M. Rathinam, D. T. Gillespie, The numerical stability of leaping methods for stochastic simulation of chemically reacting systems, Journal of Chemical Physics 121 (2004) 12169 – 12178.
[9] Y. Cao, D. T. Gillespie, L. R. Petzold, The slow-scale stochastic simulation algorithm, The Journal of Chemical Physics 122 (2005) 014116.
[10] Y. Cao, D. T. Gillespie, L. R. Petzold, Multiscale stochastic simulation algorithm with stochastic partial equilibrium assumption for chemically reacting systems, Journal of Computational Physics 206 (2005) 395–411.
[11] Y. Cao, D. T. Gillespie, L. R. Petzold, Accelerated stochastic simulation of the stiff enzyme-substrate reaction, Journal of Chemical Physics 123 (2005) 144917.
[12] E. L. Haseltine, J. B. Rawlings, Approximate simulation of coupled fast and slow reactions for stochastic chemical kinetics., Journal of Chemical Physics 117 (2002) 6959 – 6969.
[13] H. Salis, Y. Kaznessis, Accurate hybrid stochastic simulation of a system of coupled chemical or biochemical reactions., Journal of Chemical Physics 122 (2005) 54103 – 1.
[14] J. Goutsias, Quasiequilibrium approximation of fast reaction kinetics in stochastic biochemical systems, Journal of Chemical Physics 122 (2005) 184102.
[15] W. E, D. Liu, E. Vanden-Eijnden, Nested stochastic simulation algorithms for chemical kinetic systems with disparate rates, Journal of Chemical Physics 123 (2005) 194107.
[16] A. Moraes, R. Tempone, P. Vilanova, A multilevel adaptive reaction-splitting simulation method for stochastic reaction networks, SIAM Journal on Scientific Computing 38 (2016) A2091–A2117.
[17] C. V. Rao, A. P. Arkin, Stochastic chemical kinetics and the quasi-steady-state assumption: application to the Gillespie algorithm, J. Chem. Phys. 118 (2003) 4999–5010.
[18] D. A. Goussis, Quasi steady state and partial equilibrium approximations: Their relation and their validity, Combustion Theory and Modelling 16 (2012) 869–926.

[19] J. K. Kim, K. Josić, M. R. Bennett, The Validity of Quasi-Steady-State Approximations in Discrete Stochastic Simulations, Biophysical Journal 107 (2014) 783–793.

[20] N. A. Sinitsyn, N. Hengartner, I. Nemenman, Adiabatic coarse-graining and simulations of stochastic biochemical networks, Proceedings of the National Academy of Sciences 106 (2009) 10546–10551.

[21] D. T. Gillespie, The chemical langevin equation., Journal of Chemical Physics 113 (2000) 297 – 306.

[22] D. T. Gillespie, The chemical langevin and fokker-planck equations for the reversible isomerization reaction, Journal of Physical Chemistry A 106 (2002) 5063 – 5071.

[23] D. J. Higham, Modeling and simulating chemical reactions, SIAM Review 50 (2008) 347–368.

[24] D. Higham, R. Khanin, Chemical Master versus Chemical Langevin for First-Order Reaction Networks, The Open Applied Math Journal 2 (2008) 59–79.

[25] V. Sotiropoulos, M. Contou-Carrere, P. Daoutidis, Y. N. Kaznessis, Model reduction of multiscale chemical langevin equations: A numerical case study, IEEE/ACM Transactions on Computational Biology and Bioinformatics 6 (2009) 470–482.

[26] M. N. Contou-Carrere, P. Daoutidis, Decoupling of fast and slow variables in chemical langevin equations with fast and slow reactions, in: 2006 American Control Conference, 2006, pp. 6 pp.–.

[27] P. Thomas, A. V. Straube, R. Grima, The slow-scale linear noise approximation: an accurate, reduced stochastic description of biochemical networks under timescale separation conditions, BMC Systems Biology 6 (2012) 39.

[28] X. Han, M. Valorani, H. Najm, Explicit Time Integration of the Stiff Chemical Langevin Equations using Computational Singular Perturbation, Journal of Chemical Physics (2019). In press.

[29] S. H. Lam, Singular perturbation for stiff equations using numerical methods, in: C. Casci (Ed.), Recent Advances in the Aerospace Sciences, Plenum Press, New York, 1985, p. 3.

[30] S. H. Lam, D. A. Goussis, Understanding complex chemical kinetics with computational singular perturbation, Proc. Comb. Inst. 22 (1988) 931–941.

[31] S. H. Lam, Using CSP to Understand Complex Chemical Kinetics, Combustion Science and Technology 89 (1993) 375–404.

[32] S. H. Lam, D. A. Goussis, The CSP Method for Simplifying Kinetics, International Journal of Chemical Kinetics 26 (1994) 461–486.

[33] S. H. Lam, D. A. Goussis, Computational Singular Perturbation; Theory and Applications, Report 1986-MAE, Princeton Univ., 1991.

[34] S. H. Lam, D. A. Goussis, The Analytic Foundation of CSP, Report 1800-MAE, Princeton Univ., 1991.

[35] M. Valorani, D. A. Goussis, H. N. Najm, Using CSP to Analyze Computed Reactive Flows, 8th SIAM Int. Conf. On Numerical Combustion, Amelia Island, FL, 2000.

[36] M. Valorani, D. A. Goussis, H. N. Najm, Using CSP to Analyze Computed Reacting Flows, 2000. Eighth International Conference on Numerical Combustion, SIAM.

[37] M. Valorani, D. Goussis, H. N. Najm, Enhanced CSP Diagnostic Tools to Analyze Reacting Flows, in: Ninth Int. Conf. on Numerical Combustion, Sorrento, Italy, 2002.

[38] M. Valorani, H. N. Najm, D. Goussis, CSP Analysis of a Transient Flame-Vortex Interaction: Time Scales and Manifolds, Combustion and Flame 134 (2003) 35–53.

[39] D. A. Goussis, M. Valorani, F. Creta, H. N. Najm, Inertial Manifolds with CSP, in: K. Bathe (Ed.), Computational Fluid and Solid Mechanics 2003, volume 2, Elsevier Science, Cambridge, MA, 2003, pp. 1951–1954.

[40] M. Valorani, D. A. Goussis, F. Creta, H. N. Najm, Higher Order Corrections in the Approximation of Low Dimensional Manifolds and the Construction of Simplified Problems with the CSP Method, J. Comput. Phys. 209 (2005) 754–786.

