A BOUNDARY PRESERVING NUMERICAL SCHEME FOR THE WRIGHT-FISHER MODEL

I. S. STAMATIOU

Abstract. We are interested in the numerical approximation of non-linear stochastic differential equations (SDEs) with solution in a certain domain. Our goal is to construct explicit numerical schemes that preserve that structure. We generalize the semi-discrete method Halidias N. and Stamatiou I.S. (2016), On the numerical solution of some non-linear stochastic differential equations using the Semi-Discrete method, Computational Methods in Applied Mathematics,16(1) and propose a numerical scheme, for which we prove a strong convergence result, to a class of SDEs that appears in population dynamics and ion channel dynamics within cardiac and neuronal cells. We furthermore extend our scheme to a multidimensional case.

1. Introduction

Let $T > 0$ and $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbb{P})$ be a complete probability space and let $W_{t,\omega}: [0,T] \times \Omega \rightarrow \mathbb{R}$ be a one-dimensional Wiener process adapted to the filtration $\{\mathcal{F}_t\}_{0 \leq t \leq T}$. We are interested in the numerical approximation of the following scalar stochastic differential equation (SDE),

$$x_t = x_0 + \int_0^t (k_1 - k_2 x_s)ds + k_3 \int_0^t \sqrt{x_s(1-x_s)}dW_s,$$

where $k_i > 0, i = 1,2,3$. A boundary classification result, see Appendix A, implies that $0 < x_t < 1$ a.s. when $x_0 \in (0,1)$ and $0 < k_1 < k_2$. We therefore aim for a numerical scheme which apart from strongly converging to the true solution of (1.1), produces values in the same domain, i.e. in $(0,1)$. In other words, we are interested in numerical schemes that have an eternal life time.

**Definition 1.1** [Eternal Life time of numerical solution] Let $D \subseteq \mathbb{R}^d$ and consider a process $(X_t)$ well defined on the domain $D$, with initial condition $X_0 \in D$ and such that

$$\mathbb{P}(\{\omega \in \Omega : X(t,\omega) \in D\}) = 1,$$

for all $t > 0$. A numerical solution $(Y_n)_{n \in \mathbb{N}}$ has an eternal life time if

$$\mathbb{P}(Y_{n+1} \in D | Y_n \in D) = 1.$$

In [1] the main interest is in the domain $D = \mathbb{R}_+$. Moreover, it is clear that the Euler-Maruyama scheme has always a finite life time.
The proposed semi-discrete (SD) iterative scheme for the numerical approximation of (1.1) reads
\begin{equation}
    y_{t_{n+1}}^{SD} = \sin^2 \left( \frac{k_3}{2} \Delta W_n + \arcsin(\sqrt{y_n}) \right),
\end{equation}
where
\begin{align*}
    y_n &:= y_{t_n} + \left( k_1 - \frac{(k_3)^2}{4} \right) + y_n \left( \frac{(k_3)^2}{2} - k_2 \right) \cdot \Delta \\
\end{align*}
and \( \Delta W_n := W_{t_{n+1}} - W_{t_n} \), are the increments of the Wiener process and the discretization step \( \Delta \) is such that (1.2) is well-defined. By construction, the SD scheme (1.2) possesses an eternal life time. To get (1.2) we use an additive semi-discretization of the drift coefficient.

Briefly saying, a part of the SDE is discretized in a certain way such that the resulting SDE to be solved has an analytical solution (see details in Section 2). This is also a special feature of the method, since in the derivation of it, instead of an algebraic equation a new SDE has to be solved. The SD method can also reproduce the Euler scheme. The semi-discrete method was originally proposed in [2].

An attempt in that direction, i.e. in constructing explicit numerical schemes with an eternal life time, has been made in [3] where a class of one-dimensional SDEs with non-negative solutions is treated, which covers cases like that of the Heston 3/2-model, a popular model in the field of financial mathematics which is super-linear. The case of sub-linearities is also treated in [4] where the domain is still \( \mathbb{R}_+ \).

The purpose of this paper is to generalize further the method to preserve the structure of the original SDE. In the previous works, the suggested schemes preserve positivity; all the quantities appearing belong to the field of finance and are meant to be non-negative. The application that motivated us now, is used in population dynamics to describe fluctuations in gene frequency of reproducing individuals among finite populations [5] and in a different setting for the description of the random behavior of ion channels within cardiac and neuronal cells (cf. [6], [7], [8] and references therein). We are able in that case to preserve the domain of the original process. In fact, in applications in biology we have to solve systems of SDEs. The extension of the Wright-Fisher model to the multidimensional case has been proposed in [9] and [10]. Here, we will treat a three-state system as in [6, Sec. 6] given by the following system of SDEs
\begin{align*}
    X_t^{(1)} &= X_0^{(1)} + \int_0^t (k_1^{(1,1)} + k_1^{(1,2)} X_s^{(2)} - k_2^{(1)} X_s^{(1)}) ds + k_3^{(1,1)} \int_0^t \! \sqrt{X_s^{(1)}} X_s^{(2)} dW_s^{(1)} + k_3^{(1,2)} \int_0^t \! \sqrt{X_s^{(1)}} (1 - X_s^{(1)} - X_s^{(2)}) dW_s^{(2)}, \\
    X_t^{(2)} &= X_0^{(2)} + \int_0^t (k_1^{(2,1)} + k_1^{(2,2)} X_s^{(1)} - k_2^{(2)} X_s^{(2)}) ds + k_3^{(2,1)} \int_0^t \! \sqrt{X_s^{(1)}} X_s^{(2)} dW_s^{(1)} + k_3^{(2,2)} \int_0^t \! \sqrt{X_s^{(2)}} (1 - X_s^{(1)} - X_s^{(2)}) dW_s^{(2)},
\end{align*}
(1.4)
where \( X^{(i)} \) is the proportion of alleles or channels in state \( i, i = 1, 2, \) and \( 1 - X^{(1)} - X^{(2)} \) is the proportion in state 3.

In Section 2 we provide the setting and the main goal which concerns the mean-square convergence of the proposed structure-preserving SD scheme (1.2) for the approximation
of a modification of (1.1) with dynamics described by $\hat{W}$ (see (2.6)). We also discuss the multidimensional case (1.3)-(1.4).

In Section 3 we treat a more general class of SDEs. We further extend the analysis of the semi-discrete method introduced in [3]. We cover the sub-linear diffusion case and show as in [3, Th. 2.1] the strong convergence of the proposed numerical scheme to the true solution.

Section 4 is devoted to numerical experiments. The proofs of all the results are given in the sections to follow, that is in Sections 5 and 6.

2. The setting and the main goal.

Consider the partition $0 = t_0 < t_1 < \ldots < t_N = T$ with uniform discretization step $\Delta = T/N$ and the following process

$$y_{t_n}^{SD} = y_{t_n} + \int_{t_n}^{t_{n+1}} \left( k_1 - \frac{(k_3)^2}{4} + y_{t_n} \left( \frac{(k_3)^2}{2} - k_2 \right) \right) ds + \int_{t_n}^{t} \frac{(k_3)^2}{4} (1 - 2y_s) ds$$

$$+ k_3 \int_{t_n}^{t} \sqrt{y_s(1 - y_s)} \text{sgn}(z_s) dW_s,$$

for $t \in (t_n, t_{n+1}]$, with $y_0 = x_0$ a.s.

$$y_n := y_{t_n} + \left( k_1 - \frac{(k_3)^2}{4} + y_{t_n} \left( \frac{(k_3)^2}{2} - k_2 \right) \right) \cdot \Delta$$

and

$$z_t = \sin \left( k_3 \Delta W_n^t + 2 \arcsin(\sqrt{y_n}) \right),$$

where $\Delta W_n^t := W_t - W_{t_n}$. Process (2.1) has jumps at nodes $t_n$ of order $\Delta$ and the solution in each step is given by, see Appendix B,

$$y_{t_n}^{SD} = \sin^2 \left( \frac{k_3}{2} \Delta W_n^t + \arcsin(\sqrt{y_n}) \right),$$

which has the pleasant feature that $y_{t_n}^{SD} \in (0, 1)$ when $y_0 \in (0, 1)$. Process (2.4) is well defined when $0 < y_{t_n} < 1$, i.e. when

$$0 < y_{t_n} + \left( k_1 - \frac{(k_3)^2}{4} + y_{t_n} \left( \frac{(k_3)^2}{2} - k_2 \right) \right) \cdot \Delta < 1.$$

Therefore, we assume the following condition for the well-posedness of the SD scheme (2.4).

**Assumption 2.2** Let the discretization step $\Delta$ be such that (2.5) holds. $\square$

**Remark 2.3** Note that in general the discretization step $\Delta$ satisfying (2.5) is a r.v. depending on $\omega$. The $\omega$-dependence is inherited through the increments $\Delta W_n(\omega)$ which in turn affect the sequence $(y_{t_n})_{n \in \mathbb{N}}$. Nevertheless under the assumptions on the parameters considered later on the step $\Delta$ is not a r.v. but a fixed sufficiently small number. $\square$

Now, we consider the process

$$\hat{W}_t := \int_{0}^{t} \text{sgn}(z_s) dW_s,$$
which is a martingale with quadratic variation \(\langle \hat{W}_t, \hat{W}_t \rangle = t\) and thus a standard Brownian motion w.r.t. its own filtration, justified by Lévy’s theorem [11, Th. 3.16, p.157] and consequently (2.1) becomes

\[
y_t^{SD} = y_n + \int_{t_n}^{t} \frac{(k_3)^2}{4} (1 - 2y_s) ds + k_3 \int_{t_n}^{t} \sqrt{y_s(1 - y_s)} d\hat{W}_s,
\]

Moreover, consider

\[
\hat{x}_t = x_0 + \int_{0}^{t} (k_1 - k_2 \hat{x}_s) ds + k_3 \int_{0}^{t} \sqrt{\hat{x}_s(1 - \hat{x}_s)} d\hat{W}_s.
\]

The process \((x_t)\) of (1.1) and the process \((\hat{x}_t)\) of (2.8) have the same distribution. Our main goal is to deduce an estimate of the form

\[
\lim_{\Delta \downarrow 0} E \sup_{0 \leq t \leq T} |y_t^{SD} - x_t|^2 = 0.
\]

In Theorem 2.4 below, we deduce that

\[
\lim_{\Delta \downarrow 0} E \sup_{0 \leq t \leq T} |y_t^{SD} - \hat{x}_t|^2 = 0.
\]

By a simple application of the triangle inequality we deduce an analogous result for the unique solution of (1.1), i.e. \(\lim_{\Delta \downarrow 0} E \sup_{0 \leq t \leq T} |y_t^{SD} - x_t|^2 = 0\). We present in Appendix C the details. To simplify notation we write \(\hat{W}, (\hat{x}_t)\) as \(W, (x_t)\) respectively.

