https://doi.org/10.15388/vu.thesis.319
https://orcid.org/0000-0003-3102-1587

VILNIUS UNIVERSITY

Gabrielė Mongirdaitė

Weak Approximations of the Wright–Fisher process and Related Problems

DOCTORAL DISSERTATION

Natural sciences,
mathematics (N 001)

VILNIUS 2022

Doctoral dissertation was prepared between 2017 and 2022 at Vilnius University.

Academic supervisor:

prof. habil. dr. Vigirdas Mackevičius
(Vilnius University, Natural sciences, Mathematics, N 001)

https://doi.org/10.15388/vu.thesis.319
https://orcid.org/0000-0003-3102-1587

VILNIAUS UNIVERSITETAS

Gabrielė Mongirdaitė

Wright–Fisher proceso silpnosios aproksimacijos ir susiję uždaviniai

DAKTARO DISERTACIJA

Gamtos mokslai,
matematika (N 001)

VILNIUS 2022

Disertacija rengta 2017–2022 metais Vilniaus universitete.

Mokslinis vadovas:

prof. habil. dr. Vigirdas Mackevičius
(Vilniaus universitetas, gamtos mokslai, matematika, N 001)

Contents

List of Figures

Notation and Abbreviations

Notations	Descriptions
$\mathbb{N}$	The set of positive integers $\{1, 2, \ldots\}$.
$\mathbb{N}_0$	The set of nonnegative integers, $\mathbb{N} \bigcup \{0\}$.
$\mathbb{R}$	The set of real numbers $(-\infty, \infty)$.
$\mathbb{R}_+$	The set of positive real numbers $(0, \infty)$.
$\overline{\mathbb{R}}_+$	The set of nonnegative real numbers $[0, \infty)$.
$\mathbb{D}$	The domain of the solution of SDE. The domain of the solution of the WF equation is $[0, 1]$.
$C^\infty(\mathbb{D})$	The set of infinitely differentiable functions $f : \mathbb{D} \to \mathbb{R}$.
$C_0^\infty(\mathbb{D})$	The set of functions $f : \mathbb{D} \to \mathbb{R}$ of class C^∞ with a compact support.
$C_{\mathrm{pol}}^\infty(\mathbb{D})$	The set of functions $f : \mathbb{D} \to \mathbb{R}$ of class C^∞ with all partial derivatives of polynomial growth.
$C_{\mathrm{lin}}(\mathbb{D})$	The set of functions $f : \mathbb{D} \to \mathbb{R}$ of linear growth.
$O(h^n)$	A function of polynomial growth with respect to h^n, i.e., we write $g(x, h) = O(h^n)$ if for some $C > 0$, $k \in \mathbb{N}$, and $h_0 > 0$, $\|g(x, h)\| \leqslant C(1 + \|x\|^k)h^n, \quad x \geqslant 0, \quad 0 < h \leqslant h_0$.
$\mathcal{O}(h^n)$	A function of polynomial growth with respect to h^n when the function g is expressed in terms of another function $f \in C_{\mathrm{pol}}^\infty(\mathbb{D})$ and the constants C, h_0, and k depend only on a good sequence for f.
$\lfloor x \rfloor$	The integer part of a number x.
Δ^h	The (equidistant) discretization of a time interval.
$\mathbb{E}X$	The mean of a random variable X.
$\mathcal{N}(a, \sigma^2)$	The normal distribution with mean a and variance σ^2.
B_t, $\tilde{B}_t$	Standard Brownian motions (Wiener processes).
$\hat{X}$	A discretization of a random variable X.
$f^{(i)}(z)$	The ith derivative of a function f of a real variable z.
sgn	$\operatorname{sgn}(x) = 1$ for $x \geq 0$ and $\operatorname{sgn}(x) = -1$ for $x < 0$
SDE	Stochastic differential equation.

Notations	Descriptions
PDE	Parabolic partial differential equation.
ODE	Ordinary differential equation.
CIR	Cox–Ingersoll–Ross.
WF	Wright–Fisher.
CKLS	Chan–Karolyi–Longstaff–Sanders.

Chapter 1

Introduction

In this chapter, we present our research topic, aim and applied methods, novelty of main results, the list of published papers, and the list of conferences where our results were presented.

1.1 Research topic

The main topics of the thesis are the following:

1. Construction of easy-to-implement weak first- and second-order approximations for the Wright–Fisher equation

$$X_t^x = x + \int_0^t (a - bX_s^x)\,\mathrm{d}s + \sigma \int_0^t \sqrt{X_s^x(1 - X_s^x)}\,\mathrm{d}B_s, \quad x \in [0,1], \tag{1.1.1}$$

   with parameters $0 \le a \le b$ and $\sigma > 0$. Here B is a standard Brownian motion (Wiener process). It is well known that equation (1.1.1) has a unique strong solution X^x, which remains in the interval $[0,1]$, i.e., $\mathbb{P}\{X_t^x \in [0,1]\ \forall t \geqslant 0\} = 1$ [4].

2. Regularity of the solutions of the backward Kolmogorov PDEs related to Wright–Fisher, CIR, and general Stratonovich equations. Such a regularity is needed for a rigorous proof that a potential ("candidate") first- or second-order weak approximation of an SDE is indeed a weak approximation of the corresponding order.

1.2 Aim and difficulties

The aim of the thesis has been to construct simple, yet effective first- and second-order weak approximations for the solution of the WF model that use only generation of discrete random variables at each approximation step. In addition, we have focused on proving the regularity of solutions of the backward Kolmogorov equations for the WF equation and, in addition, for the CIR and general Stratonovich equations with square-root diffusion coefficient since this fact is essential in rigorous proofs of the convergence rates of weak approximations of SDEs.

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 [34]) 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 [4, Chap. 6] constructed a weak second-order approximation of the WF process by using its connection with the CIR [11] process and the earlier constructed approximations of the latter (Alfonsi [3]). In comparison with the numerical scheme of Alfonsi [4, Prop. 6.1.13, Algs. 6.1 and 6.2], our algorithm is direct and, in addition, much simpler and easier to implement. In our constructions, we follow some ideas of Lileika and Mackevičius [29, 30]. 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 [29, 30]).

1.3 Methods

Methods of calculus, stochastic calculus, probability theory, statistics, and functional analysis are applied in the thesis. Numerical experiments were simulated using the programming language Python. The figures were generated using Python and the computing environment Maple. The same software was also used for solving equalities and inequalities.

1.4 Actuality and novelty

The Wright–Fisher process was originally used to model gene frequencies, i.e., the proportions of genes in a population. In recent years Wright–Fisher and Jacobi processes have started appearing in the finance applications. These processes are restricted to a finite interval, and due to this, they seem appropriate to model dynamic bounded variables such as a regime probability or a default probability. However, closed-form solutions of the WF or Jacobi models are unknown therefore a need to approximate them arises.

An existing weak second-order approximation of the WF process [4] relies heavily on the WF process connection with the CIR process and the earlier constructed approximations of the latter. However, we propose schemes that are constructed directly for the WF process and are easier to implement. The accuracy of the approximation is shown not only by the rigorous proofs but also by the simulation examples, which are not present for the existing approximation constructed by Alfonsi.

When it comes to the regularity of solutions of the backward Kolmogorov equations, known results rely on explicit formulas of density [2] or partial differential equation theory [13]. In this thesis, we give probabilistic proofs of the regularity of solutions for the WF equation and for the CIR equation without relying on existing transition density formulas. This enabled us to extend our method to general Stratonovich equations with square-root diffusion coefficient.

1.5 Main results

We managed to construct simple and effective first- and second-order weak approximations for the solution of the Wright–Fisher model. These discretization schemes use only generation of discrete random variables at each approximation step. They are presented in the theorems below.[1]

Theorem 1.1. *Let*

$$D_t^x = D(x,t) = \begin{cases} x\mathrm{e}^{-bt} + \frac{a}{b}\left(1-\mathrm{e}^{-bt}\right), & 0 \le a \le b \ne 0, \\ x, & a = b = 0, \end{cases} \tag{1.5.1}$$

and let the random variable $\hat{S}_h^x$ take the values

$$\hat{x}_{1,2} := \begin{cases} x_{1,2}(x,h) \text{ with prob. } p_{1,2} = \frac{x}{2x_{1,2}(x,h)}, & x \in [0,1/2], \\ 1 - x_{1,2}(1-x,h) \text{ with prob. } p_{1,2} = \frac{1-x}{2x_{1,2}(1-x,h)}, & x \in (1/2,1], \end{cases} \tag{1.5.2}$$

where

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

Then the one-step approximation $\hat{X}_h^x$ defined by the composition

$$\hat{X}_h^x := D(\hat{S}_h^x, h),\ x \in [0,1],\ h > 0,$$

defines a first-order weak approximation for the Wright–Fisher equation (1.1.1)*.*

[1] For definitions, see Chapter 3.

Theorem 1.2. *Let*

$$
\begin{aligned}
x_1 &= x_1(x,h) = x + \frac{3(1-x)\sigma^2 h}{4} + \frac{2x\sigma^2 h}{3} - \sqrt{3x(1-x)\sigma^2 h},\\
x_2 &= x_2(x,h) = x + \frac{17x\sigma^2 h}{12},\\
x_3 &= x_3(x,h) = x + \frac{3(1-x)\sigma^2 h}{4} + \frac{2x\sigma^2 h}{3} + \sqrt{3x(1-x)\sigma^2 h},\\
y_{1,2} &= y_{1,2}(x,h) = x + (1-x)\sigma^2 h \mp \sqrt{(x+(1-x)\sigma^2 h)(1-x)\sigma^2 h},
\end{aligned}
$$

and let us define the random variable $\hat{S}_h^x$ as follows:

$$
\hat{S}_h^x := \begin{cases}
x_{1,2,3}(x,h) \text{ with probabilities } p_{1,2,3}(x,h) \text{ and} \\
\quad 0 \text{ with probability } p_0(x,h) = 1 - (p_1+p_2+p_3)(x,h), \\
\hfill x \in (\frac{\sigma^2 h}{3}, \frac{1}{2}], \\
1 - x_{1,2,3}(1-x,h) \text{ with prob. } p_{1,2,3}(1-x,h) \text{ and} \\
\quad 1 \text{ with probability } p_0(x,h) = 1 - (p_1+p_2+p_3)(1-x,h), \\
\hfill x \in (\frac{1}{2}, 1 - \frac{\sigma^2 h}{3}), \\
y_{1,2}(x,h) \text{ with probabilities } \tilde{p}_{1,2}(x,h) := \frac{x}{2y_{1,2}(x,h)}, \\
\hfill x \in [0, \frac{\sigma^2 h}{3}], \\
1 - y_{1,2}(1-x,h) \text{ with probabilities } \tilde{p}_{1,2}(1-x,h), \\
\hfill x \in [1 - \frac{\sigma^2 h}{3}, 1],
\end{cases} \tag{1.5.3}
$$

where

$$
\begin{aligned}
p_1(x,h) &= \frac{\hat{m}_1 x_2 x_3 - \hat{m}_2 x_2 - \hat{m}_2 x_3 + \hat{m}_3}{x_1(x_1-x_3)(x_1-x_2)},\\
p_2(x,h) &= -\frac{\hat{m}_1 x_1 x_3 - \hat{m}_2 x_1 - \hat{m}_2 x_3 + \hat{m}_3}{x_2(x_1-x_2)(x_2-x_3)},\\
p_3(x,h) &= \frac{\hat{m}_1 x_1 x_2 - \hat{m}_2 x_1 - \hat{m}_2 x_2 + \hat{m}_3}{x_3(x_2-x_3)(x_1-x_3)},
\end{aligned} \tag{1.5.4}
$$

with

$$
\begin{aligned}
\hat{m}_1 &= x,\\
\hat{m}_2 &= x^2 + \sigma^2 h x(1-x)(1 - \tfrac{1}{2}\sigma^2 h),\\
\hat{m}_3 &= x^3 + \tfrac{3}{2}x(\sigma^2 h)^2(3x^2 - 4x + 1) - 3x\sigma^2 h(x^2 - x).
\end{aligned}
$$

Then the one-step approximation $\hat{X}_h^x$ defined by the composition

$$
\hat{X}_h^x = D(\hat{S}(D(x,h/2),h),h/2), \; x \in [0,1], \; h > 0, \tag{1.5.5}
$$

defines a second-order weak approximation of the Wright–Fisher equation (1.1.1).

The PDE with initial condition

$$\begin{cases} \partial_t u(t,x) = Au(t,x), & x \geq 0,\ t \in [0,T], \\ u(0,x) = f(x), & x \geq 0, \end{cases} \tag{1.5.6}$$

is called the backward Kolmogorov equation related to the SDE

$$X_t^x = x + \int_0^t b(X_s^x)\,\mathrm{d}s + \int_0^t \tilde{\sigma}(X_s^x)\,\mathrm{d}B_s, \quad t \geq 0, \quad x \in \mathbb{D} \subset \mathbb{R}, \tag{1.5.7}$$

where $Af = bf' + \frac{1}{2}\tilde{\sigma}^2 f''$ is the generator of the solution of Eq. (1.5.7). We prove the regularity of solutions of the backward Kolmogorov equations for the WF equation and, in addition, for the CIR and general Stratonovich equations with square-root diffusion coefficient. Such a regularity is needed for rigorous proofs that potential ("candidate") weak approximations are indeed weak approximations of the corresponding order.

Theorem 1.3. *Let*

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

and let X_t^x be a Wright–Fisher process. If $f \in C_*^\infty[0,1]$, then

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

is a C^∞ function that solves

$$\partial_t u(t,x) = Au(t,x). \tag{1.5.8}$$

Theorem 1.3 is a known result proved by Epstein and Mazzeo [13] for $f \in C^\infty[0,1]$ using methods of partial differential equation theory. We present a probabilistic proof, which is much simpler and straightforward.

We denote by $C^\infty_{\mathrm{pol}}(\mathbb{D})$ the functions $f \in C^\infty(\mathbb{D})$ such that

$$|f^{(n)}(x)| \leq C_n(1 + |x|^{k_n}), \quad x \in \mathbb{D}, \quad n \in \mathbb{N}_0 := \{0,1,2,\dots\},$$

for some sequence $(C_n, k_n) \in \mathbb{R}_+ \times \mathbb{N}_0$. Following Alfonsi [3], we say that such a sequence $\{(C_n,k_n), n \in \mathbb{N}_0\}$ is a *good* sequence for f. In addition, by $C_{\mathrm{lin}}(\mathbb{D})$ we denote functions $f \in C(\mathbb{D})$ such that

$$|f(x)| \leq C(1 + |x|), \quad x \in \mathbb{D}.$$

Theorem 1.4. *Let*

$$X_t(x) = X_t^x = x + \int\limits_0^t \theta(\kappa - X_s^x)\,\mathrm{d}s + \int\limits_0^t \sigma\sqrt{X_s^x}\,\mathrm{d}B_s, \quad t \in [0,T],$$

be the CIR process with coefficients $\theta, \kappa, \sigma > 0$ *satisfying the condition* $\sigma^2 \le 4\theta\kappa$ *and starting at* $x \ge 0$. *Let* $f \in C^q_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$ *for some* $q \ge 4$. *Then the function*

$$u(t,x) := \mathbb{E}f(X_t(x)), \quad x \ge 0, t \in [0,T],$$

is l *times continuously differentiable in* $x \ge 0$ *and* l' *times continuously differentiable in* $t \in [0,T]$ *for* $l, l' \in \mathbb{N}$ *such that* $2l + 4l' \le q$. *Moreover, there exist constants* $C \ge 0$ *and* $k \in \mathbb{N}$, *depending only on a good set* $\{(C_i, k_i),\ i = 0, 1, \dots, q\}$ *for* f, *such that*

$$\left|\partial_x^j \partial_t^i u(t,x)\right| \leqslant C(1 + x^k), \quad x \ge 0,\ t \in [0,T], \tag{1.5.9}$$

for $j = 0, 1, \dots, l,\ i = 0, 1, \dots, l'$. *In particular,* $u(t,x)$ *is a* (*classical*) *solution of the Kolmogorov backward equation for* $(t,x) \in [0,T] \times \overline{\mathbb{R}}_+$.

As a consequence, if $f \in C^\infty_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$, *then* $u(t,x)$ *is infinitely differentiable on* $[0,T] \times \overline{\mathbb{R}}_+$, *and estimate* (1.5.9) *holds for all* $i, j \in \mathbb{N}$ *with* C *and* k *depending on* (i,j) *and a good sequence* $\{(C_i, k_i),\ i \in \mathbb{N}_0\}$ *for* f.

Theorem 1.4 is a known result proved by Alfonsi [2] by using the complex analytical formula for the transition density of the CIR process. We prove it without relying on this density. Our method allows us to generalize the result to Stratonovich equations with square-root diffusion coefficient:

Theorem 1.5. *Let* $X_t(x) = X_t^x$, $t \ge 0$, *be the process satisfying the Stratonovich SDE*

$$\mathrm{d}X_t = \sqrt{X_t a(X_t)} \circ \mathrm{d}B_t, \qquad X_0 = x \ge 0. \tag{1.5.10}$$

In addition, let $f \in C^q_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$ *for some* $q \geqslant 4$, *let* $a \in C^{2l-1}_{\mathrm{pol}}(\overline{\mathbb{R}}_+) \cap C_{\mathrm{lin}}(\overline{\mathbb{R}}_+)$ *for some* $l \geqslant 1$, *and suppose that* $0 < C_0 \leqslant a(x)$, $x \geqslant 0$. *Then the function*

$$u(t,x) := \mathbb{E}f(X_t(x)), \quad x \geqslant 0,\ t \in [0,T],$$

is l *times continuously differentiable in* $x \geqslant 0$ *and* l' *times continuously differentiable in* $t \in [0,T]$ *for* $l, l' \in \mathbb{N}$ *such that* $2l + 4l' \leqslant q$. *Moreover, there exist constants* $C \ge 0$ *and* $k \in \mathbb{N}$, *depending only on a good set* $\{(C_i, k_i),\ i = 0, 1, \dots, q\}$ *for* f *such that*

$$\left|\partial_x^j \partial_t^i u(t,x)\right| \leqslant C(1+x^k), \quad x \geq 0, \tag{1.5.11}$$

for $j = 0, 1, \ldots, l, i = 0, 1, \ldots, l'$. *In particular,* $u(t,x)$ *is a (classical) solution of the Kolmogorov backward equation for* $(t,x) \in [0,T] \times \overline{\mathbb{R}}_+$.

As a consequence, if $f \in C^\infty_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$ *and* $a \in C^\infty_{\mathrm{pol}}(\overline{\mathbb{R}}_+) \cap C_{\mathrm{lin}}(\overline{\mathbb{R}}_+)$*, then* $u(t,x)$ *is infinitely differentiable on* $[0,T] \times \overline{\mathbb{R}}_+$*, and estimate* (1.5.11) *holds for all* $i, j \in \mathbb{N}$ *with* C *and* k *depending on* (i,j) *and a good sequence* $\{(C_i, k_i),\ i \in \mathbb{N}_0\}$ *for* f.

1.6 Publications

1. V. Mackevičius, G. Mongirdaitė, *On backward Kolmogorov equation related to CIR process*, Modern Stoch. Theory Appl. **5**(1), 113–127 (2018).

2. G. Mongirdaitė, V. Mackevičius, *Weak approximations of Wright–Fisher equation*, Lietuvos matematikos rinkinys **62**: 23–26 (2021).

3. V. Mackevičius, G. Mongirdaitė, *Weak approximations of the Wright–Fisher process*, Mathematics **10**, 125 (2022).

1.7 Conferences

The results of the thesis-related studies were presented in the following conferences:

- 12th International Vilnius Conference on Probability Theory and Mathematical Statistics, Vilnius, Lithuania, 2018–07–03.

- 58th Conference of Lithuanian Mathematical Society (LMS), Vilnius, Lithuania, 2017–06–21.

- 60th Conference of LMS, Vilnius, Lithuania, 2019–06–19.

- 61st Conference of LMS, Šiauliai, Lithuania, 2020–12–04.

- 62nd Conference of LMS, Vilnius, Lithuania, 2021–06–16.

1.8 Structure of the thesis

The thesis is organized as follows. In Chapter 2, we give an overview of related results obtained by other authors. Preliminaries and definitions are provided in Chapter 3. In Section 4.1, we construct a first-order weak approximation of the WF model, and in Section 4.2, we construct a second-order weak approximation of the WF model. In the same sections, we summarize the constructed first- and second-order algorithms and illustrate them by numerical simulation results. In Section 5, we prove the regularity of solutions of backward Kolmogorov equations for SDEs with square-root diffusion coefficients: for WF, CIR, and general Stratonovich square-root diffusions (Sections 5.1, 5.2, and 5.3, respectively). We provide conclusions of the thesis in Chapter 6, and in the Appendix (Chapter 7), we provide additional calculations.

1.9 Acknowledgments

First and foremost I am extremely grateful to my scientific adviser prof. Vigirdas Mackevičius for his invaluable advice, continuous support, and patience during my PhD studies. Your guidance, experience and empathy encouraged me to continue this sometimes rough path.

I would also like to thank prof. Kęstutis Kubilius and prof. Remigijus Leipus for the careful reading of the thesis manuscript and their useful remarks.

Finally, special thanks goes to my family and my significant other Mindaugas. You encouraged me to embark on this journey and were here for me to celebrate success and support during difficult times.

Chapter 2

A short historical overview

2.1 A deep dive into the Wright–Fisher process

The Wright–Fisher process was originally intended to model gene frequencies, i.e., the proportions of genes in a population. Before the introduction of the WF process, the change in gene frequencies was seen as a deterministic process [17]. However, such assumptions are valid only when the population is big enough and the surrounding environment is constant or changes in a deterministic way [23]. Wright [41] summarized all factors that cause secular changes in mutation rates, conditions of selection, size and structure of the population, and the possibilities of ingression from other populations. The list is as follows:

1. Systematic change
   - Pressure of recurrent mutation.
   - Pressure of immigration and crossbreeding.
   - Pressure of intragroup selection.
2. Random fluctuations
   - From accidents of sampling.
   - From fluctuations in the systematic pressures.
3. Nonrecurrent change

- Nonrecurrent mutation.
- Nonrecurrent hybridization.
- Nonrecurrent selective incidents.
- From nonrecurrent extreme reduction in numbers.

Let us introduce the mathematical rationale behind the WF process [21, Chap. 15, Sec. 2], [14, Chap. 10]. Consider a population of N individuals composed of two types A and a. Suppose that the current state (the number of A-types) is i, and therefore the other $N-i$ individuals are of a-type. The next generation is produced subject to the influence of mutation, selection, and sampling forces. We stipulate that mutation converts at birth an A-type to an a-type and an a-type to an A-type with probabilities α and β, respectively. Given the parental population comprised of i A-types and $N-i$ a-types, the expected fraction of A-types after mutation is $(i/N)(1-\alpha)+(1-i/N)\beta$ and of the a-types is $(i/N)\alpha+(1-i/N)(1-\beta)$. We next stipulate that the relative survival abilities of the two types A and a in contributing to the next generation are in the ratio of $1+s$ to 1, where s is small and positive. Thus type A is selectively superior to type a. Taking account of these mutation and selection forces, the expected fraction of mature A-types before reproduction is

$$p_i = \frac{(1+s)[i(1-\alpha)+(N-i)\beta]}{(1+s)[i(1-\alpha)+(N-i)\beta]+[i\alpha+(N-i)(1-\beta)]}. \tag{2.1.1}$$

The Wright–Fisher model postulates that the composition of the next generation is determined through N binomial trials, where the probability of producing an A-type offspring on each trial is p_i as given in (2.1.1). Thus the population process $\{X(t) =$ number of A-types in the tth generation$\}$ evolves as a Markov chain governed by the transition probability matrix with elements

$$P_{ij} = \binom{N}{j} p_i^j \left(1-p_i\right)^{N-j}. \tag{2.1.2}$$

Kimura [22] showed that in large populations the discrete WF process can be closely approximated by a continuous-time continuous-space diffusion process of the form (1.1.1).