[41] M. Valorani, F. Creta, D. A. Goussis, J. C. Lee, H. N. Najm, An Automatic Procedure for the Simplification of Chemical Kinetics Mechanisms based on CSP, in: K. Bathe (Ed.), Computational Fluid and Solid Mechanics 2005, Elsevier Science, 2005, pp. 900–904.

[42] M. Valorani, F. Creta, D. A. Goussis, J. C. Lee, H. N. Najm, Chemical Kinetics Simplification via CSP, Combustion and Flame 146 (2006) 29–51.

[43] M. Valorani, F. Creta, F. Donato, H. N. Najm, D. A. Goussis, Skeletal Mechanism Generation and Analysis for *n*-heptane with CSP, Proc. Comb. Inst. 31 (2007) 483–490.

[44] M. Valorani, F. Creta, F. Donato, H. N. Najm, D. A. Goussis, Skeletal mechanism generation and analysis for *n*-heptane with csp, in: 31^{st} Symposium on Combustion, Heidelberg, Germany, 2006.

[45] M. Valorani, F. Creta, F. Donato, H. N. Najm, D. A. Goussis, A csp-based skeletal mechanism generation procedure: Auto-ignition and premixed laminar flames in *n*-heptane/air mixtures, in: ECCOMAS CFD 2006, Delft, Holland, 2006.

[46] A. Zagaris, H. G. Kaper, T. J. Kaper, Analysis of the CSP Reduction Method for Chemical Kinetics, in: SIAM Conference on Applications of Dynamical Systems, May 27-31, 2003 at Snowbird, Utah, 2003.

[47] A. Zagaris, H. G. Kaper, T. J. Kaper, Analysis of the CSP Reduction Method for Chemical Kinetics, Nonlinear Sci. 14 (2004) 59–91.

[48] M. Hadjinicolaou, D. A. Goussis, Asymptotic solutions of stiff pdes with the csp method: the reaction-diffusion equation, SIAM J. Sci. Comput. 20 (1999) 781–810.

[49] M. Salloum, A. Alexanderian, O. Le Maître, H. N. Najm, O. Knio, Simplified CSP Analysis of a Stiff Stochastic ODE System, Computer Methods in Applied Mechanics and Engineering 217-220 (2012) 121–138.

[50] J. C. Lee, H. N. Najm, S. Lefantzi, J. Ray, M. Frenklach, M. Valorani, D. A. Goussis, A CSP and Tabulation Based Adaptive Chemistry Model, Combustion Theory and Modeling 11 (2007) 73–102.

[51] J. Prager, H. N. Najm, M. Valorani, D. A. Goussis, Skeletal Mechanism Generation with CSP and Validation for Premixed n-Heptane Flames, Proc. Comb. Inst. 32 (2008) 509–517.

[52] L. Arnold, Stochastic differetial equations: theory and applications, John Wiley & Sons, Inc., 1974.

[53] X. Mao, Stochastic differential equations and applications, Horwood, New York, 1997.

[54] P. E. Kloeden, E. Platen, Numerical solution of stochastic differential equations, volume 23 of *Applications of Mathematics (New York)*, Springer-Verlag, Berlin, 1992.

[55] B. Oksendal, Stochastic differential equations: An Introduction with Applications, Springer-Verlag, Berlin Heidelberg, 2003.

[56] D. J. Higham, X. Mao, A. M. Stuart, Strong convergence of euler-type methods for nonlinear stochastic differential equations, SIAM J. Numer. Anal. 40 (2002) 1041–1063.

[57] D. Higham, Modeling and Simulating Chemical Reactions, SIAM Review 50 (2008) 347–368.

[58] L. Hu, X. Li, X. Mao, Convergence rate and stability of the truncated euler-maruyama method for stochastic differential equations, Journal of Computational and Applied Mathematics 337 (2018) 274 – 289.
[59] Y. Hu, Semi-implicit euler-maruyama scheme for stiff stochastic equations, in: H. Körezliouglu, B. Oksendal, A. S. Üstünel (Eds.), Stochastic Analysis and Related Topics V, Birkhäuser Boston, Boston, MA, 1996, pp. 183–202.
[60] A. Abdulle, T. Li, S-rock methods for stiff itô sdes, Commun. Math. Sci. 6 (2008) 845 – 868.
[61] Z. Yin, S. Gan, An error corrected euler–maruyama method for stiff stochastic differential equations, Applied Mathematics and Computation 256 (2015) 630 – 641.
[62] G. N. Milstein, E. Platen, H. Schurz, Balanced implicit methods for stiff stochastic systems, SIAM J. Num. Anal. 35 (1998) 1010 – 1019.
[63] T. Tian, K. Burrage, Implicit taylor methods for stiff stochastic differential equations, Applied Numerical Mathematics 38 (2001) 167 – 185.
[64] W. Wang, S. Gan, D. Wang, A family of fully implicit milsterin methods for stiff stochastic differential equations with multiplicative noise, BIT Num. Math 52 (2012) 741 – 772.
[65] C. Ta, D. Wang, Q. Nie, An integration factor method for stochastic and stiff reaction–diffusion systems, Journal of Computational Physics 295 (2015) 505 – 522.
[66] Implicit simulation methods for stochastic chemical kinetics, Journal of Applied Analysis and Computation 5 (2015) 420 – 452.
[67] A. Abdulle, Explicit methods for stiff stochastic differential equations, in: B. Engquist, O. Runborg, Y.-H. R. Tsai (Eds.), Numerical Analysis of Multiscale Computations, Springer Berlin Heidelberg, Berlin, Heidelberg, 2012, pp. 1–22.
[68] P. Kloeden, A. Neuenkirch, Convergence of numerical methods for stochastic differential equations in mathematical finance, 2013, pp. 49–80.
[69] P. E. Kloeden, E. Platen, Higher-order implicit strong numerical schemes for stochastic differential equations, Journal of Statistical Physics 66 (1992) 283–314.
[70] T. Li, A. Abdulle, W. E, Effectiveness of implicit methods for stiff stochastic differential equations, Commun. Comput. Phys. 3 (2008) 295 – 307.
[71] D. T. Gillespie, The chemical langevin equation, J. Chem. Phys. 113 (2000) 297 – 306.
[72] R. Reams, Hadamard inverses, square roots and products of almost semidefinite matrices, Linear Algebra and its Applications 288 (1999) 35–43.