**Theorem 2.4** [Strong convergence] Let Assumption 2.2 hold. Then, the semi-discrete scheme (2.1) converges strongly in the mean-square sense to the true solution of (1.1), that is

\[
\lim_{\Delta \downarrow 0} E \sup_{0 \leq t \leq T} |y_t^{SD} - x_t|^2 = 0.
\]

As already noted in Remark 2.3, in order to apply Theorem 2.4 we have to find a sufficiently small step-size \(\Delta\) such that (2.5) holds, i.e.

\[
0 < y_{tn}(1 + \beta \Delta) + \alpha \Delta < 1.
\]

where

\[
\alpha := k_1 - \frac{(k_3)^2}{4}, \quad \beta := \frac{(k_3)^2}{2} - k_2.
\]

To simplify the conditions on \(\alpha, \beta, \Delta\), when necessary, we may adopt the following procedure. We consider a perturbation of order \(\Delta\) in the initial condition of (2.7), that is

\[
\tilde{y}_t^{SD} = \tilde{y}_n + \int_{t_n}^{t} \frac{(k_3)^2}{4} (1 - 2\tilde{y}_s) ds + k_3 \int_{t_n}^{t} \sqrt{\tilde{y}_s(1 - \tilde{y}_s)} d\hat{W}_s,
\]

for \(t \in (t_n, t_{n+1}]\), with \(\tilde{y}_0 = x_0\) a.s. and

\[
\tilde{y}_n := \frac{y_{tn}(1 + \beta \Delta) + \alpha \Delta}{1 + (\alpha + \beta)\Delta}.
\]

The following result is a consequence of Theorem 2.4.
Proposition 2.5 [Strong convergence] Let \((k_3)^2 < 2k_2\) and \(\Delta < -1/\beta\), where \(\beta\) is given by (2.10). Then, the semi-discrete scheme (2.11) converges strongly in the mean-square sense to the true solution of (1.1), that is

\[
\lim_{\Delta \to 0} \mathbb{E} \sup_{0 \leq t \leq T} |\tilde{y}^{SD}_t - x_t|^2 = 0.
\]

Now, we turn to the approximation of the solution of system (1.3)-(1.4). This time we also discretize the diffusion coefficient in a multiplicative way such that the resulting SDEs for each component can be solved analytically. In particular we consider the following processes

\[
SDY^{(1)}_t = Y^{(1)}_{t_n} + \int_{t_n}^{t_{n+1}} \left( k_{1,1}^{(1,1)} + k_{1,2}^{(1,2)} Y^{(2)}_{t_n} - \frac{(k_{3,1}^{(1,1)})^2 Y^{(2)}_{t_n} + (k_{3,2}^{(1,2)})^2 (1 - Y^{(1)}_{t_n} - Y^{(2)}_{t_n})}{4(1 - Y^{(1)}_{t_n})} \right) dt
\]

\[
+ \int_{t_n}^{t_{n+1}} Y^{(1)}_{t_n} \left( \frac{(k_{3,1}^{(2,1)})^2 Y^{(2)}_{t_n} + (k_{3,2}^{(2,2)})^2 (1 - Y^{(1)}_{t_n} - Y^{(2)}_{t_n})}{2(1 - Y^{(1)}_{t_n})} - k_{2}^{(1)} \right) dt
\]

\[
+ \int_{t_n}^{t} \frac{(k_{3,1}^{(2,1)})^2 Y^{(2)}_{t_n} + (k_{3,2}^{(2,2)})^2 (1 - Y^{(1)}_{t_n} - Y^{(2)}_{t_n})}{4(1 - Y^{(1)}_{t_n})} (1 - 2Y^{(1)}_{s}) ds
\]

\[
+ k_{3}^{(1,1)} \sqrt{\frac{Y^{(2)}_{t_n}}{1 - Y^{(1)}_{t_n}}} \int_{t_n}^{t} \sqrt{Y^{(1)}_{s} (1 - Y^{(1)}_{s})} \text{sgn}(z^{(1)}_{s}) dW^{(1)}_{s}
\]

\[
+k_{3}^{(1,2)} \sqrt{\frac{1 - Y^{(1)}_{t_n} - Y^{(2)}_{t_n}}{1 - Y^{(1)}_{t_n}}} \int_{t_n}^{t} \sqrt{Y^{(1)}_{s} (1 - Y^{(1)}_{s})} \text{sgn}(z^{(1)}_{s}) dW^{(2)}_{s},
\]

(2.13)

\[
SDY^{(2)}_t = Y^{(2)}_{t_n} + \int_{t_n}^{t_{n+1}} \left( k_{1,1}^{(2,1)} + k_{1,2}^{(2,2)} Y^{(1)}_{t_n} - \frac{(k_{3,1}^{(2,1)})^2 Y^{(1)}_{t_n} + (k_{3,2}^{(2,2)})^2 (1 - Y^{(1)}_{t_n} - Y^{(2)}_{t_n})}{4(1 - Y^{(2)}_{t_n})} \right) dt
\]

\[
+ \int_{t_n}^{t_{n+1}} Y^{(2)}_{t_n} \left( \frac{(k_{3,1}^{(3,1)})^2 Y^{(1)}_{t_n} + (k_{3,2}^{(3,2)})^2 (1 - Y^{(1)}_{t_n} - Y^{(2)}_{t_n})}{2(1 - Y^{(2)}_{t_n})} - k_{2}^{(2)} \right) dt
\]

\[
+ \int_{t_n}^{t} \frac{(k_{3,1}^{(3,1)})^2 Y^{(1)}_{t_n} + (k_{3,2}^{(3,2)})^2 (1 - Y^{(1)}_{t_n} - Y^{(2)}_{t_n})}{4(1 - Y^{(2)}_{t_n})} (1 - 2Y^{(2)}_{s}) ds
\]

\[
+ k_{3}^{(2,1)} \sqrt{\frac{Y^{(1)}_{t_n}}{1 - Y^{(2)}_{t_n}}} \int_{t_n}^{t} \sqrt{Y^{(2)}_{s} (1 - Y^{(2)}_{s})} \text{sgn}(z^{(2)}_{s}) dW^{(1)}_{s}
\]

\[
+k_{3}^{(2,3)} \sqrt{\frac{1 - Y^{(1)}_{t_n} - Y^{(2)}_{t_n}}{1 - Y^{(2)}_{t_n}}} \int_{t_n}^{t} \sqrt{Y^{(2)}_{s} (1 - Y^{(2)}_{s})} \text{sgn}(z^{(2)}_{s}) dW^{(3)}_{s},
\]

(2.14)

for \(t \in (t_n, t_{n+1}], \) with \(Y^{(i)}_{0} = X^{(i)}_{0}, i = 1, 2, 3\) a.s.

\[
z_{t}^{(1)} = \sin \left( k_{3,1}^{(1,1)} \sqrt{\frac{Y^{(2)}_{t_n}}{1 - Y^{(1)}_{t_n}}} \Delta W^{(1)}_{t} + k_{3,2}^{(1,2)} \sqrt{\frac{1 - Y^{(1)}_{t_n} - Y^{(2)}_{t_n}}{1 - Y^{(1)}_{t_n}}} \Delta W^{(2)}_{t} + 2 \arcsin(\sqrt{y^{(1)}_{t_n}}) \right),
\]
Theorem 2.6 

To conclude, we obtain the following result.

We set \( Y_{tn} = 1 - \varepsilon, Y_{tn} = \varepsilon / 2 \), for \( i, j = 1, 2 \) with \( i \neq j \) whenever \( Y_{tn} > 1 - \varepsilon \) for a tolerance \( \varepsilon > 0 \). Finally \( Y_{tn} = 1 - Y_{tn}^{(1)} - Y_{tn}^{(2)} \).

Processes (2.13) and (2.14) have jumps at nodes \( t_n \) of order \( \Delta \) and their solution in each step is given respectively by, see Appendix B,

\[
(2.15)\]

\[SDY_t^{(1)} = \sin^2 \left( k_{13}^{(1,1)} \sqrt{Y_{tn}^{(1)}} \Delta W_t^{(1)} + k_{13}^{(1,2)} \frac{1 - Y_{tn}^{(1)} - Y_{tn}^{(2)}}{1 - Y_{tn}^{(2)}} \Delta W_t^{(2)} + \arcsin(\sqrt{y_n^{(1)}}) \right),\]

and

\[
(2.16)\]

\[SDY_t^{(2)} = \sin \left( k_{13}^{(2,1)} \sqrt{Y_{tn}^{(1)}} \Delta W_t^{(1)} + k_{13}^{(2,3)} \frac{1 - Y_{tn}^{(1)} - Y_{tn}^{(2)}}{1 - Y_{tn}^{(2)}} \Delta W_t^{(3)} + \arcsin(\sqrt{y_n^{(2)}}) \right),\]

which has the pleasant feature that \( Y_t^{SD} \in (0, 1)^3 \) when \( Y_0 \in (0, 1)^3 \). Processes (2.15) and (2.16) are well defined when \( 0 < y_n^{(i)} < 1, i = 1, 2 \).

Working as before, considering this time the processes

\[
(2.17)\]

\[\widetilde{W}_t^{(i)} := \int_0^t \text{sgn}(z_s^{(i)})dW_s^{(i)},\]

we conclude to the following result.

**Theorem 2.6** [Strong convergence] Let the discretization step be such that \( 0 < y_n^{(i)} < 1, i = 1, 2 \). Then, the semi-discrete scheme (2.13)-(2.14) converges strongly in the mean-square sense to the true solution of (1.3)-(1.4), that is

\[
\lim_{\Delta t \to 0} E \sup_{0 \leq t \leq T} ||SDY_t - X_t||_2 = 0,
\]

where \( ||x||_2 = \sqrt{\sum_{i=1}^d x_i^2} \) for a \( d \)-dimensional vector \( x \). \( \square \)

We apply the above results in Section 4 and prove them in Section 6.

### 3. An extension of the semi-discrete method.

Throughout, let \( T > 0 \) and \((\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0 \leq t \leq T}, \mathbb{P})\) be a complete probability space, meaning that the filtration \( \{\mathcal{F}_t\}_{0 \leq t \leq T} \) satisfies the usual conditions, i.e. is right continuous and \( \mathcal{F}_0 \) includes all \( \mathbb{P} \)-null sets. Let \( W_{t, \omega} : [0, T] \times \Omega \to \mathbb{R} \) be a one-dimensional Wiener process adapted to the filtration \( \{\mathcal{F}_t\}_{0 \leq t \leq T} \). Consider the following stochastic differential equation (SDE),

\[
(3.1)\]

\[x_t = x_0 + \int_0^t a(s, x_s)ds + \int_0^t b(s, x_s)dW_s, \quad t \in [0, T],\]
where the coefficients \(a, b : [0, T] \times \mathbb{R} \to \mathbb{R}\) are measurable functions such that (3.1) has a unique strong solution and \(x_0\) is independent of all \(\{W_t\}_{0 \leq t \leq T}\). SDE (3.1) has non-autonomous coefficients, i.e. \(a(t, x), b(t, x)\) depend explicitly on \(t\).

