mathematics

Article

Weak Approximations of the Wright–Fisher Process

Vigirdas Mackevičius † and Gabrielė Mongirdaitė *,†

Faculty of Mathematics and Informatics, Institute of Mathematics, Vilnius University, Naugarduko 24, 03225 Vilnius, Lithuania; vigirdas.mackevicius@mif.vu.lt
* Correspondence: gabriele.mongirdaite@mif.vu.lt
† These authors contributed equally to this work.

Abstract: In this paper, we construct first- and second-order weak split-step approximations for the solutions of the Wright–Fisher equation. The discretization schemes use the generation of, respectively, two- and three-valued random variables at each discretization step. The accuracy of constructed approximations is illustrated by several simulation examples.

Keywords: weak approximations; split-step; Wright–Fisher equation; Jacobi equation

MSC: 60H35, 65C30

check for updates

Citation: Mackevičius, V.; Mongirdaitė, G. Weak Approximations of the Wright–Fisher Process. *Mathematics* **2022**, *10*, 125. https://doi.org/10.3390/math10010125

Academic Editor: Manuel Alberto M. Ferreira

Received: 8 December 2021
Accepted: 30 December 2021
Published: 1 January 2022

1. Introduction

We are interested in weak first- and second-order approximations for the Wright–Fisher equation

$$X_t^x = x + \int_0^t (a - bX_s^x)\, ds + \sigma \int_0^t \sqrt{X_s^x(1 - X_s^x)}\, dB_s, \quad x \in [0,1], \tag{1}$$

with parameters $0 \le a \le b$ and $\sigma > 0$. The Wright–Fisher (WF) process (a solution of Equation (1)) is well defined in $[0,1]$ and models the gene frequencies in a population. The main problem in developing numerical methods for "square-root" diffusions is that the diffusion coefficient has unbounded derivatives near "singular" points (in our case, 0 and 1), and therefore standard methods (see, e.g., Milstein and Tretyakov [1]) are not applicable; typically, discretization schemes involving (explicitly or implicitly) the derivatives of the coefficients usually lose their accuracy near singular points, especially for large σ.

Alfonsi [2] (Chap. 6) constructed a weak second-order approximation of the WF process by using its connection with the Cox–Ingersoll–Ross (CIR) [3] process and the earlier constructed approximations of the latter (Alfonsi [4]). The main result of this paper is a direct construction of first- and second-order weak split-step approximations of the WF processes by discrete random variables. We believe that in comparison with the numerical scheme of Alfonsi [2] (Prop. 6.1.13, Algs. 6.1 and 6.2), our algorithm is much simpler and easier to implement. In our construction, we follow some ideas of Lileika and Mackevičius [5,6]. However, we had to overcome a serious additional challenge (in comparison with CIR or CKLS processes): two "singular" points, 0 and 1, of the diffusion coefficient make it essentially more difficult to ensure that the approximations take values in $[0,1]$ (instead of $[0,+\infty)$ as in [5,6]).

The paper is organized as follows. In Section 2, we recall some definitions and results. In Sections 3 and 4, we construct first- and second-order approximations for the WF equation by two- and three-valued discrete random variables, respectively. The main results of these sections are presented as Theorems 4 and 5. We illustrate the accuracy of our approximations by several simulation examples. In Section 5, we prove an auxiliary result on the smoothness of solutions of the corresponding backward Kolmogorov PDE equation. Tedious technical calculations have been performed using Maple and Python.

2. Preliminaries

In this section, we give some known definitions adapted to our context of the WF process defined by Equation (1).

Having a fixed time interval $[0, T]$, consider an equidistant time discretization $\Delta^h = \{ih, i = 0, 1, \ldots, [T/h], h \in (0, T]\}$, where $[a]$ is the integer part of a. By a discretization scheme (or approximation) of Equation (1) we mean any family of discrete-time homogeneous Markov chains $\hat{X}^h = \left\{\hat{X}^h(x,t), x \in [0,1], t \in \Delta^h\right\}$ in [0, 1] with initial values $\hat{X}^h(x, 0) = x$ and one-step transition probabilities $p^h(x, \mathrm{d}z)$, $x \in [0, 1]$. For convenience, we only consider steps $h = T/n$, $n \in \mathbb{N}$. For brevity, we sometimes write $\hat{X}^x_t$ or $\hat{X}(x, t)$ instead of $\hat{X}^h(x, t)$. Note that because of the Markovity, a one-step approximation $\hat{X}^x_h$ of the scheme completely defines the distribution of the whole discretization scheme $\hat{X}^x_t$, so that we only need to construct the former. Therefore, we will abbreviate one-step approximations as approximations. As usual, $\mathbb{N}$ and $\mathbb{R}$ are the sets of natural and real numbers, $\overline{\mathbb{N}} := \mathbb{N} \cup \{0\}$, and $\mathbb{R}_+ := [0, \infty)$.

We will write $g(x, h) = O(h^n)$ if, for some $C > 0$ and $h_0 > 0$,

$$|g(x,h)| \le Ch^n, \quad x \in [0,1], \quad 0 < h \le h_0.$$

Definition 1 (c.f. [4], Def. 1.3, [6], Def. 1). *A discretization scheme $\hat{X}^h$ is a weak νth-order approximation for the solution $(X^x_t, t \in [0, T])$ of Equation (1) if for every $f \in C^\infty[0, 1]$,*

$$\left|\mathbb{E}f(X^x_T) - \mathbb{E}f\left(\hat{X}^x_T\right)\right| = O(h^\nu).$$

Definition 2 (c.f. [4], Def. 1.8, [6], Defs. 2, 3). *The νth-order remainder of a discretization scheme $\hat{X}^x_t$ for X^x_t is the operator $R^h_v : C^\infty[0, 1] \to C[0, 1]$ defined by*

$$R^h_\nu f(x) := \mathbb{E}f\left(\hat{X}^x_h\right) - \left[f(x) + \sum_{k=1}^{\nu} \frac{A^k f(x)}{k!} h^k\right], x \in [0,1], h > 0,$$

where A is the generator of X^x_t,

$$Af(x) = (a - bx)f'(x) + \frac{1}{2}\sigma^2 x(1-x)f''(x).$$

A discretization scheme $\hat{X}^x_t$ is a potential νth-order weak approximation of Equation (1) if

$$R^h_\nu f(x) = O(h^{\nu+1})$$

for all $f \in C^\infty[0, 1]$ and $x \in [0, 1]$.

The following two theorems ensure that a potential νth-order weak approximation is in fact a νth-order weak approximation (in the sense of Definition 1). Note that the requirement of uniformly bounded moments (see, e.g., [4]) is obviously satisfied by our approximations since they take values in [0, 1].

Theorem 1 (see Theorem 1.19 of [4]). *Let $\hat{X}^h$ be a discretization scheme with transition probabilities $p^h(x, \mathrm{d}z)$ on $[0, 1]$ that starts from $\hat{X}^x_0 = x \in [0, 1]$. We assume that*

1. *the scheme is a potential weak νth-order discretization scheme for the operator A.*
2. *$f \in C^\infty[0, 1]$ is a function such that $u(t, x) = \mathbb{E}f(X^x_{T-t})$ defined on $[0, T] \times [0, 1]$ solves $\partial_t u(t, x) = -Au(t, x)$ for $(t, x) \in [0, T] \times [0, 1]$.*

Then $|\mathbb{E}f(\hat{X}^x_T) - \mathbb{E}f(X^x_T)| = O(h^\nu)$.

Theorem 2 (see Theorem 6.1.12 of [2]). *Let $f \in C^\infty[0, 1]$. Then*

$$\widetilde{u}(t,x) := \mathbb{E}f(X^x_t), \quad (t,x) \in \mathbb{R}_+ \times [0,1],$$

is a C^∞ function that solves

$$\partial_t \widetilde{u}(t,x) = A\widetilde{u}(t,x). \tag{2}$$

We split Equation (1) into the deterministic part

$$\mathrm{d}D_t^x = (a - bD_t^x)\mathrm{d}t, D_0^x = x \in [0,1], \tag{3}$$

and the stochastic part

$$\mathrm{d}S_t^x = \sigma\sqrt{S_t^x(1 - S_t^x)}\mathrm{d}B_t, \quad S_0^x = x \in [0,1]. \tag{4}$$

The solution of the deterministic part is positive for all $(x,t) \in [0,1] \times (0,T]$, namely:

$$D_t^x = D(x,t) = \begin{cases} x\mathrm{e}^{-bt} + \frac{a}{b}\left(1 - \mathrm{e}^{-bt}\right), & 0 \le a \le b \ne 0, \\ x, & a = b = 0. \end{cases} \tag{5}$$

The solution of the stochastic part is not explicitly known. However, suppose that $\hat{S}_t^x$ is a discretization scheme for the stochastic part. We define the first-order composition $\hat{X}_t^x$ of the latter with the solution of the deterministic part as a Markov chain that has the transition probability in one step equal to the distribution of the random variable

$$\hat{X}^h(x,h) := D(\hat{S}(x,h),h). \tag{6}$$

Similarly, the second-order composition is defined by

$$\hat{X}^h(x,h) := D\left(\hat{S}\left(D\left(x,\tfrac{h}{2}\right),h\right),\tfrac{h}{2}\right). \tag{7}$$

Theorem 3 (see [4], Thm. 1.17)**.** *Let $\hat{S}_t^x$ be a potential first- or second-order approximation of the stochastic part of the WF equation. Then, compositions (6) and (7) define, respectively, a first- or second-order approximation $\hat{X}_t^x$ of the WF Equation (1).*

From this theorem, it follows that to construct a first- or second-order weak approximation, we only need to construct a first- or second-order approximation of the stochastic part, respectively.

Remark 1. *For various applications, we may be interested in similar processes with values in $[\alpha,\beta]$ satisfying the equation*

$$\mathrm{d}\widetilde{X}_t = (\tilde{a} - b\widetilde{X}_t)\ \mathrm{d}t + \sigma\sqrt{(\widetilde{X}_t - \alpha)(\beta - \widetilde{X}_t)}\ \mathrm{d}B_t, \quad \widetilde{X}_0 \in [\alpha,\beta], \tag{8}$$

which is well defined when $b\alpha \le \tilde{a} \le b\beta$. A popular choice is the Jacobi process with $\alpha = -1$ and $\beta = 1$. Process (8) can be obtained from the WF process by the affine transformation $\widetilde{X}_t = \alpha + (\beta - \alpha)X_t$ $(\tilde{a} = a(\beta - \alpha))$. Clearly, by the same transformation we can get weak approximations for (8) from weak approximations for the WF process.