Appendix A. Details of proofs in convergence analysis

Throughout the analysis in this section, the notations of c_T and C_T are used for generic constants dependent on T but not Δt, that may change from line to line.

Appendix A.1. Proof of Lemma 1

We first consider the case with $p = 2$. Note that due to the equation (3.14) and Itô isometry we have

$$\mathbb{E}\left[\left|\hat{X}(s) - x_{n_s}\right|^2\right] \le M\mathbb{E}\left[\sum_{k=1}^{M}\Big(\int_{t_{n_s}}^{s} g_k(\hat{X}(\tau))\mathrm{d}W_k(\tau)\Big)^2\right] \le M\sum_{k=1}^{M}\mathbb{E}\left[\int_{t_{n_s}}^{s} g_k^2(\hat{X}(\tau))\mathrm{d}\tau\right], \quad s \ge 0.$$

Then by (4.8) we have

$$g_k^2(\hat{X}(\tau)) \le 2\left(L_k^2\left|\hat{X}(\tau) - x_{n_s}\right|^2 + 2g_k^2(0) + 2L_k^2 x_{n_s}^2\right), \quad \tau \in [t_s, s),$$

and thus due to Assumption (**A2**) we have

$$\begin{aligned}\mathbb{E}\left[\left|\hat{X}(s) - x_{n_s}\right|^2\right] &\le 2M\sum_{k=1}^{M}\int_{t_{n_s}}^{s}\left(L_k^2\mathbb{E}\left[\left|\hat{X}(\tau) - x_{n_s}\right|^2\right] + 2g_k^2(0) + 2L_k^2\mathbb{E}\left[x_{n_s}^2\right]\right)\mathrm{d}\tau \\ &\le 2M\sum_{k=1}^{M}L_k^2\int_{t_{n_s}}^{s}\mathbb{E}\left[\left|\hat{X}(\tau) - x_{n_s}\right|^2\right]\mathrm{d}\tau + 4M\sum_{k=1}^{M}\left(g_k^2(0) + L_k^2\Lambda_T\right)\Delta t, \quad s \ge 0.\end{aligned}$$

It then follows from Gronwall's inequality that

$$\mathbb{E}\left[\left|\hat{X}(s) - x_{n_s}\right|^2\right] \le 4M\sum_{k=1}^{M}\left(g_k^2(0) + L_k^2\Lambda_T\right)\Delta t e^{2M\sum_{k=1}^{M}L_k^2\Delta t} \le c_T\Delta t, \quad s \ge 0, \tag{A.1}$$

where c_T is a constant depends on M, Λ_T, T, L_k and $g_k(0)$ for $k = 1, \cdots, M$, but independent of Δt.

For $p > 2$, using again equation (3.14), Itô isometry, Hölder's inequality, and Cauchy-Schwarz, we have

$$\begin{aligned}\mathbb{E}\left[\left|\hat{X}(s) - x_{n_s}\right|^p\right] &\leq c_{p,M}\mathbb{E}\left[\sum_{k=1}^{M}\Big(\int_{t_{n_s}}^{s} g_k(\hat{X}(\tau))\mathrm{d}W_k(\tau)\Big)^p\right] \leq c_{p,M}\sum_{k=1}^{M}\mathbb{E}\left[\left(\int_{t_{n_s}}^{s} g_k^2(\hat{X}(\tau))\mathrm{d}\tau\right)^{p/2}\right] \\ &\leq \cdots\cdots \leq c_{p,M}(\Delta t)^{p/2-1}\mathbb{E}\left[\int_{t_{n_s}}^{s} g_k^p(\hat{X}(\tau))\mathrm{d}\tau\right], \quad s \geq 0,\end{aligned}$$

where here and below $c_{p,M}$ is a generic constant depending on p and M that may change from line to line. On the other hand by by (4.8) we have

$$g_k^p(\hat{X}(\tau)) \leq c_{p,M}\left(L_k^2\left|\hat{X}(\tau) - x_{n_s}\right|^2 + 2g_k^2(0) + 2L_k^2 x_{n_s}^2\right)^{p/2} \leq \cdots \leq c_{p,M}\left(L_k^p\left|\hat{X}(\tau) - x_{n_s}\right|^p + g_k^p(0) + L_k^p x_{n_s}^p\right),$$

and thus

$$\begin{aligned}\mathbb{E}\left[\left|\hat{X}(s) - x_{n_s}\right|^p\right] &\leq c_{p,M}(\Delta t)^{p/2-1}\sum_{k=1}^{M}\int_{t_{n_s}}^{s}\left(L_k^p\mathbb{E}\left[\left|\hat{X}(\tau) - x_{n_s}\right|^p\right] + g_k^p(0) + L_k^p\mathbb{E}\left[x_{n_s}^p\right]\right)\mathrm{d}\tau \\ &\leq c_{p,M}(\Delta t)^{p/2-1}\sum_{k=1}^{M}L_k^p\int_{t_{n_s}}^{s}\mathbb{E}\left[\left|\hat{X}(\tau) - x_{n_s}\right|^p\right]\mathrm{d}\tau + c_{p,M}(\Delta t)^{p/2}\sum_{k=1}^{M}\left(g_k^p(0) + L_k^p\Lambda_T\right).\end{aligned}$$

Similar to (A.1), applying Gronwall's Lemma to the above inequality results in

$$\mathbb{E}\left[\left|\hat{X}(s) - x_{n_s}\right|^p\right] \leq c_{p,M}(\Delta t)^{p/2}\sum_{k=1}^{M}\left(g_k^p(0) + L_k^p\Lambda_T\right)e^{(\Delta t)^{p/2}\sum_{k=1}^{M}L_k^p}.$$

which implies the desired assertion for all every number $p \geq 2$ with $C_{p,T} = c_{p,M}\sum_{k=1}^{M}\left(g_k^p(0) + L_k^p\Lambda_T\right)e^{\sum_{k=1}^{M}L_k^p}$, and the assumption without loss of generality that $\Delta t \leq 1$. The proof is complete.