To be more precise, we assume the existence of a predictable stochastic process \(x : [0, T] \times \Omega \to \mathbb{R}\) such that ([12, Def. 2.1]),

\[
\{a(t, x_t)\} \in \mathcal{L}^1([0, T]; \mathbb{R}), \quad \{b(t, x_t)\} \in \mathcal{L}^2([0, T]; \mathbb{R})
\]

and

\[
P\left[ x_t = x_0 + \int_0^t a(s, x_s) ds + \int_0^t b(s, x_s) dW_s \right] = 1, \quad \text{for every } t \in [0, T].
\]

SDEs of the form (3.1) have rarely explicit solutions, thus numerical approximations are necessary for simulations of the paths \(x_t(\omega)\), or for approximation of functionals of the form \(\mathbb{E} F(x)\), where \(F : C([0, T], \mathbb{R}) \to \mathbb{R}\). We are interested in strong approximations (mean-square) of (3.1), in the case of super- or sub-linear drift and diffusion coefficients and cover cases not included in the previous work [3].

The purpose of this section is to further generalize the semi-discrete (SD) method covering cases of sub-linear diffusion coefficients such as the Cox-Ingersoll-Ross model (CIR) or the Constant Elasticity of Variance model (CEV) (cf. [3, (1.2) and (1.4)]), where also an additive discretization is considered.

**Assumption 3.7** Let \(f_1(s, x) : [0, T] \times \mathbb{R} \to \mathbb{R}\) and \(f_2(s, r, x, y), g(s, r, x, y) : [0, T]^2 \times \mathbb{R}^2 \to \mathbb{R}\) be such that \(f_1(s, x) + f_2(s, s, x, x) = a(s, x), g(s, s, x, x) = b(s, x)\), where \(f_1, f_2, g\) satisfy the following conditions

\[
f_1(s, x) \leq C(1 + |x|^l), \quad (\text{Polynomial Growth}),
\]

for some appropriate \(0 < l\) (we take \(0 < l \leq p/2\) in Theorem 3.8)

\[
|f_2(s_1, r_1, x_1, y_1) - f_2(s_2, r_2, x_2, y_2)| \leq C_R\left(|s_1 - s_2| + |r_1 - r_2| + |x_1 - x_2| + |y_1 - y_2|\right)
\]

and

\[
|g(s_1, r_1, x_1, y_1) - g(s_2, r_2, x_2, y_2)| \leq C_R\left(|s_1 - s_2| + |r_1 - r_2| + |x_1 - x_2| + |y_1 - y_2| + |x_1 - x_2|^q\right),
\]

for any \(R > 0\) such that \(|x_1| \vee |x_2| \vee |y_1| \vee |y_2| \leq R\), where the positive parameter \(q \in (0, \frac{1}{2})\) the quantity \(C_R\) depends on \(R\) and \(x \vee y\) denotes the maximum of \(x, y\). (By the fact that we want the problem (3.1) to be well-posed and by the conditions on \(f_1, f_2\) and \(g\) we get that \(f_1, f_2, g\) are bounded on bounded intervals.)

Let the equidistant partition \(0 = t_0 < t_1 < ... < t_N = T\) and \(\Delta = T/N\). We propose the following semi-discrete numerical scheme

\[
y_t = y_n + \int_{t_n}^{t} f_2(t_n, s, y_{tn}, y_s) ds + \int_{t_n}^{t} g(t_n, s, y_{tn}, y_s) dW_s, \quad t \in (t_n, t_{n+1}],
\]

where we assume that for every \(n \leq N - 1\), (3.2) has a unique strong solution and \(y_0 = x_0\) a.s

\[y_n = y_{tn} + f_1(t_n, y_{tn}) \cdot \Delta.
\]

In order to compare with the exact solution \(x_t\), which is a continuous time process, we consider the following interpolation process of the semi-discrete approximation, in a compact
form,

\begin{equation}
    y_t = y_0 + f_1(t_0, y_0) \cdot \Delta + \int_0^t f_2(\hat{s}, s, y_s, y_\hat{s}) ds + \int_0^t g(\hat{s}, s, y_s, y_\hat{s}) dW_s,
\end{equation}

where \( \hat{s} = t_n \) when \( s \in [t_n, t_{n+1}) \). Process (3.3) has jumps at nodes \( t_n \). The first and third variable in \( f_2, g \) denote the discretized part of the original SDE. We observe from (3.3) that in order to solve for \( y_t \), we have to solve an SDE and not an algebraic equation, thus in this context, we cannot reproduce implicit schemes, but we can reproduce the Euler scheme if we choose \( f_1 = 0, f_2(s, r, x, y) = a(s, x) \) and \( g(s, r, x, y) = b(s, x) \).

The numerical scheme (3.3) converges to the true solution \( x_t \) of SDE (3.1) and this is stated in the following, which is our main result.

**Theorem 3.8** [Strong convergence] Suppose Assumption 3.7 holds and (3.2) has a unique strong solution for every \( n \leq N - 1 \), where \( x_0 \in L^p(\Omega, \mathbb{R}) \). Let also

\[
    \mathbb{E}(\sup_{0 \leq t \leq T} |x_t|^p) \vee \mathbb{E}(\sup_{0 \leq t \leq T} |y_t|^p) < A,
\]

for some \( p > 2 \) and \( A > 0 \). Then the semi-discrete numerical scheme (3.3) converges to the true solution of (3.1) in the \( L^2 \)-sense, that is

\begin{equation}
    \lim_{\Delta \to 0} \mathbb{E} \sup_{0 \leq t \leq T} |y_t - x_t|^2 = 0.
\end{equation}

---

### 4. Numerics

Here, we make numerical tests to study the strong convergence of the proposed semi-discrete methods (2.1) and (2.11) for the Wright-Fisher model described by the Itô SDE (1.1) with \( k_1 = A, k_2 = A + B, k_3 = C \), where the parameters \( A, B \) and \( C \) are positive and \( C = \sqrt{2k_2/(N_r - 1)} \). We take as initial condition the steady state of the deterministic part, that is \( x_0 = A/(A + B) \). This setting has been used for the approximation of ion channels within cardiac and neuronal cells, see [13, Sec. 2.1]; the ion channel occupies one of two positions (open and closed states) with transition rates \( A \) and \( B \) respectively and \( N_r \) is the total number of ion channels within a cell, see [6, (2.3)]. We consider two set of parameters

- **SET I**: \((A, B, N_r) = (1, 2, 100)\),
- **SET II**: \((A, B, N_r) = (7.0064, 0.0204, 100)\),

as in [6, Sec. 6 and 7.1], where the Balance Implicit Split Step (BISS) method is suggested [6, (4.8)]

\begin{equation}
    y_{n+1}^{\text{BISS}} = y_n + (A - (A + B)y_n) \Delta + \frac{C\sqrt{y_n(1 - y_n)}}{1 + d^1(y_n)|\Delta W_n|} (1 - (A + B)\Delta),
\end{equation}

where \( \Delta \) is the step-size of the equidistant discretization of the interval \([0,1]\), the control function \( d^1 \) is given by

\[
    d^1(y) = \begin{cases} 
        C\sqrt{(1 - \varepsilon)/\varepsilon} & \text{if } y < \varepsilon, \\
        C\sqrt{(1 - y)/y} & \text{if } \varepsilon \leq y < 1/2, \\
        C\sqrt{y/(1 - y)} & \text{if } 1/2 \leq y \leq 1 - \varepsilon, \\
        C\sqrt{(1 - \varepsilon)/\varepsilon} & \text{if } y > 1 - \varepsilon,
    \end{cases}
\]
and

$$\varepsilon = \min\{A\Delta, B\Delta, 1 - A\Delta, 1 - B\Delta\}.$$ 

The hybrid (HYB) scheme as proposed in [13, (11)] is the result of a splitting method and reads

$$y_{n+1}^{HYB} = \frac{\alpha}{\beta}(e^{\beta \Delta} - 1) + e^{\beta \Delta} \sin^2 \left(\frac{C}{2} \Delta W_n + \arcsin(\sqrt{y_n})\right).$$

It works only for the parameter SET I since we have to assume that

$$\frac{a}{a+b} \in \left(\frac{1}{2(N_r - 1)}, 1 - \frac{1}{2(N_r - 1)}\right).$$

Finally, the proposed semi-discrete (SD) scheme reads

$$y_{n+1}^{SD} = \sin^2 \left(\frac{C}{2} \Delta W_n + \arcsin(\sqrt{y_n(1 + \beta \Delta) + \alpha \Delta})\right),$$

for SET I and

$$y_{n+1}^{SD} = \sin^2 \left(\frac{C}{2} \Delta W_n + \arcsin\left(\frac{y_n(1 + \beta \Delta) + \alpha \Delta}{1 + (\alpha + \beta \Delta)}\right)\right),$$

for parameter SET II. The parameters of SET I are chosen in a way that the probability of the Euler-Maruyama (EM) scheme

$$y_{n+1}^{EM} = (A - (A + B)y_n)\Delta + C \sqrt{y_n(1 - y_n)} \Delta W_n,$$

leaving the interval $[0, 1]$ is very small, whereas in the case of SET II this probability is high. The paths of the solutions of EM exiting the boundaries 0 and 1 are only a few in the first case and one may reject them. Nevertheless, since such an approach induces bias to the solution obtained by the EM method (much more evident in the second case) we choose only to compare our method with BISS and HYB.

We estimate the endpoint $L^2$-norm $\varepsilon = \sqrt{\mathbb{E}|y^{(\Delta)}(T) - x_T|^2}$, of the difference between the numerical scheme evaluated at step size $\Delta$ and the exact solution of (1.1). To do so, we compute $M$ batches of $L$ simulation paths, where each batch is estimated by $\hat{\varepsilon}_j = \frac{1}{L} \sum_{i=1}^{L} |y_{i,j}^{(\Delta)}(T) - y_{i,j}^{(ref)}(T)|^2$ and the Monte Carlo estimator of the error is

$$\hat{\varepsilon} = \sqrt{\frac{1}{ML} \sum_{j=1}^{M} \sum_{i=1}^{L} |y_{i,j}^{(\Delta)}(T) - y_{i,j}^{(ref)}(T)|^2}$$

and requires $M \cdot L$ Monte Carlo sample paths. The reference solution is calculated using the method at a very fine time grid, $\Delta = 2^{-13}$. We have shown in Theorem 2.4 and Proposition 2.5 that the SD numerical schemes converge strongly to the exact solution, so we use the SD method as a reference solution, and the HYB method when applicable. The BISS method converges in the $L^1$-norm to the true solution [6, Th. 5.1], so we choose not to consider it as a reference solution even though we conjecture that a similar technique may be used to show an $L^2$-convergence result.