3. First-Order Weak Approximation of Wright–Fisher Equation

3.1. Approximation of the Stochastic Part

Let us construct an approximation for the stochastic part of the WF equation, that is, the solution S_t^x of Equation (1) with $a = b = 0$. A two-valued discrete random variable $\hat{S}_h^x$ taking values $x_1, x_2 \in [0,1]$ with probabilities p_1, p_2 is a first-order weak approximation if (see [5] and references therein)

$$p_1 + p_2 = 1, \tag{9}$$

$$\mathbb{E}\hat{S}_h^x = x_1p_1 + x_2p_2 = m_1 := \mathbb{E}S_h^x = x, \tag{10}$$
$$\mathbb{E}(\hat{S}_h^x)^2 = x_1^2p_1 + x_2^2p_2 = m_2 + O(h^2), \tag{11}$$
$$\mathbb{E}(\hat{S}_h^x - x)^3 = (x_1 - x)^3p_1 + (x_2 - x)^3p_2 = O(h^2), \tag{12}$$
$$\mathbb{E}(\hat{S}_h^x - x)^4 = (x_1 - x)^4p_1 + (x_2 - x)^4p_2 = O(h^2), \tag{13}$$

where the second moment $m_2 = \mathbb{E}(S_h^x)^2$ can be calculated by Lemma 6 with $a = b = 0$:

$$m_2 = m_2(x,h) = x^2e^{-\sigma^2h} + x(1 - e^{-\sigma^2h}) \tag{14}$$
$$= x^2 + x(1-x)\sigma^2h + O(h^2),\ x \in [0,1]. \tag{15}$$

One of the solutions to the equation system (9)–(11) is (see [5])

$$x_{1,2} = \frac{m_2}{m_1} \mp \sqrt{\frac{m_2(m_2 - m_1^2)}{m_1^2}},$$
$$p_{1,2} = \frac{x}{2x_{1,2}}.$$

Therefore, in our case, we get

$$x_{1,2} = xe^{-\sigma^2h} + 1 - e^{-\sigma^2h} \tag{16}$$
$$\mp \sqrt{\frac{xe^{-\sigma^2h} + (1 - e^{-\sigma^2h})}{x}\left(x^2e^{-\sigma^2h} + x(1 - e^{-\sigma^2h}) - x^2\right)}$$
$$= xe^{-\sigma^2h} + 1 - e^{-\sigma^2h} \mp \sqrt{(xe^{-\sigma^2h} + 1 - e^{-\sigma^2h})(1-x)(1 - e^{-\sigma^2h})}. \tag{17}$$

Since $1 - e^{-\sigma^2h} = \sigma^2h + O(h^2)$, to simplify the expressions, we may try to replace $1 - e^{-\sigma^2h}$ by σ^2h and, instead of (17), use

$$x_{1,2} = x_{1,2}(x,h) = x(1 - \sigma^2h) + \sigma^2h \mp \sqrt{(x(1-\sigma^2h) + \sigma^2h)(1-x)\sigma^2h}$$
$$= x + (1-x)\sigma^2h \mp \sqrt{(x + (1-x)\sigma^2h)(1-x)\sigma^2h}. \tag{18}$$

In Lemma 1, we will check that after this replacement, $x_{1,2}$ and $p_{1,2}$ still satisfy (9)–(13). Unfortunately, for the values of x near 1, the values of x_2 are slightly greater than 1 (as well as those defined by (17)), which is unacceptable. We overcome this problem by using the symmetry of the solution of the stochastic part with respect to the point $\frac{1}{2}$; to be precise, $S_t^x \overset{d}{=} 1 - S_t^{1-x}$ for all $x \in [0,1]$ ($\overset{d}{=}$ means equality in distribution). Therefore, in the interval $[0,1/2]$, we can use the values $x_{1,2}$ defined by (18), whereas in the interval $(1/2,1]$, we use the values corresponding to the process $1 - S_t^{1-x}$, that is,

$$\tilde{x}_{1,2} = \tilde{x}_{1,2}(x,h) := 1 - x_{1,2}(1-x,h)$$
$$= x - x\sigma^2h \pm \sqrt{(1 - x + x\sigma^2h)x\sigma^2h} \tag{19}$$

with probabilities $\tilde{p}_{1,2} = \frac{1-x}{2x_{1,2}(1-x,h)}$. Thus, we obtain the acceptable (i.e., with values in [0, 1]) approximation $\hat{S}_h^x$ taking the values

$$\hat{x}_{1,2} := \begin{cases} x_{1,2}(x,h) \text{ with probabilities } p_{1,2} = \frac{x}{2x_{1,2}(x,h)}, & x \in [0,1/2], \\ 1 - x_{1,2}(1-x,h) \text{ with probabilities } p_{1,2} = \frac{1-x}{2x_{1,2}(1-x,h)}, & x \in (1/2,1]. \end{cases} \tag{20}$$

Lemma 1. *The values $\hat{x}_{1,2}$ defined by (20) satisfy conditions (9)–(13), and $\hat{x}_{1,2} \in [0,1]$.*

Proof. We first check that $x_{1,2}$ defined by (18) obtain values from the interval [0, 1] when $x \in [0,1/2]$ and h is sufficiently small ($0 < h \leq h_0$ with $h_0 > 0$ independent from x):

$$\begin{aligned}
x_1 &= x + (1-x)\sigma^2 h - \sqrt{(x + (1-x)\sigma^2 h)(1-x)\sigma^2 h} \geq 0 \\
&\Leftrightarrow x + (1-x)\sigma^2 h \geq \sqrt{(x + (1-x)\sigma^2 h)(1-x)\sigma^2 h} \\
&\Leftrightarrow x + (1-x)\sigma^2 h \geq (1-x)\sigma^2 h \\
&\Leftrightarrow x \geq 0; \\
x_2 &= x + (1-x)\sigma^2 h + \sqrt{(x + (1-x)\sigma^2 h)(1-x)\sigma^2 h} \leq 1 \\
&\Leftrightarrow \sqrt{(x + (1-x)\sigma^2 h)(1-x)\sigma^2 h} \leq (1-x)(1-\sigma^2 h) \\
&\Leftrightarrow x\sigma^2 h + (1-x)(\sigma^2 h)^2 \leq (1-x)(1-\sigma^2 h)^2 \\
&\Leftrightarrow x\sigma^2 h + (1-x)(\sigma^2 h)^2 \leq (1-x)(1-2\sigma^2 h + (\sigma^2 h)^2) \\
&\Leftrightarrow x\sigma^2 h \leq (1-x)(1-2\sigma^2 h) \\
&\Leftrightarrow x\sigma^2 h + 1 - x - 2\sigma^2 h \geq 0.
\end{aligned}$$

If $x \in [0,1/2]$, then

$$x\sigma^2 h + 1 - x - 2\sigma^2 h \geq 1/2 - 2\sigma^2 h \geq 0, \text{ where } 0 < h \leq h_0 := \frac{1}{4\sigma^2}. \tag{21}$$

Thus $0 \leq x_1 < x_2 \leq 1$ for $x \in [0,1/2]$ and $0 < h \leq h_0 = 1/4\sigma^2$. So, if $x \in (1/2,1]$, then $1 - x \in [0,1/2)$, and according to (19), instead of $x_{1,2}$, we can take $\tilde{x}_{1,2} = 1 - x_{1,2}(1-x,h)$ for $0 < h \leq h_0$. Thus, as we have just checked, we have $0 \leq x_{1,2}(1-x,h) \leq 1$, that is, $0 \leq \tilde{x}_{1,2} \leq 1$ for $x \in (1/2,1]$ and $0 < h \leq h_0$.

Now we check conditions (9)–(13) for $x_{1,2}$:

$$\begin{aligned}
p_1 + p_2 &= \frac{x}{2x_1} + \frac{x}{2x_2} \\
&= \frac{2x(x+(1-x)\sigma^2 h)}{2((x^2+2x(1-x)\sigma^2 h+(1-x)^2(\sigma^2 h)^2-(x(1-x)\sigma^2 h+(1-x)^2(\sigma^2 h)^2)} \\
&= \frac{2x(x+(1-x)\sigma^2 h)}{2((x^2+x(1-x)\sigma^2 h)} = 1; \\
x_1 p_1 + x_2 p_2 &= x_1 \frac{x}{2x_1} + x_2 \frac{x}{2x_2} = x, \\
x_1^2 p_1 + x_2^2 p_2 &= x_1^2 \frac{x}{2x_1} + x_2^2 \frac{x}{2x_2} = \frac{x}{2}(x_1 + x_2) \\
&= \frac{x}{2} \cdot 2(x + (1-x)\sigma^2 h) = x^2 + x(1-x)\sigma^2 h \\
&= m_2 + O(h^2); \\
(x_1 - x)^3 p_1 + (x_2 - x)^3 p_2 &= 2x(1-x)^2(\sigma^2 h)^2 = O(h^2), \\
(x_1 - x)^4 p_1 + (x_2 - x)^4 p_2 &= x(1-x)^2(x + 4(1-x)\sigma^2 h)(\sigma^2 h)^2 = O(h^2).
\end{aligned}$$

The last two equalities were obtained by using the Python SymPy package. The conditions for $\tilde{x}_{1,2}$ follow automatically from the symmetry. □

For the initial Equation (1) we obtain an approximation $\hat{X}_h^x$ by the "split-step" procedure defined by (6):

$$\hat{X}_h^x := \hat{S}_h^x e^{-bh} + \tfrac{a}{b}(1 - e^{-bh}). \tag{22}$$

Now we can state our first main result.