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

$$\mathrm{d}\widetilde{X}_t = (\tilde{a} - b\widetilde{X}_t)\,\mathrm{d}t + \sigma\sqrt{(\tilde{X}_t - \alpha)(\beta - \tilde{X}_t)}\,\mathrm{d}B_t, \quad \tilde{X}_0 \in [\alpha, \beta], \quad (2.1.3)$$

which is well defined when $b\alpha \leq \tilde{a} \leq b\beta$. A popular choice is the Jacobi process with $\alpha = -1$ and $\beta = 1$. Process (2.1.3) can be obtained from the WF process by the affine transformation

$$\widetilde{X}_t = \alpha + (\beta - \alpha)X_t, \quad \tilde{a} = a(\beta - \alpha). \quad (2.1.4)$$

For the rest of this chapter, all Jacobi SDE parameters are marked with tilde, that is,

$$\widetilde{X}_t^x = \tilde{x} + \int_0^t (\tilde{a} - \tilde{b}\widetilde{X}_s^x)\,ds + \tilde{\sigma}\int_0^t \sqrt{(1 - (\widetilde{X}_s^x)^2)}\,dB_s, \quad x \in [-1, 1]. \quad (2.1.5)$$

In more recent years the Wright–Fisher and Jacobi processes found their place in finance applications. It seems only natural as these processes are restricted to a finite interval, and due to this, they seem appropriate to model dynamic bounded variables such as a regime probability or a default probability. For instance, Delbaen and Shirakawa [12] presented a new interest rate dynamics model where the interest rates fluctuate in a bounded region, whereas Ethier and Kurtz [31] proposed a modified Jacobi process to evaluate risk premium of the stochastic correlation. The multivariate Jacobi process has been studied by Gourieroux and Jasiak [16], who suggested using it for modeling smooth transitions between alternative regimes in continuous time. The most recent research regarding Wright–Fisher process applications in finance focused on credit default swaps [6] and LIBOR market [5].

2.2 One-step approximations

CIR approximation schemes employing the properties of variance

The well-known Cox–Ingersoll–Ross (CIR) process [11] is the solution of the SDE

$$X_t^x = x + \int_0^t \theta(\kappa - X_s^x)\,\mathrm{d}s + \int_0^t \sigma\sqrt{X_s^x}\,\mathrm{d}B_s, \quad t \in [0,T], \tag{2.2.1}$$

with parameters $\theta, \kappa, \sigma > 0$ and $x \geq 0$. The Ninomiya and Victoir approximation scheme [36], starting from x with a time step t, for the CIR equation is described in [4]

$$\begin{aligned}\hat{X}_t^x = e^{-\frac{\theta t}{2}} &\left(\sqrt{(\theta\kappa - \sigma^2/4)\,\zeta_\theta(t/2) + e^{-\frac{\theta t}{2}}x} + \frac{\sigma}{2}\sqrt{t}N\right)^2 \\ &+ \left(\theta\kappa - \sigma^2/4\right)\zeta_\theta(t/2),\end{aligned} \tag{2.2.2}$$

where

$$\begin{aligned}\zeta_\theta(t) &= \int_0^t e^{-\theta s} ds = \begin{cases} \frac{1}{\theta}\left(1 - e^{-\theta t}\right) \text{ if } \theta \neq 0, \\ t \text{ if } \theta = 0, \end{cases} \\ N &\sim \mathcal{N}(0,1),\end{aligned}$$

is only appropriately defined when $\sigma^2 \leqslant 4\theta\kappa$.

Alfonsi [3, Thm. 2.8] modifies the previous scheme and suggests a second-order weak approximation that is well defined without restriction on the parameters. First, he defines the threshold

$$\begin{aligned}K_2(t) = \mathbb{1}_{\{\sigma^2 > 4\theta\kappa\}} e^{\frac{\theta t}{2}} &\left(\left(\frac{\sigma^2}{4} - \theta\kappa\right)\zeta_\theta(t/2)\right. \\ &\left.+ \left[\sqrt{e^{\frac{\theta t}{2}}\left[\left(\frac{\sigma^2}{4} - \theta\kappa\right)\zeta_\theta(t/2)\right]} + \frac{\sigma}{2}\sqrt{3t}\right]^2\right).\end{aligned} \tag{2.2.3}$$

Then the other nonnegative $\hat{X}_t^x$ scheme such that

$$\begin{cases} \forall i \in \{1,2\}, \mathbb{E}\left[(\hat{X}_t^x)^i\right] = \mathbb{E}\left[(X_t^x)^i\right], \\ \forall q \in \mathbb{N}, \exists C_q > 0, \forall t \in [0,1], x \in [0, K_2(t)), \mathbb{E}\left[(\hat{X}_t^x)^q\right] \leqslant C_q t^q, \end{cases}$$

is used in the interval $[0, K_2(t))$. The algorithm of the discretization scheme is presented in Algorithm 1.

Algorithm 1: Algorithm for the second-order scheme of the CIR with time step $t > 0$, U being sampled uniformly on $[0,1]$.

if $x \geqslant K_2(t)$ **then**

$$X = e^{-\frac{\theta t}{2}} \left(\sqrt{(\theta\kappa - \sigma^2/4)\,\zeta_\theta(t/2) + e^{-\frac{\theta t}{2}} x} + \tfrac{\sigma}{2}\sqrt{t}Y \right)^2$$
$$+ \left(\theta\kappa - \sigma^2/4\right)\zeta_\theta(t/2),$$

else

$p = \frac{1-\sqrt{1-\tilde{u}_1(t,x)^2/\tilde{u}_2(t,x)}}{2}$,

if $U < p$ **then**

$X = \frac{\tilde{u}_1(t,x)}{2p}$,

else

$X = \frac{\tilde{u}_1(t,x)}{2(1-p)}$.

Here

$$\mathbb{P}(Y = \sqrt{3}) = \mathbb{P}(Y = -\sqrt{3}) = \frac{1}{6}, \mathbb{P}(Y = 0) = \frac{2}{3}, \tag{2.2.4}$$
$$\tilde{u}_1(t,x) = xe^{-\theta t} + \theta\kappa\zeta_\theta(t), \tag{2.2.5}$$
$$\tilde{u}_2(t,x) = \tilde{u}_1(t,x)^2 + \sigma^2\zeta_\theta(t)\left[\theta\kappa\zeta_\theta(t)/2 + xe^{-\theta t}\right]. \tag{2.2.6}$$

WF and Jacobi connection with CIR process

Alfonsi [1,4] establishes a connection between CIR (and squared-Bessel) processes and one-dimensional Wright–Fisher and Jacobi diffusions.

Proposition 2.1. *Let B^1 and B^2 be two independent real Brownian motions. Let $b_1, b_2, z_1, z_2 \geqslant 0$ and $\sigma > 0$ be such that $\sigma^2 \leqslant 2(b_1 + b_2)$ and $z_1 + z_2 > 0$. We consider the following CIR processes:*

$$Z_t^i = z_i + b_i t + \int_0^t \sigma\sqrt{Z_s^i}\mathrm{d}B_s^i, i = 1, 2.$$

Then $Y_t = Z_t^1 + Z_t^2$ is a CIR process that never reaches 0, and we define

$$t \geqslant 0, \quad X_t = \frac{Z_t^1}{Y_t}, \quad \phi(t) = \int_0^t \frac{1}{Y_s}\mathrm{d}s.$$

Then ϕ is bijective on $\overline{\mathbb{R}}_+$, and the process $\left(X_{\phi^{-1}(t)}, t \geqslant 0\right)$ is a Wright–Fisher diffusion with parameters $a = b_1, b = b_1 + b_2$, and σ that is independent of $(Y_t, t \geqslant 0)$.

Second-order schemes for WF and Jacobi processes

Alfonsi [4] presents two approaches that lead to a second-order scheme for the Jacobi process without any restriction on the parameters. However, none of them is illustrated by simulation examples.

The first method relies on the Ninomiya–Victoir scheme. The infinitesimal generator of $\tilde{X}$ is given by

$$\tilde{L}f(x) = (\tilde{a} - \tilde{b}x)f'(x) + \frac{\tilde{\sigma}^2}{2}\left(1 - x^2\right) f''(x),\ x \in [-1, 1],$$

for C^2 functions $f : [-1, 1] \to \mathbb{R}$. The generator is split into $\tilde{L} = \tilde{L}_1 + \tilde{L}_2$ with

$$\tilde{L}_1 f(x) = \left(\tilde{a} - \left(\tilde{b} - \frac{\tilde{\sigma}^2}{2}\right) x\right) f'(x), \tag{2.2.7}$$

$$\tilde{L}_2 f(x) = -\frac{\tilde{\sigma}^2}{2} x f'(x) + \frac{\tilde{\sigma}^2}{2}\left(1 - x^2\right) f''(x). \tag{2.2.8}$$

The SDE associated with $\tilde{L}_2$ is equal to

$$\begin{aligned} X_t &= \sin\left(y + \tilde{\sigma}W_t\right), \quad y \in [-\pi/2, \pi/2] \Rightarrow \\ \mathrm{d}X_t &= \tilde{\sigma}\cos\left(y + \tilde{\sigma}W_t\right)\mathrm{d}W_t - \frac{\tilde{\sigma}^2}{2}X_t\mathrm{d}t = \tilde{\sigma}\sqrt{1 - X_t^2}\mathrm{d}W_t' - \frac{\tilde{\sigma}^2}{2}X_t\mathrm{d}t, \end{aligned}$$

with $\mathrm{d}W_t' = \left[\mathbb{1}_{\cos(y+\tilde{\sigma}W_t)\geqslant 0} - \mathbb{1}_{\cos(y+\tilde{\sigma}W_t)<0}\right]\mathrm{d}W_t$. The ODE associated with $\tilde{L}_1$ is linear and is solved by $\xi(t, x) = xe^{-\left(\tilde{b}-\frac{\tilde{\sigma}^2}{2}\right)t} + \tilde{a}\zeta_{\tilde{b}-\frac{\tilde{\sigma}^2}{2}}(t)$. This solution stays in $[-1, 1]$ if and only if

$$\tilde{a} - \left(\tilde{b} - \frac{\tilde{\sigma}^2}{2}\right) \leqslant 0 \text{ and } \tilde{a} + \left(\tilde{b} - \frac{\tilde{\sigma}^2}{2}\right) \geqslant 0.$$

Thus Alfonsi proposes the following steps:

1. First, use the Ninomiya and Victoir scheme

$$\xi(t/2, \sin(\arcsin(\xi(t/2, x)) + \tilde{\sigma}\sqrt{t}Y)),$$

where Y is given by (2.2.4). This scheme is well defined on the interval$[-1+K(t), 1-K'(t)]$, with $K(t), K'(t) \geqslant 0$ that satisfy $K(t) \underset{t\to 0}{=} O(t)$ and $K'(t) \underset{t\to 0}{=} O(t)$.

2. Use a scheme valued in $[-1, 1]$ that matches the two first moments of the Jacobi process when the starting point is either in the interval $[-1, -1+K(t))$ or $(1-K'(t), 1]$.

However, from the experience we know that defining $K(t), K'(t)$ as well as a scheme to be used near the singularity points seem to be more complicated than it looks from the first glance.

The second method relies on the fact that CIR and Wright–Fisher diffusions are closely related. This method reuses the high-order schemes developed for the CIR process.

Proposition 2.2. *Let $\xi(t,x) = xe^{-\tilde{b}t} + \tilde{a}\zeta_{\tilde{b}}(t)$ for $x \in [-1,1]$ and $t \geqslant 0$. The scheme defined by*

$$\hat{X}_t^x = \xi\left(t/2, p\left(\hat{Z}_{\phi(t)}^{1,\xi(t/2,x)}, \hat{Z}_{\phi(t)}^{2,\xi(t/2,x)}\right)\right),$$

where

$$\begin{aligned} &\hat{Z}_t^{1,x} = x + \tilde{\sigma}\sqrt{t}Y \quad \textit{with } Y \textit{ defined by } (2.2.4), \\ &\hat{Z}_t^{2,x} \textit{ is sampled independently according to the second-order} \\ &\qquad \textit{scheme of the CIR process, and} \\ &\phi(t) = \frac{-1+\sqrt{1+6\tilde{\sigma}^2 t}}{3\tilde{\sigma}^2}, \end{aligned}$$

is a second-order scheme for the Jacobi process.

This proposition is summarized in Algorithm 2:

Algorithm 2: Algorithm for the second-order scheme of the Jacobi process starting from x with time step $t > 0$.

1. $X = xe^{-\tilde{b}t/2} + \tilde{a}\zeta_{\tilde{b}}(t/2)$.
2. $Z_1 = X + \tilde{\sigma}\sqrt{\phi(t)}\tilde{N}$, with $N \sim \mathcal{N}(0,1)$.
3. Sample independently Z_2 by using Algorithm 1 with time step $\phi(t)$, starting point $1 - X^2$, and with CIR parameters $\kappa = 0, \theta = 0$ and $\sigma = 2\tilde{\sigma}$.
4. $X = \frac{Z_1}{\sqrt{(Z_1)^2 + Z_2}}$.
5. $X = Xe^{-\tilde{b}t/2} + \tilde{a}\zeta_{\tilde{b}}(t/2)$.

By using Algorithm 2 Alfonsi provides an algorithm for the second–order scheme for the Wright–Fisher process.

Algorithm 3: Algorithm for the second-order scheme of the WF process starting from x with time step $t > 0$.

1. Sample X with Algorithm 2 starting from $2x - 1$ and parameters $\tilde{a} = 2a - b, \tilde{b} = b, \tilde{\sigma} = \sigma$.
2. $X = (X + 1)/2$.

We believe that despite a brief form, these algorithms are pretty difficult to simulate given their dependency on each other.

Overview of split-step schemes for other SDEs

The split-step method (also called the splitting technique) used in this thesis has been suggested by Higham et al. [19] and reviewed in [35].

Mackevičius [33] provides first- and second–order weak approximations for the CIR process. Here we provide the formulation only for the second order.

Theorem 2.3. *Let $a = \sigma^2$, and let the discretization scheme $\hat{X}^x_t$ be defined by composition*

$$\hat{X}^h(x,h) := D\Big(\hat{S}\left(D\Big(x, \frac{h}{2}\Big), h\right), \frac{h}{2}\Big),$$

where the three-valued random variables $\hat{S}^x_h$ take the values x_1, x_2, and x_0 with probabilities p_1, p_2, and $p_0 = 1 - p_1 - p_2$ defined as follows:

- *If $x \geqslant 2ah$, then*

$$x_1 = x + \frac{s - \sqrt{\Delta}}{2}, \quad x_2 = x + \frac{s + \sqrt{\Delta}}{2}, \quad x_0 = x,$$

$$p_1 = \frac{2xah}{\sqrt{\Delta}(\sqrt{\Delta}-s)}, \quad p_2 = \frac{2xah}{\sqrt{\Delta}(\sqrt{\Delta}+s)},$$

where

$$s = \frac{3ah}{2}, \quad \Delta = \frac{21}{4}(ah)^2 + 12xah.$$

- *If* $0 < x < 2ah$, *then*

$$x_1 = \frac{s-\sqrt{\Delta}}{2}, \quad x_2 = \frac{s+\sqrt{\Delta}}{2}, \quad x_0 = 0,$$
$$p_1 = \frac{x(2ah-s-\sqrt{\Delta})}{\sqrt{\Delta}(\sqrt{\Delta}-s)}, \quad p_2 = \frac{x(2ah-s+\sqrt{\Delta})}{\sqrt{\Delta}(\sqrt{\Delta}+s)},$$

where

$$s = \frac{4x^2+9xah+3(ah)^2}{2x+ah},$$
$$\Delta = \frac{ah\left(16x^3+33x^2ah+18xa^2h^2+3a^3h^3\right)}{(2x+ah)^2}.$$

Then $\hat{X}_t^x$ *is a second-order discretization scheme for the CIR equation.*

Similarly, Lenkšas and Mackevičius [26, 27] constructed first- and second-order schemes for the Heston model [18], whereas Lileika and Mackevičius [29, 30] constructed first- and second-order schemes for the CKLS model [8].

2.3 Regularity of solutions of Kolmogorov backward equation

In this section, we provide an overview of various trials to prove the regularity of the solutions of Kolmogorov backward equation. Such a regularity is needed for a rigorous proof that a potential ("candidate") weak approximation is indeed a weak approximation of the corresponding order.

Alfonsi [2] proves the regularity for the CIR process using available analytical formula for the transition density

$$p(t,x,z) = \sum_{i=0}^{\infty} \frac{e^{-\lambda_t x/2}\left(\lambda_t x/2\right)^i}{i!} \frac{c_t/2}{\Gamma(i+v/2)} \left(\frac{c_t z}{2}\right)^{i-1+v/2} e^{-c_t z/2},$$

where $c_t = \frac{4\theta}{\sigma^2\left(1-e^{-\theta t}\right)}$, $v = 4\theta\kappa/\sigma^2$ and $\lambda_t = c_t e^{-\theta t}$. However, except for a few models in finance, e.g. Black and Scholes (1973) [7], Vasicek (1977) [40], Cox, Ingersoll and Ross (1985) [11] and Cox (1975) [10], the transition density of the SDE is not usually available in a closed form.

Epstein and Mazzeo in [13] prove the regularity for the WF process using methods of partial differential equation theory. However, such a method, to the best of our knowledge, is not generalizable for a broader class of square root diffusions and is overall long and complex.

Gabrielli in [15] made an effort to generalize the regularity result for the affine type stochastic process. One-dimensional affine stochastic process has the following form:

$$\mathrm{d}X_t^x = (a + bX_t^x)\mathrm{d}t + \sqrt{c + dX_t^x}\mathrm{d}B_t. \tag{2.3.1}$$

For instance, CIR, Coupled CIR, CIR with jumps or Heston models can be written in such a form. However, the Wright-Fisher process is not affine. Gabrielli result is provided below.

Theorem 2.4 (see [15, Thm. 4.23]). *Let $f \in C^\infty_{\text{pol}}$. Then, the function $u : \overline{\mathbb{R}}_+ \times \mathbb{D} \to \mathbb{R}$ defined by $u(t,x) = \mathbb{E}^x\left[f\left(X_t\right)\right]$ is smooth, with all derivatives satisfying the following property:*

$$\textit{for all } (t,x) \in [0,T] \times \mathbb{D}, \quad \left|\partial^\alpha_{(t,x)} u(t,x)\right| \leqslant K_\alpha(T)\left(1 + |x|^{2\eta_\alpha(T)}\right),$$

where $K_\alpha(T)$ and $\eta_\alpha(T)$ are positive constants depending on the time horizon T and the order of derivative α.

To our view, the proof of this theorem has a significant flaw. It is based on the incorrect equality

$$\partial_t \partial_x^{\bar{\alpha}} u(t,x) = A\partial_x^{\bar{\alpha}} u(t,x),$$

as the differentiation operator and generator of the Markov process do not commute. Therefore we cannot clearly state that regularity is proved for all affine stochastic processes.

Chapter 3

Preliminaries

In this chapter, we provide all definitions and describe techniques used to construct discretization schemes.

3.1 Preliminaries and definitions

In this section, we give some definitions for the general one-dimensional stochastic differential equation

$$X_t^x = x + \int_0^t b(X_s^x)\,\mathrm{d}s + \int_0^t \tilde{\sigma}(X_s^x)\,\mathrm{d}B_s, \quad t \geq 0, \quad x \in \mathbb{D} \subset \mathbb{R}. \qquad (3.1.1)$$

To avoid ambiguity, we indicate functions with the supplementary symbol $\tilde{}$ if the same letter is used for a function and a constant, for example, we denote by $\tilde{\sigma}$ the diffusion coefficient in the general equation (3.1.1) and by σ the constant in the WF equation (1.1.1).

We assume that the equation has a unique weak solution X_t^x such that $\mathbb{P}(X_t^x \in \mathbb{D}, t \geq 0) = 1$ for all $x \in \mathbb{D}$. For example, for Eq. (1.1.1), $\mathbb{D} = [0, 1]$.

Having a fixed time interval $[0, T]$, consider an equidistant time discretization $\Delta^h = \{ih, i = 0, 1, \ldots, \lfloor T/h \rfloor, h \in (0, T]\}$, where $\lfloor a \rfloor$ is the integer part of a. By a discretization scheme of Eq. (3.1.1) we mean a family of discrete-time homogeneous Markov chains $\hat{X}^h = \{\hat{X}^h(x, t), x \in \mathbb{D}, t \in \Delta^h\}$ with initial values $\hat{X}^h(x, 0) = x$ and one-step transition probabilities $p^h(x, \mathrm{d}z)$, $x \in \mathbb{D}$. For convenience, we only consider steps

$h = T/n$, $n \in \mathbb{N}$. We shortly write $\hat{X}_t^x$ or $\hat{X}(x,t)$ instead of $\hat{X}^h(x,t)$. Note that because of the Markovity, a one-step approximation $\hat{X}_h^x$ of the scheme completely defines the distribution of the whole discretization scheme $\hat{X}_t^x$, so that we only need to construct the former.

We denote by $C^\infty(\mathbb{D})$ the space of C^∞ functions $f : \mathbb{D} \to \mathbb{R}$, by $C_0^\infty(\mathbb{D})$ the functions $f \in C^\infty(\mathbb{D})$ with compact support in $\mathbb{D}$, and by $C_{\text{pol}}^\infty(\mathbb{D})$ the functions $f \in C^\infty(\mathbb{D})$ such that

$$|f^{(n)}(x)| \le C_n(1 + |x|^{k_n}), \quad x \in \mathbb{D}, n \in \mathbb{N}_0 := \{0, 1, 2, \dots\}, \tag{3.1.2}$$

for some sequence $(C_n, k_n) \in \mathbb{R}_+ \times \mathbb{N}_0$. Following Alfonsi [3], we say that such a sequence $\{(C_n, k_n), n \in \mathbb{N}_0\}$ is a *good* sequence for f. Finally, by $C_{\text{lin}}(\mathbb{D})$ we denote functions $f \in C(\mathbb{D})$ such that

$$|f(x)| \le C(1 + |x|), \quad x \in \mathbb{D}. \tag{3.1.3}$$

We will write $g(x,h) = O(h^n)$ if for some $C > 0$, $k \in \mathbb{N}$, and $h_0 > 0$,

$$|g(x,h)| \le C(1 + |x|^k)h^n, \quad x \ge 0, \quad 0 < h \le h_0.$$

If, in particular, the function g is expressed in terms of another function $f \in C_{\text{pol}}^\infty(\mathbb{R})$ and the constants C, k, and h_0 only depend on a *good* sequence for f, then we will, instead, write, $g(x,h) = \mathcal{O}(h^n)$.

The following PDE is called the backward Kolmogorov equation, with initial condition

$$\begin{cases} \partial_t u(t,x) = Au(t,x), & x \ge 0,\ t \in [0,T], \\ u(0,x) = f(x), & x \ge 0, \end{cases} \tag{3.1.4}$$

where $Af = bf' + \frac{1}{2}\tilde{\sigma}^2 f''$ is the generator of the solution of Eq. (3.1.1). If the coefficients $b, \tilde{\sigma}$ and the initial function f are sufficiently "good", then the function $u = u(t,x) := \mathbb{E} f(X_t^x)$ is a (classical) solution of PDE (3.1.4). From this by Itô's formula it follows that the random process $M_t^x := u(T - t, X_t^x)$, $t \in [0,T]$, is a martingale with mean $\mathbb{E}\, M_t^x = f(x)$ satisfying the final condition $M_T^x = f(X_T^x)$. This fact is essential in rigorous proofs of the convergence rates of weak approximations of SDEs. The higher the convergence rate, the greater smoothness of the coefficients and final condition is to be assumed to get a sufficient smoothness of the solution u of (3.1.4). The question of the existence of smooth classical solutions of the backward Kolmogorov equation is more complicated than it might seem from the first sight. General results typically require smoothness and polynomial growth of several higher-order derivatives of the coefficients; we refer to the book by Kloeden and Platen [24, Thm. 4.8.6 on p. 153].

Definition 3.1. A discretization scheme $\hat{X}^h$ is a weak νth-order approximation for the solution $(X_t^x, t \in [0,T])$ of Eq. (3.1.1) if for every $f \in C_0^\infty(\mathbb{D})$, there exists $C > 0$ such that

$$|\mathbb{E}f(X_T^x) - \mathbb{E}f(\hat{X}_T^x)| \leq Ch^\nu, \; h > 0.$$

Definition 3.2. Suppose $Af \in C_{\text{pol}}^\infty(\mathbb{D})$ for all $f \in C_{\text{pol}}^\infty(\mathbb{D})$, that is, $b, \tilde{\sigma}^2 \in C_{\text{pol}}^\infty(\mathbb{D})$. The νth-order remainder of a discretization scheme $\hat{X}_t^x$ for X_t^x is the operator $R_\nu^h : C_{\text{pol}}^\infty(\mathbb{D}) \to C(\mathbb{D})$ defined by

$$R_\nu^h f(x) := \mathbb{E}f(\hat{X}_h^x) - \left[f(x) + \sum_{k=1}^{\nu} \frac{A^k f(x)}{k!} h^k\right], \; x \in \mathbb{D}, \; h > 0. \quad (3.1.5)$$

A discretization scheme $\hat{X}_t^x$ is a local νth-order weak approximation of Eq. (3.1.1) if

$$R_\nu^h f(x) = O(h^{\nu+1}), \; h \to 0,$$

for all $f \in C_{\text{pol}}^\infty(\mathbb{D})$ and $x \in \mathbb{D}$.