We simulate $100 \cdot 100 = 10^4$ paths, where the choice of the number of Monte Carlo paths is adequately large, so as not to significantly hinder the mean-square errors. We compute the approximation error (4.5) with 98%-confidence intervals. We present the results for the parameter SET I and II in a log₂-log₂ scale in Figures 1 - 3 and Tables 1 - 3 respectively.
| Step \(\Delta\) | SD     | BISS   | HYB   |
|-----------------|--------|--------|-------|
| \(2^{-3}\)      | 0.008997 | 0.020613 | 0.009737 |
| \(2^{-4}\)      | 0.004216 | 0.012167 | 0.004937 |
| \(2^{-5}\)      | 0.002072 | 0.007558 | 0.002515 |
| \(2^{-6}\)      | 0.001015 | 0.004750 | 0.001248 |
| \(2^{-7}\)      | 0.000509 | 0.003080 | 0.000619 |
| \(2^{-8}\)      | 0.000257 | 0.001248 | 0.000324 |

Table 1. Error and step size of SD, BISS and Hybrid schemes for (1.1) with parameter SET I and HYB as a reference solution.

| Step \(\Delta\) | SD     | BISS   |
|-----------------|--------|--------|
| \(2^{-3}\)      | 0.009030 | 0.020507 |
| \(2^{-4}\)      | 0.004210 | 0.011952 |
| \(2^{-5}\)      | 0.002054 | 0.007588 |
| \(2^{-6}\)      | 0.000999 | 0.004791 |
| \(2^{-7}\)      | 0.000498 | 0.003123 |
| \(2^{-8}\)      | 0.000243 | 0.002069 |
| \(2^{-9}\)      | 0.000120 | 0.001410 |
| \(2^{-10}\)     | 0.000057 | 0.000967 |
| \(2^{-11}\)     | 0.000026 | 0.000668 |
| \(2^{-12}\)     | 0.000021 | 0.000459 |

Table 2. Error and step size of SD, BISS and Hybrid schemes for (1.1) with parameter SET I and SD as a reference solution.

| Step \(\Delta\) | SD     | BISS   |
|-----------------|--------|--------|
| \(2^{-3}\)      | 0.004860 | 0.009378 |
| \(2^{-4}\)      | 0.002181 | 0.008626 |
| \(2^{-5}\)      | 0.001169 | 0.008074 |
| \(2^{-6}\)      | 0.000653 | 0.007477 |
| \(2^{-7}\)      | 0.000376 | 0.007035 |
| \(2^{-8}\)      | 0.000231 | 0.006556 |
| \(2^{-9}\)      | 0.000138 | 0.006381 |
| \(2^{-10}\)     | 0.000083 | 0.006171 |
| \(2^{-11}\)     | 0.000051 | 0.005977 |
| \(2^{-12}\)     | 0.000030 | 0.005898 |

Table 3. Error and step size of SD and BISS schemes for (1.1) with parameter SET II and SD as a reference solution.
Furthermore, we study the strong convergence of the proposed semi-discrete methods (2.13)-(2.14) for the Wright-Fisher model described by the system of Itô SDEs (1.3)-(1.4) with

\[ k_1^{(1,1)} = A_3, k_1^{(1,2)} = A_2 - A_3, k_1^{(2,1)} = A_1, k_1^{(2,2)} = B_1 - A_1, k_2^{(1)} = B_3 + B_1 + A_3, k_2^{(2)} = A_2 + B_2 + A_1, k_3^{(1,1)} = -C_1, k_3^{(1,2)} = C_2, k_3^{(2,1)} = C_1, k_3^{(2,2)} = -C_3. \]

This setting has been used...
for the approximation of the proportion of alleles or channels in a 3 state system and we consider the set of parameters

- SET III: \((A_1, A_2, A_3, B_1, B_2, B_3, C_1, C_2, C_3) = (1, 2, 3, 1.2, 2.3, 3.4, 0.1271, 0.1798, 0.1291)\), as in [6, Sec. 6], where the Balance Implicit Split Step (BISS) method is suggested; they use the splitting method, solving first the stochastic system with a balanced implicit scheme

\[
\begin{align*}
\text{BISS} Y^{(1)}_{n+1} &= \frac{Y^{(1)}_n - C_1 \sqrt{Y^{(1)}_n Y^{(2)}_n} \Delta W^{(1)}_n + C_2 \sqrt{Y^{(1)}_n (1 - Y^{(1)}_n - Y^{(2)}_n)} \Delta W^{(2)}_n}{1 + D_1(Y_n) + D_2(Y_n) + D_3(Y_n)} \\
&\quad + \frac{Y^{(1)}_n (1 + D_1(Y_n) + D_2(Y_n) + D_3(Y_n))}{1 + D_1(Y_n) + D_2(Y_n) + D_3(Y_n)}, \\
\text{BISS} Y^{(2)}_{n+1} &= \frac{Y^{(2)}_n + C_1 \sqrt{Y^{(1)}_n Y^{(2)}_n} \Delta W^{(1)}_n - C_3 \sqrt{Y^{(2)}_n (1 - Y^{(1)}_n - Y^{(2)}_n)} \Delta W^{(3)}_n}{1 + D_1(Y_n) + D_2(Y_n) + D_3(Y_n)} \\
&\quad + \frac{Y^{(2)}_n (1 + D_1(Y_n) + D_2(Y_n) + D_3(Y_n))}{1 + D_1(Y_n) + D_2(Y_n) + D_3(Y_n)}.
\end{align*}
\]

(4.6, 4.7)

and then the deterministic part with the one-step EM scheme; here the control functions \(D_i, i = 1, \ldots, 3\), are given by

\[
D_1(Y) = \begin{cases} 
C_1 \left( \sqrt{\frac{Y^{(2)}_n}{Y^{(1)}_n}} + \sqrt{\frac{Y^{(1)}_n}{Y^{(2)}_n}} \right) |\Delta W^{(1)}_n| & \text{if } \varepsilon < Y^{(1)} \leq Y^{(2)}, \\
C_1 \left( \sqrt{1 - \frac{Y^{(1)}_n - Y^{(2)}_n}{Y^{(2)}_n}} + \sqrt{\frac{Y^{(1)}_n}{Y^{(2)}_n}} \right) |\Delta W^{(1)}_n| & \text{if } \varepsilon < Y^{(2)} < Y^{(1)},
\end{cases}
\]
\[ D_2(Y) = \begin{cases} 
  C_2 \left( \frac{\sqrt{1-Y(1)^2} - Y(2)^2}{Y(1)} + \frac{\varepsilon}{\sqrt{Y(1)(1-Y(1)^2) - Y(2)^2)}} \right) |\Delta W_n^{(2)}| & \text{if } 2Y(1) + Y(2)^2 < 1, \\
  C_2 \left( \frac{\sqrt{1-Y(1)^2} - Y(2)^2}{Y(1)} + \frac{\varepsilon}{\sqrt{Y(1)(1-Y(1)^2) - Y(2)^2)}} \right) |\Delta W_n^{(2)}| & \text{if } 2Y(1) + Y(2)^2 \geq 1, 
\end{cases} \]

\[ D_3(Y) = \begin{cases} 
  C_3 \left( \frac{\sqrt{1-Y(1)^2} - Y(2)^2}{Y(2)} + \frac{\varepsilon}{\sqrt{Y(2)(1-Y(1)^2) - Y(2)^2)}} \right) |\Delta W_n^{(3)}| & \text{if } 2Y(2) + Y(1)^2 < 1, \\
  C_3 \left( \frac{\sqrt{1-Y(1)^2} - Y(2)^2}{Y(2)} + \frac{\varepsilon}{\sqrt{Y(2)(1-Y(1)^2) - Y(2)^2)}} \right) |\Delta W_n^{(3)}| & \text{if } 2Y(2) + Y(1)^2 \geq 1; 
\end{cases} \]

In case \( Y(i) \leq \varepsilon \) then set \( Y(i) = \varepsilon, i = 1, 2 \) and if \( Y(1) + Y(2)^2 \geq 1 - \varepsilon \) then set \( Y(1) + Y(2)^2 = 1 - \varepsilon \).

The proposed semi-discrete scheme reads, see (2.15)-(2.16)

\[ SDY_{n+1}^{(1)} = \sin^2 \left( \frac{C_1}{2} \sqrt{\frac{Y_n^{(2)}}{1 - Y_n^{(1)}}} \Delta W_n^{(1)} + \frac{C_2}{2} \sqrt{\frac{1 - Y_n^{(1)} - Y_n^{(2)}}{1 - Y_n^{(2)}}} \Delta W_n^{(2)} + \arcsin(\sqrt{y_n^{(1)}}) \right), \]

\[ SDY_{n+1}^{(2)} = \sin \left( \frac{C_1}{2} \sqrt{\frac{Y_n^{(1)}}{1 - Y_n^{(2)}}} \Delta W_n^{(1)} - \frac{C_3}{2} \sqrt{\frac{1 - Y_n^{(1)} - Y_n^{(2)}}{1 - Y_n^{(2)}}} \Delta W_n^{(3)} + \arcsin(\sqrt{y_n^{(2)}}) \right), \]

where \( \Delta \) is the step-size of the equidistant discretization of the interval \([0, 1]\). Note that for the parameter SET III \( 0 < y_n^{(i)} < 1, i = 1, 2 \), for any \( \Delta < \min\{1/k_2^{(1)}, 1/k_2^{(2)}\} = 1/(B_3 + B_1 + A_3) \). The parameters of SET III are chosen in a way that the probability of the Euler-Maruyama (EM) scheme

\[ EMY_{n+1}^{(1)} = Y_n^{(1)} + (A_3 + (A_2 - A_3)Y_n^{(2)} - (B_3 + B_1 + A_3)Y_n^{(1)}) \Delta \\
- C_1 \sqrt{Y_n^{(1)}Y_n^{(2)}} \Delta W_n^{(1)} + C_2 \sqrt{Y_n^{(1)}(1 - Y_n^{(1)} - Y_n^{(2)})} \Delta W_n^{(2)}, \]

\[ EMY_{n+1}^{(2)} = Y_n^{(2)} + (A_1 + (B_1 - A_1)Y_n^{(1)} - (A_2 + B_2 + A_1)Y_n^{(2)}) \Delta \\
+ C_1 \sqrt{Y_n^{(1)}Y_n^{(2)}} \Delta W_n^{(1)} - C_3 \sqrt{Y_n^{(2)}(1 - Y_n^{(1)} - Y_n^{(2)})} \Delta W_n^{(3)}, \]

leaving the region \([0, 1]^3\) is very small. The paths of the solutions of EM exiting the boundaries 0 and 1 are extremely few when we take as initial condition the steady state of the deterministic part of the system (which for the parameter SET III is away from the boundaries) and consider a very fine time discretization (\( \Delta = 2^{-13} \)). Thus, we take as the exact solution the EM method and reject the paths, if any, outside the region \([0, 1]^3\). We compare our method with BISS and present the results in Figure 4.

We make the following comments.

- The performance of the HYB and SD schemes for parameter Set I is quite similar with SD producing smaller errors. They both perform better than BISS. Nevertheless, HYB is not boundary preserving for parameter SET II. The performance of EM and SD schemes is almost identical for parameter Set III. They both perform better than BISS. Nevertheless, EM is not boundary preserving (we just used it in this experiment for comparative reasons as in [6, Sec. 6].)
Figure 4. Convergence of SD, EM and BISS methods applied to (1.3)-(1.4) with parameter SET III with EM as a reference solution.