Theorem 4. *Let $\hat{X}_t^x$ be the discretization scheme defined by one-step approximation* (22)*. Then, $\hat{X}_t^x$ is a first-order weak approximation of Equation* (1) *for functions $f \in C^\infty[0,1]$.*

3.2. Algorithm

In this section, we provide an algorithm for calculating $\hat{X}_{(i+1)h}$ given $\hat{X}_{ih} = x$ at each simulation step i:

1. Draw a uniform random variable U from the interval [0, 1].
2. If $x \le \frac{1}{2}$, then
   - calculate x_1, x_2 according to (18),

   else
   - calculate x_1, x_2 according to (18) with $x := 1 - x$,
   - $x_{1,2} := 1 - x_{1,2}$.
3. Calculate $p_{1,2} := \frac{x}{2x_{1,2}(x,h)}$.
4. If $U < p_1$, then $\hat{S} := x_1$ else $\hat{S} := x_2$.
5. Calculate (see (6) and (22))

$$\hat{X}_{(i+1)h} = D(\hat{S}, h) = \hat{S}e^{-bh} + \tfrac{a}{b}(1 - e^{-bh}).$$

3.3. Simulation Examples

We illustrate our approximation for the test functions $f(x) = x^5$ and $f(x) = \mathrm{e}^{-x}$. Since we do not explicitly know the moments $\mathbb{E}\mathrm{e}^{-X_t^x}$, we use the approximate equality $\mathrm{e}^{-x} \approx 1 - x + \frac{x^2}{2} - \frac{x^3}{6} + \frac{x^4}{24} - \frac{x^5}{120}$. We have chosen the parameters of the WF equation so that the fifth moment of X_t^x is nonmonotonic as a function of t to see how the approximated fifth moment "follows" the bends of the true one as t varies. In Figures 1–3, we compare $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t (left plots) and as functions of a discretization step h (right plots) in terms of the relative error $\left|1 - \frac{\mathbb{E}f(\hat{X}_t^x)}{\mathbb{E}f(X_t^x)}\right|$. As expected, the approximations agree with exact values pretty well. Note an impressive match between the approximated and true values of $\mathbb{E}\mathrm{e}^{-X_t^x}$ in Figure 3 even for rather large discretization step h.

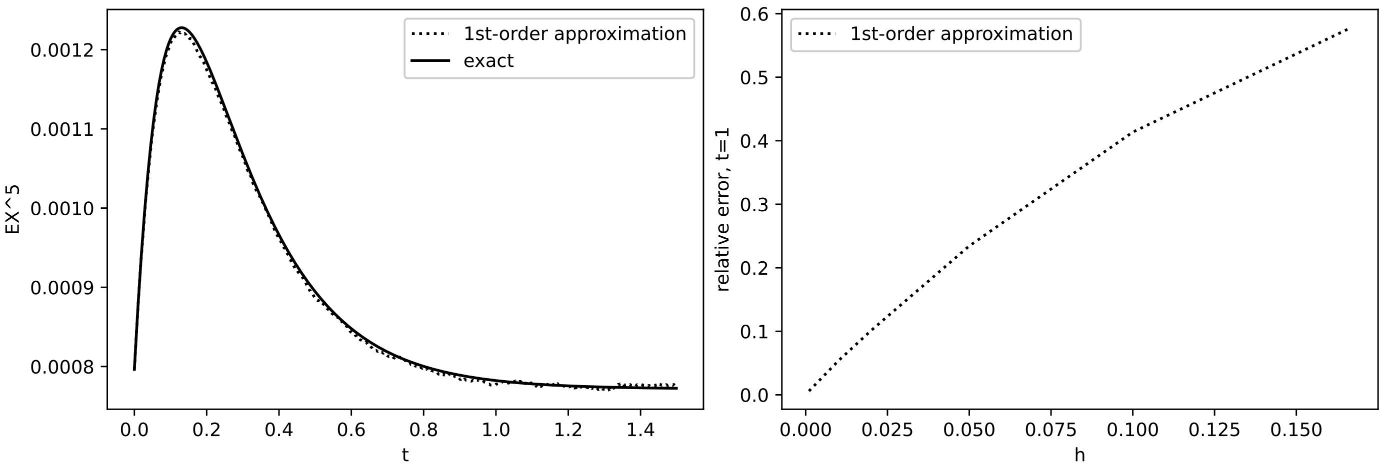


Figure 1. Comparison of $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t and h for $f(x) = x^5$: $x = 0.24$, $\sigma^2 = 0.6$, $a = 0.8$, $b = 5$, the number of iterations $N = 1{,}000{,}000$. **Left**: $h = 0.001$; **Right**: the relative error at $t = 1$.

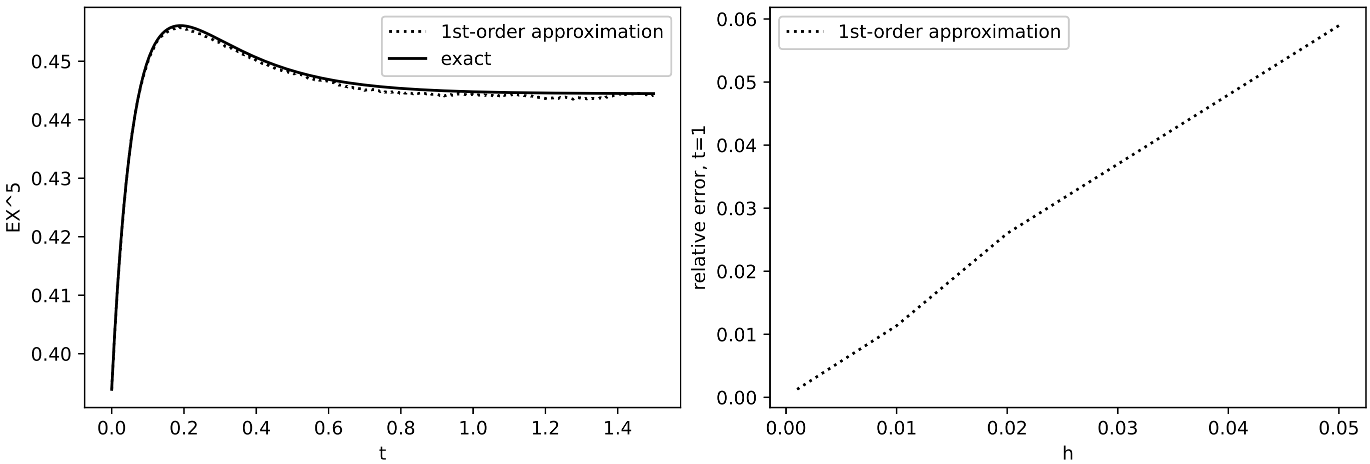


Figure 2. Comparison of $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t and h for $f(x) = x^5$: $x = 0.83$, $\sigma^2 = 2,\ a = 4,\ b = 5,\ N = 1{,}000{,}000$. **Left**: $h = 0.001$; **Right**: the relative error at $t = 1$.

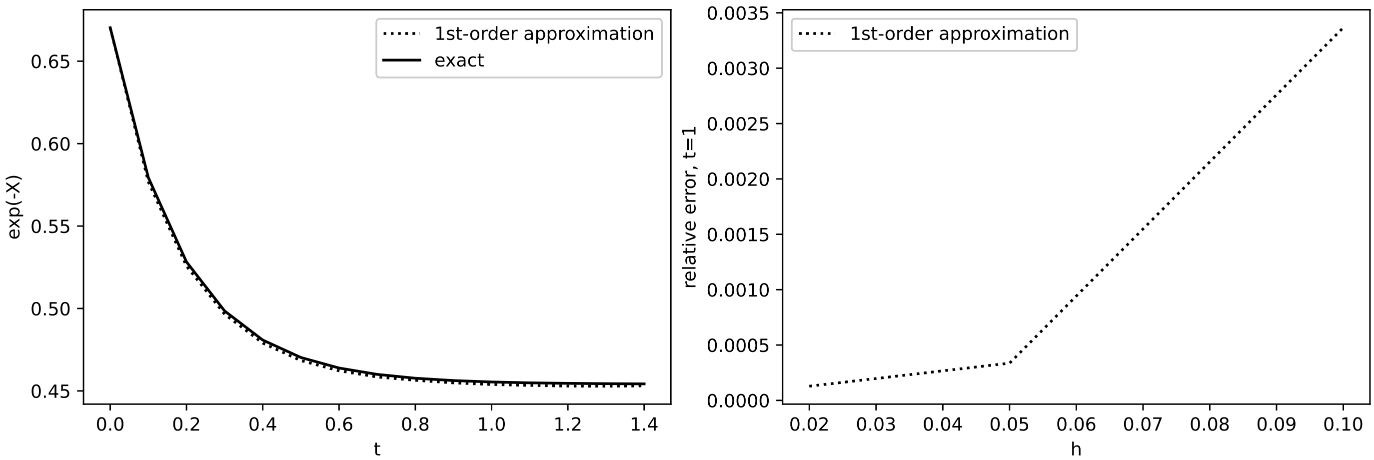


Figure 3. Comparison of $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t and h for $f(x) = \mathrm{e}^{-x}$: $x = 0.4$, $\sigma^2 = 1.6,\ a = 4,\ b = 5,\ N = 100{,}000$. **Left**: $h = 0.1$; **Right**: the relative error at $t = 1$.

4. Second-Order Weak Approximation of Wright–Fisher Equation

4.1. Approximation of the Stochastic Part

Let $\hat{S}_h^x$ be any discretization scheme. Applying Taylor's formula to $f \in C^\infty[0,1]$, we have

$$\begin{aligned}\mathbb{E}f(\hat{S}_h^x) &= f(x) + f'(x)\mathbb{E}\big(\hat{S}_h^x - x\big) + \frac{f''(x)}{2}\mathbb{E}\big(\hat{S}_h^x - x\big)^2 + \frac{f'''(x)}{6}\mathbb{E}\big(\hat{S}_h^x - x\big)^3 \\ &+ \frac{f^{(4)}(x)}{4!}\mathbb{E}\big(\hat{S}_h^x - x\big)^4 + \frac{f^{(5)}(x)}{5!}\mathbb{E}\big(\hat{S}_h^x - x\big)^5 \\ &+ \frac{1}{5!}\mathbb{E}\int_x^{\hat{S}_h^x} f^{(6)}(s)\big(\hat{S}_h^x - s\big)^5\,\mathrm{d}s.\end{aligned}$$

The generator A_0 and its square of the stochastic part are

$$\begin{aligned}A_0 f(x) &= \frac{1}{2}\sigma^2 x(1-x)f''(x), \\ A_0^2 f(x) &= -\frac{1}{2}\sigma^4 x(1-x)f''(x) + \frac{1}{2}\sigma^4 x(1-x)(1-2x)f'''(x) \\ &+ \frac{1}{4}\sigma^4 x^2(1-x)^2 f^{(4)}(x).\end{aligned}$$