Remark 3.1. Iterating the Dynkin formula

$$\mathbb{E}f(X_h^x) = f(x) + \int_0^h \mathbb{E}Af(X_s^x)\mathrm{d}s,$$

we have

$$\mathbb{E}f(X_h^x) = f(x) + \sum_{k=1}^{\nu} \frac{A^k f(x)}{k!} h^k$$
$$+ \int_0^h \int_0^{s_1} \cdots \int_0^{s_\nu} \mathbb{E}A^{\nu+1} f(X_{s_{\nu+1}}^x)\mathrm{d}s_{\nu+1} \cdots \mathrm{d}s_2 \mathrm{d}s_1, \quad (3.1.6)$$

which motivates Definition 3.2: If $A^{\nu+1}f$ behaves "well" (e.g., $b, \tilde{\sigma}^2, f \in C_0^\infty(\mathbb{D})$, and $\mathbb{E}A^{\nu+1}f$ is bounded), then for the "one-step" νth-order weak approximation scheme $\hat{X}_h^x$, we have

$$|\mathbb{E}f(X_h^x) - \mathbb{E}f(\hat{X}_h^x)| = O(h^{\nu+1}),\ h \to 0. \tag{3.1.7}$$

We may expect that in "good" cases, a local νth-order weak discretization scheme is a νth-order (global) approximation. Rigorous statements require a certain uniformity of (3.1.7) with respect to f and regularity of A.

Definition 3.3. A discretization scheme $\hat{X}_t^x$ is a potential νth-order weak approximation for Eq. (3.1.1) if for every $f \in C_{\mathrm{pol}}^\infty(\mathbb{D})$,

$$|R_\nu^h f(x)| = \mathcal{O}(h^{\nu+1}).$$

Definition 3.4. A discretization scheme $\hat{X}_t^x = \hat{X}^h(x,t),\ h > 0,$ has uniformly bounded moments if there exists $h_0 > 0$ such that

$$\sup_{0<h\le h_0} \sup_{t\in\Delta^h} \mathbb{E}(|\hat{X}^h(x,t)|^n) < +\infty,\ n \in \mathbb{N},\ x \in \mathbb{D}.$$

We say that a potential νth-order weak approximation is a *strongly* potential νth-order weak approximation if it has uniformly bounded moments.

The following two theorems ensure that a potential νth-order weak approximation for the Wright–Fisher equation is in fact a νth-order weak approximation (in the sense of Definition 3.1). Note that the requirement of uniformly bounded moments (see, e.g., [3]) is obviously satisfied by our approximations since they take values in [0, 1].

Theorem A (see [3, Thm. 1.19])**.** *Let $\hat{X}^h$ be a discretization scheme with transition probabilities $p^h(x, \mathrm{d}z)$ on $[0,1]$ that starts from $\hat{X}_0^x = x \in [0,1]$. 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 B (see [4, Thm. 6.1.12])**.** *Let* $f \in C^\infty[0,1]$. *Then*

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

is a C^∞ *function that solves*

$$\partial_t \tilde{u}(t,x) = A\tilde{u}(t,x). \tag{3.1.8}$$

3.2 Split-step technique for the WF model

We split Equation (1.1.1) into the deterministic part

$$\mathrm{d}D^x_t = (a - bD^x_t)\,\mathrm{d}t, \quad D^x_0 = x \in [0,1], \tag{3.2.1}$$

and the stochastic part

$$\mathrm{d}S^x_t = \sigma\sqrt{S^x_t(1-S^x_t)}\mathrm{d}B_t, \quad S^x_0 = x \in [0,1]. \tag{3.2.2}$$

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

$$D^x_t = D(x,t) = \begin{cases} x\mathrm{e}^{-bt} + \frac{a}{b}\left(1 - \mathrm{e}^{-bt}\right), & 0 \le a \le b \ne 0, \\ x, & a = b = 0. \end{cases} \tag{3.2.3}$$

The solution of the stochastic part is not explicitly known. However, suppose that $\hat{S}^x_t$ is a discretization scheme for the stochastic part. We define the first-order composition $\hat{X}^x_t$ 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

$$\hat{X}^h(x,h) := D(\hat{S}(x,h),h). \tag{3.2.4}$$

Similarly, the second-order composition is defined by

$$\hat{X}^h(x,h) := D\left(\hat{S}\left(D\left(x,\tfrac{h}{2}\right),h\right),\tfrac{h}{2}\right). \tag{3.2.5}$$

Theorem C (see [3, 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* (3.2.4) *and* (3.2.5) *define, respectively, a first- or second-order approximation $\hat{X}_t^x$ of the WF Equation* (1.1.1).

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 3.2. Clearly, by the transformation (2.1.4) we can get weak approximations for the similar to the WF processes (2.1.3) from weak approximations for the WF process.

3.3 Moment matching technique for the WF model

First-order

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

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

Let us denote $z = \sigma^2 h$ for brevity. The generator A_0 of the stochastic part $\mathrm{d}S_t^x = \sigma\sqrt{S_t^x(1 - S_t^x)}\mathrm{d}B_t$ is

$$A_0 f(x) = \frac{1}{2}\sigma^2 x(1-x)f''(x).$$

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

$$\begin{aligned}
R_1^h f(x) &= \mathbb{E}f(\hat{S}_h^x) - [f(x) + A_0 f(x)h] \\
&= f'(x)\mathbb{E}\big(\hat{S}_h^x - x\big) \\
&\quad + \frac{f''(x)}{2}\Big[\mathbb{E}\big(\hat{S}_h^x - x\big)^2 - zx(1-x)\Big] \\
&\quad + \frac{f'''(x)}{6}\mathbb{E}\big(\hat{S}_h^x - x\big)^3 + r_1(x,h),\ x \geqslant 0,\ h > 0,
\end{aligned}$$

where

$$|r_1(x,h)| = \frac{1}{6}\left|\mathbb{E}\int_x^{\hat{S}_h^x} f^{(4)}(s)\big(\hat{S}_h^x - s\big)^3 \,\mathrm{d}s\right| \leqslant \frac{1}{24}\max_{s\in[0,1]}\left|f^{(4)}(s)\right|\mathbb{E}\big(\hat{S}_h^x - x\big)^4.$$

This expression shows that $\hat{S}_h^x$ is a potential first-order approximation of the stochastic part (3.2.2) if

$$\mathbb{E}\big(\hat{S}_h^x - x\big) = O(h^2), \tag{3.3.1}$$

$$\mathbb{E}\big(\hat{S}_h^x - x\big)^2 = zx(1-x) + O(h^2), \tag{3.3.2}$$

$$\mathbb{E}\big(\hat{S}_h^x - x\big)^3 = O(h^2), \tag{3.3.3}$$

$$\mathbb{E}\big(\hat{S}_h^x - x\big)^4 = O(h^2). \tag{3.3.4}$$

Converting the central moments of $\hat{S}_h^x$ to noncentral moments, from (3.3.1)–(3.3.2) we get

$$\mathbb{E}\big(\hat{S}_h^x\big)^i = \hat{m}_i + O(h^2), \quad i = 1,2, \tag{3.3.5}$$

where the "moments" (further we call them approximate moments) $\hat{m}_i = \hat{m}_i(x,h)$, $x \geq 0, h > 0$, $i = 1,2$, are defined as

$$\begin{aligned} \hat{m}_1 &= x, \\ \hat{m}_2 &= x^2 + zx(1-x). \end{aligned} \tag{3.3.6}$$

Second-order

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 square of the generator A_0 of the stochastic part is

$$A_0^2 f(x) = -\frac{1}{2}\sigma^4 x(1-x)f''(x) + \frac{1}{2}\sigma^4 x(1-x)(1-2x)f'''(x)$$

$$+\frac{1}{4}\sigma^4 x^2(1-x)^2 f^{(4)}(x).$$

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

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

where

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

This expression shows that $\hat{S}_h^x$ is a potential second-order approximation of the stochastic part (3.2.2) if

$$\mathbb{E}\big(\hat{S}_h^x - x\big) = O(h^3), \tag{3.3.7}$$

$$\mathbb{E}\big(\hat{S}_h^x - x\big)^2 = zx(1-x)\Big(1-\frac{1}{2}z\Big) + O(h^3), \tag{3.3.8}$$

$$\mathbb{E}\big(\hat{S}_h^x - x\big)^3 = \frac{3}{2}z^2 x(1-x)(1-2x) + O(h^3), \tag{3.3.9}$$

$$\mathbb{E}\big(\hat{S}_h^x - x\big)^4 = 3z^2(x(1-x))^2 + O(h^3), \tag{3.3.10}$$

$$\mathbb{E}\big(\hat{S}_h^x - x\big)^5 = O(h^3), \tag{3.3.11}$$

$$\mathbb{E}\big(\hat{S}_h^x - x\big)^6 = O(h^3). \tag{3.3.12}$$

Converting the central moments of $\hat{S}_h^x$ to noncentral moments, from (3.3.7)–(3.3.12) we get

$$\mathbb{E}\big(\hat{S}_h^x\big)^i = \hat{m}_i + O(h^3), \quad i = 1, \dots, 6, \tag{3.3.13}$$

where

$$
\begin{aligned}
\hat{m}_1 &= x, \\
\hat{m}_2 &= x^2 + zx(1-x)(1-\tfrac{1}{2}z), \\
\hat{m}_3 &= x^3 + \tfrac{3}{2}xz^2(3x^2-4x+1) - 3xz(x^2-x), \\
\hat{m}_4 &= x^4 + 9x^2z^2(2x^2-3x+1) - 6x^2z(x^2-x), \\
\hat{m}_5 &= x^5 + 10x^3z^2(5x^2-8x+3) - 10x^3z(x^2-x), \\
\hat{m}_6 &= x^6 + \tfrac{75}{2}x^4z^2(3x^2-5x+2) - 15x^4z(x^2-x).
\end{aligned}
\tag{3.3.14}
$$

Note that (3.3.6) and (3.3.14) refer to respectively first- and second–order "moments". We use the same notation, since (3.3.6) "moments" are used only in the context of the first-order approximation, whereas (3.3.14) only for the second one.

Chapter 4

Weak approximations of the Wright–Fisher process

4.1 A first-order approximation

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 (1.1.1) with $a = b = 0$. In [33] (see also [29]), it is shown that 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

$$p_1 + p_2 = 1, \tag{4.1.1}$$

$$\mathbb{E}\hat{S}_h^x = x_1 p_1 + x_2 p_2 = m_1 := \mathbb{E}S_h^x = x, \tag{4.1.2}$$

$$\mathbb{E}(\hat{S}_h^x)^2 = x_1^2 p_1 + x_2^2 p_2 = m_2 := \mathbb{E}(S_h^x)^2, \tag{4.1.3}$$

$$\mathbb{E}(\hat{S}_h^x - x)^3 = (x_1 - x)^3 p_1 + (x_2 - x)^3 p_2 = O(h^2), \tag{4.1.4}$$

$$\mathbb{E}(\hat{S}_h^x - x)^4 = (x_1 - x)^4 p_1 + (x_2 - x)^4 p_2 = O(h^2), \tag{4.1.5}$$

where the second moment $m_2 = \mathbb{E}(S_h^x)^2$ can be calculated by Lemma 5.3 with $a = b = 0$:

$$\begin{aligned} m_2 = m_2(x,h) &= x^2 e^{-\sigma^2 h} + x(1 - e^{-\sigma^2 h}) && (4.1.6)\\ &= x^2 + x(1-x)\sigma^2 h + O(h^2) && (4.1.7)\\ &= \hat{m}_2 + O(h^2),\ x \in [0,1]. \end{aligned}$$

One of the solutions to the equation system (4.1.1)–(4.1.3) is (see [29])

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

Therefore, in our case, we get

$$
\begin{aligned}
x_{1,2} = {}& xe^{-\sigma^2 h} + 1 - e^{-\sigma^2 h} \\
&\mp \sqrt{\frac{xe^{-\sigma^2 h} + (1 - e^{-\sigma^2 h})}{x}\left(x^2 e^{-\sigma^2 h} + x(1 - e^{-\sigma^2 h}) - x^2\right)} \\
= {}& xe^{-\sigma^2 h} + 1 - e^{-\sigma^2 h} \mp \sqrt{(xe^{-\sigma^2 h} + 1 - e^{-\sigma^2 h})(1 - x)(1 - e^{-\sigma^2 h})}.
\end{aligned}
\tag{4.1.8}
$$

Since $1 - e^{-\sigma^2 h} = \sigma^2 h + O(h^2)$, to simplify the expressions, we may try to replace $1 - e^{-\sigma^2 h}$ by $\sigma^2 h$ and, instead of (4.1.8), use

$$
\begin{aligned}
x_{1,2} = x_{1,2}(x, h) &= x(1 - \sigma^2 h) + \sigma^2 h \mp \sqrt{(x(1 - \sigma^2 h) + \sigma^2 h)(1 - x)\sigma^2 h} \\
&= x + (1 - x)\sigma^2 h \mp \sqrt{(x + (1 - x)\sigma^2 h)(1 - x)\sigma^2 h}.
\end{aligned}
\tag{4.1.9}
$$

In Lemma 4.1, we will check that after this replacement, $x_{1,2}$ and $p_{1,2}$ still satisfy (4.1.1)–(4.1.5). Unfortunately, for the values of x near 1, the values of x_2 are slightly greater than 1 (as well as those defined by (4.1.8)), 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, $S_t^x \stackrel{d}{=} 1 - S_t^{1-x}$ for all $x \in [0, 1]$ ($\stackrel{d}{=}$ means equality in distribution). Therefore, in the interval $[0, 1/2]$, we can use the values $x_{1,2}$ defined by (4.1.9), whereas in the interval $(1/2, 1]$, we use the values corresponding to the process $1 - S_t^{1-x}$, that is,

$$
\begin{aligned}
\tilde{x}_{1,2} = \tilde{x}_{1,2}(x, h) &:= 1 - x_{1,2}(1 - x, h) \\
&= x - x\sigma^2 h \pm \sqrt{(1 - x + x\sigma^2 h)x\sigma^2 h}
\end{aligned}
\tag{4.1.10}
$$

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

$$\hat{x}_{1,2} := \begin{cases} x_{1,2}(x,h) \text{ with prob. } p_{1,2} = \frac{x}{2x_{1,2}(x,h)}, & x \in [0,1/2], \\ 1 - x_{1,2}(1-x,h) \text{ with prob. } p_{1,2} = \frac{1-x}{2x_{1,2}(1-x,h)}, & x \in (1/2,1]. \end{cases} \tag{4.1.11}$$

Lemma 4.1. *The values $\hat{x}_{1,2}$ defined by* (4.1.11) *satisfy conditions* (4.1.1)–(4.1.5) *for $h \leqslant h_0 := \frac{1}{4\sigma^2}$; moreover, $\hat{x}_{1,2} \in [0,1]$.*

Proof. We first check that $x_{1,2}$ defined by (4.1.9) obtain values from the interval [0, 1] for $x \in [0,1/2]$:

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

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

$$x\sigma^2 h + 1 - x - 2\sigma^2 h \geq 1/2 - 2\sigma^2 h \geq 0 \text{ for } 0 < h \leq \frac{1}{4\sigma^2}. \tag{4.1.12}$$

Thus $0 \leq x_1 < x_2 \leq 1$ for $x \in [0,1/2]$ and $0 < h \leq h_0$. So, if $x \in (1/2,1]$, then $1-x \in [0,1/2)$, and according to (4.1.10), 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 (4.1.1)–(4.1.5) for $x_{1,2}$:

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

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 (1.1.1) we obtain an approximation $\hat{X}_h^x$ by the "split-step" procedure defined by (3.2.4):

$$\hat{X}_h^x := \hat{S}_h^x e^{-bh} + \tfrac{a}{b}(1 - e^{-bh}). \tag{4.1.13}$$

Now we can state the main result for the first-order weak approximation of the WF process.

Theorem 4.2. *Let $\hat{X}_t^x$ be the discretization scheme defined by one-step approximation* (4.1.13). *Then, $\hat{X}_t^x$ is a first-order weak approximation of Equation* (1.1.1) *for functions $f \in C^\infty[0,1]$.*

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:

Algorithm 4: Algorithm for the first-order scheme of the WF.

Draw a uniform random variable U from the interval [0, 1].
if $x \leqslant \frac{1}{2}$ **then**
  calculate x_1, x_2 according to (4.1.9),
else
  calculate x_1, x_2 according to (4.1.9) with $x := 1 - x$,
  $x_{1,2} := 1 - x_{1,2}$.
Calculate $p_{1,2} := \frac{x}{2x_{1,2}(x,h)}$.
if $U < p_1$ **then**
  $\hat{S} := x_1$,
else
  $\hat{S} := x_2$.
Calculate (see (3.2.4) and (4.1.13))

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

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 4.1–4.3, we compare $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t (top plots) and as functions of a discretization step h (bottom 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 4.3 even for rather large discretization step h.

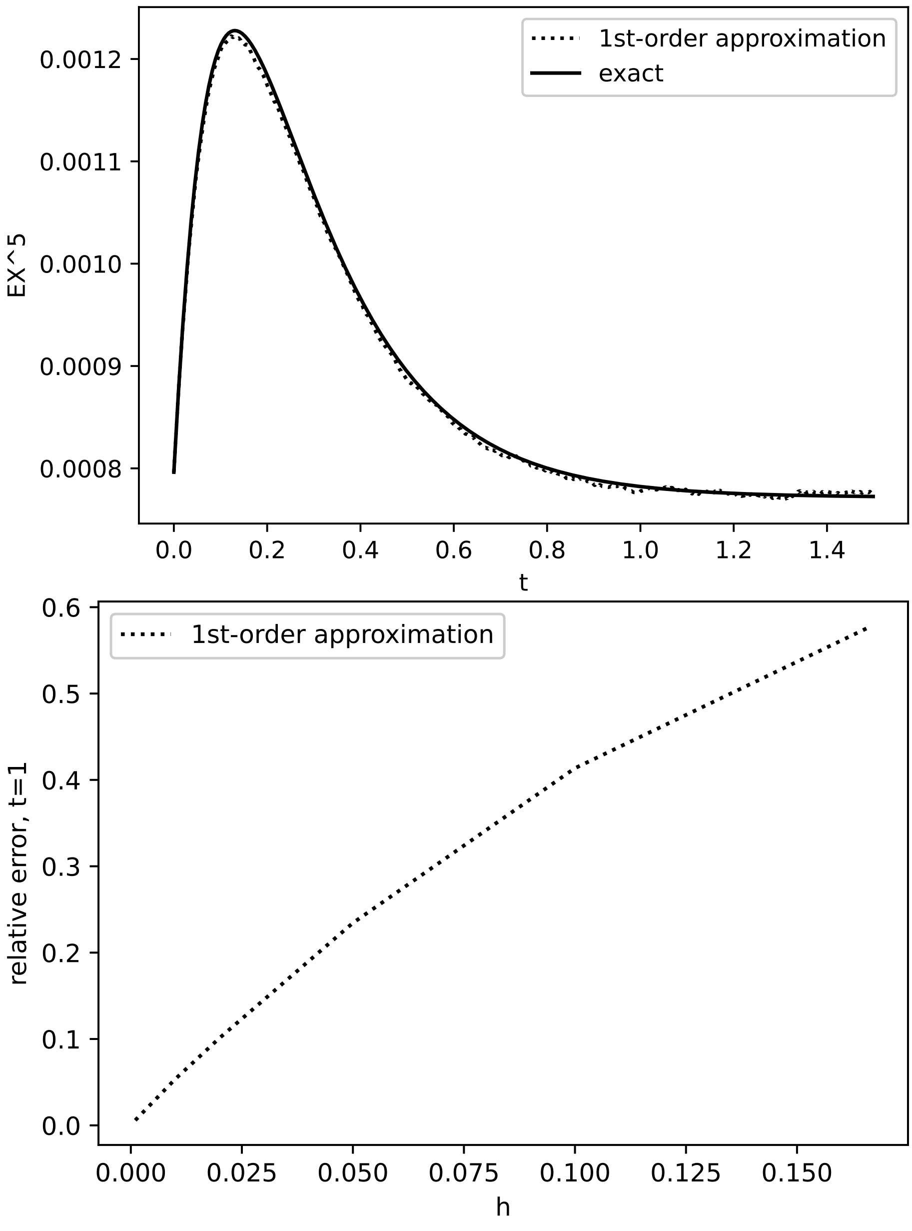

Figure 4.1: Comparison of $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t and h for $f(x) = x^5$: $x = 0.24$, $\sigma^2 = 0.6$, $a = 0.8$, $b = 5$, the number of iterations $N = 1{,}000{,}000$. Top: $h = 0.001$; bottom: the relative error at $t = 1$.

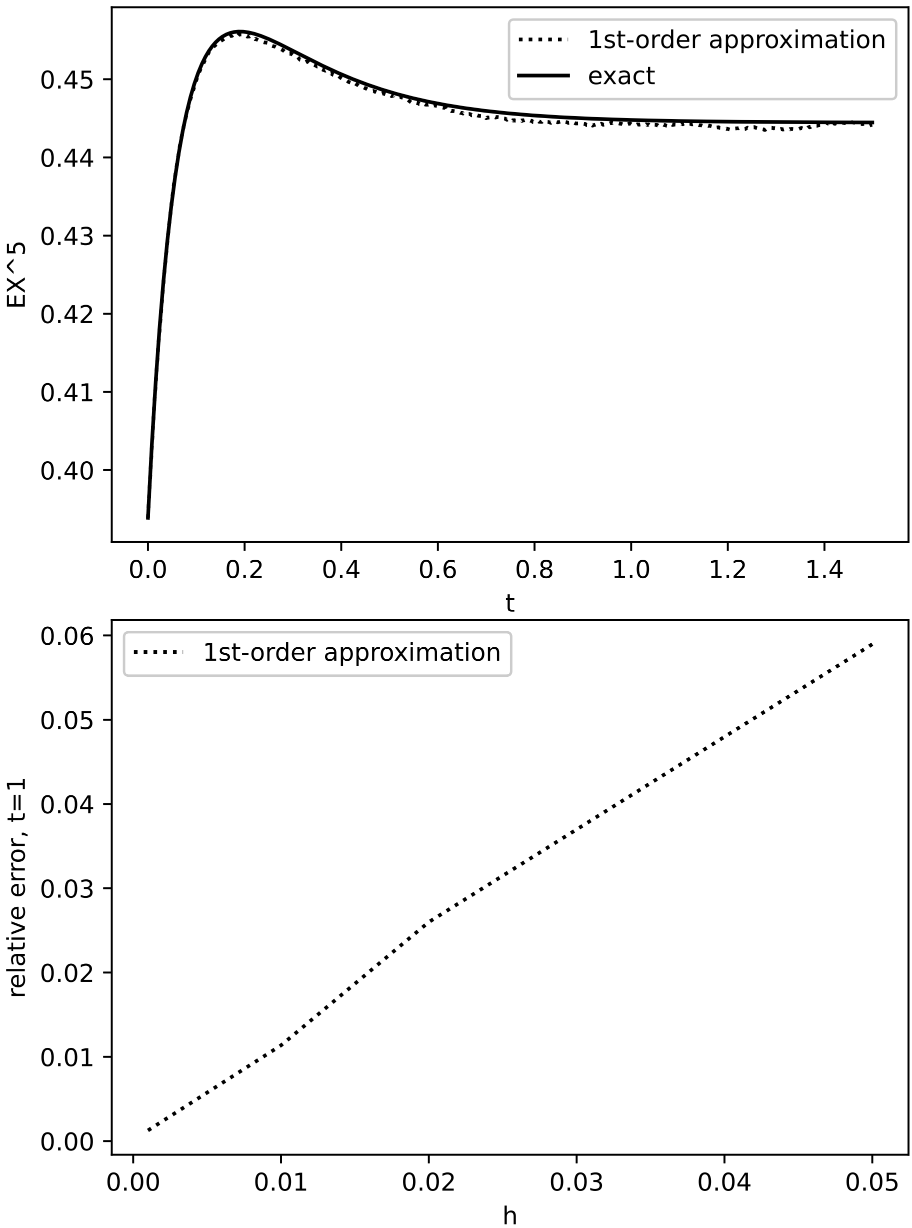

Figure 4.2: Comparison of $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t and h for $f(x) = x^5$: $x = 0.83$, $\sigma^2 = 2$, $a = 4$, $b = 5$, $N = 1{,}000{,}000$. Top: $h = 0.001$; bottom: the relative error at $t = 1$.