- The numerical results suggest that the SD schemes converge in the mean-square sense with order 1 for parameter SET I and SET III and at least 1/2 for parameter SET II.
- The proposed SD schemes perform better w.r.t. the computational time required to achieve a desired level of accuracy, since there is no need to calculate a control function.
- For the implementation of the SD method, we have to assume that \( \Delta < -1/(\alpha + \beta) = 0.1423 \), for the parameters of SET I, \( \Delta < -1/\beta = 0.1437 \), for the parameters of SET II, and \( \Delta < 1/(B_1 + B_3 + A_3) = 1/7.6 \), for the parameters of SET III, so the step-size \( \Delta = 1/8 \) is sufficient. There is not a step-size restriction in the BISS method; nevertheless we propose the SD method since it is fast and more accurate.

5. **Proof of Theorem 3.8.**

We split the proof is three steps. First, we prove a general estimate of the error of the SD method for any \( p > 0 \). Then, we show the \( \mathcal{L}^1 \)-convergence of the semi-discrete method and finally the \( \mathcal{L}^2 \)-convergence (3.4). We denote the indicator function of a set \( A \) by \( \mathbb{I}_A \). The quantity \( C_R \) may vary from line to line and it may depend apart from \( R \) on other quantities, like time \( T \) for example, which are all constant, in the sense that we don’t let them grow to infinity.

**Lemma 5.9** [Error bound for the semi-discrete scheme] Let the assumption of Theorem 3.8 hold. Let \( R > 0 \), and set the stopping time \( \theta_R = \inf\{t \in [0, T] : |y_t| > R \text{ or } |y_t| > R\} \). Then the following estimate holds

\[
\mathbb{E}|y_{s\wedge\theta_R} - \tilde{y}_{s\wedge\theta_R}|^p \leq C_R \Delta^{p/2},
\]

where \( \Delta \) is the step-size of the SD method.
for any \( p > 0 \), where \( C_R \) does not depend on \( \Delta \), implying \( \sup_{s \in [t_{n_s}, t_{n_{s+1}}]} \mathbb{E}|y_s \wedge \theta_R - \bar{y}_{s \wedge \theta_R}|^p = O(\Delta^{p/2}) \), as \( \Delta \downarrow 0 \).

**Proof of Lemma 5.9.** We fix a \( p \geq 2 \). Let \( n_s \) integer such that \( s \in [t_{n_s}, t_{n_{s+1}}) \). It holds that

\[
|y_{s \wedge \theta_R} - \bar{y}_{s \wedge \theta_R}|^p = \left| \int_{t_{n_s} \wedge \theta_R}^{s \wedge \theta_R} f_2(\hat{u}, u, y_{\hat{u}}, y_u) du + \int_{t_{n_s} \wedge \theta_R}^{s \wedge \theta_R} g(u, y_{\hat{u}}, y_u) dW_u \right|^p \leq 2^{p-1} \int_{t_{n_s} \wedge \theta_R}^{s \wedge \theta_R} |f_2(\hat{u}, u, y_{\hat{u}}, y_u)|^p du + 2^{p-1} \int_{t_{n_s} \wedge \theta_R}^{s \wedge \theta_R} |g(u, y_{\hat{u}}, y_u)|^p dW_u \leq C_R \Delta^p + 2^{p-1} \int_{t_{n_s} \wedge \theta_R}^{s \wedge \theta_R} |g(u, y_{\hat{u}}, y_u)|^p dW_u,
\]

where we have used Cauchy-Schwarz inequality and Assumption 3.7 for the function \( f_2 \). Taking expectations in the above inequality gives

\[
\mathbb{E}|y_{s \wedge \theta_R} - \bar{y}_{s \wedge \theta_R}|^p \leq C_R \Delta^p + 2^{p-1} \mathbb{E} \left| \int_{t_{n_s} \wedge \theta_R}^{t_{n_{s+1}} \wedge \theta_R} g(u, y_{\hat{u}}, y_u) dW_u \right|^p \leq C_R \Delta^p + 2^{p-1} \mathbb{E} \left( \frac{p+1}{2(p-1)} \right)^{p/2} \left| \int_{t_{n_s} \wedge \theta_R}^{t_{n_{s+1}} \wedge \theta_R} |g(u, y_{\hat{u}}, y_u)|^2 du \right|^{p/2} \leq C_R \Delta^p + C_R \Delta^{p/2},
\]

where in the third step we have used the Burkholder-Davis-Gundy (BDG) inequality [12, Th. 1.7.3], [11, Th. 3.3.28] on the diffusion term, in the last step Assumption 3.7 for the function \( g \). Thus,

\[
\lim_{\Delta \downarrow 0} \sup_{s \in [t_{n_s}, t_{n_{s+1}}]} \mathbb{E}|y_{s \wedge \theta_R} - \bar{y}_{s \wedge \theta_R}|^p = O(\Delta^{p/2}),
\]

which justifies the \( O(\Delta^{p/2}) \) notation. Now for \( 0 < p < 2 \) we have that

\[
\mathbb{E}|y_{s \wedge \theta_R} - \bar{y}_{s \wedge \theta_R}|^p \leq \left( \mathbb{E}|y_{s \wedge \theta_R} - \bar{y}_{s \wedge \theta_R}|^2 \right)^{p/2} \leq C_R \Delta^{p/2},
\]

where we have used Jensen inequality for the concave function \( \phi(x) = x^{p/2} \).

In the next result we estimate the \( L^1 \)-error using the Yamada-Watanabe approach. We denote the difference \( E_t := y_t - x_t \).

**Proposition 5.10** ([Convergence of the semi-discrete scheme in \( L^1 \)] Let the assumptions of Theorem 3.8 hold. Let \( R > 0 \), and set the stopping time \( \theta_R = \inf\{t \in [0, T] : |y_t| > \)

\[
\text{for any } p > 0, \text{ where } C_R \text{ does not depend on } \Delta, \text{ implying } \sup_{s \in [t_{n_s}, t_{n_{s+1}}]} \mathbb{E}|y_s \wedge \theta_R - \bar{y}_{s \wedge \theta_R}|^p = O(\Delta^{p/2}), \text{ as } \Delta \downarrow 0.
\]
We have that
\[
(5.1) \quad \sup_{0 \leq t \leq T} \mathbb{E} |\mathcal{E}_{t \wedge \Theta_0}| \leq \left( C_R \sqrt{\Delta} + \frac{C_R}{m \epsilon_m} \Delta^q + \frac{C_R}{m} \epsilon_m + \epsilon_m^{-1} \right) e^{a_{R,m,q} t},
\]
for any \( m \geq 1 \), where
\[
e_m = e^{-m(m+1)/2}, \quad a_{R,m,q} := C_R + \frac{C_R}{m} + \frac{C_R}{m(\epsilon_m)^{1/2}}
\]
and \( C_R \) does not depend on \( \Delta \). It holds that \( \lim_{m \uparrow \infty} \epsilon_m = 0 \).

**Proof of Proposition 5.10.** Let the non-increasing sequence \( \{ \epsilon_m \}_{m \in \mathbb{N}} \) with \( \epsilon_m = e^{-m(m+1)/2} \) and \( \epsilon_0 = 1 \). We introduce the following sequence of smooth approximations of \( |x| \), (method of Yamada and Watanabe, [14])
\[
\phi_m(x) = \int_{0}^{[x]} dy \int_{0}^{y} \psi_m(u) du,
\]
where the existence of the continuous function \( \psi_m(u) \) with \( 0 \leq \psi_m(u) \leq 2/(mu) \) and support in \( (\epsilon_m, \epsilon_m-1) \) is justified by \( \int_{\epsilon_m}^{\epsilon_m-1} (du/u) = m \). The following relations hold for \( \phi_m \in C^2(\mathbb{R}, \mathbb{R}) \) with \( \phi_m(0) = 0 \),
\[
|x| - \epsilon_m^{-1} \leq \phi_m(x) \leq |x|, \quad |\phi'_m(x)| \leq 1, \quad x \in \mathbb{R},
\]
\[
|\phi''_m(x)| \leq \frac{2}{m|x|}, \quad \text{when } \epsilon_m < |x| < \epsilon_m^{-1} \text{ and } |\phi'_m(x)| = 0 \text{ otherwise.}
\]
We have that
\[
(5.2) \quad \mathbb{E} |\mathcal{E}_{t \wedge \Theta_0}| \leq \epsilon_m^{-1} + \mathbb{E} \phi_m(\mathcal{E}_{t \wedge \Theta_0}).
\]
Applying Itô’s formula to the sequence \( \{ \phi_m \}_{m \in \mathbb{N}} \), we get
\[
\phi_m(\mathcal{E}_{t \wedge \Theta_0}) = \int_{0}^{t \wedge \Theta_0} \phi'_m(\mathcal{E}_s)(f_2(\cdot, s, y_s, y_s) - f_2(s, s, x_s, x_s) - f_1(s, x)) ds + M_t
\]
\[
+ \frac{1}{2} \int_{0}^{t \wedge \Theta_0} \phi''_m(\mathcal{E}_s)(g(\cdot, s, y_s, y_s) - g(s, s, x_s, x_s))^2 ds
\]
\[
\leq \int_{0}^{t \wedge \Theta_0} C_R \left( |y_s - x_s| + |\mathcal{E}_s| + |\hat{s} - s| \right) ds + M_t
\]
\[
+ C_R \int_{0}^{t \wedge \Theta_0} \frac{1}{m |\mathcal{E}_s|} \left( |y_s - y_s|^2 + |\mathcal{E}_s|^2 + |\hat{s} - s|^2 + \left( |y_s - y_s|^{2q} + |\mathcal{E}_s|^{2q} \right) \right) ds
\]
\[
\leq C_R \int_{0}^{t \wedge \Theta_0} |y_s - y_s| ds + C_R \int_{0}^{t \wedge \Theta_0} |\mathcal{E}_s| ds + C_R \int_{0}^{t \wedge \Theta_0} |\hat{s} - s| ds + M_t
\]
\[
+ C_R \int_{0}^{t \wedge \Theta_0} |y_s - y_s|^2 + |y_s - y_s|^{2q} + |\mathcal{E}_s|^2 + |\mathcal{E}_s|^{2q} + |\hat{s} - s|^2 ds
\]
\[
\leq C_R \int_{0}^{t \wedge \Theta_0} |y_s - y_s| ds + C_R \frac{|t/\Delta - 1|}{m \epsilon_m} \int_{0}^{t \wedge \Theta_0} \left( |y_s - y_s|^2 + |y_s - y_s|^{2q} \right) ds + (C_R + C_R/m) \int_{0}^{t \wedge \Theta_0} |\mathcal{E}_s| ds
\]
\[
+ C_R \frac{|t/\Delta - 1|}{m} \sum_{k=0}^{[t/\Delta - 1]} \int_{t_k}^{t_{k+1} \wedge \Theta_0} |k - s|^2 ds + C_R \sum_{k=0}^{[t/\Delta - 1]} \int_{t_k}^{t_{k+1} \wedge \Theta_0} |k - s| ds + M_t,
\]
where in the second step we have used Assumption 3.7 for the functions \( f_1, f_2, g \) the subadditivity property of \( h(x) = x^{2q} \), and the properties of \( \phi_m \) and

\[
M_t := \int_0^{t \wedge \theta_R} \phi_m'(\mathcal{E}_u)(g(\bar{u}, u, y_u, y_u) - g(u, u, x_u, x_u))dW_u.
\]