Thus, the second-order remainder of the discretization scheme $\hat{S}_h^x$ is

$$\begin{aligned}
R_2^h f(x) &= \mathbb{E} f(\hat{S}_h^x) - \left[f(x) + A_0 f(x) h + A_0^2 f(x) \frac{h^2}{2} \right] \\
&= f'(x) \mathbb{E}(\hat{S}_h^x - x) \\
&\quad + \frac{f''(x)}{2} \left[\mathbb{E}(\hat{S}_h^x - x)^2 - \sigma^2 x(1-x) h \left(1 - \tfrac{1}{2}\sigma^2 h\right) \right] \\
&\quad + \frac{f'''(x)}{6} \left[\mathbb{E}(\hat{S}_h^x - x)^3 - \frac{3}{2}\sigma^4 h^2 x(1-x)(1-2x) \right] \\
&\quad + \frac{f^{(4)}(x)}{4!} \left[\mathbb{E}(\hat{S}_h^x - x)^4 - 3\sigma^4 (x(1-x)h)^2 \right] \\
&\quad + \frac{f^{(5)}(x)}{5!} \mathbb{E}(\hat{S}_h^x - x)^5 + r_2(x,h), \ x \ge 0, \ h > 0,
\end{aligned}$$

where

$$|r_2(x,h)| = \frac{1}{5!} \left| \mathbb{E} \int_x^{\hat{S}_h^x} f^{(6)}(s) (\hat{S}_h^x - s)^5 \, \mathrm{d}s \right| \le \frac{1}{6!} \max_{s \in [0,1]} |f^{(6)}(s)| \mathbb{E}(\hat{S}_h^x - x)^6.$$

This expression shows that $\hat{S}_h^x$ is a potential second-order approximation of the stochastic part (4) if

$$\mathbb{E}(\hat{S}_h^x - x) = O(h^3), \tag{23}$$
$$\mathbb{E}(\hat{S}_h^x - x)^2 = \sigma^2 x(1-x) h \left(1 - \frac{1}{2}\sigma^2 h\right) + O(h^3), \tag{24}$$
$$\mathbb{E}(\hat{S}_h^x - x)^3 = \frac{3}{2}\sigma^4 h^2 x(1-x)(1-2x) + O(h^3), \tag{25}$$
$$\mathbb{E}(\hat{S}_h^x - x)^4 = 3\sigma^4 (x(1-x)h)^2 + O(h^3), \tag{26}$$
$$\mathbb{E}(\hat{S}_h^x - x)^5 = O(h^3), \tag{27}$$
$$\mathbb{E}(\hat{S}_h^x - x)^6 = O(h^3). \tag{28}$$

Let us denote $z = \sigma^2 h$ for brevity. Converting the central moments of $\hat{S}_h^x$ to noncentral moments, from (23)–(28) we get

$$\mathbb{E}(\hat{S}_h^x)^i = \hat{m}_i + O(h^3), \quad i = 1, \dots, 6, \tag{29}$$

where

$$\begin{aligned}
\hat{m}_1 &= x, \\
\hat{m}_2 &= x^2 + zx(1-x)(1 - \tfrac{1}{2}z), \\
\hat{m}_3 &= x^3 + \tfrac{3}{2} xz^2 (3x^2 - 4x + 1) - 3xz(x^2 - x), \\
\hat{m}_4 &= x^4 + 9x^2 z^2 (2x^2 - 3x + 1) - 6x^2 z(x^2 - x), \\
\hat{m}_5 &= x^5 + 10x^3 z^2 (5x^2 - 8x + 3) - 10x^3 z(x^2 - x), \\
\hat{m}_6 &= x^6 + \tfrac{75}{2} x^4 z^2 (3x^2 - 5x + 2) - 15x^4 z(x^2 - x).
\end{aligned} \tag{30}$$

Our aim is to construct a potential second-order approximation for the WF equation by discrete random variables at each generation step. Therefore, we look for approximations $\hat{S}_h^x$ taking three values x_1, x_2, x_3 from the interval $[0,1]$ with probabilities p_1, p_2, p_3 satisfying the following conditions:

$$p_1 + p_2 + p_3 = 1, \tag{31}$$

$$x_1^i p_1 + x_2^i p_2 + x_3^i p_3 = \hat{m}_i + O(h^3), \quad i = 1, \dots, 6. \tag{32}$$

In anticipation, note that when solving system (31)–(32), a serious challenge is ensuring the first equality $p_1 + p_2 + p_3 = 1$. A simple way out of this situation is relaxing the latter by the inequality

$$p_1 + p_2 + p_3 \le 1 \tag{33}$$

and, at the same time, allowing $\hat{S}_h^x$ to take the additional trivial value 0 with probability $p_0 = 1 - (p_1 + p_2 + p_3)$. Notice that this does not change Equation (32) in any way.

Solving the system

$$x_1^i p_1 + x_2^i p_2 + x_3^i p_3 = \hat{m}_i, \; i = 1, 2, 3,$$

with respect to x_1, x_2, x_3, we obtain (cf. [6])

$$\begin{aligned} p_1 &= \frac{\hat{m}_1 x_2 x_3 - \hat{m}_2 x_2 - \hat{m}_2 x_3 + \hat{m}_3}{x_1(x_1 - x_3)(x_1 - x_2)}, \\ p_2 &= -\frac{\hat{m}_1 x_1 x_3 - \hat{m}_2 x_1 - \hat{m}_2 x_3 + \hat{m}_3}{x_2(x_1 - x_2)(x_2 - x_3)}, \\ p_3 &= \frac{\hat{m}_1 x_1 x_2 - \hat{m}_2 x_1 - \hat{m}_2 x_2 + \hat{m}_3}{x_3(x_2 - x_3)(x_1 - x_3)}. \end{aligned} \tag{34}$$

Note that, here, differently from [6], we used approximate "moments" $\hat{m}_i$ instead of the true moments m_i. This eventually allows us to get simpler expressions because $\hat{m}_i$ are polynomials in x and z.

Now we have to find $x_{1,2,3}$ that, together with $p_{1,2,3}$ defined by Equations (34), satisfy the remaining conditions

$$\begin{cases} x_1^4 p_1 + x_2^4 p_2 + x_3^4 p_3 - \hat{m}_4 = O(h^3), \\ x_1^5 p_1 + x_2^5 p_2 + x_3^5 p_3 - \hat{m}_5 = O(h^3), \\ x_1^6 p_1 + x_2^6 p_2 + x_3^6 p_3 - \hat{m}_6 = O(h^3). \end{cases} \tag{35}$$

Motivated by the first-order approximation (20) and [6], we look for $x_{1,2,3}$ of the following form:

$$x_1 = x + zA_1(1-x) + A_2 xz - \sqrt{(z(1-x)(Bx + Cz(1-x)))}, \tag{36}$$

$$x_2 = x + A_3 xz, \tag{37}$$

$$x_3 = x + zA_1(1-x) + A_2 xz + \sqrt{(z(1-x)(Bx + Cz(1-x)))}, \tag{38}$$

with unknown parameters $A_1, A_2, A_3, B, C \ge 0$.

4.2. Calculation of the Parameters

Substituting (36)–(38) into the left-hand sides of (35), we have (for technical calculations, using Maple and Python)

$$\begin{aligned} x_1^4 p_1 + x_2^4 p_2 + x_3^4 p_3 - \hat{m}_4 = \Big[&(BA_3 + B + 2A_1 - 2A_2 - A_3 - 6)x^4 \\ &+ \big(A_3 - 2B - A_3 B - 4A_1 + 2A_2 + \tfrac{21}{2}\big)x^3 \\ &+ \big(B + 2A_1 - \tfrac{9}{2}\big)x^2\Big] z^2 + O(h^3), \end{aligned} \tag{39}$$

$$\begin{aligned} x_1^5 p_1 + x_2^5 p_2 + x_3^5 p_3 - \hat{m}_5 = \Big[&(8A_1 + (4B - 4)A_3 + 5B - 8A_2 - 27)x^5 \\ &+ ((-4B + 4)A_3 - 10B + 8A_2 - 16A_1 + 48)x^4 \\ &+ (8A_1 + 5B - 21)x^3\Big] z^2 + O(h^3), \end{aligned} \tag{40}$$

$$x_1^6 p_1 + x_2^6 p_2 + x_3^6 p_3 - \hat{m}_6 = \Big[(20A_1 + (10B - 10)A_3 + 15B - 20A_2 - 75)x^6 + \big((-10B + 10)A_3 - 30B + 20A_2 - 40A_1 + 135\big)x^5 + (15B + 20A_1 - 60)x^4\Big]z^2 + O(h^3). \quad (41)$$

To ensure equalities (35), we need to choose A_1, A_2, A_3, B such that expressions at z^2 would be equal to 0. Equating the coefficients at the lowest powers of x to zero, we get the system for the parameters A_1 and B:

$$\begin{cases} B + 2A_1 - \frac{9}{2} &= 0, \\ 8A_1 + 5B - 21 &= 0, \\ 15B + 20A_1 - 60 &= 0. \end{cases}$$

Although the system contains three equations with respect to two unknowns, it has the solution $A_1 = \frac{3}{4}, B = 3$. Substituting these values back to Equations (39)–(41), we get the relation $A_3 = A_2 + \frac{3}{4}$, which makes all the expressions at z^2 vanish. Summarizing, we have that $x_{1,2,3}$ of the form (36)–(38) and $p_{1,2,3}$ defined by (34) satisfy all of Equation (32), provided that the parameters satisfy the following relations:

$$A_1 = \frac{3}{4},\ A_2 \geq 0,\ A_3 = A_2 + \frac{3}{4},\ B = 3,\ C \geq 0. \quad (42)$$

4.3. Positivity of the Solution

Now we would like to choose the values of free parameters A_2 and C so that all $x_1, x_2, x_3, p_1, p_2, p_3$ are positive and $p_1 + p_2 + p_3 \leq 1$. We first consider the latter restriction.