Appendix A.2. Proof of Lemma 2

First, by Hölder's inequality and the Lipschitz condition on f,

$$\sup_{0\leq t\leq T}\mathcal{E}_1^2(t) \leq T\sup_{0\leq t\leq T}\int_0^t\left|f(X(s)) - f(\tilde{x}(s))\right|^2\mathrm{d}s \leq TL_f^2\sup_{0\leq t\leq T}\int_0^t\left|X(s) - \tilde{x}(s)\right|^2\mathrm{d}s.$$

Taking the expectation of the above inequality and using Doob's maximal inequality gives

$$\begin{aligned}\mathbb{E}\left[\sup_{0\leq t\leq T}\mathcal{E}_1^2(t)\right] &\leq 4TL_f^2\mathbb{E}\left[\int_0^T\left|X(s) - \tilde{x}(s)\right|^2\mathrm{d}s\right] \\ &\leq 4TL_f^2\left(\int_0^T\mathbb{E}\Big[\sup_{0\leq t\leq s}\mathcal{E}_0^2(t)\Big]\mathrm{d}s + \int_0^T\mathbb{E}\left[\left|y(s) - \tilde{x}(s)\right|^2\right]\mathrm{d}s\right).\end{aligned} \tag{A.2}$$

Using (4.9) to obtain the term $y(s) - \tilde{x}(s)$, and squaring it, gives

$$\begin{aligned}|y(s) - \tilde{x}(s)|^2 &\leq (m+1)\left(f^2(x_{n_s})(\Delta t)^2 + \sum_{k=M+1}^{m}g_k^2(x_{n_s})\left(W_k(s) - W_k(t_{n_s})\right)^2\right) \\ &\quad +(m+1)\sum_{k=1}^{M}\left(\int_{t_{n_s}}^{s}\left(g_k(x_{n_s}) + g_k'(x_{n_s})(\hat{X}(\tau) - x_{n_s})\mathrm{d}W_k(\tau)\right)^2.\end{aligned}$$

Then taking the expectation of the above inequality, and using (4.8) and Itô isometry, we deduce

$$\begin{aligned}
\mathbb{E}\left[|y(s)-\tilde{x}(s)|^2\right] &\leq 2(m+1)\left((\Delta t)^2(f^2(0)+L_f^2\mathbb{E}[x_{n_s}^2])+\sum_{k=M+1}^{m}\mathbb{E}\Big[(g_k^2(0)+L_k^2x_{n_s}^2)(W_k(s)-W_k(t_{n_s}))^2\Big]\right)\\
&\quad +2(m+1)\sum_{k=1}^{M}\mathbb{E}\left[\int_{t_{n_s}}^{s}\Big(g_k^2(x_{n_s})+(g_k'(x_{n_s}))^2(\hat{X}(\tau)-x_{n_s})^2\Big)\mathrm{d}\tau\right]\\
&\leq 2(m+1)\left((\Delta t)^2(f^2(0)+L_f^2\mathbb{E}[x_{n_s}^2])+\sum_{k=1}^{m}\Big((g_k^2(0)+L_k^2\mathbb{E}[x_{n_s}^2])\Delta t\Big)\right)\\
&\quad +2(m+1)\mathbb{E}\left[\int_{t_{n_s}}^{s}(g_k'(x_{n_s}))^2(\hat{X}(\tau)-x_{n_s})^2\mathrm{d}\tau\right]. \qquad \text{(A.3)}
\end{aligned}$$

Using Hölder's inequality, Lemma 1, Assumption (**A2**) and (**A3**), the last term of (A.3) satisfies

$$\begin{aligned}
\mathbb{E}\left[\int_{t_{n_s}}^{s}(g_k'(x_{n_s}))^2(\hat{X}(\tau)-x_{n_s})^2\mathrm{d}\tau\right] &\leq \Big(\mathbb{E}[(g_k'(x_{n_s}))^4]\Big)^{1/2}\left(\mathbb{E}\Big[\Big(\int_{t_{n_s}}^{s}(\hat{X}(\tau)-x_{n_s})^2\mathrm{d}\tau\Big)^2\Big]\right)^{1/2}\\
&\leq \lambda^2\Big(\mathbb{E}[(1+|x_{n_s}|^\gamma)^4]\Big)^{1/2}\left(\Delta t\int_{t_{n_s}}^{s}\mathbb{E}\Big[(\hat{X}(\tau)-x_{n_s})^4\Big]\mathrm{d}\tau\right)^{1/2}\leq c_T(\Delta t)^2, \qquad \text{(A.4)}
\end{aligned}$$

where c_T depends on T, λ, γ, Λ_T, L_k and $g_k(0)$ for $k=1,\cdots,m$, but independent of Δt.

Now inserting (A.4) into (A.3), using Assumption (**A2**), and integrating from 0 to T gives

$$\int_0^T\mathbb{E}\left[|y(s)-\tilde{x}(s)|^2\right]\mathrm{d}s\leq 2(m+1)T\left((\Delta t)^2\big(f^2(0)+L_f^2\Lambda_T+c_T\big)+\Delta t\sum_{k=1}^{m}\Big(g_k^2(0)+L_k^2\Lambda_T\Big)\right). \qquad \text{(A.5)}$$

Consequently, (A.2) can be further estimated to satisfy

$$\mathbb{E}\left[\sup_{0\leq t\leq T}\mathcal{E}_1^2(t)\right]\leq 4TL_f^2\int_0^T\mathbb{E}\Big[\sup_{0\leq t\leq s}\mathcal{E}_0^2(t)\Big]\mathrm{d}s+2(m+1)T^2L_f^2\Big((\Delta t)^2c_1+\Delta t c_2\Big),$$

with $c_1=f^2(0)+L_f^2\Lambda_T+c_T$ and $c_2=\sum_{k=1}^{m}(g_k^2(0)+L_k^2\Lambda_T)$. Setting $C_T=2(m+1)T^2L_f^2\max\{c_1,c_2\}$ implies the desired assertion. The proof is complete.