Using the estimate

\[
\frac{C_R}{m} \int_0^{t \wedge \theta_R} |\mathcal{E}_s|^{2q-1}ds \leq \frac{C_R}{m|\mathcal{E}_s|} \int_0^{t \wedge \theta_R} \left( q|\mathcal{E}_s|^2(e_m)^{\frac{2q-2q}{q}} + (1-q)(e_m)^2 \right) ds 
\]

\[
\leq \frac{C_R}{m(e_m)^{\frac{2q-2q}{q}}} \int_0^{t \wedge \theta_R} |\mathcal{E}_s|ds + \frac{C_R}{m} e_m,
\]

we get

\[
\phi_m(\mathcal{E}_{t \wedge \theta_R}) \leq C_R \int_0^{t \wedge \theta_R} |y_s - y_s|ds + \frac{C_R}{m e_m} \int_0^{t \wedge \theta_R} (|y_s - y_s|^2 + |y_s - y_s|^{2q}) ds 
\]

\[
+ \left( C_R + \frac{C_R}{m} + \frac{C_R}{m(e_m)^{\frac{2q-2q}{q}}} \right) \int_0^{t \wedge \theta_R} |\mathcal{E}_s|ds + \frac{C_R}{m} e_m + \frac{C_R}{m} \Delta^2 + C_R \Delta + M_t.
\]

Taking expectations in the above inequality yields

\[
\mathbb{E}\phi_m(\mathcal{E}_{t \wedge \theta_R}) \leq C_R \Delta + C_R \sqrt{\Delta} + \frac{C_R}{m e_m} (\Delta^2 + \Delta + \Delta^q) + \frac{C_R}{m} e_m + \left( C_R + \frac{C_R}{m} + \frac{C_R}{m(e_m)^{\frac{2q-2q}{q}}} \right) \int_0^{t \wedge \theta_R} \mathbb{E}|\mathcal{E}_s|ds,
\]

where we have used Lemma 5.9 and the fact that \( \mathbb{E}M_t = 0 \). Thus (5.2) becomes

\[
\mathbb{E}|\mathcal{E}_{t \wedge \theta_R}| \leq C_R \sqrt{\Delta} + \frac{C_R}{m e_m} \Delta^q + \frac{C_R}{m} e_m + \Delta^q + \left( C_R + \frac{C_R}{m} + \frac{C_R}{m(e_m)^{\frac{2q-2q}{q}}} \right) \int_0^{t \wedge \theta_R} \mathbb{E}|\mathcal{E}_s|ds
\]

\[
\leq \left( C_R \sqrt{\Delta} + \frac{C_R}{m e_m} \Delta^q + \frac{C_R}{m} e_m + \Delta^q \right) e^{a_{R,m,q} t},
\]

where in the last step we have used the Gronwall inequality ([15, (7)]) and \( a_{R,m,q} = C_R + \frac{C_R}{m} + \frac{C_R}{m(e_m)^{\frac{2q-2q}{q}}} \). Taking the supremum over all \( 0 \leq t \leq T \) gives (5.1).

\[
\square
\]

**Convergence of the semi-discrete scheme in \( L^2 \).** Let the events \( \Omega \) be defined by \( \Omega_R := \{ \omega \in \Omega : \sup_{0 \leq t \leq T} |x_t| \leq R, \sup_{0 \leq t \leq T} |y_t| \leq R \} \) and the stopping time \( \theta_R = \{ \inf t \in [0, T] : \)

\[
\text{Relevant notes:}
\]

| Footnote 1 | The function \( h(u) = \phi_m'(\mathcal{E}_u)(g(\bar{u}, u, y_u, y_u) - g(u, u, x_u, x_u)) \) belongs to the space \( M^2([0, t \wedge \theta_R]; \mathbb{R}) \) of real-valued measurable \( \mathcal{F}_t \)-adapted processes such that \( \mathbb{E}\int_0^{t \wedge \theta_R} |h(u)|^2du < \infty \) thus ([12, Th. 1.5.8]) implies \( \mathbb{E}M_t = 0 \). |
\[ |y_t| > R \text{ or } |x_t| > R \] for some \( R > 0 \) big enough. We have that

\[
\mathbb{E} \sup_{0 \leq t \leq T} |\mathcal{E}_t|^2 = \mathbb{E} \sup_{0 \leq t \leq T} |\mathcal{E}_t|^2 \mathbb{I}_{\Omega_R} + \mathbb{E} \sup_{0 \leq t \leq T} |\mathcal{E}_t|^2 \mathbb{I}_{(\Omega_R)^c} \\
\leq \mathbb{E} \sup_{0 \leq t \leq T} |\mathcal{E}_{t \wedge \theta_R}|^2 + \left( \mathbb{E} \sup_{0 \leq t \leq T} |\mathcal{E}_t|^p \right)^{2/p} \left( \mathbb{P}(\Omega_R)^c \right)^{(p-2)/p} \\
\leq \mathbb{E} \sup_{0 \leq t \leq T} |\mathcal{E}_{t \wedge \theta_R}|^2 + \left( \mathbb{E} \sup_{0 \leq t \leq T} |\mathcal{E}_t|^p \right)^{2/p} \left( \mathbb{P}(\Omega_R)^c \right)^{(p-2)/p} \\
\leq \mathbb{E} \sup_{0 \leq t \leq T} |\mathcal{E}_{t \wedge \theta_R}|^2 + 4 \cdot A^{2/p} \left( \mathbb{P}(\Omega_R)^c \right)^{(p-2)/p},
\]

(5.3)

where \( p > 2 \) is such that the moments of \( |x_t|^p \) and \( |y_t|^p \) are bounded by the constant \( A \). We want to estimate each term of the right hand side of (5.3). It holds that

\[
\mathbb{P}(\Omega_R^c) \leq \mathbb{P}(\sup_{0 \leq t \leq T} |y_t| > R) + \mathbb{P}(\sup_{0 \leq t \leq T} |x_t| > R) \\
\leq (\mathbb{E} \sup_{0 \leq t \leq T} |y_t|^k) R^{-k} + (\mathbb{E} \sup_{0 \leq t \leq T} |x_t|^k) R^{-k},
\]

for any \( k \geq 1 \) where the first step comes from the subadditivity of the measure \( \mathbb{P} \) and the second step from Markov inequality. Thus for \( k = p \) we get

\[
\mathbb{P}(\Omega_R^c) \leq 2AR^{-p}.
\]

We estimate the difference \( |\mathcal{E}_{t \wedge \theta_R}|^2 = |y_{t \wedge \theta_R} - x_{t \wedge \theta_R}|^2 \). Itô’s formula implies that

\[
|\mathcal{E}_{t \wedge \theta_R}|^2 = \int_0^{t \wedge \theta_R} 2f(\hat{s}, s, \hat{y}_s, y_s) - f_2(s, s, x_s, x_s) - f_1(s, x_s) |\mathcal{E}_s| + (g(\hat{s}, s, y_s, y_s) - g(s, s, x_s, x_s))^2 \, ds \\
+ |f_1(t_0, y_{t_0})|^2 + \int_0^{t \wedge \theta_R} 2|\mathcal{E}_s| (g(\hat{s}, s, y_s, y_s) - g(s, s, x_s, x_s)) \, dW_s \\
\leq \int_0^{t \wedge \theta_R} |f_2(\hat{s}, s, \hat{y}_s, y_s) - f_2(s, s, x_s, x_s)|^2 ds + \int_0^{t \wedge \theta_R} |\mathcal{E}_s|^2 ds + 2 \int_0^{t \wedge \theta_R} |f_1(s, x_s)| |\mathcal{E}_s| ds + 2M_t \\
+C(1 + |x_0|^2) \Delta^2 + \int_0^{t \wedge \theta_R} |g(\hat{s}, s, y_s, y_s) - g(s, s, x_s, x_s)|^2 \, ds,
\]

where \( M_t := \int_0^{t \wedge \theta_R} |\mathcal{E}_s| (g(\hat{s}, s, y_s, y_s) - g(s, s, x_s, x_s)) \, dW_s \). It holds that

\[
\mathbb{E} \sup_{0 \leq t \leq T} |M_t| \leq 2\sqrt{32} \cdot \mathbb{E} \int_0^{t \wedge \theta_R} |\mathcal{E}_s|^2 (g(\hat{s}, s, y_s, y_s) - g(s, s, x_s, x_s))^2 \, ds \\
\leq \mathbb{E} \sqrt{\sup_{0 \leq s \leq T} |\mathcal{E}_{s \wedge \theta_R}|^2} \cdot 128 \int_0^{t \wedge \theta_R} (g(\hat{s}, s, y_s, y_s) - g(s, s, x_s, x_s))^2 \, ds \\
\leq \frac{1}{2} \mathbb{E} \sup_{0 \leq s \leq T} |\mathcal{E}_{s \wedge \theta_R}|^2 + 64 \mathbb{E} \int_0^{t \wedge \theta_R} (g(\hat{s}, s, y_s, y_s) - g(s, s, x_s, x_s))^2 \, ds,
\]
thus we get that

\begin{equation}
\mathbb{E} \sup_{0 \leq t \leq T} |\mathcal{E}_{t \land \theta_R}|^2 \leq 2\mathbb{E} \sup_{0 \leq t \leq T} \int_0^{t \land \theta_R} |f(\bar{s}, s, y_\bar{s}, y_s) - f(s, s, x_s, x_s)|^2 ds + C_R \int_0^{t \land \theta_R} \mathbb{E} \sup_{0 \leq t \leq s} |\mathcal{E}_t|^2 ds \\
+ C \Delta^2 + 130 \cdot \mathbb{E} \int_0^{T \land \theta_R} |g(\bar{s}, s, y_\bar{s}, y_s) - g(s, s, x_s, x_s)|^2 ds.
\end{equation}

(5.4)

Assumption 3.7 implies that

\[ \int_0^{t \land \theta_R} |f_2(\bar{s}, s, y_\bar{s}, y_s) - f_2(s, s, x_s, x_s)|^2 ds \leq \int_0^{T \land \theta_R} C_R \left( |y_s - y_{\bar{s}}|^2 + |\mathcal{E}_s|^2 + |\bar{s} - s|^2 \right) ds \]

Moreover, it holds that

\[ \int_0^{t \land \theta_R} |\bar{s} - s|^2 ds \leq \sum_{k=0}^{t/\Delta - 1} \int_{t_k}^{t_{k+1} \land \theta_R} |t_k - s|^2 ds. \]