Lemma 2. *We have* $p_1 + p_2 + p_3 \leq 1$ *if*

$$A_2 \geq \frac{(3 + 2\sqrt{2})^{\frac{1}{3}}}{4} + \frac{1}{4(3 + 2\sqrt{2})^{\frac{1}{3}}} \approx 0.58883 \quad (43)$$

and

$$0 \leq C \leq \frac{3(32A_2^3 - 6A_2 - 3)}{16A_2^2(16A_2^2 + 24A_2 + 9)}.$$

Proof. We have

$$p_1 + p_2 + p_3 = \frac{N}{D}$$

with numerator

$$N = 64x^2 + \big((192A_2 + 144)x^2 - 96x\big)z + \big((192A_2^2 - 64C + 96A_2 + 108)x^2 + (128C - 144)x - 64C + 36\big)z^2 + \big(48 + (-96A_2 + 24)x^2 + (96A_2 - 72)x\big)z^3$$

and denominator

$$D = 64x^2 + \big((192A_2 + 144)x^2 - 96x\big)z + \big((192A_2^2 - 64C + 96A_2 + 108)x^2 + (128C - 144)x - 64C + 36\big)z^2 + \big((64A_2^3 - 64CA_2 - 48A_2^2 - 48C - 36A_2 + 27)x^2 + (128CA_2 + 96A_2^2 + 96C - 54)x - 64A_2C - 48C + 36A_2 + 27\big)z^3.$$

The numerator and denominator differ only by the coefficients at z^3. Thus, it suffices to show that their difference $D - N$ is nonnegative, that is,

$$\begin{aligned}(64A_2^3 - 48A_2^2 + (-64C + 60)A_2 - 48C + 3)x^2 \\ + (96A_2^2 + (128C - 96)A_2 + 96C + 18)x \\ + (-64C + 36)A_2 - 48C - 21 =: a_1x^2 + a_2x + a_3 \geq 0. \end{aligned} \tag{44}$$

Inequality (44) is satisfied for all $x \in \mathbb{R}$ if

$$a_1 > 0 \text{ and } a_3 \geq \frac{a_2^2}{4a_1}. \tag{45}$$

Solving inequality (45), we get

$$A_2 > 0, \quad C \leq \frac{3(32A_2^3 - 6A_2 - 3)}{16(16A_2^2 + 24A_2 + 9)A_2^2}.$$

Since C must be nonnegative, we obtain condition (43) for A_2. □

Remark 2. *We observe that possible values of C are rather small (see Figure 4). Therefore, to simplify the expressions for x_1, x_2, x_3, we simply take $C = 0$ and $A_2 = \frac{2}{3}$.*

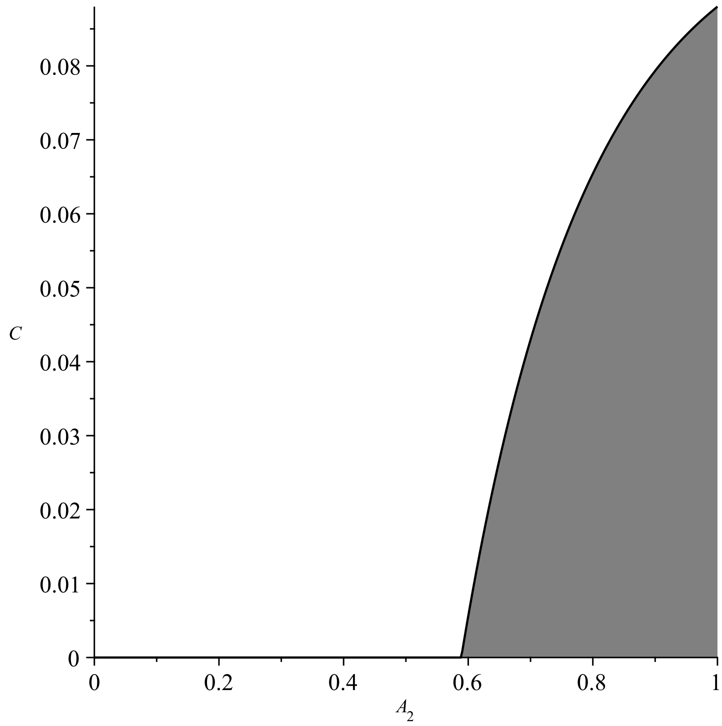


Figure 4. Possible values of A_2 and C.

We have now arrived at the following expressions for x_1, x_2, x_3:

$$x_1 = x + \frac{3(1-x)z}{4} + \frac{2xz}{3} - \sqrt{3x(1-x)z}, \tag{46}$$

$$x_2 = x + \frac{17xz}{12}, \tag{47}$$

$$x_3 = x + \frac{3(1-x)z}{4} + \frac{2xz}{3} + \sqrt{3x(1-x)z}. \tag{48}$$

However, the game is not over. At this point, we only have that x_1, x_2, x_3 defined by (46)–(48), together with p_1, p_2, p_3 defined by (34), satisfy conditions (32) and (33). From numerical calculations it appears that for "small" x, it happens that $x_1 > x_2$ and thus $p_1, p_2 < 0$. Moreover, on the other hand, for "not small" x and "large" h, it happens that $x_3 > 1$. We can see a typical situation in Figure 5 with $z = \sigma^2 h = \frac{1}{5}$, where for small x, p_1 and p_2 take values outside the interval $[0, 1]$, whereas $x_3 > 1$ for x near $\frac{1}{2}$.

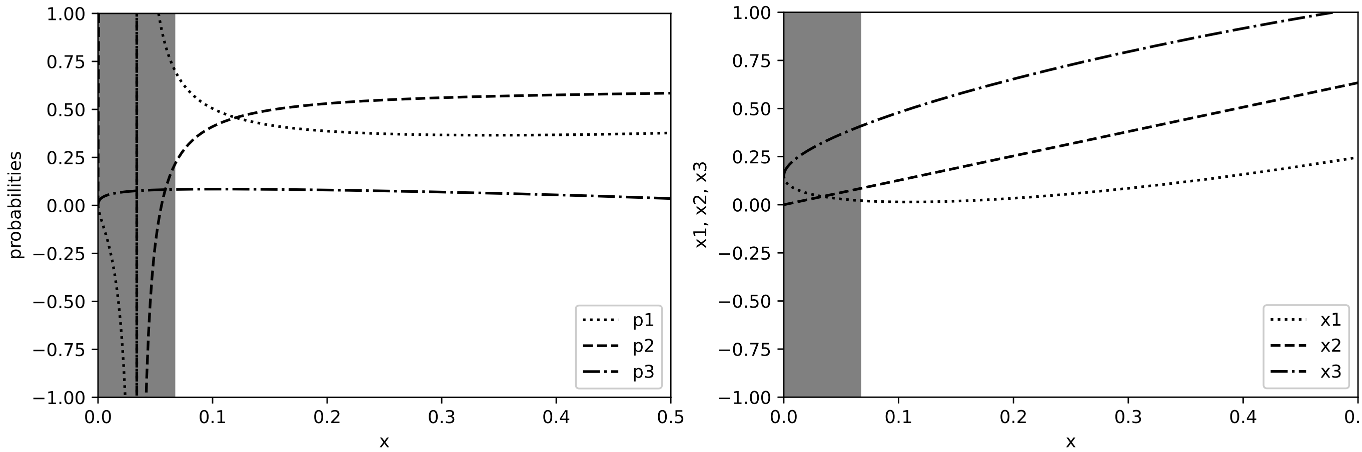


Figure 5. Graphs of $p_{1,2,3}$ (**Left**) and $x_{1,2,3}$ (**Right**) as functions of x with fixed z. Gray area shows the region where first-order approximation is used to avoid negative probabilities. Parameters: $K = \frac{1}{3}, z = \frac{1}{5}$.

Due to these reasons, similarly to [4], for small x below the threshold Kz (with some fixed $K > 0$), we will switch to the first-order approximation (20), which behaves as a second-order one for such x. We also have to consider $z \le z_0$, where z_0 is to be sufficiently small to ensure that $x_3 \le 1$. To be precise, for $0 \le x \le Kz$, $0 < z \le z_0$, we will use scheme (20), whereas for $Kz \le x \le \frac{1}{2}$, $0 < z \le z_0$, we will use scheme (46)–(48) together with (34); finally, for $x \in (\frac{1}{2}, 1]$, we will use the symmetry $S_t^x \overset{d}{=} 1 - S_t^{1-x}$ as in the first-order approximation. The following lemmas justify such a switch for $K = \frac{1}{3}$ and $z_0 = \frac{1}{6}$.

Lemma 3. *The first-order approximation* (20) *in the region* $x \le K\sigma^2 h$ (*with arbitrary fixed* $K > 0$) *satisfies conditions* (29). *In other words, in this region, it behaves as a second-order approximation.*

Proof. We prove equalities (29) in the region $x \le Kz = K\sigma^2 h$, where $\hat{S}_x^h$ and $\hat{m}_i$, $i = 1, \dots, 6$, are defined by (20) and (30), respectively:

$$\begin{aligned}
\mathbb{E}(\hat{S}_h^x)^2 - \hat{m}_2 &= x_1^2 p_1 + x_2^2 p_2 - \hat{m}_2 \\
&= x^2 + x(1-x)z - (x^2 + zx(1-x)(1 - \tfrac{1}{2}z)) \\
&= \tfrac{1}{2}x(1-x)z^2 = O(h^3), \\
\mathbb{E}(\hat{S}_h^x)^3 - \hat{m}_3 &= x_1^3 p_1 + x_2^3 p_2 - \hat{m}_3 \\
&= x^3 - 3xz(x^2 - x) + 2z^2x^2(x-1) \\
&\quad - (x^3 + \tfrac{3}{2}xz^2(3x^2 - 4x + 1) - 3xz(x^2 - x)) \\
&= \tfrac{1}{2}xz^2(5x^2 - 8x + 3) = O(h^3), \\
\mathbb{E}(\hat{S}_h^x)^4 - \hat{m}_4 &= x_1^4 p_1 + x_2^4 p_2 - \hat{m}_4 = x^4 - 6x^2 z(x^2 - x) \\
&\quad + x^2z^2(9x^2 - 10x + 1) - 4x^3z^3(x-1) \\
&\quad - (x^4 + 9x^2z^2(2x^2 - 3x + 1) - 6x^2z(x^2 - x)) \\
&= x^2z^2(-4x^2z - 9x^2 + 4xz + 17x - 8) = O(h^4), \\
\mathbb{E}(\hat{S}_h^x)^5 - \hat{m}_5 &= x_1^5 p_1 + x_2^5 p_2 - \hat{m}_5 = x^5 - 10x^3z(x^2 - x)
\end{aligned}$$

$$
\begin{aligned}
&+5x^3z^2(5x^2-6x+1)+4x^3z^3(-6x^2+7x-1)\\
&+8x^4z^4(x-1)\\
&-(x^5+10x^3z^2(5x^2-8x+3)-10x^3z(x^2-x))\\
&=x^3z^2(x^2(8z^2-24z-25)-2x(4z^2-14z-25)-4z-25)=O(h^5),\\
\mathbb{E}(S_h^x)^6-\hat{m}_6 &= x_1^6p_1+x_2^6p_2-\hat{m}_6 = x^6-15x^4z(x^2-x)\\
&+5x^4z^2(11x^2-14x+3)\\
&+x^3z^3(-85x^3+111x^1-27x+1)\\
&+12x^4z^4(5x^2-6x+1)-16x^5z^5(x-1)\\
&-(x^6+\tfrac{75}{2}x^4z^2(3x^2-5x+2)-15x^4z(x^2-x))\\
&=\tfrac{1}{2}x^3z^2(x-1)(-32x^2z^3+120x^2z^2-170x^2\,z\\
&-115x^2-24xz^2-170xz+120x-2z)=O(h^6). \quad \square
\end{aligned}
$$

Lemma 4. *For $z \in [0, \frac{1}{6}]$ and $x \in [0, \frac{1}{2}]$, x_1, x_2, x_3 defined in (46)–(48) take values in the interval $[0, 1]$.*

Proof. Obviously, $x_2 \in [0, 1]$. Thus, we focus on x_1 and x_3. Since $x_1 \le x_3$, it suffices to prove that $x_1 \ge 0$ and $x_3 \le 1$.