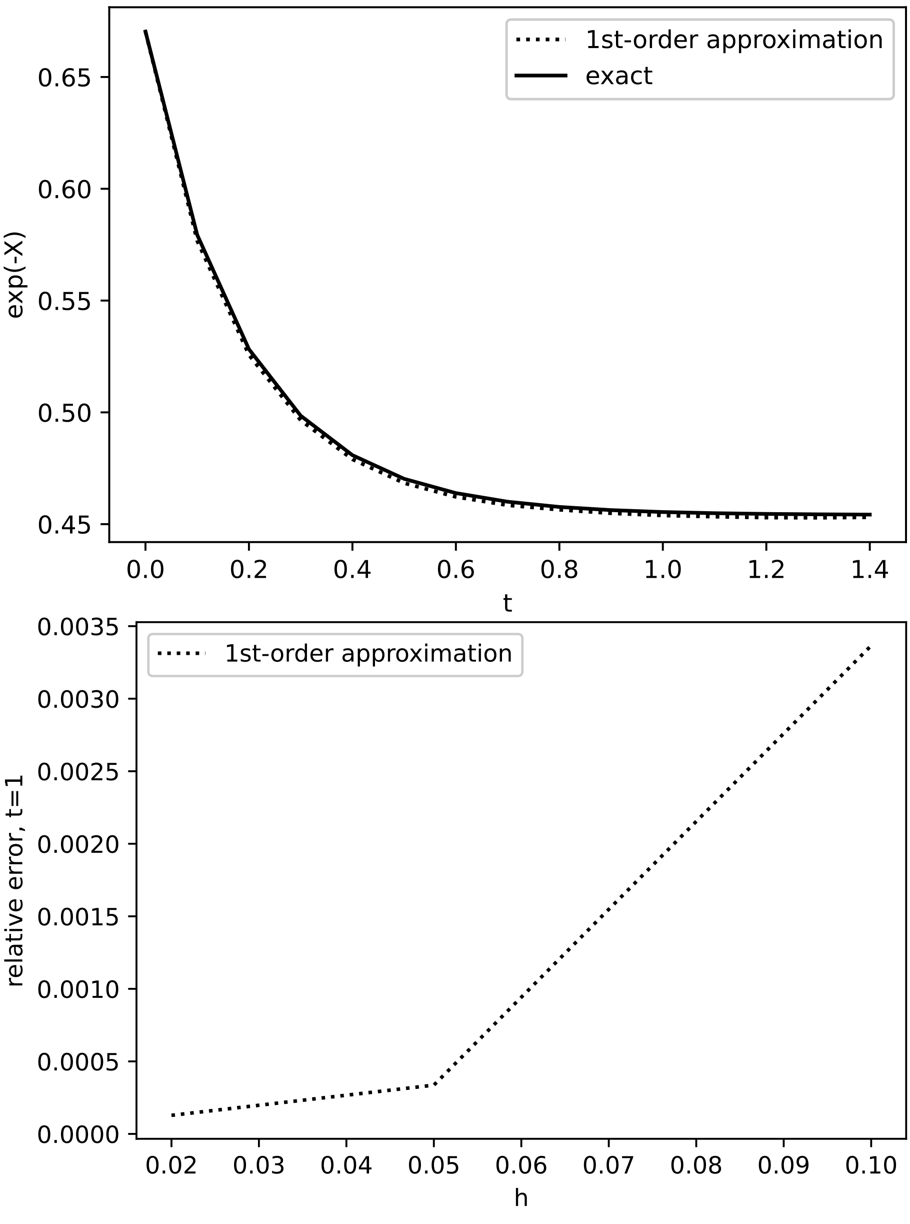

Figure 4.3: Comparison of $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t and h for $f(x) = \mathrm{e}^{-x}$: $x = 0.4,\ \sigma^2 = 1.6,\ a = 4,\ b = 5,\ N = 100{,}000$. Top: $h = 0.1$; bottom: the relative error at $t = 1$.

4.2 A second-order approximation

Approximation of the stochastic part

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:

$$p_1 + p_2 + p_3 = 1, \tag{4.2.1}$$

$$x_1^i p_1 + x_2^i p_2 + x_3^i p_3 = \hat{m}_i + O(h^3), \quad i = 1, \ldots, 6. \tag{4.2.2}$$

In anticipation, note that when solving system (4.2.1)–(4.2.2), 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

$$p_1 + p_2 + p_3 \le 1 \tag{4.2.3}$$

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 (4.2.2) 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. [30])

$$\begin{aligned}
p_1 := p_1(x,h) &= \frac{\hat{m}_1 x_2 x_3 - \hat{m}_2 x_2 - \hat{m}_2 x_3 + \hat{m}_3}{x_1(x_1 - x_3)(x_1 - x_2)}, \\
p_2 := p_2(x,h) &= -\frac{\hat{m}_1 x_1 x_3 - \hat{m}_2 x_1 - \hat{m}_2 x_3 + \hat{m}_3}{x_2(x_1 - x_2)(x_2 - x_3)}, \\
p_3 := p_3(x,h) &= \frac{\hat{m}_1 x_1 x_2 - \hat{m}_2 x_1 - \hat{m}_2 x_2 + \hat{m}_3}{x_3(x_2 - x_3)(x_1 - x_3)}.
\end{aligned} \tag{4.2.4}$$

Note that, here, differently from [30], we used the approximate "moments" $\hat{m}_i$ instead of the true moments $m_i = \mathbb{E}(S_h^x)^i$. 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 (4.2.4), satisfy the remaining conditions

$$\begin{cases} x_1^4 p_1 + x_2^4 p_2 + x_3^4 p_3 - \hat{m}_4 = O(h^3), \\ x_1^5 p_1 + x_2^5 p_2 + x_3^5 p_3 - \hat{m}_5 = O(h^3), \\ x_1^6 p_1 + x_2^6 p_2 + x_3^6 p_3 - \hat{m}_6 = O(h^3). \end{cases} \tag{4.2.5}$$

Motivated by the first-order approximation (4.1.11) and [30], we look for $x_{1,2,3}$ of the following form:

$$x_1 = x + zA_1(1-x) + A_2xz - \sqrt{(z(1-x)(Bx + Cz(1-x)))}, \tag{4.2.6}$$

$$x_2 = x + A_3xz, \tag{4.2.7}$$

$$x_3 = x + zA_1(1-x) + A_2xz + \sqrt{(z(1-x)(Bx + Cz(1-x)))}, \tag{4.2.8}$$

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

Calculation of the parameters

Substituting (4.2.6)–(4.2.8) into the left-hand sides of (4.2.5), we have (for technical calculations, using Maple and Python)

$$\begin{aligned} x_1^4 p_1 + x_2^4 p_2 + x_3^4 p_3 - \hat{m}_4 = \Big[&(BA_3 + B + 2A_1 - 2A_2 - A_3 - 6)x^4 \\ &+ \big(A_3 - 2B - A_3B - 4A_1 + 2A_2 + \tfrac{21}{2}\big)x^3 \\ &+ \big(B + 2A_1 - \tfrac{9}{2}\big)x^2\Big]z^2 + O(h^3), \end{aligned} \tag{4.2.9}$$

$$\begin{aligned} x_1^5 p_1 + x_2^5 p_2 + x_3^5 p_3 - \hat{m}_5 = \Big[&(8A_1 + (4B-4)A_3 + 5B - 8A_2 - 27)x^5 \\ &+ \big((-4B+4)A_3 - 10B \\ &\quad + 8A_2 - 16A_1 + 48\big)x^4 \\ &+ (8A_1 + 5B - 21)x^3\Big]z^2 + O(h^3), \end{aligned} \tag{4.2.10}$$

$$\begin{aligned} x_1^6 p_1 + x_2^6 p_2 + x_3^6 p_3 - \hat{m}_6 = \Big[&(20A_1 + (10B-10)A_3 + 15B \\ &\quad - 20A_2 - 75)x^6 \\ &+ \big((-10B+10)A_3 - 30B + 20A_2 \\ &\quad - 40A_1 + 135\big)x^5 \\ &+ (15B + 20A_1 - 60)x^4\Big]z^2 + O(h^3). \end{aligned} \tag{4.2.11}$$

To ensure equalities (4.2.5), 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 (4.2.9)–(4.2.11), 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 (4.2.6)–(4.2.8) and $p_{1,2,3}$ defined by (4.2.4) satisfy all of Equation (4.2.2), provided that the parameters satisfy the following relations:

$$A_1 = \frac{3}{4},\ A_2 \geq 0,\ A_3 = A_2 + \frac{3}{4},\ B = 3,\ C \geq 0. \tag{4.2.12}$$

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 $p_1 + p_2 + p_3 \leqslant 1$. We first consider the latter restriction.

Lemma 4.3. *We have* $p_1 + p_2 + p_3 \leqslant 1$ *if*

$$A_2 \geqslant \frac{(3+2\sqrt{2})^{\frac{1}{3}}}{4} + \frac{1}{4(3+2\sqrt{2})^{\frac{1}{3}}} \approx 0.58883 \tag{4.2.13}$$

and

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

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

$$\begin{aligned}
D = {} & 64x^2 + \big((192A_2 + 144)x^2 - 96x\big)z \\
& + \big((192A_2^2 - 64C + 96A_2 + 108)x^2 + (128C - 144)x - 64C + 36\big)z^2 \\
& + \big((64A_2^3 - 64CA_2 - 48A_2^2 - 48C - 36A_2 + 27\big)x^2 \\
& \quad + \big(128CA_2 + 96A_2^2 + 96C - 54)x \\
& \qquad - 64A_2C - 48C + 36A_2 + 27\big)z^3.
\end{aligned}$$

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,

$$\begin{aligned}
& (64A_2^3 - 48A_2^2 + (-64C + 60)A_2 - 48C + 3)x^2 \\
& \quad + (96A_2^2 + (128C - 96)A_2 + 96C + 18)x \\
& \quad + (-64C + 36)A_2 - 48C - 21 =: a_1x^2 + a_2x + a_3 \geqslant 0.
\end{aligned} \tag{4.2.14}$$

Inequality (4.2.14) is satisfied for all $x \in \mathbb{R}$ if

$$a_1 > 0 \text{ and } a_3 \geqslant \frac{a_2^2}{4a_1}. \tag{4.2.15}$$

Solving inequality (4.2.15), we get

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

Since C must be nonnegative, we obtain condition (4.2.13) for A_2. $\square$

Remark 4.4. We observe that possible values of C are rather small (see Figure 4.4). Therefore, to simplify the expressions for x_1, x_2, x_3, we simply take $C = 0$ and $A_2 = \frac{2}{3}$.

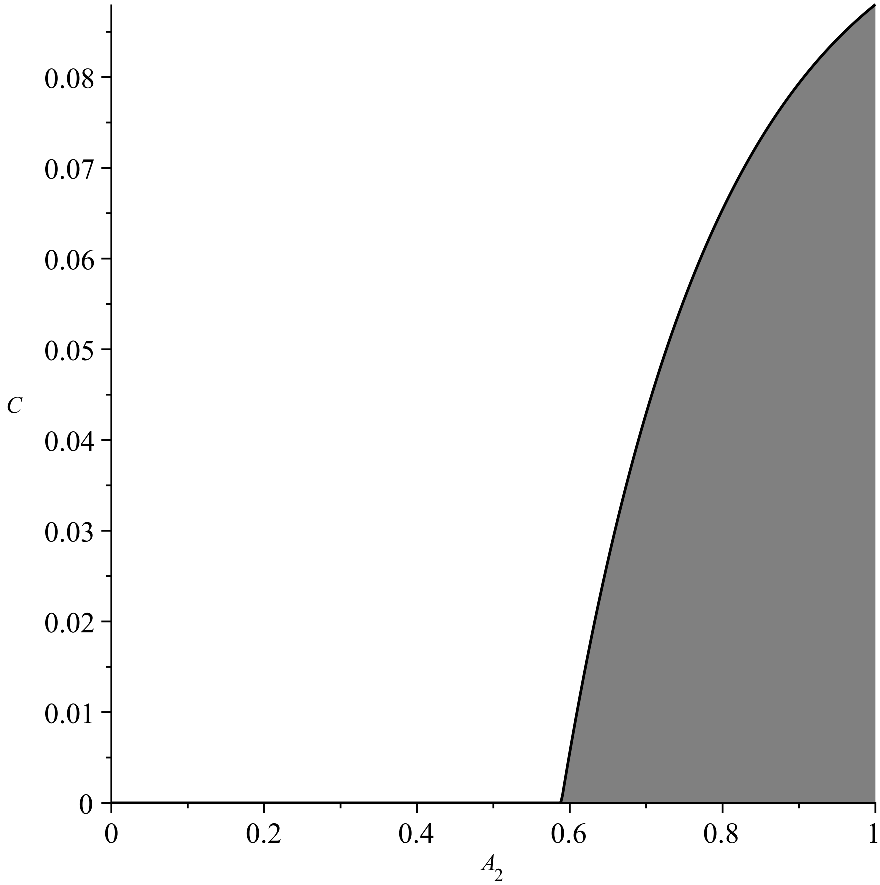

Figure 4.4: Possible values of A_2 and C.

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

$$x_1 = x + \frac{3(1-x)z}{4} + \frac{2xz}{3} - \sqrt{3x(1-x)z}, \tag{4.2.16}$$

$$x_2 = x + \frac{17xz}{12}, \tag{4.2.17}$$

$$x_3 = x + \frac{3(1-x)z}{4} + \frac{2xz}{3} + \sqrt{3x(1-x)z}. \tag{4.2.18}$$

However, at this point, we only have that x_1, x_2, x_3 defined by (4.2.16)–(4.2.18), together with p_1, p_2, p_3 defined by (4.2.4), satisfy conditions (4.2.2) and (4.2.3). 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 4.5 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 $\frac{1}{2}$.

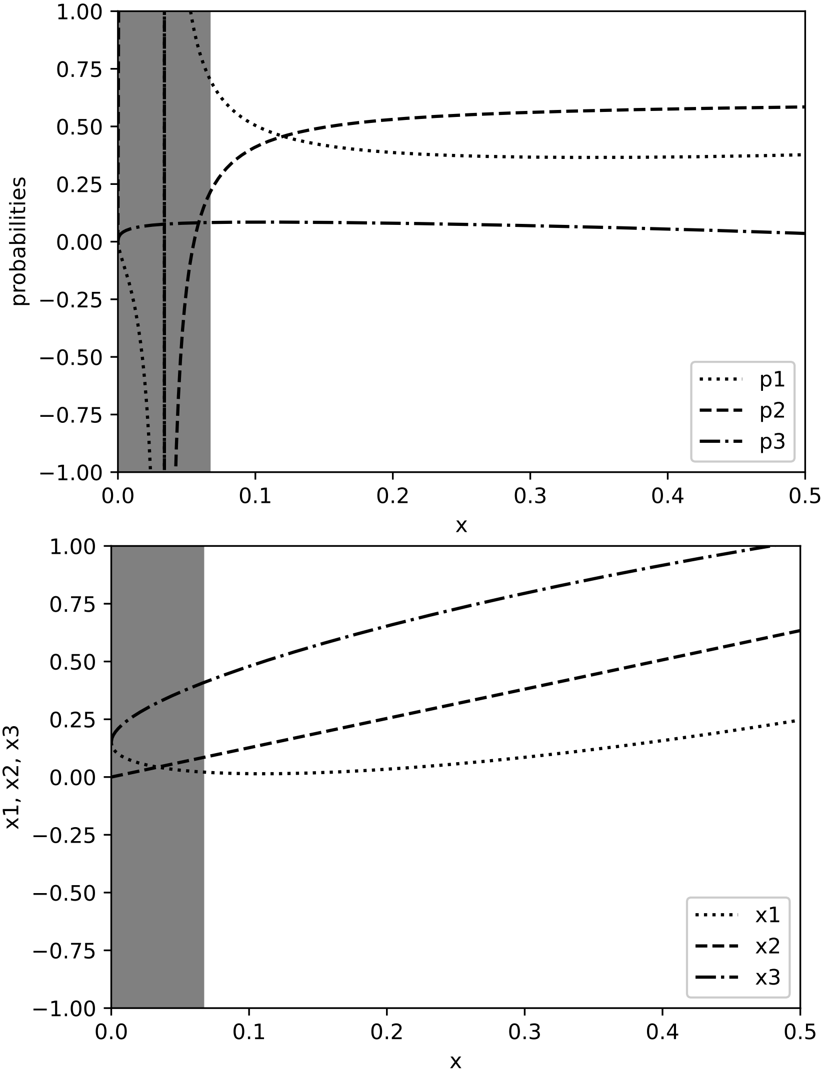

Figure 4.5: Graphs of $p_{1,2,3}$ (top) and $x_{1,2,3}$ (bottom) as functions of x with fixed z. Gray area shows the region where first-order approximation is used to avoid negative probabilities. Parameters: $K = \frac{1}{3}, z = \frac{1}{5}$.

Due to these reasons, similarly to [3] and [33], for small x below the threshold Kz (with some fixed $K > 0$), we will switch to the first-order approximation (4.1.11), 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 (4.1.11), whereas for $Kz \le x \le \frac{1}{2}$, $0 < z \le z_0$, we will use scheme (4.2.16)–(4.2.18) together with (4.2.4); 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 4.5. *The first-order approximation* (4.1.11) *in the region* $x \leqslant K\sigma^2 h$ *(with arbitrary fixed* $K > 0$*) satisfies conditions* (3.3.13). *In other words, in this region, it behaves as a second-order approximation.*

Proof. We prove equalities (3.3.13) in the region $x \leqslant Kz = K\sigma^2 h$, where $\hat{S}_x^h$ and $\hat{m}_i$, $i = 1, \dots, 6$, are defined by (4.1.11) and (3.3.14), respectively:

$$
\begin{aligned}
\mathbb{E}(\hat{S}_h^x)^2 - \hat{m}_2 &= x_1^2 p_1 + x_2^2 p_2 - \hat{m}_2 \\
&= x^2 + x(1-x)z - (x^2 + zx(1-x)(1 - \tfrac{1}{2}z)) \\
&= \tfrac{1}{2}x(1-x)z^2 = O(h^3), \\
\mathbb{E}(\hat{S}_h^x)^3 - \hat{m}_3 &= x_1^3 p_1 + x_2^3 p_2 - \hat{m}_3 \\
&= x^3 - 3xz(x^2 - x) + 2z^2x^2(x-1) \\
&\qquad - (x^3 + \tfrac{3}{2}xz^2(3x^2 - 4x + 1) - 3xz(x^2 - x)) \\
&= \tfrac{1}{2}xz^2(5x^2 - 8x + 3) = O(h^3), \\
\mathbb{E}(\hat{S}_h^x)^4 - \hat{m}_4 &= x_1^4 p_1 + x_2^4 p_2 - \hat{m}_4 = x^4 - 6x^2z(x^2 - x) \\
&\qquad + x^2z^2(9x^2 - 10x + 1) - 4x^3z^3(x-1) \\
&\qquad - (x^4 + 9x^2z^2(2x^2 - 3x + 1) - 6x^2z(x^2 - x)) \\
&= x^2z^2(-4x^2z - 9x^2 + 4xz + 17x - 8) = O(h^4), \\
\mathbb{E}(\hat{S}_h^x)^5 - \hat{m}_5 &= x_1^5 p_1 + x_2^5 p_2 - \hat{m}_5 = x^5 - 10x^3z(x^2 - x) \\
&\qquad + 5x^3z^2(5x^2 - 6x + 1) + 4x^3z^3(-6x^2 + 7x - 1) \\
&\qquad + 8x^4z^4(x-1) \\
&\qquad - (x^5 + 10x^3z^2(5x^2 - 8x + 3) - 10x^3z(x^2 - x)) \\
&= x^3z^2\big(x^2(8z^2 - 24z - 25)
\end{aligned}
$$

$$
\begin{aligned}
&-2x(4z^2-14z-25)-4z-25\big)=O(h^5),\\
\mathbb{E}(S_h^x)^6-\hat m_6 &= x_1^6p_1+x_2^6p_2-\hat m_6 = x^6-15x^4z(x^2-x)\\
&\quad+5x^4z^2(11x^2-14x+3)\\
&\quad+x^3z^3(-85x^3+111x^1-27x+1)\\
&\quad+12x^4z^4(5x^2-6x+1)-16x^5z^5(x-1)\\
&\quad-\big(x^6+\tfrac{75}{2}x^4z^2(3x^2-5x+2)-15x^4z(x^2-x)\big)\\
&=\tfrac12 x^3z^2(x-1)(-32x^2z^3+120x^2z^2-170x^2\ z\\
&\quad-115x^2-24xz^2-170xz+120x-2z)=O(h^6).\square
\end{aligned}
$$

Lemma 4.6. *For $z\in[0,\frac16]$ and $x\in[0,\frac12]$, x_1,x_2,x_3 defined in* (4.2.16)–(4.2.18) *take values in the interval* $[0,1]$.

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

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

$$\Big(x+\frac{3(1-x)z}{4}+\frac{2xz}{3}\Big)^2-3x(1-x)z\geq 0.$$

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

$$
\begin{aligned}
&\Big(x+\frac{3yz}{4}+\frac{2xz}{3}\Big)^2-3xyz\\
&\qquad=x^2+\frac{9}{16}y^2z^2+\frac49x^2z^2-\frac32xyz+\frac43x^2z+xyz^2\geq 0.
\end{aligned}
$$

We will prove the stronger inequality

$$x^2+\frac{9}{16}y^2z^2+\frac43x^2z-\frac32xyz\geqslant 0,$$

which after substitution $y=1-x$ becomes

$$\Big(1+\frac{17}{6}z+\frac{9}{16}z^2\Big)x^2-\Big(\frac32z+\frac98z^2\Big)x+\frac{9z^2}{16}\geqslant 0. \tag{4.2.19}$$

The discriminant of the quadratic polynomial (4.2.19) in x is

$$D=\Big(\frac32z+\frac98z^2\Big)^2-4\Big(1+\frac{17}{6}z+\frac{9}{16}z^2\Big)\cdot\frac{9z^2}{16}=-3z^3,$$

which is negative for all $z > 0$. This means that the left-hand side (4.2.19) is positive and thus $x_1 > 0$ for all $x \in [0,1]$ and $z \geq 0$ except for $x = z = 0$, where $x_1 = 0$.

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

$$\begin{aligned} x_3 &= x + \frac{3(1-x)z}{4} + \frac{2xz}{3} + \sqrt{3x(1-x)z} \\ &\leq x + \frac{1}{8}(1-x) + \frac{1}{18} + \sqrt{\tfrac{x(1-x)}{2}} \\ &\leq \frac{1}{8} + \frac{7}{8} \cdot \frac{1}{2} + \frac{1}{18} + \frac{1}{\sqrt{8}} \approx 0.972 < 1. \end{aligned}$$ $\square$

Lemma 4.7. *For $x \in (\frac{z}{3}, \frac{1}{2}]$ and $z \leqslant \frac{1}{6}$, we have $p_1, p_2, p_3 \in [0,1]$.*

Proof. From Lemma 4.3, we already have that $p_1 + p_2 + p_3 \leqslant 1$. Therefore, it suffices to prove that $p_1, p_2, p_3 \geqslant 0$. 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 Figures 4.6–4.8, where the 3D graphs of p_1, p_2, p_3 as functions of (x, z) are plotted in the domain $\{(x,z) : z/3 \leq x \leq 1/2,\ 0 \leq z \leq 1/6\}$. $\square$

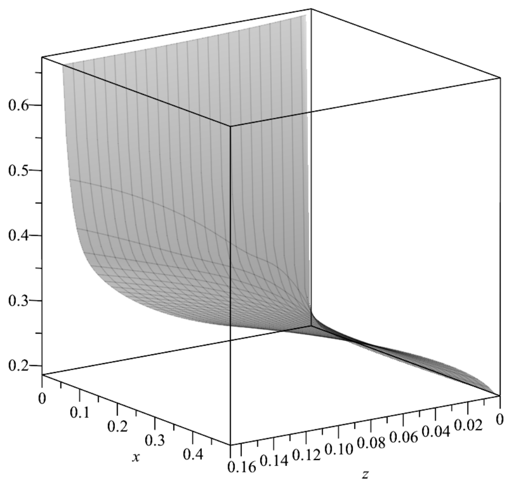

Figure 4.6: Graph of p_1 as a function of x and z.

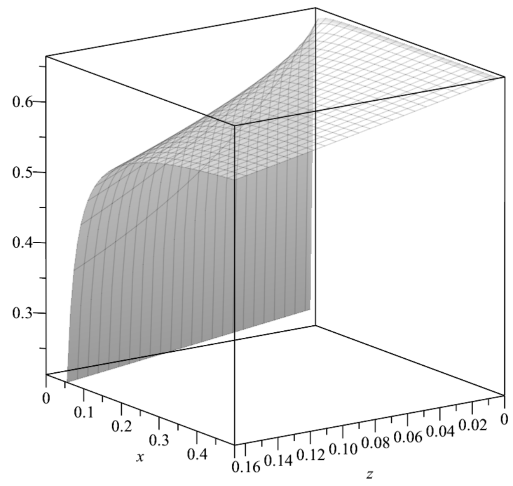

Figure 4.7: Graph of p_2 as a function of x and z.