Appendix A.3. Proof of Lemma 3

First by Cauchy-Schwarz, Doob's martingale maximal inequality, Itô's isometry, and Assumption (**A1**) we have

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\leq t\leq T}\mathcal{E}_2^2(t)\right] &\leq (m-M)\mathbb{E}\left[\sup_{0\leq t\leq T}\sum_{k=M+1}^{m}\left|\int_0^t(g_k(X(s))-g_k(\tilde{x}(s)))\mathrm{d}W_k(s)\right|^2\right]\\
&\leq 4(m-M)\mathbb{E}\left[\sum_{k=M+1}^{m}\left|\int_0^T(g_k(X(s))-g_k(\tilde{x}(s)))\mathrm{d}W_k(s)\right|^2\right]\\
&= 4(m-M)\mathbb{E}\left[\sum_{k=M+1}^{m}\int_0^T\big(g_k(X(s))-g_k(\tilde{x}(s))\big)^2\mathrm{d}s\right]\\
&\leq 4(m-M)\sum_{k=M+1}^{m}L_k^2\mathbb{E}\left[\int_0^T\big|X(s)-\tilde{x}(s)\big|^2\mathrm{d}s\right].
\end{aligned}$$

Similar to (A.2) in Lemma 2,

$$\mathbb{E}\left[\int_0^T\big|X(s)-\tilde{x}(s)\big|^2\mathrm{d}s\right]\leq\int_0^T\mathbb{E}\Big[\sup_{0\leq t\leq s}\mathcal{E}_0^2(t)\Big]\mathrm{d}s+\mathbb{E}\left[\int_0^T\big|y(s)-\tilde{x}(s)\big|^2\mathrm{d}s\right]$$

and then it follows from the estimate (A.5) that

$$\mathbb{E}\left[\int_0^T \left|X(s)-\tilde{x}(s)\right|^2 \mathrm{d}s\right] \le \int_0^T \mathbb{E}\Big[\sup_{0\le t\le s}\mathcal{E}_0^2(t)\Big]\mathrm{d}s + 2(m+1)T\sum_{k=M+1}^{m} L_k^2\left((\Delta t)^2 c_1 + \Delta t c_2\right),$$

where c_1 and c_2 are the same as in Lemma 2. Setting $C_T = 4(m-M)(2m+1)T\sum_{k=M+1}^{m} L_k^2 \max\{c_1, c_2\}$ implies the desired assertion. The proof is complete.

Appendix A.4. Proof of Lemma 4

First, it follows from Cauchy-Schwarz, Doob's martingale maximal inequality, and Itô's isometry that

$$\mathbb{E}\left[\sup_{0\le t\le T}\mathcal{E}_3^2(t)\right] \le 4M\sum_{k=1}^{M}\int_0^T \mathbb{E}\left[\Big(g_k(X(s)) - g_k(\tilde{x}(s)) - g_k'(\tilde{x}(s))(\hat{X}(s)-\tilde{x}(s))\Big)^2\right]\mathrm{d}s. \tag{A.6}$$

Note that due to Rolle's theorem, and then by Assumption (**A4**), for every $s\in\mathbb{R}$ there exists y_s between $\hat{X}(s)$ and x_{n_s} such that

$$\left|g_k(\hat{X}(s)) - g_k(x_{n_s}) - g_k'(\tilde{x}(s))(\hat{X}(s)-x_{n_s})\right| = \left|(g'(y_s) - g_k'(x_{n_s}))(\hat{X}(s)-x_{n_s})\right| \le D_k(\hat{X}(s)-x_{n_s})^2. \tag{A.7}$$

Writing $g_k(X(s)) - g_k(\tilde{x}(s))$ in (A.6) as $g_k(X(s)) - g_k(\hat{X}(s)) + g_k(\hat{X}(s)) - g_k(x_{n_s})$, it then follows from (A.7), Lemma 1, and Assumption (**A1**) that

$$\begin{aligned}
\mathbb{E}\left[\sup_{0\le t\le T}\mathcal{E}_3^2(t)\right] &\le 8M\sum_{k=1}^{M}\left(\int_0^T \mathbb{E}\Big[\big(g_k(X(s)) - g_k(\hat{X}(s))\big)^2\Big]\mathrm{d}s + D_k^2\int_0^T \mathbb{E}\Big[(\hat{X}(s)-x_{n_s})^4\Big]\mathrm{d}s\right)\\
&\le 8M\sum_{k=1}^{M}\left(L_k^2\int_0^T \mathbb{E}\left[\big(X(s)-\hat{X}(s)\big)^2\right]\mathrm{d}s + D_k^2 T C_{4,T}(\Delta t)^2\right),
\end{aligned} \tag{A.8}$$

where $C_{4,T}$ is the constant in Lemma 1 which is independent of Δt.

It remains to estimate $\mathbb{E}\left[\left|X(s)-\hat{X}(s)\right|^2\right]$. In fact, by equations (2.1) and (2.5), Cauchy inequality and Hölder inequality we have

$$\begin{aligned}
\mathbb{E}\left[\left|X(s)-\hat{X}(s)\right|^2\right] &\le (m+1)\mathbb{E}\left[\left|\int_{t_{n_s}}^{s} f(X(\tau))\mathrm{d}\tau\right|^2\right] + (m+1)\sum_{k=M+1}^{m}\mathbb{E}\left[\left|\int_{t_{n_s}}^{s} g_k(X(\tau))\mathrm{d}W_k(\tau)\right|^2\right]\\
&\quad +(m+1)\sum_{k=1}^{M}\mathbb{E}\left[\left|\int_{t_{n_s}}^{s}\big(g_k(X(\tau)) - g_k(\hat{X}(\tau))\big)\mathrm{d}W_k(\tau)\right|^2\right]\\
&\le (m+1)\left(\Delta t\int_{t_{n_s}}^{s}\mathbb{E}[f^2(X(\tau))]\mathrm{d}\tau + \sum_{k=M+1}^{m}\int_{t_{n_s}}^{s}\mathbb{E}[g_k^2(X(\tau))]\mathrm{d}\tau\right)\\
&\quad +(m+1)\sum_{k=1}^{M}\int_{t_{n_s}}^{s}\mathbb{E}[(g_k(X(\tau)) - g_k(\hat{X}(\tau)))^2]\mathrm{d}\tau.
\end{aligned} \tag{A.9}$$

By Assumption (**A1**), $f^2(X(\tau)) \le 2f^2(x_{n_s}) + 2L_f^2|X(\tau)-x_{n_s}|^2$ and $g_k^2(X(\tau)) \le 2g_k^2(x_{n_s}) + 2L_k^2|X(\tau)-x_{n_s}|^2$ and thus

$$\int_{t_{n_s}}^{s}\mathbb{E}[f^2(X(\tau))]\mathrm{d}\tau \le 2\Delta t\left(2f^2(0) + 2L_f^2\Lambda_T + L_f^2\mathbb{E}\left[\sup_{0\le t\le s}\mathcal{E}_0^2(t)\right]\right), \tag{A.10}$$

$$\int_{t_{n_s}}^{s}\mathbb{E}[g_k^2(X(\tau))]\mathrm{d}\tau \le 2\Delta t\left(2g_k^2(0) + 2L_k^2\Lambda_T + L_k^2\mathbb{E}\left[\sup_{0\le t\le s}\mathcal{E}_0^2(t)\right]\right). \tag{A.11}$$