The condition $x_1 \ge 0$ is equivalent to the inequality

$$\left(x+\frac{3(1-x)z}{4}+\frac{2xz}{3}\right)^2-3x(1-x)z \ge 0.$$

By denoting $y = 1 - x > 0$, this becomes

$$\left(x+\frac{3yz}{4}+\frac{2xz}{3}\right)^2-3xyz = x^2+\frac{9}{16}y^2z^2+\frac{4}{9}x^2z^2-\frac{3}{2}xyz+\frac{4}{3}x^2z+xyz^2 \ge 0.$$

We will prove the stronger inequality

$$x^2+\frac{9}{16}y^2z^2+\frac{4}{3}x^2z-\frac{3}{2}xyz \ge 0,$$

which after substitution $y = 1 - x$ becomes

$$\left(1+\frac{17}{6}z+\frac{9}{16}z^2\right)x^2-\left(\frac{3}{2}z+\frac{9}{8}z^2\right)x+\frac{9z^2}{16} \ge 0. \tag{49}$$

The discriminant of the quadratic polynomial (49) in x is

$$D=\left(\frac{3}{2}z+\frac{9}{8}z^2\right)^2-4\left(1+\frac{17}{6}z+\frac{9}{16}z^2\right)\cdot\frac{9z^2}{16}=-3z^3,$$

which is negative for all $z > 0$. This means that the left-hand side (49) is positive and thus $x_1 > 0$ for all $x \in [0, 1]$ and $z \ge 0$ except for $x = z = 0$, where $x_1 = 0$.

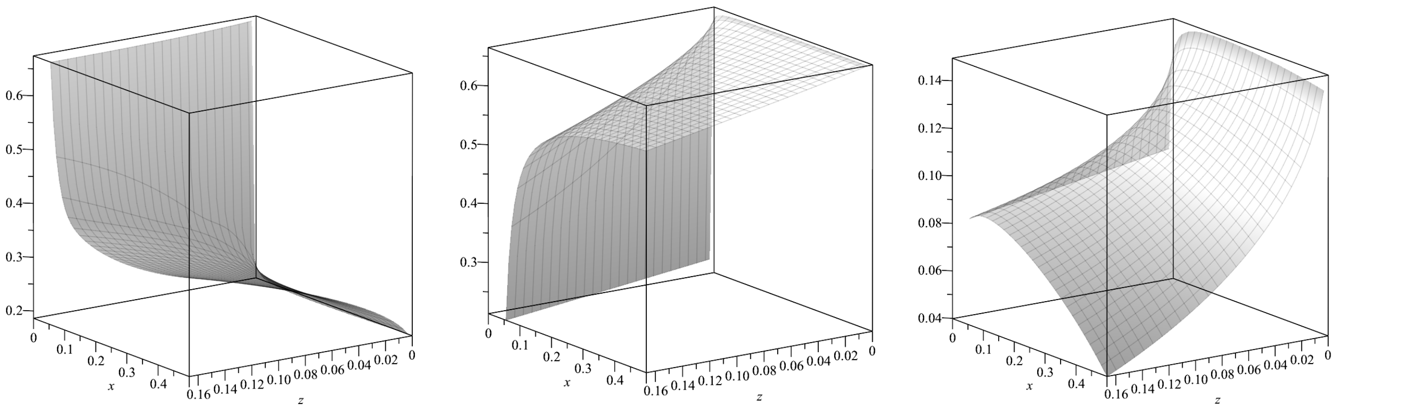
Let us now prove that $x_3 \le 1$. For $z \in [0, \frac{1}{6}]$ and $x \in [0, \frac{1}{2}]$, we have

$$
\begin{aligned}
x_3 &= x+\frac{3(1-x)z}{4}+\frac{2xz}{3}+\sqrt{3x(1-x)z}\\
&\le x+\frac{1}{8}(1-x)+\frac{1}{18}+\sqrt{\tfrac{x(1-x)}{2}}
\end{aligned}
$$

$$\leq \frac{1}{8} + \frac{7}{8} \cdot \frac{1}{2} + \frac{1}{18} + \frac{1}{\sqrt{8}} \approx 0.972 < 1. \quad \square$$

Lemma 5. *For $x \in (\frac{z}{3}, \frac{1}{2}]$ and $z \leq \frac{1}{6}$, we have $p_1, p_2, p_3 \in [0,1]$.*

Proof. From Lemma 2, we already have that $p_1 + p_2 + p_3 \leq 1$. Therefore, it suffices to prove that $p_1, p_2, p_3 \geq 0$. Because of the complex expressions of p_1, p_2, p_3, we prefer to show this graphically by using the Maple function plot3d. See Figure 6, where the 3D graphs of p_1, p_2, p_3 as functions of (x, z) are plotted in the domain $\{(x,z) : z/3 \leq x \leq 1/2,\ 0 \leq z \leq 1/6\}$. $\square$



(**a**) (**b**) (**c**)

Figure 6. Graphs of p_1, p_2, p_3 as functions of x and z. (**a**) p_1, (**b**) p_2, (**c**) p_3.

4.4. The Second Main Result

Now let us summarize the results of this section. For clarity, recall the main notations:

$$x_1 = x_1(x,h) = x + \frac{3(1-x)\sigma^2 h}{4} + \frac{2x\sigma^2 h}{3} - \sqrt{3x(1-x)\sigma^2 h}, \tag{50}$$

$$x_2 = x_2(x,h) = x + \frac{17x\sigma^2 h}{12}, \tag{51}$$

$$x_3 = x_3(x,h) = x + \frac{3(1-x)\sigma^2 h}{4} + \frac{2x\sigma^2 h}{3} + \sqrt{3x(1-x)\sigma^2 h}. \tag{52}$$

To distinguish the functions $x_{1,2,3}$ from $x_{1,2}$ given by (18), here we denote the latter by

$$y_{1,2} = y_{1,2}(x,h) = x + (1-x)\sigma^2 h \mp \sqrt{(x + (1-x)\sigma^2 h)(1-x)\sigma^2 h}. \tag{53}$$

Using the symmetry $S_t^x \stackrel{d}{=} 1 - S_t^{1-x}$ for $x \in [0,1]$, we define the approximation of the stochastic part of the WF equation as follows:

$$\hat{S}_h^x := \begin{cases} x_{1,2,3}(x,h) \text{ with probabilities } p_{1,2,3} \text{ (34) and} & \\ \quad 0 \text{ with probability } p_0 = 1 - (p_1 + p_2 + p_3), & x \in (\frac{\sigma^2 h}{3}, \frac{1}{2}], \\ 1 - x_{1,2,3}(1-x,h) \text{ with prob. } p_{1,2,3}(1-x,h) \text{ and} & \\ \quad 1 \text{ with probability } p_0 = 1 - (p_1 + p_2 + p_3), & x \in (\frac{1}{2}, 1 - \frac{\sigma^2 h}{3}), \\ y_{1,2}(x,h) \text{ with probabilities } \tilde{p}_{1,2}(x,h) := \frac{x}{2y_{1,2}(x,h)}, & x \in [0, \frac{\sigma^2 h}{3}], \\ 1 - y_{1,2}(1-x,h) \text{ with probabilities } \tilde{p}_{1,2}(1-x,h), & x \in [1 - \frac{\sigma^2 h}{3}, 1]. \end{cases} \tag{54}$$

Now in view of Theorem 3 and Lemmas 2–5, we can state the main result on the second-order approximation of the WF process.

Theorem 5. *Let $\hat{X}_t^x$ be the discretization scheme defined by one-step approximation*

$$\hat{X}_h^x = D(\hat{S}(D(x,h/2),h),h/2), \tag{55}$$

where $D(x,h)$ is defined by (5), and $\hat{S}(x,h) = \hat{S}_h^x$ is defined by (54). Then, $\hat{X}_t^x$ is a second-order weak approximation of Equation (1).

4.5. Algorithm for Second-Order Approximation

In this section, we provide an algorithm for calculating $\hat{X}_{(i+1)h}$ given $\hat{X}_{ih} = x$ at each simulation step i:

1. Draw a uniform random variable U from the interval [0, 1].
2. $x := D(x,h/2)$ (where D is given by (5))
3. If $x \le \frac{1}{2}$, then
   - 3.1. if $x > \frac{\sigma^2 h}{3}$, then
     $x_0 := 0$,
     calculate x_1, x_2, x_3 according to (50)–(52),
     calculate p_1, p_2, p_3 according to (34),
     if $U < p_1$ then $\hat{S} := x_1$ else if $U < p_1 + p_2$ then $\hat{S} := x_2$
     else if $U < p_1 + p_2 + p_3$ then $\hat{S} := x_3$ else $\hat{S} := x_0$

     else
     calculate y_1, y_2 according to (53),
     $p_{1,2} := \frac{x}{y_{1,2}(x,h)}$,
     if $U < p_1$ then $\hat{S} := y_1$ else $\hat{S} := y_2$

   else
   - 3.2. do step 3.1 with $x := 1 - x$, $x_{0,1,2,3} := 1 - x_{0,1,2,3}$,
     $y_{1,2} := 1 - y_{1,2}$.
4. $\hat{X}_{(i+1)h} := D(\hat{S}, h/2)$.