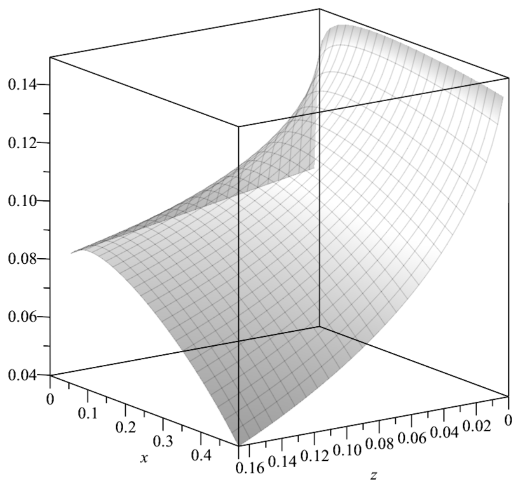

Figure 4.8: Graphs of p_3 as a function of x and z.

The second main result

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

$$x_1 = x_1(x,h) = x + \frac{3(1-x)\sigma^2 h}{4} + \frac{2x\sigma^2 h}{3} - \sqrt{3x(1-x)\sigma^2 h}, \quad (4.2.20)$$

$$x_2 = x_2(x,h) = x + \frac{17x\sigma^2 h}{12}, \quad (4.2.21)$$

$$x_3 = x_3(x,h) = x + \frac{3(1-x)\sigma^2 h}{4} + \frac{2x\sigma^2 h}{3} + \sqrt{3x(1-x)\sigma^2 h}. \quad (4.2.22)$$

To distinguish the functions $x_{1,2,3}$ from $x_{1,2}$ given by (4.1.9), here we denote the latter by

$$y_{1,2} = y_{1,2}(x,h) = x + (1-x)\sigma^2 h \mp \sqrt{(x+(1-x)\sigma^2 h)(1-x)\sigma^2 h}. \quad (4.2.23)$$

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:

$$\hat{S}_h^x := \begin{cases} x_{1,2,3}(x,h) \text{ with probabilities } p_{1,2,3} \text{ (4.2.4) and} \\ \quad 0 \text{ with probability } p_0 = 1-(p_1+p_2+p_3), \\ \hfill x \in (\frac{\sigma^2 h}{3}, \frac{1}{2}], \\ 1 - x_{1,2,3}(1-x,h) \text{ with prob. } p_{1,2,3}(1-x,h) \text{ and} \\ \quad 1 \text{ with probability } p_0 = 1-(p_1+p_2+p_3), \\ \hfill x \in (\frac{1}{2}, 1-\frac{\sigma^2 h}{3}), \\ y_{1,2}(x,h) \text{ with probabilities } \tilde{p}_{1,2}(x,h) := \frac{x}{2y_{1,2}(x,h)}, \\ \hfill x \in [0, \frac{\sigma^2 h}{3}], \\ 1 - y_{1,2}(1-x,h) \text{ with probabilities } \tilde{p}_{1,2}(1-x,h), \\ \hfill x \in [1-\frac{\sigma^2 h}{3}, 1]. \end{cases} \quad (4.2.24)$$

Now in the view of Theorem C and Lemmas 4.3–4.7, we can state the main result on the second-order approximation of the WF process.

Theorem 4.8. *Let $\hat{X}_t^x$ be the discretization scheme defined by one-step approximation*

$$\hat{X}_h^x = D(\hat{S}(D(x,h/2),h),h/2), \quad (4.2.25)$$

where $D(x,h)$ is defined by (3.2.3), *and $\hat{S}(x,h) = \hat{S}_h^x$ is defined by* (4.2.24). *Then, $\hat{X}_t^x$ is a second-order weak approximation of Equation* (1.1.1).

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:

Algorithm 5: Algorithm for the second-order scheme of the WF.

Draw a uniform random variable U from the interval $[0, 1]$.
$x := D(x, h/2)$ (where D is given by (3.2.3)).
if $x \leqslant \frac{1}{2}$ **then**
    if $x > \frac{\sigma^2 h}{3}$ **then**
        $x_0 := 0$,
        calculate x_1, x_2, x_3 according to (4.2.20)–(4.2.22),
        calculate p_1, p_2, p_3 according to (4.2.4),
        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 (4.2.23),
        $p_{1,2} := \frac{x}{y_{1,2}(x,h)}$,
        if $U < p_1$ **then**
            $\hat{S} := y_1$,
        else
            $\hat{S} := y_2$,
else
    do if-step with $x := 1 - x$, $x_{0,1,2,3} := 1 - x_{0,1,2,3}$,
    $y_{1,2} := 1 - y_{1,2}$,
$\hat{X}_{(i+1)h} := D\big(\hat{S}, h/2\big)$.

Simulation examples

We illustrate our approximation for the test function $f(x) = x^5$. In Figures 4.9–4.13, we compare the moments $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t (top plots, $h = 0.01$ or $h = 0.02$) and as functions of a discretization step h (bottom plots, $t = 1$) 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 ($2 \leqslant \sigma^2 \leqslant 6$) volatility. In comparison with the simulation results for the first-order approximation, 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.

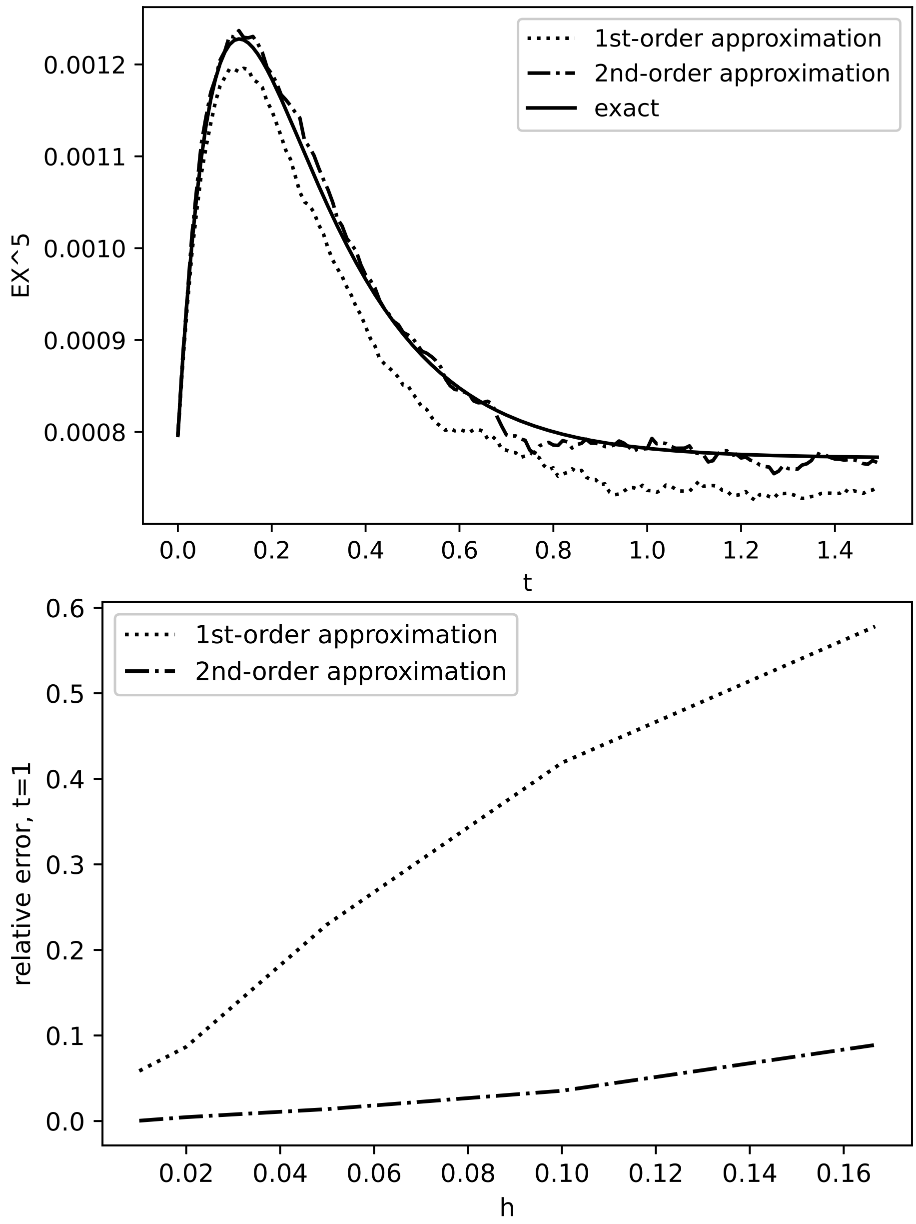

Figure 4.9: Comparison of $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t and h for $f(x) = x^5$: $x = 0.24$, $\sigma^2 = 0.6$, $a = 0.8$, $b = 5$, the number of iterations $N = 100{,}000$. Top: $h = 0.01$; bottom: the relative error at $t = 1$.

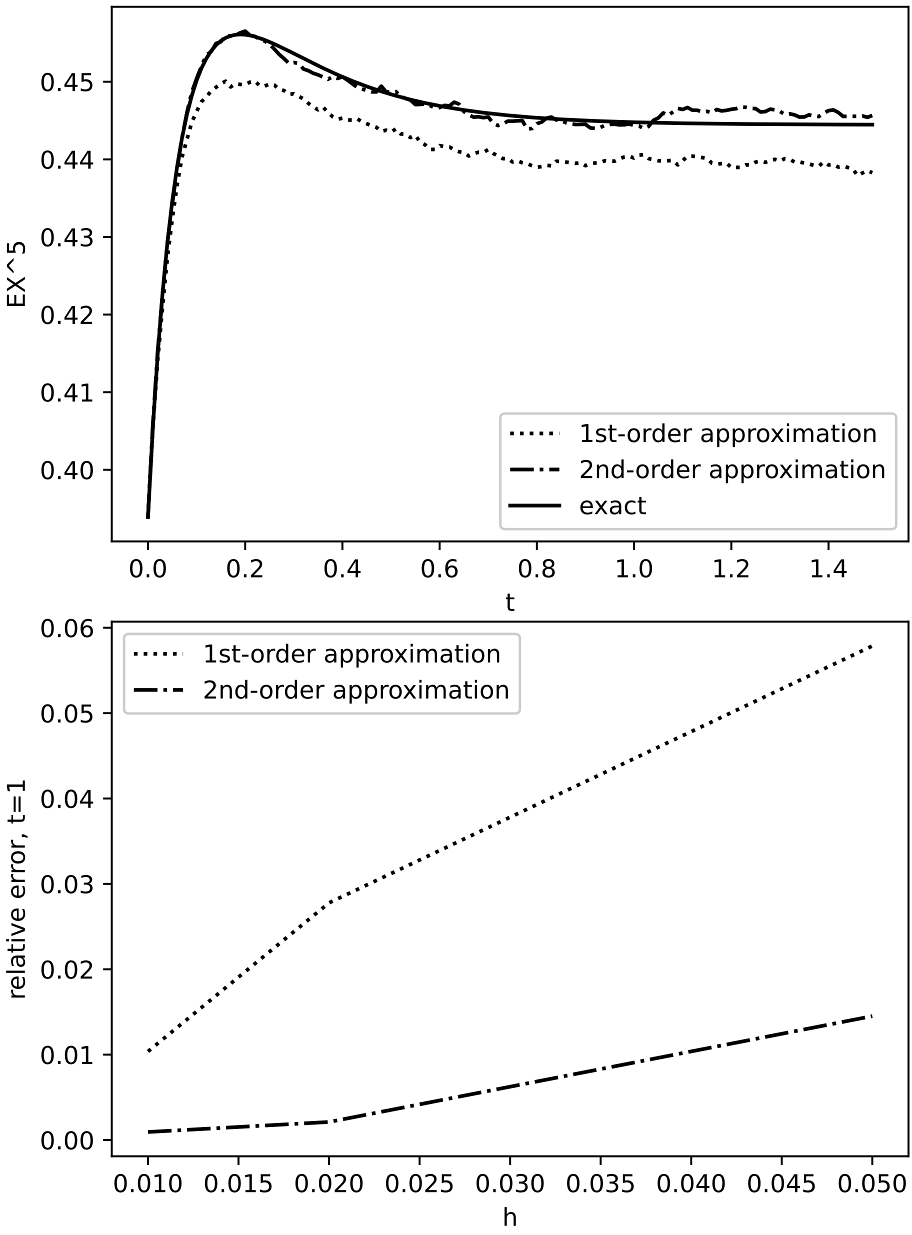

Figure 4.10: Comparison of $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t and h for $f(x) = x^5$: $x = 0.83$, $\sigma^2 = 2$, $a = 4$, $b = 5$, the number of iterations $N = 100{,}000$. Top: $h = 0.01$; bottom: the relative error at $t = 1$.

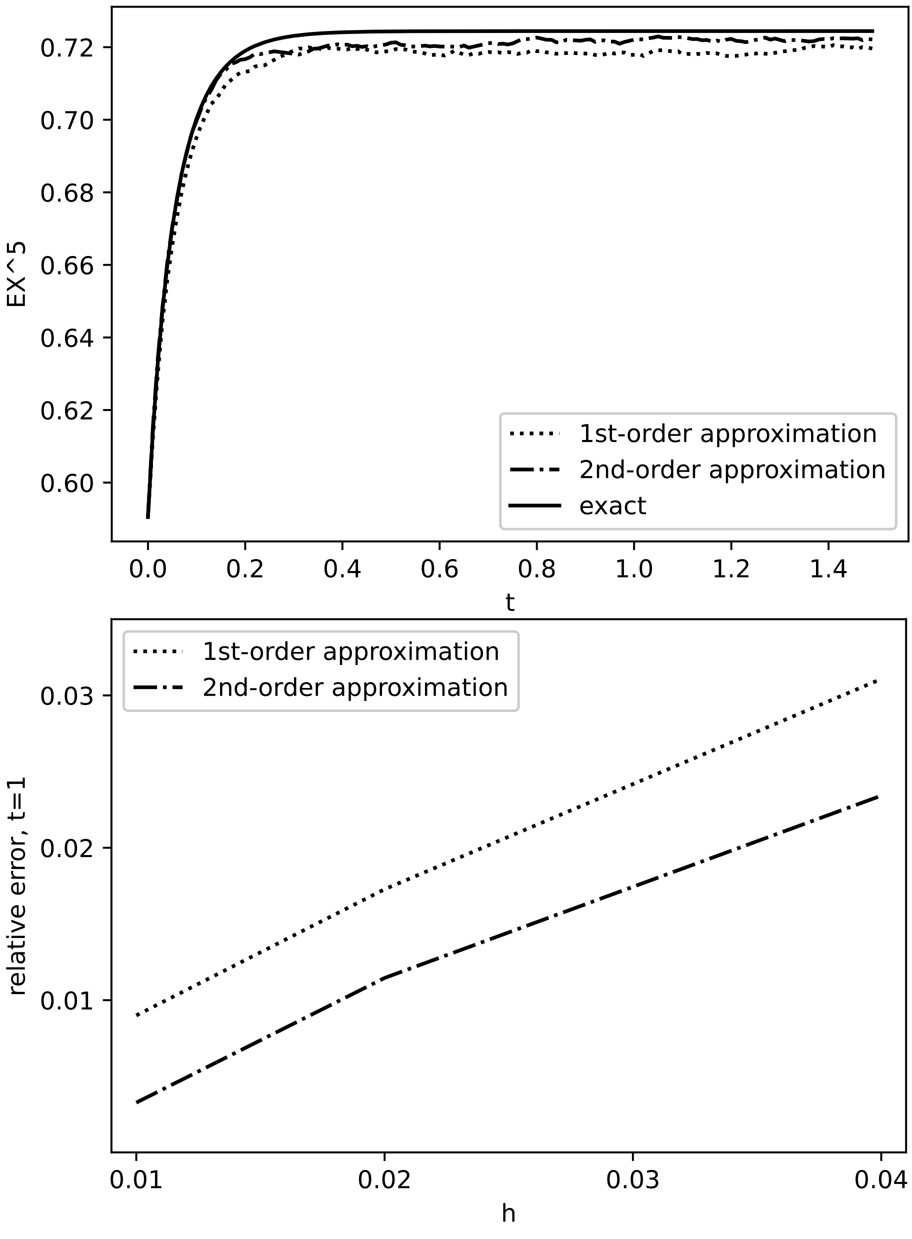

Figure 4.11: Comparison of $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t and h for $f(x) = x^5$: $x = 0.9$, $\sigma^2 = 4$, $a = 4.5$, $b = 5$, the number of iterations $N = 100{,}000$. Top: $h = 0.01$; bottom: the relative error at $t = 1$.

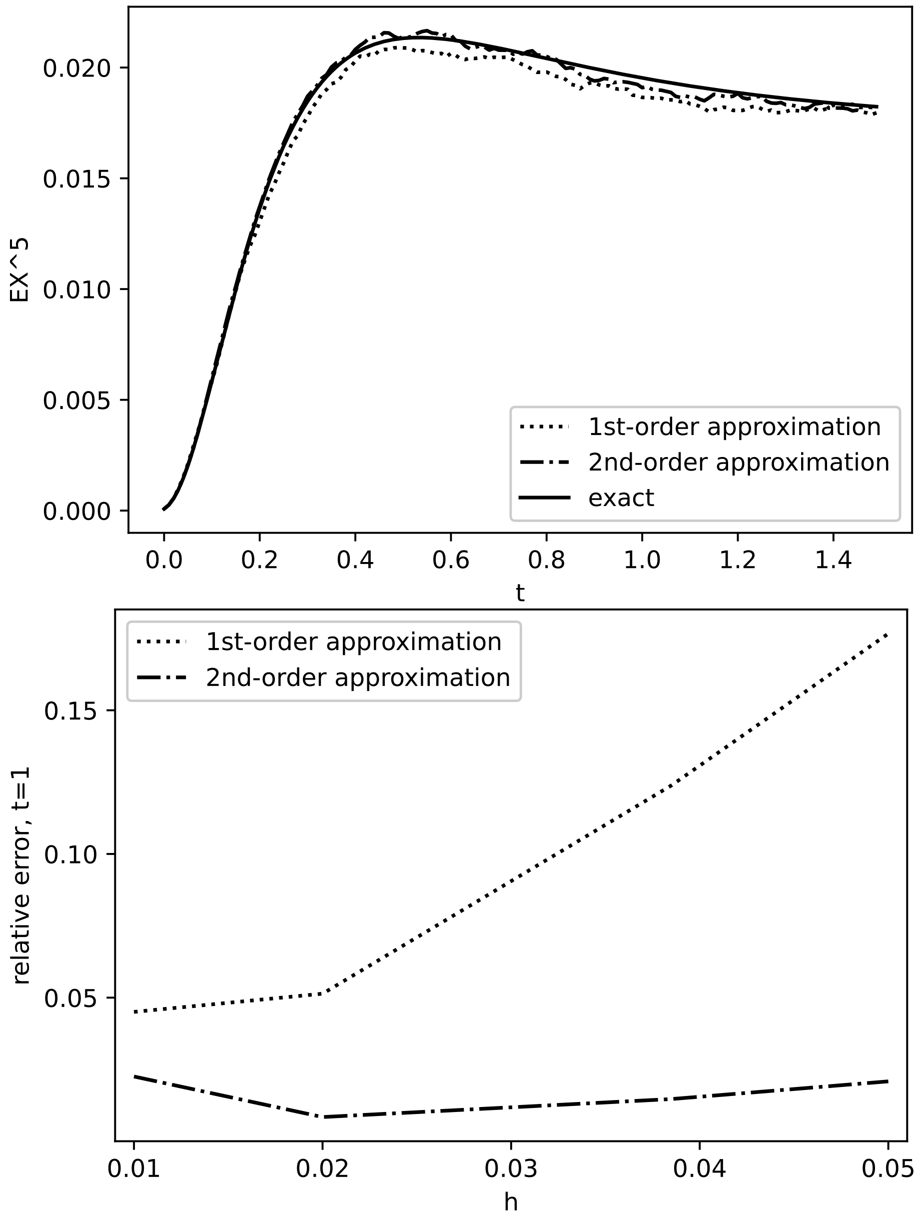

Figure 4.12: Comparison of $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t and h for $f(x) = x^5$: $x = 0.15$, $\sigma^2 = 3$, $a = 0.2$, $b = 2$, the number of iterations $N = 100{,}000$. Top: $h = 0.01$; bottom: the relative error at $t = 1$.

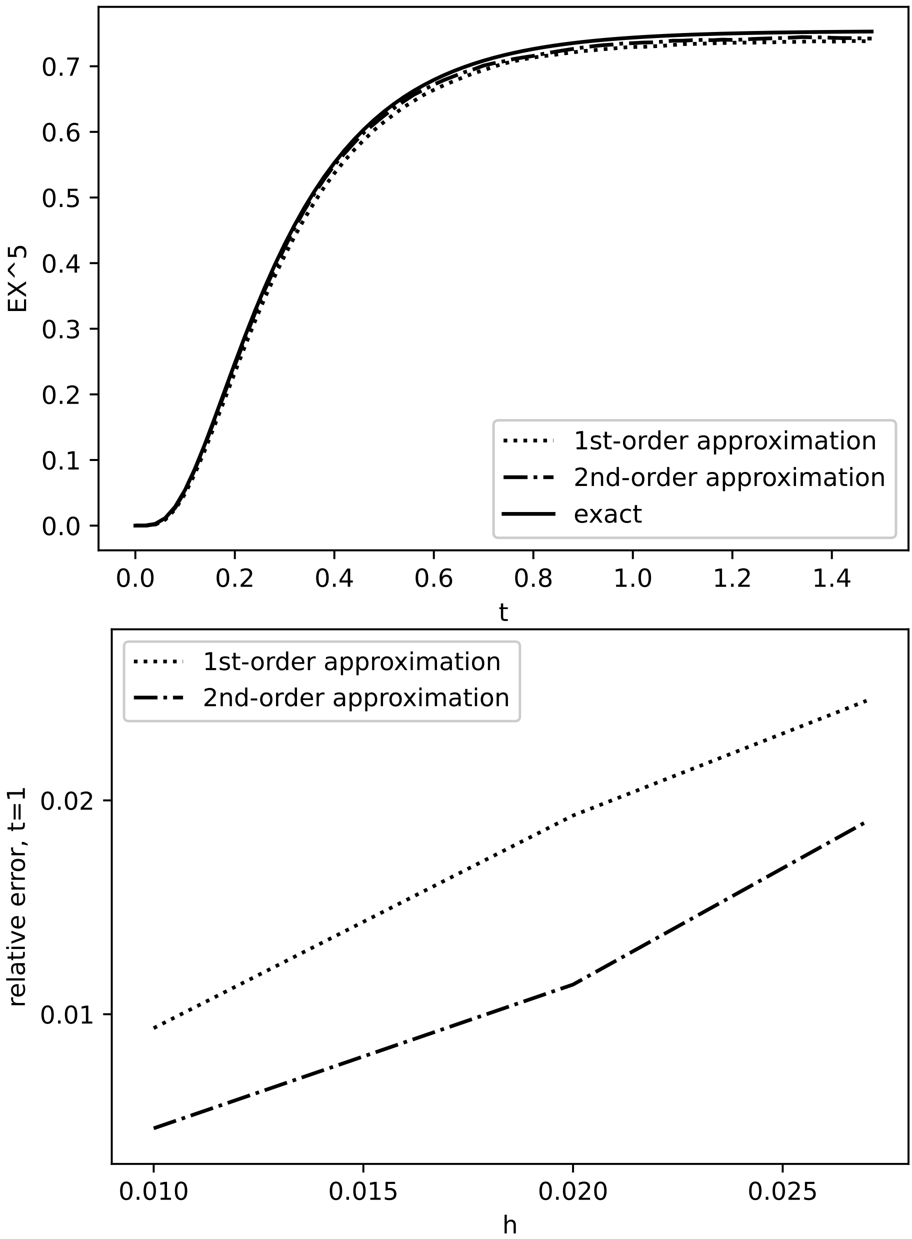

Figure 4.13: Comparison of $\mathbb{E}f(\hat{X}_t^x)$ and $\mathbb{E}f(X_t^x)$ as functions of t and h for $f(x) = x^5$: $x = 0.005$, $\sigma^2 = 6$, $a = 4.5$, $b = 5$, the number of iterations $N = 100{,}000$. Top: $h = 0.02$; bottom: the relative error at $t = 1$.

Chapter 5

Regularity of solutions of Kolmogorov backward equation related to square-root diffusions

The regularity of solutions of backward Kolmogorov equations (as in Theorem B) is needed in proving that a potential ("candidate") weak approximation of the solution to a SDE is, indeed, a true weak approximation of the corresponding order. In this chapter, we prove three variants of Theorem B for SDEs with square-root diffusion coefficient: for WF, CIR, and general Stratonovich diffusions (Theorems 5.2, 5.8, and 5.10, respectively).