Inserting (A.10)–(A.11) into (A.9) and using Assumption (**A1**) again we obtain

$$\mathbb{E}\left[\left|X(s)-\hat{X}(s)\right|^2\right] \le c_1\left[\sup_{0\le t\le s}\mathcal{E}_0^2(t)\right] + c_2 + (m+1)\sum_{k=1}^{M} L_k^2 \int_{t_{n_s}}^{s} \mathbb{E}[|X(\tau)-\hat{X}(\tau)|^2]\mathrm{d}\tau,$$

where due to Assumption (**A2**)

$$c_1 = 2(m+1)\Delta t\left(\Delta t L_f^2 + \sum_{M+1}^{m} L_k^2\right), \quad c_2 = 2(m+1)\Delta t\left(\Delta t + 2\Delta t(f^2(0)+L_f^2\Lambda_T) + 2\sum_{M+1}^{m}(g_k^2(0)+L_k^2\Lambda_T)\right).$$

It then follows from Gronwall's inequality that

$$\begin{aligned}\mathbb{E}\left[\left|X(s)-\hat{X}(s)\right|^2\right] &\le c_1\left[\sup_{0\le t\le s}\mathcal{E}_0^2(t)\right] + c_2 + (m+1)\sum_{k=1}^{M} L_k^2\int_{t_{n_s}}^{s}\left(c_1\Big[\sup_{0\le t\le \tau}\mathcal{E}_0^2(t)\Big]+c_2\right)e^{(m+1)\sum_{k=1}^{M}L_k^2(s-\tau)}\mathrm{d}\tau \\ &\le \left(c_1\left[\sup_{0\le t\le s}\mathcal{E}_0^2(t)\right]+c_2\right)e^{(m+1)\sum_{k=1}^{M}L_k^2\Delta t} \le c\Delta t\left[\sup_{0\le t\le s}\mathcal{E}_0^2(t)\right] + c_T\Delta t, \end{aligned} \tag{A.12}$$

where $c = 2(m+1)(\Delta t L_f^2 + \sum_{k=M+1}^{M} L_k^2)e^{(m+1)\sum_{k=1}^{M}L_k^2}$ and c_T is a generic constant dependent on Λ_T, m, $f^2(0)$, $g_k^2(0)$, L_f^2, L_k^2 but independent of Δt. Finally, inserting (A.12) into (A.8) results in the desired assertion by setting

$$C = 16M(m+1)\left(\Delta t L_f^2 + \sum_{k=M+1}^{M} L_k^2\right)e^{(m+1)\sum_{k=1}^{M}L_k^2}\sum_{k=1}^{M} L_k^2. \tag{A.13}$$

The proof is complete.

Appendix A.5. Proof of Lemma 6

Notice that $y(t_{n+1}) - x_{n+1} = Y_n(t_{n+1}) - \eta(t_{n+1})$, in which $\eta(t)$ is defined in (4.3). Hence to estimate the second term on the right hand side of (4.5), we also consider integral representation of the piecewise continuous process $\eta(t)$

$$\eta(t) = \phi_n(t)\sum_{k=1}^{M}\left(b_{k,n}\Big(-a_{k,n}\int_{t_n}^{t}\mathrm{d}s + \int_{t_n}^{t}\mathrm{d}W_k(s)\Big)\right), \quad t\in[t_n,t_{n+1}), \tag{A.14}$$

where $\phi_n(t)$ is defined in (3.17). Then

$$|\eta(t)-Y_n(t)| = \phi_n(t)\sum_{k=1}^{M}\left|b_{k,n}\Big(-a_{k,n}\int_{t_n}^{t}(\phi_n^{-1}(s)-1)\mathrm{d}s + \int_{t_n}^{t}(\phi_n^{-1}(s)-1)\mathrm{d}W_k(s)\Big)\right|, \quad t\in[t_n,t_{n+1}),$$

and it follows from Cauchy-Bunyakovsky-Schwarz inequality and Hölder's inequality that

$$\begin{aligned}\Big(\mathbb{E}|\eta(t)-Y_n(t)|\Big)^2 &\le 2\mathbb{E}\left[\phi_n^2(t)\right]\mathbb{E}\left[\sum_{k=1}^{M}\left|b_{k,n}\Big(-a_{k,n}\int_{t_n}^{t}(\phi_n^{-1}(s)-1)\mathrm{d}s + \int_{t_n}^{t}(\phi_n^{-1}(s)-1)\mathrm{d}W_k(s)\Big)\right|^2\right] \\ &\le 2M\mathbb{E}\left[\phi_n^2(t)\right]\sum_{k=1}^{M}\mathbb{E}\left[b_{k,n}^2\left(a_{k,n}^2\Delta t\int_{t_n}^{t}(\phi_n^{-1}(s)-1)^2\mathrm{d}s + \Big(\int_{t_n}^{t}(\phi_n^{-1}(s)-1)\mathrm{d}W_k(s)\Big)^2\right)\right]. \end{aligned} \tag{A.15}$$

Noting that at each t, $\phi_n(t)$ follows a log-normal distribution, i.e.,

$$\ln\phi_n(t) \sim \mathcal{N}\left(-\frac{1}{2}\sum_{k=1}^{M}a_{k,n}^2(t-t_n), \sum_{k=1}^{M}a_{k,n}^2(t-t_n)\right), \tag{A.16}$$

then by using Assumptions (**A3**) and (**A2**), the first term on the right hand side of (A.15) satisfies

$$\begin{aligned}\mathbb{E}\left[\phi_n^2(t)\right] = e^{\sum_{k=1}^M a_{k,n}^2(t-t_n)} &\leq 1+\sum_{k=1}^M \mathbb{E}\left[(g_k'(x_n))^2\right]\Delta t + O(\Delta t^2)\\ &\leq 1+\lambda^2\sum_{k=1}^M \left(\mathbb{E}[(1+|x_n|^\gamma)^4]\right)^{1/2}\Delta t + O(\Delta t^2) \leq c. \end{aligned} \tag{A.17}$$