4.6. Simulation Examples

We illustrate our approximation for the test function $f(x) = x^5$. In Figures 7 and 8, we compare the moments $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t (left plots, $h = 0.01$) and as functions of a discretization step h (right plots) in terms of the relative error. We observe that with a rather small number of iterations, the second-order approximation agrees with the exact values pretty well. These specific examples have been chosen to illustrate the behavior of approximations with small ($\sigma^2 = 0.6$) and high ($\sigma^2 = 2$) volatility. In comparison with the simulation results for the first-order approximation (Section 3.3), we see that to get a similar accuracy, we can use the second-order approximation with a significantly smaller number of iterations N and larger step size h, which in turn requires significantly less computation time.

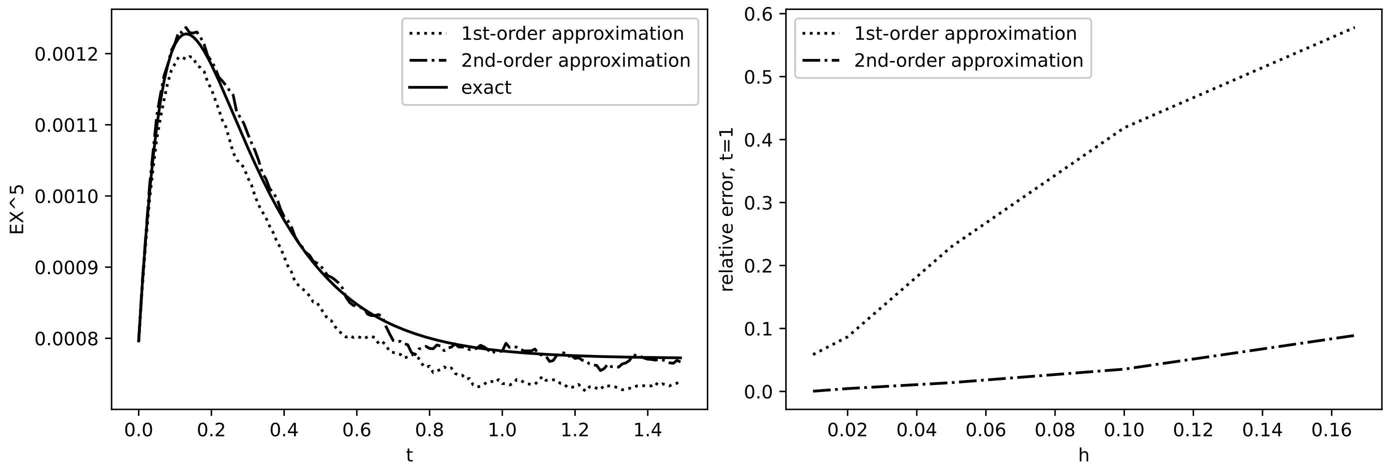


Figure 7. Comparison of $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t and h for $f(x) = x^5$: $x = 0.24$, $\sigma^2 = 0.6$, $a = 0.8$, $b = 5$, the number of iterations $N = 100{,}000$. **Left**: $h = 0.01$; **Right**: the relative error at $t = 1$.

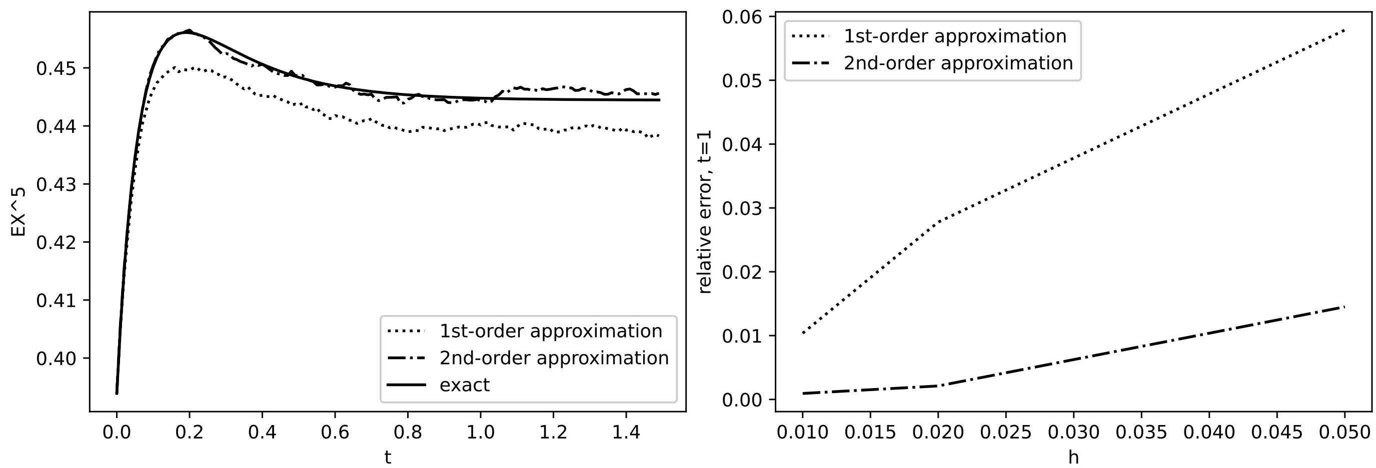


Figure 8. Comparison of $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t and h for $f(x) = x^5$: $x = 0.83$, $\sigma^2 = 2$, $a = 4$, $b = 5$, the number of iterations $N = 100{,}000$. **Left**: $h = 0.01$; **Right**: the relative error at $t = 1$.

5. Probabilistic Proof of Regularity of Solutions of the Kolmogorov Backward Equation

Theorem B is in fact Theorem 1.19 of [4] stated based on the results of [7], which are proved by methods of partial differential equation theory. Here, we provide a significantly simpler *probabilistic* proof of the theorem for a rather wide subclass of $C^\infty[0,1]$, which practically includes all functions interesting for applications, for example, polynomials or exponentials.

Definition 3. *We denote by* $C_*^\infty[0,1]$ *the class of infinitely differentiable functions on* $[0,1]$ *with "not too fast" growing derivatives:*

$$C_*^\infty[0,1] := \Big\{ f \in C^\infty[0,1] : \limsup_{k\to\infty} \frac{1}{k!} \max_{x\in[0,1]} |f^{(k)}(x)| = 0 \Big\}.$$

Every $f \in C_*^\infty[0,1]$ is the sum of its (uniformly convergent) Taylor series:

$$f(x) = \sum_{k=0}^{\infty} c_k x^k, \; x \in [0,1], \tag{56}$$

where $c_k = f^{(k)}(0)/k!$, $k \in \overline{\mathbb{N}}$. This easily follows from the Lagrange error bound for Taylor series.

Remark 3. *Clearly, every $f \in C_*^\infty[0,1]$ is a real analytic function; see* [8].

Denote $m_k(x,t) := \mathbb{E}(X_t^x)^k$, $k \in \overline{\mathbb{N}}$. Then, from (56), we formally have

$$\widetilde{u}(t,x) = \mathbb{E}f(X_t^x) = \sum_{k=0}^\infty c_k m_k(x,t),\ x \in [0,1],\ t \geq 0. \tag{57}$$

If $\widetilde{u}$ is infinitely continuously differentiable, then it satisfies Equation (2) (see, e.g., [9] (Thm. 8.1.1)). Therefore, it suffices to show that

(1) the moments $m_k(x,t)$ are infinitely continuously differentiable and
(2) all formal partial derivatives of the series in (57),

$$\sum_{k=0}^\infty c_k \partial_t^p \partial_x^q m_k(x,t), \tag{58}$$

converge uniformly for $(x,t) \in [0,1] \times [0,T]$ (for any fixed $T > 0$).

Lemma 6. *The moments of the WF process X_t^x satisfy the following recurrence relation:*

$$m_1(x,t) = \begin{cases} xe^{-bt} + \frac{a}{b}(1-e^{-bt}), & 0 \leq a \leq b \neq 0, \\ x, & a = b = 0, \end{cases} \tag{59}$$

$$m_k(x,t) = e^{-b_k t}\left(x^k + a_k \int_0^t e^{b_k s} m_{k-1}(x,s)\,ds\right),\ k \geq 2, \tag{60}$$

where $b_k = kb + k(k-1)\frac{\sigma^2}{2}$, $a_k = ka + k(k-1)\frac{\sigma^2}{2}$.

In particular, by induction on k it follows that $m_k(x,t)$ are infinitely continuously differentiable with respect to $(x,t) \in [0,1] \times \mathbb{R}_+$.

Proof. Taking the expectations of both sides of Equation (1) and then differentiating with respect to t, we get

$$\partial_t m_1(x,t) = a - bm_1(x,t),\ m_1(x,0) = x.$$

Solvingthe latter, we get (59).