5.1 Wright–Fisher equation

Theorem B is in fact Theorem 1.19 of [3] stated based on the results of [13], 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 5.1. 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:

$$f(x) = \sum_{k=0}^\infty c_k x^k, \ x \in [0,1], \tag{5.1.1}$$

where $c_k = f^{(k)}(0)/k!$, $k \in \mathbb{N}_0$. This easily follows from the Lagrange error bound for Taylor series. Indeed, for $f \in C_*^\infty[0,1]$, by the Taylor formula with Lagrange error bound we have

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k + R_n(x), \ x \in [0,1],$$

where

$$\sup_{x\in[0,1]} |R_n(x)| \le \frac{1}{(n+1)!} \sup_{x\in[0,1]} |f^{(n+1)}(x)| \to 0, \ n \to \infty.$$

Remark 5.1. Clearly, every $f \in C_*^\infty[0,1]$ is a real analytic function; see [25].

For convenience, here we restate Theorem B for $f \in C_*^\infty[0,1]$.

Theorem 5.2. *Let* $f \in C_*^\infty[0,1]$. *Then*

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

is a C^∞ *function that solves*

$$\partial_t u(t,x) = Au(t,x). \tag{5.1.2}$$

Proof. Denote $m_k(x,t) := \mathbb{E}(X_t^x)^k$, $k \in \mathbb{N}_0$. Then from (5.1.1) we formally have

$$u(t,x) = \mathbb{E} f(X_t^x) = \sum_{k=0}^\infty c_k m_k(x,t), \ x \in [0,1], \ t \ge 0. \tag{5.1.3}$$

If u is infinitely continuously differentiable, then it satisfies Equation (5.1.2) (see, e.g., [37, Thm. 8.1.1] or [32, Cor. 10.8]). Therefore, to prove Theorem 5.2, 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* (5.1.3),

$$\sum_{k=0}^{\infty} c_k \partial_t^p \partial_x^q m_k(x,t), \tag{5.1.4}$$

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

Statements (1) and (2) are proved below in Lemmas 5.3 and 5.4, respectively. □

Lemma 5.3. *The moments of the WF process* X_t^x *satisfy the following recurrence relation:*

$$m_1(x,t) = \begin{cases} xe^{-bt} + \frac{a}{b}(1 - e^{-bt}), & 0 \le a \le b \neq 0, \\ x, & a = b = 0, \end{cases} \tag{5.1.5}$$

$$m_k(x,t) = e^{-b_k t}\Big(x^k + a_k \int_0^t e^{b_k s} m_{k-1}(x,s)\, ds\Big), \ k \ge 2, \tag{5.1.6}$$

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* $(x,t) \in [0,1] \times \overline{\mathbb{R}}_+$.

Proof. Taking the expectations of both sides of Equation (1.1.1) and then differentiating with respect to t, we get

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

Solving the latter, we get (5.1.5).

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

$$\begin{aligned}(X_t^x)^k &= x^k + k\int_0^t (X_t^x)^{k-1} \mathrm{d}X_s^x + \frac{1}{2}k(k-1)\int_0^t (X_t^x)^{k-2} \mathrm{d}\langle X^x\rangle_s \\ &= x^k + k\int_0^t (X_t^x)^{k-1}(a - bX_s^x)\,\mathrm{d}s\end{aligned}$$

$$+ k\sigma \int_0^t (X_t^x)^{k-1} \sqrt{X_s^x(1-X_s^x)}\,\mathrm{d}B_s$$

$$+ \frac{1}{2}k(k-1)\sigma^2 \int_0^t (X_t^x)^{k-2}\, X_s^x(1-X_s^x)\,\mathrm{d}s$$

$$= x^k + k \int_0^t \left(a(X_t^x)^{k-1} - b(X_s^x)^k\right) \mathrm{d}s$$

$$+ k\sigma \int_0^t (X_t^x)^{k-1} \sqrt{X_s^x(1-X_s^x)}\,\mathrm{d}B_s$$

$$+ \frac{1}{2}k(k-1)\sigma^2 \int_0^t \left((X_t^x)^{k-1} - (X_s^x)^k\right) \mathrm{d}s.$$

By taking the expectations, we get

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

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 (5.1.6). □

Lemma 5.4. *All formal partial derivatives of the series* (5.1.3),

$$\sum_{k=0}^{\infty} c_k \partial_t^p \partial_x^q m_k(x,t), \tag{5.1.7}$$

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

Proof. It is obvious that $0 \le m_k(x,t) \le 1,\ x \in [0,1], k \in \mathbb{N}_0$. 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,

$$\begin{aligned}
\partial_x m_k(x,t) &= e^{-b_k t}\Big(kx^{k-1} + a_k \int\limits_0^t e^{b_k s}\partial_x m_{k-1}(x,s)\,\mathrm{d}s\Big) \\
&\le e^{-b_k t}\Big(k + a_k(k-1)\int\limits_0^t e^{b_k s}\,\mathrm{d}s\Big) \\
&= e^{-b_k t}\Big(k + \frac{a_k}{b_k}(k-1)(e^{b_k t}-1)\Big) \\
&\le e^{-b_k t}k + k(1-e^{-b_k t}) = k,
\end{aligned}$$

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,

$$\begin{aligned}
\partial_x^l m_k(x,t) &= e^{-b_k t}\Big(k(k-1)\dots(k-l+1)x^{k-l} \\
&\qquad + a_k \int\limits_0^t e^{b_k s}\partial_x^l m_{k-1}(x,s)\,\mathrm{d}s\Big) \\
&\leqslant e^{-b_k t}\Big(k(k-1)\dots(k-l+1) \\
&\qquad + \frac{a_k}{b_k}k(k-1)\dots(k-l+1)(e^{b_k t}-1)\Big) \\
&\leqslant k(k-1)\dots(k-l+1) = (k)_l.
\end{aligned}$$

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

$$|\partial_t m_1(x,t)| = \Big|\Big(e^{-bt}(x-\frac{a}{b}) + \frac{a}{b}\Big)_t'\Big| = |-be^{-bt}(x-\frac{a}{b})|$$

$$
\begin{aligned}
&= |(a-bx)e^{-bt}| \le b,\ x\in[0,1];\\
|\partial_t m_k(x,t)| &= \Big| -b_k e^{-b_k t}\Big(x^k + a_k \int\limits_0^t e^{b_k s} m_{k-1}(x,s)\,\mathrm{d}s\Big)\\
&\qquad + e^{-b_k t} a_k e^{b_k t} m_{k-1}(x,t)\Big|\\
&\le b_k e^{-b_k t} x^k + a_k b_k e^{-b_k t}\int\limits_0^t e^{b_k s}\,\mathrm{d}s + a_k\\
&\le b_k + a_k e^{-b_k t}(e^{b_k t}-1) + a_k \le 3b_k;\\
|\partial_t^2 m_k(x,t)| &= \Big| b_k^2 e^{-b_k t}\Big(x^k + a_k \int\limits_0^t e^{b_k s} m_{k-1}(x,s)\,\mathrm{d}s\Big)\\
&\qquad - a_k b_k m_{k-1}(x,t) + a_k \partial_t m_{k-1}(x,t)\Big|\\
&\leqslant b_k^2 + b_k a_k + b_k a_k + 3a_k b_k \leqslant 6b_k^2,\\
|\partial_t^3 m_k(x,t)| &\leqslant \Big| b_k^3 e^{-b_k t}\Big(x^k + a_k \int\limits_0^t e^{b_k s} m_{k-1}(x,s)\,\mathrm{d}s\Big) + a_k b_k^2 m_{k-1}(x,t)\\
&\qquad + a_k b_k \partial_t m_{k-1}(x,t) + a_k \partial_t^2 m_{k-1}(x,t)\Big| \leqslant 12 b_k^3,
\end{aligned}
$$

and by induction

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

Finally, for all mixed partial derivatives, we have

$$
\begin{aligned}
|\partial_t^p \partial_x^q m_k(x,t)| &= \Big| \partial_t^p \partial_x^q e^{-b_k t}\Big(x^k + a_k \int\limits_0^t e^{b_k s} m_{k-1}(x,s)\,\mathrm{d}s\Big)\Big|\\
&\leqslant \Big| \partial_t^p e^{-b_k t}\Big(k(k-1)\dots(k-q+1)\\
&\qquad + a_k (k-1)(k-2)\dots(k-q)\int\limits_0^t e^{b_k s}\,\mathrm{d}s\Big)\Big|\\
&\leqslant \Big| \partial_t^p e^{-b_k t}\Big(k(k-1)\dots(k-q+1)\Big(a_k \int\limits_0^t e^{b_k s}\,\mathrm{d}s + 1\Big)\Big)\Big|\\
&= \Big| (-b_k)^p e^{-b_k t}\Big(k(k-1)\dots(k-q+1)
\end{aligned}
$$

$$\times \left(a_k \int_0^t e^{b_k s}\,\mathrm{d}s + 1\right)\Big)$$

$$+ k(k-1)\dots(k-q+1)a_k\Big|$$

$$= (b_k^p + 1)k(k-1)\dots(k-q+1)a_k$$

$$= O(k^{2p+q+2}),\ k \to \infty.$$

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 (5.1.7) converges uniformly for all $p, q \in \mathbb{N}_0$. □

5.2 CIR equation

The well-known Cox–Ingersoll–Ross (CIR) process [11] is the solution of the SDE

$$X_t^x = x + \int_0^t \theta(\kappa - X_s^x)\,\mathrm{d}s + \int_0^t \sigma\sqrt{X_s^x}\,\mathrm{d}B_s, \quad t \in [0,T], \qquad (5.2.1)$$

with parameters $\theta, \kappa, \sigma > 0$, $x \geq 0$.

Definition 5.2. For $\delta \geqslant 0$ and $x \geqslant 0$, the unique strong solution Y to the equation

$$Y_t = x + \delta t + 2\int_0^t \sqrt{Y_s}\,\mathrm{d}B_s,\ t \geqslant 0, \qquad (5.2.2)$$

is called a squared Bessel process with dimension δ, starting at x (BESQ_x^δ). We further denote it by $Y_t^\delta(x)$ or $Y^\delta(t,x)$, and also, $Y_t^\delta := Y_t^\delta(0)$.

Lemma 5.5 (see [20, Sect. 6.1])**.** *Let $B = (B^1, B^2, \dots, B^n)$ be a standard n-dimensional Brownian motion, $n \in \mathbb{N}$. Then the process*

$$R_t^2 := \|z + B_t\|^2 = \sum_{i=1}^{n}(z_i + B_t^i)^2, \quad t \geq 0,$$

where $z = (z_1, \dots, z_n) \in \mathbb{R}^n$*, coincides in distribution with* $Y_t^n(\|z\|)$*, that is, with a* $BESQ_x^n$ *random process starting at* $x = \|z\| = \sqrt{\sum_{i=1}^n z_i^2}$*. In particular,*

$$Y_t^n(x) \stackrel{d}{=} (\sqrt{x}+B_t^1)^2+\sum_{i=2}^n (B_t^i)^2 \stackrel{d}{=} (\sqrt{x}+\xi\sqrt{t})^2+Y_t^{n-1}, \quad t \geq 0, \tag{5.2.3}$$

where ξ *is a standard normal variable independent of* Y_t^{n-1}*, and* $\stackrel{d}{=}$ *means equality in distribution.*

Lemma 5.6 (see [20, Prop. 6.3.1.1])**.** *The distribution of CIR process* (5.2.1) *can be expressed in terms of a squared Bessel process as follows:*

$$X_t(x) \stackrel{d}{=} \mathrm{e}^{-\theta t} Y^\delta \left(\frac{\sigma^2}{4\theta}(\mathrm{e}^{\theta t} - 1), x \right), \quad t \geq 0, \tag{5.2.4}$$

where $\delta = 4\theta\kappa/\sigma^2$*.*

We will frequently use differentiation under the integral sign (in particular, under the expectation sign). Without special mentioning, this will be clearly justified by Lemma 5.7, which seems to be a folklore theorem; we refer to the technical report [9].

Definition 5.3. Let $(E, \mathcal{A}, \mu)$ be a measure space. Let $X \subset \mathbb{R}^k$ be an open set, and $f : X \times E \to \mathbb{R}$ be a measurable function. The function f is said to be locally integrable in X if

$$\int_K \int_E |f(x, \omega)| \mu(\mathrm{d}\omega) \mathrm{d}x < \infty$$

for all compact sets $K \subset X$.

Lemma 5.7 (Differentiation under the integral sign; see [9, Thm. 4.1])**.** *Let* $(E, \mathcal{A}, \mu)$*,* X*, and let* f *be as in Definition 5.3. Suppose that* f *has partial derivatives* $\frac{\partial f}{\partial x_i}(x, \omega)$ *for all* $(x, \omega) \in X \times E$ *and that both* f *and* $\frac{\partial f}{\partial x_i}$ *are locally integrable in* X*. Then*

$$\frac{\partial}{\partial x_i} \int_E f(x, \omega) \mu(\mathrm{d}\omega) = \int_E \frac{\partial}{\partial x_i} f(x, \omega) \mu(\mathrm{d}\omega)$$

for almost all $x \in X$. In particular, if both sides are continuous in X, then we have equality for all $x \in X$.

Alfonsi [2, Prop. 4.1], using the known expression of the transition density of CIR process by a rather complicated function series, gave an ad hoc proof that, indeed, $u = u(t,x) := \mathbb{E}f(X_t^x)$ is a solution of the PDE (3.1.4), where $Af(x) = \theta(\kappa - x)f'(x) + \frac{1}{2}\sigma^2 x f''(x)$, $x \geq 0$, is the generator of the CIR process (5.2.1). Moreover, he proved that if $f : \overline{\mathbb{R}}_+ \to \mathbb{R}$ is sufficiently smooth with partial derivatives of polynomial growth, then so is the solution u.

To have the possibility of extension of this result to other processes, it is rather natural to look for a proof that is not based on explicit expressions of the transition functions and can be extended to a wider class of "square-root-type" processes. In this section, we give such a proof in case the coefficients of Eq. (5.2.1) satisfy the condition $\sigma^2 \leq 4\theta\kappa$. The main tools are the additivity property of CIR processes and their representation in terms of squared Bessel processes. More precisely, we use, after a smooth time–space transformation, the expression of the solution of Eq. (5.2.1) in the form $X_t^x = (\sqrt{x} + B_t)^2 + Y_t$, where Y is a squared Bessel process independent from B. The main challenge is the negative powers of x appearing in the expression of $u(t,x) = \mathbb{E}f(X_t^x)$ after applying Itô's formula. To overcome it, we use a "symmetrization" trick (see Step 1 in the proof of Theorem 5.8) based on the simple fact that replacing B_t by the "opposite" Brownian motion $\bar{B}_t := -B_t$ does not change the distribution of X_t^x.

In the following two chapters, the domain $\mathbb{D} = \overline{\mathbb{R}}_+$.

Theorem 5.8 (cf. Alfonsi [2, Prop. 4.1])**.** *Let $X_t(x) = X_t^x$ be a CIR process with coefficients satisfying the condition $\sigma^2 \leq 4\theta\kappa$ and starting at $x \geq 0$. Let $f \in C^q_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$ for some $q \geq 4$. Then the function*

$$u(t,x) := \mathbb{E}f(X_t(x)), \quad x \geq 0, t \in [0,T],$$

is l times continuously differentiable in $x \geq 0$ and l' times continuously differentiable in $t \in [0,T]$ for $l, l' \in \mathbb{N}$ such that $2l + 4l' \leq q$. Moreover, *there exist constants $C \geq 0$ and $k \in \mathbb{N}$, depending only on a good set $\{(C_i, k_i),\ i = 0, 1, \ldots, q\}$ for f, such that*

$$\left|\partial_x^j \partial_t^i u(t, x)\right| \leqslant C(1 + x^k), \quad x \geq 0,\ t \in [0, T], \tag{5.2.5}$$

for $j = 0, 1, \ldots, l$, $i = 0, 1, \ldots, l'$. In particular, $u(t, x)$ is a (classical) solution of the Kolmogorov backward equation (3.1.4) for $(t, x) \in [0, T] \times \overline{\mathbb{R}}_+$.

As a consequence, if $f \in C^\infty_{\text{pol}}(\overline{\mathbb{R}}_+)$, then $u(t, x)$ is infinitely differentiable on $[0, T] \times \overline{\mathbb{R}}_+$, and estimate (5.2.5) holds for all $i, j \in \mathbb{N}$ with C and k depending on (i, j) and a good sequence $\{(C_i, k_i),\ i \in \mathbb{N}_0\}$ for f.

Proof. We first focus ourselves on the differentiability in $x \geq 0$. By Lemma 5.6 the process $X_t(x)$ can be reduced, by a space–time transformation, to the BESQ$^\delta$ process $Y_t^\delta(x)$ with $\delta = \frac{4\theta\kappa}{\sigma^2} \geq 1$. Since in (5.2.4), only bounded smooth functions of $t \in [0, T]$ are involved, it suffices to show estimate (5.2.5) for $Y_t^\delta(x)$, $t \in [0, \widetilde{T}]$, instead of $X_t(x)$, $t \in [0, T]$, with $\widetilde{T} = \frac{1}{\theta} \ln(1 + \frac{4\theta T}{\sigma^2})$. With an abuse of notation, we further write T instead of $\widetilde{T}$. We proceed by induction on l.

Step 1. Let $l = 1$. First, suppose that $\delta = n \in \mathbb{N}$. By Lemma 5.5 we have

$$Y_t^n(x) \stackrel{d}{=} (\sqrt{x} + \xi\sqrt{t})^2 + Y_t^{n-1}, \tag{5.2.6}$$

where $\xi \sim \mathcal{N}(0, 1)$ is independent of Y_t^{n-1} (in the case $n = 1$, $Y_t^0 := 0$). Denote

$$Y_t^+(x) := (\sqrt{x} + \xi\sqrt{t})^2, \qquad Y_t^-(x) := (\sqrt{x} - \xi\sqrt{t})^2.$$

Since the distributions of $Y_t^+(x)$ and $Y_t^-(x)$ coincide, we have

$$\begin{aligned}
\partial_x \mathbb{E} f(Y_t^n(x)) &= \partial_x \mathbb{E} f\left(Y_t^+(x) + Y_t^{n-1}\right) \\
&= \frac{1}{2}\left[\partial_x \mathbb{E} f\left(Y_t^+(x) + Y_t^{n-1}\right) + \partial_x \mathbb{E} f\left(Y_t^-(x) + Y_t^{n-1}\right)\right] \\
&= \frac{1}{2}\mathbb{E}\Bigg[f'\left(Y_t^+(x) + Y_t^{n-1}\right)\left(1 + \xi\sqrt{\frac{t}{x}}\right) \qquad \text{(Lemma 5.7)} \\
&\qquad + f'\left(Y_t^-(x) + Y_t^{n-1}\right)\left(1 - \xi\sqrt{\frac{t}{x}}\right)\Bigg]
\end{aligned}$$

$$
\begin{aligned}
&= \mathbb{E} f'\left(Y_t^n(x)\right) \\
&\quad + \frac{1}{2}\sqrt{\frac{t}{x}}\mathbb{E}\left\{\xi\left[f'\left(Y_t^+(x) + Y_t^{n-1}\right)\right.\right. \\
&\qquad \left.\left. - f'\left(Y_t^-(x) + Y_t^{n-1}\right)\right]\right\} \\
&= \mathbb{E} f'\left(Y_t^n(x)\right) + \frac{1}{2}\sqrt{t}\,\mathbb{E}\big(\xi g_1(x, \xi\sqrt{t}, Y_t^{n-1})\big) \qquad (5.2.7) \\
&=: P(t,x) + R(t,x), \quad x > 0, \qquad (5.2.8)
\end{aligned}
$$

where

$$
g_1(x,a,b) := \frac{f'\big((\sqrt{x}+a)^2 + b\big) - f'\big((\sqrt{x}-a)^2 + b\big)}{\sqrt{x}},
$$

$x > 0,\ a \in \mathbb{R},\ b \geq 0$. We now estimate $P(t,x)$ and $R(t,x)$ separately. By the well-known inequality

$$
\Big|\sum_{i=1}^n a_i\Big|^p \leq n^{p-1}\sum_{i=1}^n |a_i|^p \quad \text{for any } n \in \mathbb{N},\ p \geq 1,\ a_i \in \mathbb{R},\ i = 1, 2, \ldots, n, \tag{5.2.9}
$$

we have the following estimates:

$$
\begin{aligned}
\mathbb{E}(Y_t^{\pm}(x))^p &= \mathbb{E}(\sqrt{x} \pm \xi\sqrt{t})^{2p} \leq 2^{2p-1}(x^p + \mathbb{E}|\xi|^{2p}t^p) \\
&= 2^{2p-1}\Big(x^p + \frac{2^p\Gamma(p+\frac{1}{2})}{\sqrt{\pi}}t^p\Big) \\
&\leq C(1+x^p), \quad x \geq 0,\ t \in [0,T],
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}(Y_t^n)^p &= \mathbb{E}\Big(\sum_{i=1}^n |B_t^i|^2\Big)^p \leq n^{p-1}\sum_{i=1}^n \mathbb{E}|B_t^i|^{2p} \\
&= n^p\frac{2^p\Gamma(p+\frac{1}{2})}{\sqrt{\pi}}t^p \leq C, \quad t \in [0,T],
\end{aligned}
$$

and, as a consequence,

$$
\begin{aligned}
\mathbb{E}(Y_t^n(x))^p &= \mathbb{E}(Y_t^+(x) + Y_t^{n-1})^p \leq 2^{p-1}\mathbb{E}\big((Y_t^+(x))^p + \mathbb{E}(Y_t^{n-1})^p\big) \\
&\leq C(1+x^p), \quad x \geq 0,\ t \in [0,T]. \qquad (5.2.10)
\end{aligned}
$$

Now, for $P(t,x)$, we have

$$
|P(t,x)| = \mathbb{E}|f'\left(Y_t^n(x)\right)| \leqslant C_1\big(1 + \mathbb{E}(Y_t^n(x))^{k_1}\big) \leqslant C_1\big(1 + C(1+x^{k_1})\big)
$$

$$\leq C(1+x^{k_1}), \quad x \geq 0,\ t \in [0,T], \tag{5.2.11}$$

where the constant C depends only on C_1, k_1, T, and n.

At this point, we need the following technical lemma, which we will prove in the Appendix.

Lemma 5.9. *For a function $f : \overline{\mathbb{R}}_+ \to \mathbb{R}$, define the function*

$$g(x; a, b) := \frac{f\big((\sqrt{x}+a)^2+b\big) - f\big((\sqrt{x}-a)^2+b\big)}{\sqrt{x}},$$

$x > 0, a \in \mathbb{R}, b \in \overline{\mathbb{R}}_+$. If $f \in C^q_{\text{pol}}(\overline{\mathbb{R}}_+)$ for some $q = 2l+1 \in \mathbb{N}$ ($l \in \mathbb{N}_0$), then the function g is extendable to a continuous function on $\overline{\mathbb{R}}_+ \times \mathbb{R} \times \overline{\mathbb{R}}_+$ such that $g(\cdot; a, b) \in C^l_{\text{pol}}(\overline{\mathbb{R}}_+)$ for all $a \in \mathbb{R}$ and $b \in \overline{\mathbb{R}}_+$. Moreover, there exist constants $C \geq 0$ and $k \in \mathbb{N}$, depending only on a good set $\{(C_i, k_i),$ $i = 0, 1, \dots, q\}$ for f, such that

$$|\partial_x^j g(x; a, b)| \leq C|a|(1 + x^k + |a^2+b|^k), \quad x \in \overline{\mathbb{R}}_+, a \in \mathbb{R}, b \in \overline{\mathbb{R}}_+, \tag{5.2.12}$$

for all $j = 0, 1, \dots, l$.

Now consider $R(t,x)$. Applying Lemma 5.9 with f' instead of f (and thus with g_1 instead of g), we have

$$\begin{aligned}
|R(t,x)| &\leq \frac{1}{2}\sqrt{t}\mathbb{E}\left|\xi g_1(x, \xi\sqrt{t}, Y_t^{n-1})\right| \\
&\leqslant Ct\mathbb{E}\left[\xi^2(1 + x^{k_2} + |(\xi\sqrt{t})^2 + Y_t^{n-1}|^{k_2})\right] \\
&\leqslant Ct\mathbb{E}\left[\xi^2\left(1 + x^{k_2} + 2^{k-1}\left((\xi\sqrt{t})^{2k_2} + (Y_t^{n-1})^{k_2}\right)\right)\right] \\
&\leqslant C(1+x^{k_2}), \quad x \geq 0,\ t \in [0,T],
\end{aligned} \tag{5.2.13}$$

where the constant C clearly depends only on C_2, k_2, T, and n. Combining the obtained estimates, we finally obtain

$$\begin{aligned}
\left|\partial_x \mathbb{E} f(X_t(x))\right| &\leqslant C(1+x^{k_1}) + C(1+x^{k_2}) \\
&\leqslant C(1+x^k), \quad x \geq 0,\ t \in [0,T],
\end{aligned}$$

where $k = \max\{k_1, k_2\}$, and the constant C depends only on C_1, C_2, k_1, k_2, T, and n.