Then using equation (A.16) again, the two integrals on the right hand side of (A.15) satisfy respectively,

$$\mathbb{E}\left[\int_{t_n}^t (\phi_n^{-1}(s)-1)^2 \mathrm{d}s\right] \leq 2\int_{t_n}^t \mathbb{E}\left[\phi_n^{-2}(s)+1\right]\mathrm{d}s \leq 2(c+1)\Delta t, \tag{A.18}$$

$$\mathbb{E}\left[\Big(\int_{t_n}^t (\phi_n^{-1}(s)-1)\mathrm{d}W_k(s)\Big)^2\right] = \mathbb{E}\left[\int_{t_n}^t (\phi_n^{-1}(s)-1)^2\mathrm{d}s\right] \leq 2(c+1)\Delta t. \tag{A.19}$$

Inserting (A.17) – (A.19) into (A.15), and using again the boundedness of $\mathbb{E}[b_{k,n}^2]$ and $\mathbb{E}[a_{k,n}^2]$ implied by Assumptions (**A1**) – (**A3**) we obtain

$$\begin{aligned}\Big(\mathbb{E}[|\eta(t)-Y_n(t)|]\Big)^2 &\leq 2Mc\left(\Delta t\mathbb{E}[b_{k,n}^2]\mathbb{E}[a_{k,n}^2]\mathbb{E}\left[\int_{t_n}^t(\phi_n^{-1}(s)-1)^2\mathrm{d}s\right]+\mathbb{E}[b_{k,n}^2]\mathbb{E}\left[\Big(\int_{t_n}^t(\phi_n^{-1}(s)-1)\mathrm{d}W_k(s)\Big)^2\right]\right)\\ &\leq 2Mc\left((\Delta t)^2+\Delta t\right), \quad t\in[t_n,t_{n+1}),\ n=0,\cdots,N,\end{aligned}$$

in which c is a generic constant independent of Δt and may be different from line to line. Therefore

$$\max_{n=1,\cdots,N}\mathbb{E}[|y(t_n)-x_n|] \leq \sup_{0\leq t\leq T}\mathbb{E}[|\eta(t)-Y(t)|] \leq C(\Delta t)^{1/2},$$

in which c is a constant depending on M, T, Λ_T, L_k, D_k, but is independent of Δt. The proof is complete.

Appendix A.6. Proof of Theorem 2

Note that since $x(t_n)=x_n$ for $n=0,1,\cdots,N$, the weak discretization error above satisfies

$$\mathfrak{E} := \left|\mathbb{E}[\psi(x(T))]-\mathbb{E}[\psi(X(T))]\right| \leq \left|\mathbb{E}[\psi(y(T))]-\mathbb{E}[\psi(X(T))]\right| + \left|\mathbb{E}[\psi(x(T))]-\mathbb{E}[\psi(y(T))]\right|,$$

where $x(t)$ and $y(t)$ satisfy (4.2) and (4.6), respectively. Similar to the proof of strong convergence, we will first estimate $\mathfrak{E}_1 = \left|\mathbb{E}[\psi(y(T))]-\mathbb{E}[\psi(X(T))]\right|$. To that end, let $u(t,y)$ be a solution of the following Feynman-Kac partial differential equation

$$u_t(t,y)+f(y)u_y(t,y)+\frac{1}{2}u_{yy}(t,y)\sum_{k=1}^m g_k^2(y)=0 \quad \text{for } t\in[0,T],\ y\in\mathbb{R}, \quad \text{with } u(T,y)=\psi(y).$$

Applying Itô's formula to $u(t,y(t))$ with $y(t)$ satisfying (4.6) and using the above equation yields

$$\begin{aligned}\mathrm{d}u(t,y(t)) &= \Big(u_t(t,y(t))+u_y(t,y(t))f(\tilde{x}(t))\Big)\mathrm{d}t\\ &\quad+\frac{1}{2}u_{yy}(t,y(t))\left(\sum_{k=M+1}^m g_k^2(\tilde{x}(t))+\sum_{k=1}^M\Big(g_k(\tilde{x}(t))+g_k'(\tilde{x}(t))(\hat{X}(t)-\tilde{x}(t))\Big)^2\right)\mathrm{d}t\\ &\quad+u_y(t,y(t))\left(\sum_{k=M+1}^m g_k(\tilde{x}(t))\mathrm{d}W_k(t)+\sum_{k=1}^M\Big(g_k(\tilde{x}(t))+g_k'(\tilde{x}(t))(\hat{X}(t)-\tilde{x}(t))\Big)\mathrm{d}W_k(t)\right)\end{aligned}$$

$$
\begin{aligned}
\mathrm{d}u(t,y(t)) &= \left(u_y(t,y(t))\,(f(\tilde{x}(t))-f(y(t)) + \frac{1}{2}u_{yy}(t,y(t))\left(\sum_{k=M+1}^{m}\left(g_k^2(\tilde{x}(t))-g_k^2(y(t))\right)\right)\right)\mathrm{d}t \\
&\quad + \frac{1}{2}u_{yy}(t,y(t))\sum_{k=1}^{M}\left(\left(g_k(\tilde{x}(t))+g_k'(\tilde{x}(t))(\hat{X}(t)-\tilde{x}(t))\right)^2 - g_k^2(y(t))\right)\mathrm{d}t \\
&\quad + u_y(t,y(t))\left(\sum_{k=M+1}^{m} g_k(\tilde{x}(t))\mathrm{d}W_k(t) + \sum_{k=1}^{M}\left(g_k(\tilde{x}(t))+g_k'(\tilde{x}(t))(\hat{X}(t)-\tilde{x}(t))\right)\mathrm{d}W_k(t)\right).
\end{aligned}
$$