When $k \geq 2$, by Itô's formula, we have

$$\begin{aligned}
(X_t^x)^k &= x^k + k\int_0^t (X_t^x)^{k-1}\,dX_s^x + \frac{1}{2}k(k-1)\int_0^t (X_t^x)^{k-2}\,d\langle X^x\rangle_s \\
&= x^k + k\int_0^t (X_t^x)^{k-1}(a - bX_s^x)\,\mathrm{d}s + k\sigma\int_0^t (X_t^x)^{k-1}\sqrt{X_s^x(1-X_s^x)}\,\mathrm{d}B_s \\
&\quad + \frac{1}{2}k(k-1)\sigma^2\int_0^t (X_t^x)^{k-2}\,X_s^x(1-X_s^x)\,\mathrm{d}s \\
&= x^k + k\int_0^t \left(a(X_t^x)^{k-1} - b(X_s^x)^k\right)\mathrm{d}s + k\sigma\int_0^t (X_t^x)^{k-1}\sqrt{X_s^x(1-X_s^x)}\,\mathrm{d}B_s \\
&\quad + \frac{1}{2}k(k-1)\sigma^2\int_0^t \left((X_t^x)^{k-1} - (X_s^x)^k\right)\mathrm{d}s.
\end{aligned}$$

By taking the expectations, we get

$$m_k(x,t) = x^k + \int_0^t \left\{[ka + k(k-1)\tfrac{\sigma^2}{2}]m_{k-1}(x,s) - [kb + k(k-1)\tfrac{\sigma^2}{2}]m_k(x,s)\right\}\mathrm{d}s$$

$$= x^k + \int_0^t \{a_k m_{k-1}(x,s) - b_k m_k(x,s)\}\,\mathrm{d}s,$$

and thus

$$\partial_t m_k(x,t) = -b_k m_k(x,t) + a_k m_{k-1}(x,t),\ m_k(x,0) = x^k.$$

Solving the latter with respect to m_k, we arrive at (60). □

Lemma 7. *All formal partial derivatives of the series* (57)*,*

$$\sum_{k=0}^{\infty} c_k \partial_t^p \partial_x^q m_k(x,t), \tag{61}$$

converge uniformly for $(x,t) \in [0,1] \times [0,T]$ *(for any fixed* $T > 0$*).*

Proof. It is obvious that $0 \le m_k(x,t) \le 1,\ x \in [0,1], k \in \overline{\mathbb{N}}$. First, consider the derivatives with respect to x. Let us prove by induction on k that

$$\partial_x m_k(x,t) \le k,\ x \in [0,1],\ k \in \mathbb{N}.$$

For $k = 1$, we have $m_1'(x,t) = e^{-bt} \le 1$. Suppose

$$\partial_x m_{k-1}(x,t) \le k-1,\ x \in [0,1].$$

Then,

$$\begin{aligned}
\partial_x m_k(x,t) &= e^{-b_k t}\Big(k x^{k-1} + a_k \int_0^t e^{b_k s} \partial_x m_{k-1}(x,s)\,\mathrm{d}s\Big) \\
&\le e^{-b_k t}\Big(k + a_k(k-1)\int_0^t e^{b_k s}\,\mathrm{d}s\Big) = e^{-b_k t}\Big(k + \frac{a_k}{b_k}(k-1)(e^{b_k t} - 1)\Big) \\
&\le e^{-b_k t} k + k(1 - e^{-b_k t}) = k,
\end{aligned}$$

where we used the fact that $0 \le a_k \le b_k$, since $0 \le a \le b$.

Similarly, by induction on k, we can prove that

$$\partial_x^l m_k(x,t) \le (k)_l = k(k-1)\dots(k-l+1),\ x \in [0,1], k \in \mathbb{N}, l \in \mathbb{N}.$$

Indeed, for $k = 1$, $\partial_x m_1(x,t) = e^{-bt} \le 1 = (1)_1$, and $\partial_x^l m_k(x,t) = 0 = (1)_l$ for $l \ge 2$. Now suppose that for some k,

$$\partial_x^l m_{k-1}(x,t) \le (k-1)_l,\ x \in [0,1],\ l \in \mathbb{N}.$$

Then,

$$\begin{aligned}
\partial_x^l m_k(x,t) &= e^{-b_k t}\Big(k(k-1)\dots(k-l+1)x^{k-l} + a_k \int_0^t e^{b_k s} \partial_x^l m_{k-1}(x,s)\,\mathrm{d}s\Big) \\
&\le e^{-b_k t}\Big(k(k-1)\dots(k-l+1) + \frac{a_k}{b_k} k(k-1)\dots(k-l+1)(e^{b_k t} - 1)\Big) \\
&\le k(k-1)\dots(k-l+1) = (k)_l.
\end{aligned}$$

Now let us differentiate the moments with respect to t. We have

$$\begin{aligned}
|\partial_t m_1(x,t)| &= \Big|\Big(e^{-bt}(x - \frac{a}{b}) + \frac{a}{b}\Big)_t'\Big| = |-be^{-bt}(x - \frac{a}{b})| \\
&= |(a - bx)e^{-bt}| \le b,\ x \in [0,1];
\end{aligned}$$

$$
\begin{aligned}
|\partial_t m_k(x,t)| &= \Big| -b_k e^{-b_k t}\Big(x^k + a_k \int_0^t e^{b_k s} m_{k-1}(x,s)\,\mathrm{d}s\Big) + e^{-b_k t} a_k e^{b_k t} m_{k-1}(x,t)\Big| \\
&\le b_k e^{-b_k t} x^k + a_k b_k e^{-b_k t} \int_0^t e^{b_k s}\,\mathrm{d}s + a_k \\
&\le b_k + a_k e^{-b_k t}(e^{b_k t} - 1) + a_k \le 3b_k; \\
|\partial_t^2 m_k(x,t)| &= \Big| b_k^2 e^{-b_k t}\Big(x^k + a_k \int_0^t e^{b_k s} m_{k-1}(x,s)\,\mathrm{d}s\Big) \\
&\qquad - a_k b_k m_{k-1}(x,t) + a_k \partial_t m_{k-1}(x,t)\Big| \\
&\le b_k^2 + b_k a_k + b_k a_k + 3 a_k b_k \le 6 b_k^2, \\
|\partial_t^3 m_k(x,t)| &\le \Big| b_k^3 e^{-b_k t}\Big(x^k + a_k \int_0^t e^{b_k s} m_{k-1}(x,s)\,ds\Big) + a_k b_k^2 m_{k-1}(x,t) \\
&\qquad + a_k b_k \partial_t m_{k-1}(x,t) + a_k \partial_t^2 m_{k-1}(x,t)\Big| \le 12 b_k^3,
\end{aligned}
$$

and by induction

$$
|\partial_t^l m_k(x,t)| \le 3 \times 2^{l-1} b_k^l.
$$

Finally, for all mixed partial derivatives, we have

$$
\begin{aligned}
|\partial_t^p \partial_x^q m_k(x,t)| &= \Big| \partial_t^p \partial_x^q e^{-b_k t}\Big(x^k + a_k \int_0^t e^{b_k s} m_{k-1}(x,s)\,ds\Big)\Big| \\
&\le \Big| \partial_t^p e^{-b_k t}\Big(k(k-1)\dots(k-q+1) \\
&\qquad + a_k (k-1)(k-2)\dots(k-q) \int_0^t e^{b_k s}\,ds\Big)\Big| \\
&\le \Big| \partial_t^p e^{-b_k t}\Big(k(k-1)\dots(k-q+1)\Big(a_k \int_0^t e^{b_k s}\,ds + 1\Big)\Big)\Big| \\
&= \Big| (-b_k)^p e^{-b_k t}\Big(k(k-1)\dots(k-q+1)\Big(a_k \int_0^t e^{b_k s}\,ds + 1\Big)\Big) \\
&\qquad + k(k-1)\dots(k-q+1) a_k\Big| \\
&= (b_k^p + 1) k(k-1)\dots(k-q+1) a_k = O(k^{2p+q+2}),\ k \to \infty.
\end{aligned}
$$

Since $c_k = o(1/k!)$, we have that

$$
\sum_{k=1}^{\infty} c_k k^{2p+q+2} < +\infty,
$$

and by the Weierstrass M-test it follows that, indeed, the function series (61) converges uniformly for all $p, q \in \overline{\mathbb{N}}$. $\square$

6. Conclusions

We have constructed first- and second-order weak split-step approximations of the Wright–Fisher (WF) process. The approximations use generation of a two- or three-valued random variable at each discretization step. The main difficulty was ensuring that the values of approximations take values in [0, 1], the domain of the WF process. Illustrative simulations show perfect accuracy of the constructed approximations.

Author Contributions: All authors have participated equally in the development of this work, in both theoretical and computational aspects. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Conflicts of Interest: The authorsdeclare no conflict of interest.

Abbreviations

The following abbreviations are used in this paper:

WF	Wright–Fisher model
CIR	Cox–Ingersoll–Ross model
PDE	Partial differential equation
B	Brownian motion
$\mathbb{N}$	The set of positive integers $\{1, 2, \ldots\}$
$\overline{\mathbb{N}}$	The set of nonnegative integers, $\mathbb{N} \cup \{0\}$
$\mathbb{R}$	The set of real numbers
$\mathbb{R}_+$	The set of positive real numbers
$C_*^\infty[0,1]$	A subclass of $C^\infty[0,1]$, see Definition 3.
$O(h^n)$	$g(x,h) = O(h^n)$ if, for some $C > 0$ and $h_0 > 0$, $\lvert g(x,h)\rvert \le Ch^n, \quad x \in [0,1], \quad 0 < h \le h_0$.
$\hat{X}^h$	A discretization scheme of the WF process
D_t^x	The solution of the deterministic part of the WF equation
S_t^x	The solution of the stochastic part of the WF equation
$\hat{S}^h$	A discretization scheme of S_t^x
$\mathbb{E}(X)$	The mean of a random variable X
R_v^h	The νth-order remainder of a discretization scheme $\hat{X}_t^x$
A	The generator of the WF process
A_0	The generator of the stochastic part of the WF process

References

1. Milstein, G.N.; Tretyakov, M.V. *Stochastic Numerics for Mathematical Physics*; Springer: Berlin/Heidelberg, Germany, 2004.
2. Alfonsi, A. *Affine Diffusions and Related Processes: Simulation, Theory and Applications*; Bocconi & Springer Series; Springer: Cham, Switzerland, 2015.
3. Cox, J.C.; Ingersoll, J.E.; Ross, S.A. A theory of the term structure of interest rates. *Econometrica* **1985**, *53*, 385–407.
4. Alfonsi, A. High order discretization schemes for the CIR process: Application to affine term structure and Heston models. *Math. Comput.* **2010**, *79*, 209–237.
5. Lileika, G.; Mackevičius, V. Weak approximation of CKLS and CEV processes by discrete random variables. *Lith. Math. J.* **2020**, *60*, 208–224.
6. Lileika, G.; Mackevičius, V. Second-order weak approximations of CKLS and CEV processes by discrete random variables. *Mathematics* **2021**, *9*, 1337.
7. Epstein, C.L.; Mazzeo, R. Wright–Fisher diffusion in one dimension. *SIAM J. Math. Anal.* **2010**, *42*, 568–608.
8. Krantz, S.G.; Parks, H.R. *A Primer of Real Analytic Functions*, 2nd ed.; Birkhäuser: Boston, USA 2002.
9. Øksendal, B. *Stochastic Differential Equations: An Introduction with Applications*, 5th ed.; Springer: Berlin/Heidelberg, Germany; New York, NY, USA, 2000.