Now consider the general case where $\delta \geq 1$, $\delta \notin \mathbb{N}$. Note that we consider the general case only for $l = 1$ because the reasoning for higher-order derivatives is the same.

Let $n < \delta < n+1$, $n \in \mathbb{N}$. According to [20, Prop. 6.2.1.1], $Y_t^\delta(x)$ has the same distribution as the affine sum of two independent BESQ processes, namely,

$$Y_t^\delta(x) \stackrel{d}{=} \widetilde{Y}_t^{n-1}(x) + \widehat{Y}_t^{\delta-n+1}(0),$$

where $\widetilde{Y}_t^{n-1}(x)$ and $\widehat{Y}_t^{\delta-n+1}(0)$ are two independent BESQ processes of dimensions $n-1$ and $\delta-n+1$, respectively, starting at x, and 0.

Using the density of $\mathrm{BESQ}^\delta(0)$ given in [39, Prop. 3.1], we get

$$\begin{aligned}
\mathbb{E}(Y_t^{\delta-n+1}(0))^p &= \Gamma\Big(\frac{\delta-n+1}{2}\Big)2^{\frac{\delta-n+1}{2}-1}t^{-\frac{\delta-n+1}{2}}\int_0^\infty y^p y^{\delta-n}\mathrm{e}^{-y^2/2t}\,\mathrm{d}y \\
&= \Gamma\Big(\frac{\delta-n+1}{2}\Big)2^{\frac{\delta-n+1}{2}-1}t^{-\frac{\delta-n+1}{2}} \\
&\qquad \times 2^{\frac{p+\delta-n-1}{2}}t^{\frac{p+\delta-n+1}{2}}\Gamma\Big(\frac{p+\delta-n+1}{2}\Big) \\
&= 2^{\frac{p}{2}+\delta-n-1}t^{\frac{p}{2}}\Gamma\Big(\frac{\delta-n+1}{2}\Big)\Gamma\Big(\frac{p+\delta-n+1}{2}\Big) \\
&\leq C, \quad t \in [0,T].
\end{aligned}$$

Using the estimates just obtained for $\delta \in \mathbb{N}$ and $\mathbb{E}(Y_t^{\delta-n+1}(0))^p$, we have

$$\begin{aligned}
\partial_x \mathbb{E} f(Y_t^\delta(x)) &= \partial_x \mathbb{E} f\big(\widetilde{Y}_t^{n-1}(x) + \widehat{Y}_t^{\delta-n+1}(0)\big) \\
&= \frac{1}{2}\Big[\partial_x \mathbb{E} f\left(\widetilde{Y}_t^{+}(x) + \widetilde{Y}_t^{n-2} + \widehat{Y}_t^{\delta-n+1}(0)\right) \\
&\qquad + \partial_x \mathbb{E} f\left(\widetilde{Y}_t^{-}(x) + \widetilde{Y}_t^{n-2} + \widehat{Y}_t^{\delta-n+1}(0)\right)\Big],
\end{aligned}$$

where

$$\widetilde{Y}_t^{+}(x) := (\sqrt{x} + \tilde{\xi}\sqrt{t})^2, \quad \widetilde{Y}_t^{-}(x) := (\sqrt{x} - \tilde{\xi}\sqrt{t})^2, \quad \widetilde{Y}_t^{n-2} := \sum_{i=1}^{n-2}(\widetilde{B}_t^i)^2.$$

Using again the fact that the distributions of $\widetilde{Y}_t^{\pm}(x)$ coincide and proceeding as in (5.2.8), we have

$$\partial_x \mathbb{E} f(Y_t^\delta(x)) = \frac{1}{2}\Big[\partial_x \mathbb{E} f\left((\sqrt{x} + \tilde{\xi}\sqrt{t})^2 + \widetilde{Y}_t^{n-2} + \widehat{Y}_t^{\delta-n+1}(0)\right)$$

$$
\begin{aligned}
&\quad + \partial_x \mathbb{E} f\left((\sqrt{x} - \tilde{\xi}\sqrt{t})^2 + \widetilde{Y}_t^{n-2} + \widehat{Y}_t^{\delta-n+1}(0) \right) \Big] \\
&= \frac{1}{2}\Big[\mathbb{E} f'\left((\sqrt{x} + \tilde{\xi}\sqrt{t})^2 + \widetilde{Y}_t^{n-2} + \widehat{Y}_t^{\delta-n+1}(0) \right) \times \left(1 + \tilde{\xi}\sqrt{\frac{t}{x}} \right) \\
&\quad + \mathbb{E} f'\left((\sqrt{x} - \tilde{\xi}\sqrt{t})^2 + \widetilde{Y}_t^{n-2} + \widehat{Y}_t^{\delta-n+1}(0) \right) \times \left(1 - \tilde{\xi}\sqrt{\frac{t}{x}} \right) \Big] \\
&= \mathbb{E} f'(Y_t^\delta(x)) + \frac{\sqrt{t}}{2} \mathbb{E}\Big[\tilde{\xi} g_1\left(x, \tilde{\xi}\sqrt{t}, \widetilde{Y}_t^{n-2} + \widehat{Y}_t^{\delta-n+1}(0) \right) \Big] \\
&=: P_1(t,x) + R_1(t,x).
\end{aligned}
$$

Combination of estimates (5.2.9) and (5.2.10) leads to the estimate

$$
\begin{aligned}
|P_1(t,x)| = |\mathbb{E} f'(Y_t^\delta(x))| &\leqslant C_1 \mathbb{E}(1 + |Y_t^\delta(x)|^{k_1}) \\
&\leqslant C_1 \mathbb{E}\big(1 + 2^{k_1-1} \big| (\widetilde{Y}_t^{n-1}(x))^{k_1} + (\widehat{Y}_t^{\delta-n+1}(0))^{k_1} \big| \big) \\
&\leqslant C(1 + x^{k_1}), \quad x \geq 0,\ t \in [0,T],
\end{aligned}
$$

where the constant C depends only on C_1, k_1, T, and n. By Lemma 5.9, similarly to estimate (5.2.13), we have

$$
\begin{aligned}
|R_1(t,x)| &\leqslant C_2 \mathbb{E}\Big[\tilde{\xi}^2 \Big(1 + x^{k_2} + \Big| \tilde{\xi}^2 t + \widetilde{Y}_t^{n-2} + \widehat{Y}_t^{\delta-n+1}(0) \Big|^{k_2} \Big) \Big] \\
&\leqslant C(1 + x^k), \quad x \geq 0,\ t \in [0,T],
\end{aligned}
$$

where the constant C depends only on C_2, k_2, T, and n. Combining the last two estimates, we get

$$
\Big| \partial_x \mathbb{E} f(X_t(x)) \Big| \leqslant C(1 + x^k), \quad x \geq 0,\ t \in [0,T], \tag{5.2.14}
$$

where $k = \max\{k_1, k_2\}$, and the constant C depends only on C_1, C_2, k_1, k_2, and T.

Step 2. Let $l = 2$. From *Step 1* we have

$$
\partial_x \mathbb{E} f(Y_t^n(x)) = \mathbb{E} f'\left(Y_t^n(x) \right) + \frac{1}{2}\sqrt{t}\, \mathbb{E}\big(\xi g_1(x, \xi\sqrt{t}, Y_t^{n-1}) \big).
$$

Therefore,

$$
\begin{aligned}
\partial_x^2 \mathbb{E} f(Y_t^n(x)) &= \partial_x \mathbb{E} f'\left(Y_t^n(x) \right) + \frac{1}{2}\sqrt{t}\, \mathbb{E}\big(\xi \partial_x g_1(x, \xi\sqrt{t}, Y_t^{n-1}) \big) \\
&=: P_2(t,x) + R_2(t,x).
\end{aligned}
$$

From estimate (5.2.14) with f replaced by f' we obtain

$$|P_2(t,x)| \leqslant C(1+x^{k_3}), \quad x \geq 0,\ t \in [0,T], \tag{5.2.15}$$

where the constant C depends only on C_1, C_3, k_1, k_3, T, and n. For $R_2(t,x)$, applying Lemma 5.9 once more to g_1 instead of g, we get

$$\begin{aligned}
|R_2(t,x)| &\leqslant \frac{1}{2}\sqrt{t}\mathbb{E}\big|\xi\partial_x g_1(x,\xi\sqrt{t},Y_t^{n-1})\big| \\
&\leqslant Ct\mathbb{E}\big(\xi^2(1+x^k+|\xi\sqrt{t}|^k+(Y_t^{n-1})^k)\big) \\
&\leqslant C(1+x^k), \quad x \geq 0,\ t \in [0,T],
\end{aligned}$$

where the constants C and $k \in \mathbb{N}$ depend only on $\{(C_i,k_i),\, i=1,2,3,4\}$, T, and n. Combining the obtained estimates, we finally obtain

$$\left|\partial_x^2\, \mathbb{E}f(X_t(x))\right| \leqslant C(1+x^k), \quad x \geq 0,\ t \in [0,T],$$

where the constants C and $k \in \mathbb{N}$ depend only on $\{(C_i,k_i),\, i=1,2,3,4\}$, T, and n.

Step 3. Now we may continue by induction on l. Suppose that estimate (5.2.5) is valid for $l = m-1$. Let us show that it is still valid for $l = m$. The arguments are similar to those in the case $m = 2$ (Step 2). We have

$$\begin{aligned}
\partial_x^m\, \mathbb{E}f(Y_t^n(x)) &= \partial_x^{m-1}\mathbb{E}f'\,(Y_t^n(x)) + \frac{1}{2}\sqrt{t}\,\partial_x^{m-1}\mathbb{E}\big(\xi g_1(x,\xi\sqrt{t},Y_t^{n-1})\big) \\
&=: P_m(t,x) + R_m(t,x).
\end{aligned}$$

Then, similarly to estimates (5.2.11) and (5.2.15), we have

$$|P_m(t,x)| \leqslant C(1+x^{k_m}).$$

where the constant C depends only on $\{(C_i,k_i),\, i=1,3,\ldots,2m-1\}$, T, and n.

For $R_m(t,x)$, applying Lemma 5.9 to g_1 instead of g, we get

$$|R_m(t,x)| \leqslant \frac{1}{2}\sqrt{t}\mathbb{E}\left|\xi\partial_x^{m-1}g_1(x,\xi\sqrt{t},Y_t^{n-1})\right| \leqslant C(1+x^k),$$

where the constants C and $k \in \mathbb{N}$ depend only on $\{(C_i, k_i), i = 1, \dots, 2m\}$, T, and n. Combining the obtained estimates, we get

$$\left|\partial_x^m \mathbb{E} f(X_t(x))\right| \leqslant C(1 + x^k), \quad x > 0, \ t \in [0, T],$$

where the constants C and $k \in \mathbb{N}$ depend only on $\{(C_i, k_i), i = 1, \dots, 2m\}$, T, and n. Thus, Theorem 5.8 is proved for all $l \in \mathbb{N}$.

Step 4. As in Alfonsi [2, p. 28], inequality (5.2.5) for the derivatives with respect to t and mixed derivatives follows automatically by an induction on l' using that, for $l' \geq 1$ such that $4l' + 2l \leq q$,

$$\begin{aligned}\partial_x^l \partial_t^{l'} u(t,x) &= \partial_x^l \Big(\theta(\kappa - x)\partial_x \partial_t^{l'-1} u(t,x) + \frac{\sigma^2}{2} x \partial_x^2 \partial_t^{l'-1} u(t,x)\Big)\\ &= \frac{\sigma^2}{2} x \partial_x^{l+2} \partial_t^{l'-1} u(t,x) + \Big(l\frac{\sigma^2}{2} + \theta(\kappa - x)\Big)\partial_x^{l+1}\partial_t^{l'-1} u(t,x)\\ &\qquad - l\theta \partial_x^l \partial_t^{l'-1} u(t,x).\end{aligned}$$

$\square$

5.3 Stratonovich square-root SDE

In this chapter, we prove the regularity of the solution of backward Kolmogorov equation for the Stratonovich SDE

$$\mathrm{d}X_t = \sqrt{X_t a(X_t)} \circ \mathrm{d}B_t, \qquad X_0 = x \geq 0. \tag{5.3.1}$$

For self-consistency, let us briefly recall the definition and some properties of the Stratonovich integral. (For more detail, we refer to [37, Chs. 3 and 5] or [32, Ch. 8]). The Stratonovich integral of an Itô process $X_t = X_0 + \int_0^t K_s \,\mathrm{d}s + \int_0^t H_s \,\mathrm{d}B_s$ with respect another Itô process $Y_t = Y_0 + \int_0^t \tilde{K}_s \,\mathrm{d}s + \int_0^t \tilde{H}_s \,\mathrm{d}B_s$ may be defined as

$$\begin{aligned}\int_0^t X_s \circ \mathrm{d}Y_s &:= \int_0^t X_s \,\mathrm{d}Y_s + \frac{1}{2}\langle X, Y\rangle_t\\ &= \int_0^t X_s \,\mathrm{d}Y_s + \frac{1}{2}\int_0^t H_s \tilde{H}_s \,\mathrm{d}s, \ t \geq 0,\end{aligned}$$

or, in a short differential form,

$$X_t \circ \mathrm{d}Y_t := X_t \,\mathrm{d}Y_t + \frac{1}{2}\mathrm{d}\langle X, Y\rangle_t.$$

We will need the following properties of the Stratonovich integral, which are similar to the properties of ordinary (nonstochastic) integrals:

1. (Itô's formula) If $F \in C^3(\mathbb{R})$, then $\mathrm{d}F(X_t) = F'(X_t) \circ \mathrm{d}X_t$.

2. If $\mathrm{d}Z_t = X_t \circ \mathrm{d}Y_t$, then $W_t \circ \mathrm{d}Z_t = W_t X_t \circ \mathrm{d}Y_t$, or in the integral form: if $Z_t = \int_0^t X_s \circ \mathrm{d}Y_s$, then $\int_0^t W_s \circ \mathrm{d}Z_s = \int_0^t W_s X_s \circ \mathrm{d}Y_s$, provided that one of the integrals is well-defined.

3. If $\sigma \in C^2(\mathbb{R})$, then the Stratonovich SDE

$$\mathrm{d}X_t = b(X_t)\,\mathrm{d}t + \sigma(X_t) \circ \mathrm{d}B_t$$

is equivalent to the Itô SDE

$$\mathrm{d}X_t = \big(b(X_t) + \tfrac{1}{2}\,\sigma(X_t)\sigma'(X_t)\big)\,\mathrm{d}t + \sigma(X_t)\,\mathrm{d}B_t.$$

In particular, the Stratonovich equation (5.3.1) is equivalent to the Itô SDE

$$\mathrm{d}X_t = \frac{1}{4}\left(a(X_t) + X_t a'(X_t)\right)\,\mathrm{d}t + \sqrt{X_t a(X_t)}\,\mathrm{d}B_t, \qquad X_0 = x \geq 0.$$

The main result of this section is the following theorem on regularity of the solution of the backward Kolmogorov equation related to Eq. (5.3.1). In its proof, we use the ideas of the previous section. However, it is significantly more technical.

Theorem 5.10 (cf. Thm. 5.8)**.** *Let $X_t(x) = X_t^x$, $t \geq 0$, be the process satisfying the SDE* (5.3.1)*. In addition, let $f \in C^q_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$ for some $q \geqslant 4$, let $a \in C^{2l-1}_{\mathrm{pol}}(\overline{\mathbb{R}}_+) \cap C_{\mathrm{lin}}(\overline{\mathbb{R}}_+)$ for some $l \geqslant 1$, and suppose that $0 < C_0 \leqslant a(x)$, $x \geqslant 0$. Then the function*

$$u(t,x) := \mathbb{E}f(X_t(x)), \quad x \geqslant 0,\ t \in [0,T],$$

is l times continuously differentiable in $x \geqslant 0$ and l' times continuously differentiable in $t \in [0,T]$ for $l, l' \in \mathbb{N}$ such that $2l + 4l' \leqslant q$. Moreover, there exist constants $C \geq 0$ and $k \in \mathbb{N}$, depending only on a good set $\{(C_i, k_i),\ i = 0, 1, \ldots, q\}$ for f such that

$$\left|\partial_x^j \partial_t^i u(t,x)\right| \leqslant C(1 + x^k), \quad x \geq 0, \tag{5.3.2}$$

for $j = 0, 1, \ldots, l, i = 0, 1, \ldots, l'$. In particular, $u(t,x)$ is a (classical) solution of the Kolmogorov backward equation for $(t,x) \in [0,T] \times \overline{\mathbb{R}}_+$.

As a consequence, if $f \in C^\infty_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$ and $a \in C^\infty_{\mathrm{pol}}(\overline{\mathbb{R}}_+) \cap C_{\mathrm{lin}}(\overline{\mathbb{R}}_+)$, then $u(t,x)$ is infinitely differentiable on $[0,T] \times \overline{\mathbb{R}}_+$, and estimate (5.3.2) holds for all $i, j \in \mathbb{N}$ with C and k depending on (i,j) and a good sequence $\{(C_i, k_i),\ i \in \mathbb{N}_0\}$ for f.

Proof. To start with, let us write down the assumed estimates:

$$\begin{aligned}
&0 < C_0 \leqslant a(x), \quad x \geqslant 0,\\
&a(x) \leqslant C_1(1 + x), \quad x \geqslant 0,\\
&a^{(k)}(x) \leqslant C_k(1 + x^{p_k}), \quad x \geqslant 0, \quad k \geqslant 1,\\
&f^{(k)}(x) \leqslant C_k(1 + x^{p_k}), \quad x \geqslant 0, \quad k \geqslant 1.
\end{aligned}$$

To simplify the expressions, we use the same notation for estimates of functions a and f: if the functions a and f have different constants C_k, p_k, then we take their maxima and apply them for both functions.

Let us introduce the following functions:

$$G(x) = \int_0^x \frac{\mathrm{d}y}{\sqrt{y a(y)}}, \quad x \geqslant 0, \tag{5.3.3}$$

$$K(x) = G(x^2) = \int_0^{x^2} \frac{\mathrm{d}y}{\sqrt{y a(y)}} = \int_0^x \frac{2\mathrm{d}y}{\sqrt{a(y^2)}}, \quad x \in \mathbb{R}, \tag{5.3.4}$$

$$J(x) = K^{-1}(x), \quad x \in \mathbb{R}. \tag{5.3.5}$$

We will need the following derivatives of J and J^2:

$$J'(x) = \frac{1}{K'(J(x))} = \frac{1}{\frac{2}{\sqrt{a(J^2(x))}}} = \frac{\sqrt{a(J^2(x))}}{2},$$

$$
\begin{aligned}
J''(x) &= \frac{1}{4\sqrt{a(J^2(x))}} a'(J^2(x)) \cdot 2J(x)J'(x) = \frac{a'(J^2(x))J(x)}{4}, \\
(J^2)'(x) &= 2J(x)J'(x) = J(x)\sqrt{a(J^2(x))}, \\
(J^2)''(x) &= 2\left(J(x)J'(x)\right)' = 2J'^2(x) + 2J(x)J''(x) \\
&= \frac{1}{2}\left(a(J^2(x)) + J^2(x)a'(J^2(x))\right), \\
(J^2)'''(x) &= \frac{1}{2}(J^2)'(x)\left(2a'(J^2(x)) + J^2(x)a''(J^2(x))\right).
\end{aligned}
$$

Continuing, we get that the higher-order derivatives of $J^2(x)$ are of the following form:

$$
\begin{aligned}
(J^2(x))^{(n)} = F_n((J^2(x))^{(n-2)}, (J^2(x))^{(n-3)} \dots, , \\
J^2(x), a^{(n-1)}(x), a^{(n-2)}(x), \dots, a'(x)), \quad n = 1, \dots, 2l.
\end{aligned} \tag{5.3.6}
$$

where F_n has polynomial growth in all variables.

Let us check that the process $\widetilde{X}_t(x) := J^2(G(x) + B_t)$, $t \geq 0$, has the same distribution as that of $X_t(x)$. Indeed, by Itô's formula we have

$$
\begin{aligned}
\mathrm{d}\widetilde{X}_t(x) &= (J^2)'(G(x) + B_t)\,\mathrm{d}B_t + \frac{1}{2}(J^2)''(G(x) + B_t)\,\mathrm{d}t \\
&= J(G(x) + B_t)\sqrt{a(J^2(G(x) + B_t))}\,\mathrm{d}B_t \\
&+ \frac{1}{4}\left(a(J^2(G(x) + B_t)) + J^2(G(x) + B_t)a'(J^2(G(x) + B_t))\right)\mathrm{d}t \\
&= \sqrt{\widetilde{X}_t(x)a(\widetilde{X}_t(x))}\,\mathrm{sgn}(J(G(x) + B_t))\,\mathrm{d}B_t \\
&+ \frac{1}{4}\left(a(\widetilde{X}_t(x)) + \widetilde{X}_t(x)a'(\widetilde{X}_t(x))\right)\mathrm{d}t \\
&= \sqrt{\widetilde{X}_t a(\widetilde{X}_t)} \circ \mathrm{d}\widetilde{B}_t,
\end{aligned}
$$

where $\widetilde{B}_t := \int_0^t \mathrm{sgn}(J(G(x) + B_s))\mathrm{d}B_s$, $t \geq 0$, is a Brownian motion since $|\,\mathrm{sgn}(J(G(x) + B_s))| = 1$ (see Lévy's characterization of Brownian motion [28]). With some abuse in notation, we further write X instead of $\widetilde{X}$. Thus

$$
X_t(x) = J^2\left(\int_0^x \frac{\mathrm{d}y}{\sqrt{ya(y)}} + B_t\right),
$$

and we denote

$$X_t^+(x) = J^2\left(\int_0^x \frac{\mathrm{d}y}{\sqrt{ya(y)}} + B_t\right),$$

$$X_t^-(x) = J^2\left(\int_0^x \frac{\mathrm{d}y}{\sqrt{ya(y)}} - B_t\right).$$

Note that the processes $X_t^+(x)$ and $X_t^-(x)$ coincide in distribution. Furthermore, the function $d(x) = G(x)/2\sqrt{x}$, $x > 0$, can be continuously extended to the whole semiaxis $\overline{\mathbb{R}}_+ = [0, \infty)$ and is continuously differentiable there as many times as the derivative a' is. Indeed, denoting $h := 1/\sqrt{a(x)}$, $x \geqslant 0$, we have $h'(x) = -\frac{1}{2}a'(x)/a^{3/2}(x)$ and

$$\begin{aligned} d(x) = \frac{G(x)}{2\sqrt{x}} &= \frac{\int_0^x \frac{h(y)}{\sqrt{y}}\mathrm{d}y}{2\sqrt{x}} = \frac{2\int_0^x h(y)\mathrm{d}(\sqrt{y})}{2\sqrt{x}} \\ &= \frac{(\sqrt{x}h(x) - \int_0^x \sqrt{y}h'(y)\mathrm{d}y)}{\sqrt{x}} \\ &= h(x) - \int_0^x \sqrt{\tfrac{y}{x}}h'(y)\mathrm{d}y \\ &= h(x) - x\int_0^1 \sqrt{z}h'(xz)\mathrm{d}z, \quad x > 0, \end{aligned}$$

so that we can define $d(0) := \lim_{x\downarrow 0} d(x) = h(0)$. Clearly, if the function a, and thus h, is $2l-1$ times continuously differentiable, then d is $2l-2$ times continuously differentiable. Moreover, if $a \in C_{\mathrm{pol}}^{2l-1}(\overline{\mathbb{R}}_+)$, then since $0 < C_0 \leqslant a(x)$, $x \geqslant 0$, we have that also $h \in C_{\mathrm{pol}}^{2l-1}(\overline{\mathbb{R}}_+)$. If $n \leqslant 2l-2$ then

$$d'(x) = h'(x) - \int_0^1 \sqrt{z}h'(xz)\,\mathrm{d}z - x\int_0^1 \sqrt{z}h''(xz)z\,\mathrm{d}z,$$

$$\begin{aligned} d''(x) = h''(x) &- \int_0^1 \sqrt{z}h''(xz)z\,\mathrm{d}z - \int_0^1 \sqrt{z}h''(xz)z\,\mathrm{d}z \\ &- x\int_0^1 \sqrt{z}h'''(xz)z^2\,\mathrm{d}z, \end{aligned}$$

$$\dots$$

$$d^{(n)}(x) = h^{(n)}(x) - n \int_0^1 \sqrt{z} h^{(n)}(xz) z^{n-1} \,\mathrm{d}z - x \int_0^1 \sqrt{z} h^{(n+1)}(xz) z^n \,\mathrm{d}z,$$

and therefore, the function $d \in C_{\mathrm{pol}}^{2l-2}(\overline{\mathbb{R}}_+)$.