Note that due to the Feynman-Kac formula, we have $u(0,x_0)=\mathbb{E}[\psi(X(T))]$. Then, integrating the above equation from 0 to T using $u(T,x_N)=\psi(x_N)$ and taking the expectation of the resulting equation gives

$$
\begin{aligned}
\mathbb{E}[\psi(x_N)]-\mathbb{E}[\psi(y(T))] &= \mathbb{E}\left[\int_0^T\left(u_y(t,y(t))\,(f(\tilde{x}(t))-f(y(t))) + \frac{1}{2}u_{yy}(t,y(t))\sum_{k=M+1}^{m}\left(g_k^2(\tilde{x}(t))-g_k^2(y(t))\right)\right)\mathrm{d}t\right] \\
&\quad + \frac{1}{2}\mathbb{E}\left[\int_0^T u_{yy}(t,y(t))\sum_{k=1}^{M}\left(\left(g_k(\tilde{x}(t))+g_k'(\tilde{x}(t))(\hat{X}(t)-\tilde{x}(t))\right)^2 - g_k^2(y(t))\right)\mathrm{d}t\right],
\end{aligned}
$$

which implies that

$$
\mathfrak{E}_1 \le \int_0^T |\mathbb{E}[\mathrm{e}_1(t,y(t))]|\,\mathrm{d}t + \int_0^T |\mathbb{E}[\mathrm{e}_2(t,y(t))]|\,\mathrm{d}t + \int_0^T |\mathbb{E}[\mathrm{e}_3(t,y(t))]|\,\mathrm{d}t \tag{A.20}
$$

where

$$
\begin{aligned}
\mathrm{e}_1(t,y(t)) &= u_y(t,y(t))\,(f(\tilde{x}(t))-f(y(t))), \\
\mathrm{e}_2(t,y(t)) &= u_{yy}(t,y(t))\sum_{k=M+1}^{m}\left(g_k^2(\tilde{x}(t))-g_k^2(y(t))\right) \\
\mathrm{e}_3(t,y(t)) &= u_{yy}(t,y(t))\sum_{k=1}^{M}\left(\left(g_k(\tilde{x}(t))+g_k'(\tilde{x}(t))(\hat{X}(t)-\tilde{x}(t))\right)^2 - g_k^2(y(t))\right).
\end{aligned}
$$

Note that $\mathrm{e}_1(t_n,y(t_n))=\mathrm{e}_2(t_n,y(t_n))=\mathrm{e}_3(t_n,y(t_n))=0$. We next estimate each of e_1, e_2 and e_3.

First apply the Itô formula to $\mathrm{e}_1(t,x(t))$ to obtain

$$
\mathrm{d}\mathrm{e}_1(t,y(t)) = \left(\frac{\partial \mathrm{e}_1}{\partial t} + \frac{\partial \mathrm{e}_1}{\partial y}f(\tilde{x}(t)) + \frac{1}{2}\frac{\partial^2 \mathrm{e}_1}{\partial y^2}\Big(\sum_{k=M+1}^{m} g_k^2(\tilde{x}(t)) + \sum_{k=1}^{M}\left(g_k(\tilde{x}(t))+g_k'(\tilde{x}(t))(\hat{X}(t)-\tilde{x}(t))\right)^2\Big)\right)\mathrm{d}t + \frac{\partial \mathrm{e}_1}{\partial y}\sum_{k=1}^{m}\mathcal{G}_k\mathrm{d}W_k(t), \tag{A.21}
$$

where

$$
\mathcal{G}_k = \begin{cases} g_k(\tilde{x}(t))+g_k'(\tilde{x}(t))(\hat{X}(t)-\tilde{x}(t)), & \text{for } k=1,\cdots,M, \\ g_k(\tilde{x}(t)), & \text{for } k=M+1,\cdots,m. \end{cases}
$$

Integrating (A.21) from t_n to $t\in[t_n,t_{n+1})$ using $\mathrm{e}_1(t_n,y(t_n))=0$ then taking expectation of the resulting equation gives

$$
\begin{aligned}
\mathbb{E}[\mathrm{e}_1(t,y(t)] &= \int_{t_n}^{t}\left(\mathbb{E}\Big[\frac{\partial \mathrm{e}_1}{\partial t}(s,y(s))\Big] + \mathbb{E}\Big[\frac{\partial \mathrm{e}_1}{\partial y}(s,y(s))f(x_n)\Big] + \frac{1}{2}\mathbb{E}\Big[\frac{\partial^2 \mathrm{e}_1}{\partial y^2}\sum_{k=M+1}^{m} g_k^2(x_n)\Big]\right)\mathrm{d}s \\
&\quad + \frac{1}{2}\int_{t_n}^{t}\mathbb{E}\Big[\frac{\partial^2 \mathrm{e}_1}{\partial y^2}\sum_{k=1}^{M}\left(g_k(x_n)+g_k'(x_n)(\hat{X}(s)-x_n)\right)^2\Big]\mathrm{d}s.
\end{aligned}
$$

Then by Cauchy inequality, Itô isometry, and equation (2.5) we have

$$
\mathbb{E}[\mathrm{e}_1(t,y(t)] \le \int_{t_n}^{t}\left(\mathbb{E}\Big[\frac{\partial \mathrm{e}_1}{\partial t}(s,y(s))\Big] + \mathbb{E}\Big[\frac{\partial \mathrm{e}_1}{\partial x}(s,y(s))f(x_n)\Big] + \mathbb{E}\Big[\frac{\partial^2 \mathrm{e}_1}{\partial x^2}\Big(\sum_{k=1}^{m} g_k^2(x_n) + \sum_{k=1}^{M}(g'(x_n))^2\int_{t_n}^{s}\sum_{k=1}^{M} g_k^2(\hat{X}(\tau))\mathrm{d}\tau\Big)\Big]\right)\mathrm{d}s.
$$

For simplicity, assume that the functions f and g_k satisfy conditions such that all expectations appearing in the above inequality are bounded. Then there exists $C_1 > 0$ such that

$$|\mathbb{E}[\mathfrak{e}_1(t, x(t)]| \leq C_1 \Delta t. \tag{A.22}$$

Following similar analysis we can obtain that there exist $C_2 > 0$ and $C_3 > 0$ such that

$$|\mathbb{E}[\mathfrak{e}_2(t, x(t)]| \leq C_2 \Delta t, \quad |\mathbb{E}[\mathfrak{e}_3(t, x(t)]| \leq C_3 \Delta t. \tag{A.23}$$

In the end, inserting (A.22)–(A.23) into (A.20) we have that there exists $C_T > 0$ such that $\mathfrak{E}_1 \leq C_T \Delta t$.

Note that in the FPM-LP scheme, $x(t)$ is essentially an EM approximation of $y(t)$, along with an exponential approximation of a strong convergence order 3/2 implied by (4.12). It then follows immediately that there exists $C_T > 0$ such that the error

$$\mathfrak{E}_2 := \left|\mathbb{E}[\psi(x(T))] - \mathbb{E}[\psi(y(T))]\right| \leq C_T \Delta t,$$

which implies that the FPM-LP scheme has a weak order of convergence 1. The proof is complete.