Taking the supremum over all \( t \in [0, T] \) and then expectation we have

\begin{equation}
\mathbb{E} \sup_{0 \leq t \leq T} \int_0^{T \land \theta_R} |f_2(\bar{s}, s, y_\bar{s}, y_s) - f_2(s, s, x_s, x_s)|^2 ds \leq C_R \Delta + C_R \int_0^{T \land \theta_R} \mathbb{E} \sup_{0 \leq t \leq s} |\mathcal{E}_{t \land \theta_R}|^2 ds + C_R \Delta^2,
\end{equation}

where we have used Lemma 5.9 for \( p = 2 \). Using Assumption 3.7 again we get that

\[ \int_0^{T \land \theta_R} \left( g(\bar{s}, s, y_\bar{s}, y_s) - g(s, s, x_s, x_s) \right)^2 ds \leq \int_0^{T \land \theta_R} C_R \left( |y_s - y_{\bar{s}}|^2 + |\mathcal{E}_s|^2 + |\bar{s} - s|^2 + |y_s - x_s|^2 \right) ds \]

\[ \leq \int_0^{T \land \theta_R} C_R \left( |y_s - y_{\bar{s}}|^2 + |\mathcal{E}_s|^2 + |\bar{s} - s|^2 + (|y_s - y_{\bar{s}}|^2 + |\mathcal{E}_s|^2) \right) ds, \]

where we have used the subadditivity property of \( h(x) = x^{2q} \), thus taking expectation we have

\[ \mathbb{E} \int_0^{T \land \theta_R} \left( g(\bar{s}, s, y_\bar{s}, y_s) - g(s, s, x_s, x_s) \right)^2 ds \leq C_R \Delta + C_R \int_0^{T \land \theta_R} \mathbb{E} \sup_{0 \leq t \leq s} |\mathcal{E}_{t \land \theta_R}|^2 ds + C_R \Delta^2 \]

\[ + 2^{2q-1} C_R T \Delta^q + 2^{2q-1} C_R \int_0^{T \land \theta_R} (\mathbb{E} |\mathcal{E}_s|)^{2q} ds, \]

where we have applied again Lemma 5.9 for \( p = 2q \) and Jensen inequality with \( 2q < 1 \). We get the following estimate

\[ \mathbb{E} \int_0^{T \land \theta_R} \left( g(\bar{s}, s, y_\bar{s}, y_s) - g(s, s, x_s, x_s) \right)^2 ds \leq C_R \Delta^q + C_R \int_0^{T \land \theta_R} \mathbb{E} \sup_{0 \leq t \leq s} (\mathcal{E}_{t \land \theta_R})^2 ds \]

\[ + C_R \int_0^{T} (K_{R, \Delta, m, q}(s))^{2q} ds \]

(5.6) \[ \leq C_R \Delta^q + C_R (K_{R, \Delta, m, q}(T))^{2q} + C_R \int_0^{T} \mathbb{E} \sup_{0 \leq t \leq s} (\mathcal{E}_{t \land \theta_R})^2 ds, \]

where we have used Proposition 5.10 and

\[ K_{R, \Delta, m, q}(s) := \left( C_R \sqrt{\Delta} + \frac{C_R}{m e_m} \Delta^q + \frac{C_R}{m} e_m + s \right) e^{a_R, m, a_s} \]
Plugging the estimates (5.5), (5.6) into (5.4) gives
\[
\mathbb{E} \sup_{0 \leq t \leq T} |\mathcal{E}_{t\theta R}|^2 \leq C_R \Delta^q + C_R (K_{R,\Delta, m,q}(T))^{2q} + C_R \int_0^T \mathbb{E} \sup_{0 \leq s \leq s} (\mathcal{E}_{s\theta R})^2 ds
\]
\[
\leq (C_R \Delta^q + C_R (K_{R,\Delta, m,q}(T))^{2q}) e^{C_R T} \leq C_{R,\Delta,m}
\]
where we have applied the Gronwall inequality. Note that, given Theorem 3.8 we have the strong convergence result of Theorem 2.4

discretized the drift coefficient of (1.1) in an additive way. Therefore, by an application of Relation (5.3) becomes,
\[
\mathbb{E} \sup_{0 \leq t \leq T} |\mathcal{E}_t|^2 \leq C_{R,\Delta,m} + 2 \frac{3p-2}{p} AR^{2-p}
\]
\[
\leq \underbrace{C_{R,\Delta,m}}_{I_1} + \underbrace{2 \frac{3p-2}{p} AR^{2-p}}_{I_2}.
\]
Given any \( \varepsilon > 0 \), we may first choose \( R \) such that \( I_2 < \varepsilon / 2 \), then choose \( m \) big enough and \( \Delta \) small enough such that \( I_1 < \varepsilon / 2 \) a concluding \( \mathbb{E} \sup_{0 \leq t \leq T} |\mathcal{E}_t|^2 < \varepsilon \) as required to verify (3.4).

6. PROOF OF THEOREM 2.4, PROPOSITION 2.5 AND THEOREM 2.6.

In this Section we prove our main strong convergence result. First, we provide uniform moment bounds for the original SDE and the SD scheme. We remind here that for notational reasons the processes \((W_t, x_t)\) stand for \((\hat{W}_t, \hat{x}_t)\).

**Lemma 6.11.** [Moment bounds for original problem and SD approximation] Let Assumption 2.2 hold. Then
\[
\mathbb{E} \sup_{0 \leq t \leq T} |x_t|^p \bigvee \mathbb{E} \sup_{0 \leq t \leq T} |y_t|^p \leq 1,
\]
for any \( p > 0 \). \( \square \)

**Proof of Lemma 6.11.** The result is trivial since we already know that \((x_t)\) satisfying (1.1) has the property \( x_t \in D \) when \( x_0 \in D, D = (0, 1) \), by Appendix A and regarding the bounds for the SD approximation it is clear, by its form (2.4), that they are valid. \( \square \)

Now, let us rewrite the approximation process \((y_t^{SD})\)
(6.1)
\[
y_t^{SD} = y_n + \left( k_1 - \frac{(k_3)^2}{4} + y_n \left( \frac{(k_3)^2}{2} - k_2 \right) \right) \Delta + \int_{t_n}^t \frac{(k_3)^2}{4} (1 - 2y_s) ds + k_3 \int_{t_n}^t \sqrt{y_s(1 - y_s)} dW_s.
\]
In the general setting of (3.2) we have
\[
f_1(x) = k_1 - \frac{(k_3)^2}{4} + x \left( \frac{(k_3)^2}{2} - k_2 \right), \quad f_2(x) = \frac{(k_3)^2}{4} (1 - 2x), \quad g(x) = b(x) = \sqrt{x(1 - x)}.
\]
By the above representation, the form of the discretization becomes apparent. We only discretized the drift coefficient of (1.1) in an additive way. Therefore, by an application of Theorem 3.8 we have the strong convergence result of Theorem 2.4
\[
\lim_{\Delta \to 0} \mathbb{E} \sup_{0 \leq t \leq T} |y_t - x_t|^2 = 0.
\]
Now, we briefly sketch the proof of Proposition 2.5. The process (2.11) is well-defined when \( 0 < \tilde{y}_n < 1 \) or equivalently when \( (k_3)^2 < 2k_2 \) and \( \Delta < -1/\beta \) using (2.12) and (2.10).
The strong convergence result of Proposition 2.5 is a consequence of the triangle inequality and the following regularity-type result

\[ |y_{i}^{SD} \mid | y_{i}^{SD} | = \left| \sin^{2} \left( \frac{k_{3}}{2} \Delta W_{n}^{t} + \arcsin(\sqrt{y_{n}}) \right) - \sin^{2} \left( \frac{k_{3}}{2} \Delta W_{n}^{t} + \arcsin(\sqrt{y_{n}}) \right) \right| \]

\[ \leq 2 \left| \sin \left( \frac{k_{3}}{2} \Delta W_{n}^{t} + \arcsin(\sqrt{y_{n}}) \right) - \sin \left( \frac{k_{3}}{2} \Delta W_{n}^{t} + \arcsin(\sqrt{y_{n}}) \right) \right| \]

\[ \leq 2 \left| \arcsin(\sqrt{y_{n}}) - \arcsin(\sqrt{y_{n}}) \right| \]

\[ = \left| \int_{y_{n}}^{\tilde{y}_{n}} \frac{1}{\sqrt{z(1-z)}} \, dz \right| \leq |\alpha + \beta| \sup_{z \in \{y_{n}, \tilde{y}_{n}\}} \frac{1}{\sqrt{z(1-z)}} \leq C \cdot \Delta, \]

for any \( t \in (t_{n}, t_{n+1}], \) where \( C \) is finite positive.

Theorem 2.6 is an application of a slight generalization of Theorem 3.8 including multi-dimensional noise (see also [16]). Therefore, we omit the proof since one essentially repeats the steps of the proof of Theorem 3.8. The auxiliary functions in the sense of (3.2) are,

\[ f_{1}(X) = k_{1}^{(1,1)} + k_{1}^{(1,2)} X^{(2)} - \frac{(k_{3}^{(1,1)})^{2} X^{(2)} + (k_{3}^{(1,2)})^{2}(1 - X^{(1)} - X^{(2)})}{4(1 - X^{(1)})} \]

\[ + X^{(1)} \left( \frac{(k_{3}^{(1,1)})^{2} X^{(2)} + (k_{3}^{(1,2)})^{2}(1 - X^{(1)} - X^{(2)})}{2(1 - X^{(1)})} - k_{2}^{(1)} \right) \]

\[ f_{2}(X, Y^{(1)}) = \frac{(k_{3}^{(1,1)})^{2} X^{(2)} + (k_{3}^{(1,2)})^{2}(1 - X^{(1)} - X^{(2)})}{4(1 - X^{(1)})} \]

\[ g_{11}(X, Y^{(1)}) = k_{3}^{(1,1)} \sqrt{\frac{X^{(2)}}{1 - X^{(1)}}} \sqrt{Y^{(1)}(1 - Y^{(1)})} \]

\[ g_{12}(X, Y^{(1)}) = k_{3}^{(1,2)} \sqrt{\frac{1 - X^{(1)} - X^{(2)}}{1 - X^{(1)}}} \sqrt{Y^{(1)}(1 - Y^{(1))}}, \]

for the evolution of the first component (1.3), where \( X \) denotes the discretized part of the SDE, and accordingly for the second component (1.4)

\[ f_{1}(X) = k_{1}^{(2,1)} + k_{1}^{(2,2)} X^{(1)} - \frac{(k_{3}^{(2,1)})^{2} X^{(1)} + (k_{3}^{(2,3)})^{2}(1 - X^{(1)} - X^{(2)})}{4(1 - X^{(2)})} \]

\[ + X^{(2)} \left( \frac{(k_{3}^{(2,1)})^{2} X^{(1)} + (k_{3}^{(2,3)})^{2}(1 - X^{(1)} - X^{(2)})}{2(1 - X^{(2)})} - k_{2}^{(2)} \right) \]

\[ f_{2}(X, Y^{(2)}) = \frac{(k_{3}^{(2,1)})^{2} X^{(1)} + (k_{3}^{(2,3)})^{2}(1 - X^{(1)} - X^{(2)})}{4(1 - X^{(2)})} \]

\[ g_{21}(X, Y^{(2)}) = k_{3}^{(2,1)} \sqrt{\frac{X^{(1)}}{1 - X^{(2)}}} \sqrt{Y^{(2)}(1 - Y^{(2)})} \]

\[ g_{23}(X, Y^{(2)}) = k_{3}^{(2,3)} \sqrt{\frac{1 - X^{(1)} - X^{(2)}}{1 - X^{(2)}}} \sqrt{Y^{(2)}(1 - Y^{(2))}}, \]
By the above representation, the form of the discretization of (1.3) and (1.4) becomes apparent. We discretized the drift coefficient in an additive and multiplicative way and the diffusion coefficient in a multiplicative way.

Acknowledgements. The author would like to thank the anonymous referees for their helpful comments.

References

[1] H. Schurz. Numerical regularization for SDEs: Construction of nonnegative solutions. *Dynamic Systems and Applications*, 5(3):323–351, 1996.

[2] N. Halidias. Semi-discrete approximations for stochastic differential equations and applications. *International Journal of Computer Mathematics*, 89(6):780–794, 2012.

[3] N. Halidias and I.S. Stamatiou. On the Numerical Solution of Some Non-Linear Stochastic Differential Equations Using the Semi-Discrete Method. *Computational Methods in Applied Mathematics*, 16(1):105–132, 2016.

[4] N. Halidias and I.S. Stamatiou. Approximating Explicitly the Mean-Reverting CEV Process. *Journal of Probability and Statistics*, Article ID 513137, 20 pages, 2015.

[5] W. J. Ewens. *Mathematical Population Genetics 1: Theoretical Introduction*, volume 27. Springer Science & Business Media, 2012.

[6] C.E. Dangerfield, D. Kay, S. MacNamara, and K Burrage. A boundary preserving numerical algorithm for the wright-fisher model with mutation. *BIT Numerical Mathematics*, 52(2):283–304, 2012.

[7] J. H. Goldwyn, N. S. Imennov, M. Famulare, and E. Shea-Brown. Stochastic differential equation models for ion channel noise in hodgkin-huxley neurons. *Physical Review E*, 83(4):041908, 2011.

[8] C. E. Dangerfield, D. Kay, and K. Burrage. Modeling ion channel dynamics through reflected stochastic differential equations. *Physical Review E*, 85(5):051907, 2012.

[9] R.C. Griffiths. On the distribution of allele frequencies in a diffusion model. *Theoretical Population Biology*, 15(1):140 – 158, 1979.

[10] R.C. Griffiths. Allele frequencies in multidimensional wright-fisher models with a general symmetric mutation structure. *Theoretical Population Biology*, 17(1):51 – 70, 1980.

[11] I. Karatzas and S.E. Shreve. *Brownian motion and stochastic calculus*. Springer-Verlag, New York, corrected 4th printing, 1997.

[12] X. Mao. *Stochastic differential equations and applications*. Horwood Publishing, Chichester, 2nd edition, 2007.

[13] C.E. Dangerfield, D. Kay, and K. Burrage. Stochastic models and simulation of ion channel dynamics. *Procedia Computer Science*, 1(1):1587 – 1596, 2010.

[14] T. Yamada and S. Watamabe. On the uniqueness of solutions of stochastic differential equations. *Journal of Mathematics of Kyoto University*, 11(1):155–167, 1971.

[15] T.H. Gronwall. Note on the derivatives with respect to a parameter of the solutions of a system of differential equations. *Annals of Mathematics*, 20:292–296, 1919.

[16] Halidias N. Construction of positivity preserving numerical schemes for some multidimensional stochastic differential equations. *Discrete and Continuous Dynamical Systems - Series B*, 20(1):153–160, 2015.

Appendix A. Boundary Classification of One-Dimensional Time-Homogeneous SDEs.

Let us now recall some results [11, Sec. 5.5] concerning the boundary behavior of SDEs of the form,

(A.1) \[ dX_t = a(X_t)dt + b(X_t)dW_t. \]

Let \( I = (l, r) \) be an interval with \(-\infty \leq l < r \leq \infty\) and define the exit time from \( I \) to be \( S := \inf\{t \geq 0 : X_t \notin (l, r)\} \).
Let also the coefficients of (A.1) satisfy the following conditions
\[ b^2(x) > 0, \quad \forall x \in I, \quad \text{(Non Degeneracy), (ND)}, \]
\[ \forall x \in I, \exists \epsilon > 0 : \int_{x-\epsilon}^{x+\epsilon} \frac{1 + |a(y)|}{b^2(y)} dy < \infty, \quad \text{(Local Integrability), (LI)}. \]

Then for \( c \in I \), we can define the scale function
\[ (A.2) \quad s(x) := \int_c^x e^{-2 \int_c^y \frac{a(z)}{b^2(z)} dz} dy, \]
whose behavior at the endpoints of \( I \) determines the boundary behavior of \( (X_t) \) [11, Prop. 5.5.22]. In particular, we get that the dynamics (1.1) have a boundary behavior which is determined by the scale function
\[ s(x) = \int_c^x \exp \left\{ -2 \int_c^y \frac{k_1 - k_2 z}{(k_3)^2 z(1 - z)} dz \right\} dy \]
\[ = \int_c^x \exp \left\{ -2 \int_c^y \frac{k_1}{(k_3)^2} \right\} \left[ -z^{-1}(1 - z)^{-1} dz + 2 \frac{k_2}{(k_3)^2} \int_c^y (1 - z)^{-1} dz \right\} dy \]
\[ = - \int_x^c \exp \left\{ -2 \frac{k_1}{(k_3)^2} \ln(y/c) + 2 \left( \frac{k_1 - k_2}{(k_3)^2} \right) \ln(1 - y)/(1 - c) \right\} dy \]
\[ = - C \int_x^c y^{-2 \frac{k_1}{(k_3)^2}} (1 - y)^{2 \frac{k_1 - k_2}{(k_3)^2}} dy, \]
for any \( x \in I \) where \( C > 0 \). Let \( I = (0, 1) \) and take \( c = 1/2 \). We compute
\[ s(0+) = - C \int_0^{1/2} y^{-2 \frac{k_1}{(k_3)^2}} (1 - y)^{2 \frac{k_1 - k_2}{(k_3)^2}} dy = -\infty, \]
when \( k_1 > 0 \) and
\[ s(1-) = C \int_{1/2}^1 y^{-2 \frac{k_1}{(k_3)^2}} (1 - y)^{2 \frac{k_1 - k_2}{(k_3)^2}} dy = \infty, \]
when \( k_1 < k_2 \), thus by [11, Prop. 5.5.22a] we have that \( P(S = \infty) = 1 \) that is \( P(0 < x_t < 1) = 1 \).

**Appendix B. Solution process of stochastic integral equations (2.1), (2.13), (2.14).**

We will show that the process (2.4) for \( n = 0 \), is the solution of the stochastic integral equation (2.1) for \( n = 0 \), that is
\[ (B.1) \quad y_{t}^{SD} = \sin^2 \left( \frac{k_3}{2} W_t + \arcsin(\sqrt{Y_0}) \right), \]
satisfies
\[ y_{t}^{SD} = Y_0 + \int_0^t \frac{(k_3)^2}{4} (1 - 2y_s) ds + k_3 \int_0^t \sqrt{y_s(1 - y_s)} dW_s, \]
for \( t \in (0, t_1] \), with
\[ Y_0 := x_0 + \left( k_1 - \frac{(k_3)^2}{4} + x_0 \left( \frac{(k_3)^2}{2} - k_2 \right) \right) \cdot \Delta \leq 1. \]
Relations (2.6) and (2.3) yield
\[ d\hat{W}_t := \text{sgn}(z_t) dW_t, \]
where
\[ z_t = \sin \left( k_3 \Delta W + 2 \arcsin(\sqrt{Y_0}) \right). \]
The cases for \( n = 1, \ldots, N - 1 \) follow with the appropriate modifications.

We can write the increment of the Wiener process as
\[ dW_t = 0 \cdot dt + 1 \cdot dW_t, \]
and view (B.1) as a function of \( W_t \), i.e.
\[ y = V(W)_t \]
with
\[ \frac{dy}{dW} = \frac{(k_3)^2 \cos(k_3 \Delta W + 2 \arcsin(\sqrt{Y_0}))}{2} \]
\[ \frac{d^2 y}{dW^2} = \frac{(k_3)^2}{2} \cos \left( k_3 \Delta W + 2 \arcsin(\sqrt{Y_0}) \right) \]
\[ = \frac{(k_3)^2}{2} (1 - 2y). \]

Application of Itô’s formula implies
\[ dy_t = \frac{1}{2} V'(W_t) dt + V'(W_t) dW_t \]
\[ = \frac{(k_3)^2}{4} (1 - 2y_t) dt + k_3 \sqrt{y_t(1 - y_t)} d\hat{W}_t. \]

For the derivation of (2.13) and (2.14) we now write the multidimensional Wiener process as
\[ dW_t = \mathcal{O}_3 \cdot dt + \mathcal{I}_3 \cdot dW_t, \]
where \( \mathcal{O}_3 \) is the zero \( 3 \times 3 \) matrix and \( \mathcal{I}_3 \) the \( 3 \times 3 \) identity matrix and apply appropriately the multidimensional Itô formula.

**APPENDIX C. UNIFORM MOMENT ESTIMATE FOR \( |y_t - x_t|^2 \).**

In Theorem 2.4 we actually proved that \( \lim_{\Delta t \to 0} \mathbb{E} \sup_{0 \leq t \leq T} |y_t^{SD} - \hat{x}_t|^2 = 0 \). In order to finish the proof we have to find a uniform moment bound for \( |x_t - \hat{x}_t|^2 \). In particular applying the triangle inequality
\[ \lim_{\Delta t \to 0} \mathbb{E} \sup_{0 \leq t \leq T} |y_t^{SD} - x_t|^2 \leq \lim_{\Delta t \to 0} \mathbb{E} \sup_{0 \leq t \leq T} |y_t^{SD} - \hat{x}_t|^2 + \lim_{\Delta t \to 0} \mathbb{E} \sup_{0 \leq t \leq T} |x_t - \hat{x}_t|^2 \]
\[ \leq \lim_{\Delta t \to 0} \mathbb{E} \sup_{0 \leq t \leq T} |x_t - \hat{x}_t|^2 \]
thus it suffices to show
\[ \lim_{\Delta t \to 0} \mathbb{E} \sup_{0 \leq t \leq T} |x_t - \hat{x}_t|^2 = 0. \]
this follows by (1.1) and (2.8).

E-mail address: joniou@gmail.com