We assume that function J^2 has the following form

$$J^2(x) = x^2 \tilde{d}(x), \tag{5.3.7}$$

where $\tilde{d}$ behaves properly close to 0. Since $G(x) \to 0$ when $x \to 0$, we have

$$\begin{aligned} \lim_{x\to 0} \tilde{d}(x) &= \lim_{x\to 0} \frac{J^2(x)}{x^2} = \lim_{x\to 0} \frac{J^2(G(x))}{G^2(x)} = \lim_{x\to 0} \frac{x}{(2\sqrt{x} d(x))^2} \\ &= \lim_{x\to 0} \frac{1}{4d^2(x)} = \frac{a(0)}{4}. \end{aligned} \tag{5.3.8}$$

Since the distributions of $X_t^+(x)$ and $X_t^-(x)$ coincide, we have

$$\begin{aligned} X_t(x) = X_t^+(x) &= J^2(G(x) + B_t) = J^2(B_t + G(x)) \\ &\stackrel{d}{=} J^2(B_t - G(x)) =: X_t^-(x); \\ \mathbb{E} f(X_t(x)) &= \frac{1}{2}\mathbb{E}\Big(f\big(J^2(B_t + G(x))\big) + f\big(J^2(B_t - G(x))\big)\Big); \end{aligned} \tag{5.3.9}$$

$$\begin{aligned} \frac{\partial}{\partial x}\mathbb{E} f(X_t(x)) &= \frac{1}{2\sqrt{x a(x)}}\mathbb{E}\Big(f'\Big(J^2(B_t + G(x))\Big)(J^2)'(B_t + G(x)) \\ &\qquad - f'(J^2(B_t - G(x)))(J^2)'(B_t - G(x))\Big) \\ &= \frac{1}{2\sqrt{x a(x)}}\mathbb{E}\Big(f'\Big(J^2(B_t + G(x)s)\Big) \\ &\qquad \times (J^2)'(B_t + G(x)s)\Big)\Big|_{s=-1}^{1} \\ &= \frac{1}{2\sqrt{x a(x)}}\mathbb{E}\int_{-1}^{1}\Big(f'\Big(J^2(B_t + G(x)s)\Big) \\ &\qquad \times (J^2)'(B_t + G(x)s)\Big)_s' \,\mathrm{d}s \\ &= d(x)h(x)\mathbb{E}\int_{-1}^{1}\Big(f''\Big(J^2(B_t + G(x)s)\Big) \end{aligned}$$

$$
\begin{aligned}
&\times (J^2)'(B_t+G(x)s)\\
&+f'\Big(J^2(B_t+G(x)s)\Big)\\
&\times (J^2)''(B_t+G(x)s)\Big)\mathrm{d}s, \quad (5.3.10)
\end{aligned}
$$

and

$$\lim_{x\downarrow 0} d(x)h(x) = \frac{1}{a(0)}. \tag{5.3.11}$$

To prove the polynomial growth, let us analyze every component separately. We have $0 < C_0 \le a(x) \le C_1(1+x),\ x \ge 0$ and therefore using (5.3.8) and (5.3.11), we obtain

$$
\begin{aligned}
&\Big|d(x)h(x)\Big| \leqslant \frac{1}{C_0},\ \ x \geqslant 0 \text{ and}\\
&\frac{C_0}{4}x^2 \le J^2(x) \le \frac{C_1}{4}x^2,\ \ x \in \mathbb{R}.
\end{aligned}
$$

From that it follows:

$$
\begin{aligned}
|f''\big(J^2(B_t+G(x)s)\big)| &\le C_2\Big(1+\Big(\frac{C_1}{4}(B_t+G(x)s)^2\Big)^{p_2}\Big)\\
&\le C_2\Big(1+\Big(\frac{C_1}{4}(B_t+G(x))^2\Big)^{p_2}\Big)\\
&\le C_2\Big(1+\Big(\frac{C_1}{2}(B_t^2+G^2(x))\Big)^{p_2}\Big)\\
&\le C_2\Big(1+\Big(\frac{C_1}{2}\Big(B_t^2+\frac{4}{C_0}x\Big)\Big)^{p_2}\Big), \qquad (5.3.12)
\end{aligned}
$$

here we used the well known inequality, e.g. see [38, Lemma 1.2]

$$(a+b)^p \leqslant c_p(a^p+b^p),$$

where $c_p = \max\big\{1, 2^{p-1}\big\}$, and which holds with all $a, b \geqslant 0$ if $p > 0$, as well as the fact that function G is the inverse function of function J^2.

As we know that $(J^2)'(x) = J(x)\sqrt{a(J^2(x))}$, then

$$
\begin{aligned}
|(J^2)'(B_t+G(x)s)| &\le \frac{\sqrt{C_1}}{2}\left(B_t+\frac{2\sqrt{x}}{\sqrt{C_0}}\right)\sqrt{C_1\left(1+\frac{C_1}{2}\left(B_t^2+\frac{4}{C_0}x\right)\right)}\\
&\le \frac{C_1}{2}\left(B_t+\frac{2\sqrt{x}}{\sqrt{C_0}}\right)\left(1+\sqrt{\frac{C_1}{2}}\left(B_t+\frac{2\sqrt{x}}{\sqrt{C_0}}\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
&\leq \frac{C_1}{2}\Bigg(B_t + \frac{x+1}{\sqrt{C_0}} + \sqrt{\frac{C_1}{2}}B_t^2 \\
&\qquad + \frac{4}{\sqrt{C_0}}\sqrt{\frac{C_1}{2}}\left(\frac{B_t^2}{2} + \frac{x}{2}\right) + \frac{4}{C_0}\sqrt{\frac{C_1}{2}}x\Bigg) \\
&\leq C(1 + B_t + B_t^2 + x), \qquad (5.3.13)
\end{aligned}
$$

where C depends only on C_0 and C_1. Finally, since $|a'(x)| \leq C_1(1 + x^{p_1}),\ x \geq 0, p_1 \in \mathbb{N}$ we get

$$
\begin{aligned}
|(J^2)''(B_t + G(x)s)| &= \frac{1}{2}\big(a(J^2(B_t + G(x)s)) \\
&\qquad + J^2(B_t + G(x)s)a'(J^2(B_t + G(x)s))\big) \\
&\leq \frac{1}{2}\left(C_1(1 + J^2(B_t + G(x)s)) + \frac{C_1}{2}\left(B_t^2 + \frac{4}{C_0}x\right)\right. \\
&\qquad \left. \times C_1(1 + J^{2p_1}(B_t + G(x)s))\right) \\
&\leq C(1 + B_t^2 + B_t^{2p_1+2} + x^{p_1+1}), \qquad (5.3.14)
\end{aligned}
$$

where C depends only on C_0, C_1 and p_1.

Combination of estimates (5.3.12)–(5.3.14) leads to

$$
\left|\frac{\partial}{\partial x}\mathbb{E}f(X_t(x))\right| \leqslant \widetilde{C}_1(1 + x^{\tilde{k}_1}), \qquad x \geqslant 0, t \in [0, T],
$$

where the constants $\widetilde{C}_1$ and $\tilde{k}_1$ depend only on C_0, C_1 and p_1.

For the second derivative, denoting $H = H(t, x, s, u) = B_t + G(x)su$, we get

$$
\begin{aligned}
\frac{\partial^2}{\partial x^2}\mathbb{E}f(X_t(x)) &= (d(x)h(x))'\,\mathbb{E}\int_{-1}^{1}\Big(f''\big(J^2(B_t + G(x)s)\big)(J^2)'(B_t + G(x)s) \\
&\qquad + f'\big(J^2(B_t + G(x)s)\big)(J^2)''(B_t + G(x)s)\Big)\mathrm{d}s \\
&\quad + (d(x)h(x))^2\,\mathbb{E}\int_{-1}^{1}\int_{-1}^{1}\Big(f^{(iv)}\big(J^2(H)\big)((J^2)'(H))^3 \\
&\qquad + 2f'''\big(J^2(H)\big)(J^2)'(H)(J^2)''(H) \\
&\qquad + f'''\big(J^2(H)\big)(J^2)'(H)(J^2)''(H) \\
&\qquad + f''\big(J^2(H)\big)(J^2)'''(H) \\
&\qquad + f'''\big(J^2(H)\big)((J^2)'(H))^2(J^2)''(H)
\end{aligned}
$$

$$
\begin{aligned}
&+ f''\big(J^2(H)\big)((J^2)''(H))^2 \\
&+ f''\big(J^2(H)\big)(J^2)'(H)(J^2)'''(H) \\
&+ f''\big(J^2(H)\big)(J^2)'(H)(J^2)'''(H) \\
&+ f'\big(J^2(H)\big)(J^2)^{(iv)}(H)\Big)\mathrm{d}u\mathrm{d}s. \qquad (5.3.15)
\end{aligned}
$$

Previously, we have already proved that function $d(x) = \frac{G(x)}{2\sqrt{x}}$ belongs to $C^{2l-2}_{\text{pol}}(\overline{\mathbb{R}}_+)$ if $a \in C^{2l-1}_{\text{pol}}(\overline{\mathbb{R}}_+)$. Thus the first summand has the polynomial growth. The second summand has the polynomial growth too since

$$
\begin{aligned}
(J^2(x))''' &= F_3((J^2(x))', J^2(x), a''(x), a'(x)),\\
(J^2(x))^{(iv)} &= F_4((J^2(x))'', (J^2(x))', J^2(x), a'''(x), a''(x), a'(x)),
\end{aligned}
$$

and $|a''(x)| \le C_2(1 + x^{p_2})$, $|a'''(x)| \le C_3(1 + x^{p_3})$ $x \ge 0, p_2, p_3 \in \mathbb{N}$. Combining, we get

$$
\Big|\frac{\partial^2}{\partial x^2}\mathbb{E}f(X_t(x))\Big| \leqslant \widetilde{C}_2(1 + x^{\tilde{k}_2}), \qquad x \geqslant 0, t \in [0, T],
$$

where the constants $\widetilde{C}_2$ and $\tilde{k}_2$ depend only on C_0, C_1, C_2, C_3, p_1, p_2, and p_3.

Now, for any higher-order derivative $n \leqslant l$, we see that

$$
\begin{aligned}
\frac{\partial^n}{\partial x^n}\mathbb{E}f(X_t(x)) = \widetilde{F}_n\Big(&(d(x)h(x))^{(i)}, (d(x)h(x))^m,\\
&f^{(j)}\big(J^2(B_t + G(x))\big), (J^2)^{(j)}(B_t + G(x))\Big),
\end{aligned}
$$

where $\widetilde{F}_n$ has polynomial growth in all variables, $i = 1, \ldots, n-1$, $m = 1, \ldots, n$, and $j = 1, \ldots, 2n$. From this we see that if $a \in C^{2l-1}_{\text{pol}}(\overline{\mathbb{R}}_+) \cap C_{\text{lin}}(\overline{\mathbb{R}}_+)$, then

$$
\Big|\frac{\partial^l}{\partial x^l}\mathbb{E}f(X_t(x))\Big| \leqslant \widetilde{C}_l(1 + x^{\tilde{k}_l}), \qquad x \geqslant 0, t \in [0, T],
$$

where $\widetilde{C}_l$ and $\tilde{k}_l$ depend only on C_i and p_i, $i = 0, \ldots, 2l-1$.

Finally, inequality (5.3.2) for the derivatives with respect to t and mixed derivatives follows automatically by an induction on l' using that, for $l' \ge 1$ such that $4l' + 2l \le q$,

$$
\partial_x^l\partial_t^{l'}u(t,x) = \partial_x^l\Big(\frac{1}{4}(a(x) + xa'(x))\partial_x\partial_t^{l'-1}u(t,x) + \frac{1}{2}a(x)x\partial_x^2\partial_t^{l'-1}u(t,x)\Big)
$$

$$
\begin{aligned}
&= \frac{1}{4}\sum_{i=0}^{l}\binom{l}{i}a^{(l-i)}(x)\partial_x^{i+1}\partial_t^{l'-1}u(t,x) \\
&\quad + \frac{1}{4}\Big(l\sum_{i=0}^{l-1}\binom{l-1}{i}a^{(l-i)}(x)\partial_x^{i+1}\partial_t^{l'-1}u(t,x) \\
&\quad + x\sum_{i=0}^{l}\binom{l}{i}a^{(l-i+1)}(x)\partial_x^{i+1}\partial_t^{l'-1}u(t,x)\Big) \\
&\quad + \frac{1}{2}\Big(l\sum_{i=0}^{l-1}\binom{l-1}{i}a^{(l-i-1)}(x)\partial_x^{i+2}\partial_t^{l'-1}u(t,x) \\
&\quad + x\sum_{i=0}^{l}\binom{l}{i}a^{(l-i)}(x)\partial_x^{i+2}\partial_t^{l'-1}u(t,x)\Big),
\end{aligned}
$$

where we used the fact that

$$
\begin{aligned}
\partial_x^l a(x)x\partial_x f(x) = l\sum_{i=0}^{l-1}\binom{l-1}{i}a^{(l-i-1)}(x)\partial_x^{i+1}f(x) \\
+ x\sum_{i=0}^{l}\binom{l}{i}a^{(l-i)}(x)\partial_x^{i+1}f(x),
\end{aligned}
\tag{5.3.16}
$$

which is proved in the Appendix. □

Chapter 6

Conclusions

In the doctoral thesis, we construct first- and second-order weak approximations for the WF model using split-step, moments matching, and approximate moment matching techniques. Split-step technique allows us to divide the model into deterministic and stochastic parts, so that we only need to construct a discretization scheme for the stochastic part, as the deterministic part is easily solvable in an explicit way. Moment matching and approximate moments matching techniques enable us to construct discrete random variables to get weak approximations of the desired order.

Also, we provide a probabilistic proof of the regularity of solutions of the backward Kolmogorov equations for the WF equation and, in addition, for the CIR and general Stratonovich equations with square-root diffusion coefficient without relying on existing transition density formulas. Such a regularity is needed for rigorous proofs that potential ("candidate") weak approximations are indeed weak approximations of the corresponding order.

The following contributions are the main results of the thesis:

- A first-order weak approximation of the WF equation;
- A second-order weak approximation of the WF equation;
- Simulation examples of first- and second-order weak approximations of the WF equation;

- A probabilistic proof of the regularity of solutions of the backward Kolmogorov equations for the WF equation;
- A proof of the regularity of solutions of the backward Kolmogorov equations for the CIR equation without relying on its transition density. This allowed us to extend the method to the case of general Stratonovich equations with square-root diffusion coefficient.

Chapter 7

Appendix

In this chapter, we provide additional calculations which we think would only distract the reader if placed elsewhere in the text.

Appendix: Proof of Lemma 5.9

Proof. First, let $n = 5$ ($l = 2$), that is, $f \in C^5(\overline{\mathbb{R}}_+)$. Then, denoting $A := a^2 + b$, for $i = 0, \ldots, 4$, we have

$$\begin{aligned} g_i(x; a, b) &= \frac{f^{(i)}(A + x + 2a\sqrt{x}) - f^{(i)}(A + x - 2a\sqrt{x})}{\sqrt{x}} \\ &= \frac{1}{\sqrt{x}} f^{(i)}(A + x + 2a\sqrt{x}s)\Big|_{s=-1}^{s=1} \\ &= \frac{1}{\sqrt{x}} \int_{-1}^{1} f^{(i+1)}(A + x + 2a\sqrt{x}s) 2a\sqrt{x} \, \mathrm{d}s \\ &= 2a \int_{-1}^{1} f^{(i+1)}(A + x + 2a\sqrt{x}s) \, \mathrm{d}s, \quad x > 0. \end{aligned}$$

From this it follows that

$$\lim_{x \downarrow 0} g_i(x; a, b) = 2a \int_{-1}^{1} f^{(i+1)}(A) \, \mathrm{d}s = 4af^{(i+1)}(A).$$

In particular, every function g_i, $i = 0, \ldots, 4$, is continuously extendable to the whole half-line $\overline{\mathbb{R}}_+ = [0, \infty)$ by defining $g_i(0; a, b) := 4af^{(i+1)}(A)$.

Let, moreover, $f \in C^5_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$ with the estimates

$$|f^{(i)}(x)| \le C_i(1+x^{k_i}), \quad x \ge 0,\ i = 0,1,\ldots,5, \tag{7.0.1}$$

for some constants $C_i > 0$ and $k_i \in \mathbb{N}$, $i = 0,1,\ldots,5$.

Then we have the estimate

$$\begin{aligned}
|g_i(x;a,b)| &\le 2|a| \int_{-1}^{1} |f^{(i+1)}(A+x+2a\sqrt{x}s)|\,\mathrm{d}s \\
&\le 4|a|C_{i+1}\big(1+(A+x+2|a|\sqrt{x})^{k_{i+1}}\big) \\
&\le 4|a|C_{i+1}\big(1+(A+a^2+2x)^{k_{i+1}}\big) \\
&\le C|a|\big(1+A^{k_{i+1}}+x^{k_{i+1}}\big), \quad x \ge 0,
\end{aligned}$$

where C depends on C_{i+1} and k_{i+1} only.

Now let us concentrate ourselves on the derivatives of $g = g_0$ with respect to x. We have

$$\begin{aligned}
g_0'(x;a,b) &= 2a \int_{-1}^{1} f''(A+x+2a\sqrt{x}s)\Big(1+\frac{as}{\sqrt{x}}\Big)\,\mathrm{d}s \\
&= 2a \int_{-1}^{1} f''(A+x+2a\sqrt{x}s)\,\mathrm{d}s \\
&\quad + \frac{2a^2}{\sqrt{x}} \int_{-1}^{1} f''(A+x+2a\sqrt{x}s)s\,\mathrm{d}s \\
&= g_1(x;a,b) + \frac{2a^2}{\sqrt{x}} \int_{-1}^{1} \Big(f''(A+x) \\
&\qquad + \int_{0}^{s} f'''(A+x+2a\sqrt{x}u)2a\sqrt{x}\,\mathrm{d}u\Big)s\,\mathrm{d}s \\
&= g_1(x;a,b) + \frac{2a^2 f''(A+x)}{\sqrt{x}} \int_{-1}^{1} s\,\mathrm{d}s \\
&\quad + 4a^3 \int_{-1}^{1}\int_{0}^{s} f'''(A+x+2a\sqrt{x}u)\,\mathrm{d}u\, s\ \mathrm{d}s
\end{aligned}$$

$$= g_1(x;a,b) + 4a^3 \int_{-1}^{1}\int_{0}^{s} f'''(A + x + 2a\sqrt{x}u)\,\mathrm{d}u\, s\, \mathrm{d}s, \quad x > 0. \tag{7.0.2}$$

(Note that the term at the negative power of x, that is, at $1/\sqrt{x}$, vanishes since $\int_{-1}^{1} s\,\mathrm{d}s = 0$.) From this it follows that there exists the limit

$$\begin{aligned}\lim_{x\downarrow 0} g_0'(x;a,b) &= \lim_{x\downarrow 0} g_1(x;a,b) + 4a^3 \int_{-1}^{1}\int_{0}^{s} f'''(A)\,\mathrm{d}u\, s\, \mathrm{d}s \\ &= 4af''(A) + \frac{8a^3}{3} f'''(A).\end{aligned}$$

In particular, the function $g = g_0$ is continuously differentiable at $x = 0$ and thus belongs to $C^1(\overline{\mathbb{R}}_+)$ since $g_0'(0;a,b) = \lim_{x\downarrow 0} g_0'(x;a,b)$ by the Lagrange theorem.

If, moreover, $f \in C^5_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$ satisfies estimates (3.1.2) for $i \le 5$, then we have the corresponding estimate for g_0':

$$\begin{aligned}|g_0'(x;a,b)| &\le |g_1(x;a,b)| + 4|a|^3 \int_{-1}^{1}\int_{-s}^{s} |f'''(A + x + 2a\sqrt{x}u)|\,\mathrm{d}u\, |s|\, \mathrm{d}s \\ &\le C_2|a|\big(1 + A^{k_2} + x^{k_2}\big) \\ &\qquad + 4|a|^3 \int_{0}^{1}\int_{-s}^{s} C_3(1 + (A + x + 2|a|\sqrt{x}u)^{k_3})\,\mathrm{d}u\, \mathrm{d}s \\ &\le C_2|a|\big(1 + A^{k_2} + x^{k_2}\big) + 4|a|^3 C_3(1 + A^{k_3} + x^{k_3}) \\ &\le C|a|\big(1 + A^k + x^k\big), \quad x \ge 0. \end{aligned} \tag{7.0.3}$$

where C and k depend on $C_{2,3}$, $k_{2,3}$, and $A = a^2 + b$ only.

Thus, we have proved that $g = g_0 \in C^1_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$, provided that $f \in C^5_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$. (In fact, for estimate (7.0.3), it suffices that $f \in C^3_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$.) More precisely, if

$$|f^{(i)}(x)| \le C_i(1 + x^{k_i}), \quad x \ge 0,\ i = 1,2,3,$$

then

$$|g_0^{(j)}(x;a,b)| \le C(1+A^k+x^k), \quad x \ge 0,\ j=0,1,$$

where the constants $C>0$ and $k \in \mathbb{N}$ depend only on C_i and k_i, $i = 1,2,3$, and, in particular, on a good set of the function $f \in C^5_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$.

Now, let us proceed to the second derivative of g_0. From Eq. (7.0.2), denoting $\tilde{A}_u = \tilde{A}(x,u) = A + x + 2a\sqrt{x}u$, we have

$$\begin{aligned}
g_0''(x;a,b) &= g_1'(x;a,b) + 4a^3 \int_{-1}^{1}\int_{0}^{s} f^{(4)}(\tilde{A}_u)\Big(1+\frac{au}{\sqrt{x}}\Big)\,\mathrm{d}u\, s\ \mathrm{d}s \\
&= \left(2a\int_{-1}^{1} f^{(2)}(\tilde{A}_s)\,\mathrm{d}s\right)' + 4a^3\int_{-1}^{1}\int_{0}^{s} f^{(4)}(\tilde{A}_u)\,\mathrm{d}u\, s\ \mathrm{d}s \\
&\quad + \frac{4a^4}{\sqrt{x}}\int_{-1}^{1}\int_{0}^{s} f^{(4)}(\tilde{A}_u)u\,\mathrm{d}us\,\mathrm{d}s + 4a^3\int_{-1}^{1}\int_{0}^{s} f^{(4)}(\tilde{A}_u)\,\mathrm{d}us\,\mathrm{d}s \\
&\quad + \frac{4a^4}{\sqrt{x}}\int_{-1}^{1}\int_{0}^{s}\Big[f^{(4)}(A+x) + \int_{0}^{u} f^{(5)}(\tilde{A}_v)2a\sqrt{x}\mathrm{d}v\Big]u\,\mathrm{d}us\,\mathrm{d}s \\
&= 2a\int_{-1}^{1} f^{(3)}(\tilde{A}_s)\,\mathrm{d}s \\
&\quad + \frac{2a^2}{\sqrt{x}}\int_{-1}^{1}\Big[f^{(3)}(A+x) + \int_{0}^{s} f^{(4)}(\tilde{A}_u)2a\sqrt{x}\,\mathrm{d}u\Big]\,s\ \mathrm{d}s \\
&\quad + 4a^3\int_{-1}^{1}\int_{0}^{s} f^{(4)}(\tilde{A}_u)\,\mathrm{d}us\,\mathrm{d}s + \frac{4a^4}{\sqrt{x}}f^{(4)}(A)\int_{-1}^{1}\int_{0}^{s} u\,\mathrm{d}us\,\mathrm{d}s \\
&\quad + 8a^5\int_{-1}^{1}\int_{0}^{s}\int_{0}^{u} f^{(5)}(\tilde{A}_v)\mathrm{d}v\, u\ \mathrm{d}u\, s\ \mathrm{d}s \\
&= 2a\int_{-1}^{1} f^{(3)}(\tilde{A}_s)\,\mathrm{d}s + 4a^3\int_{-1}^{1}\int_{0}^{s} f^{(4)}(\tilde{A}_u)\,\mathrm{d}u\, s\ \mathrm{d}s \\
&\quad + 4a^3\int_{-1}^{1}\int_{0}^{s} f^{(4)}(\tilde{A}_u)\,\mathrm{d}u\, s\ \mathrm{d}s
\end{aligned}$$

$$+ 8a^5 \int_{-1}^{1} \int_{0}^{s} \int_{0}^{u} f^{(5)}(\tilde{A}_v) \mathrm{d}v \, u \, \mathrm{d}u \, s \, \mathrm{d}s, \quad x > 0.$$

(Note that, again, the term at the negative power of x, that is, at $1/\sqrt{x}$, vanishes since $\int_{-1}^{1} \int_0^s u \, \mathrm{d}u \, s \, \mathrm{d}s = 0$.) In particular, again by the Lagrange theorem, g_0 is twice continuously differentiable on the whole half-line $\overline{\mathbb{R}}_+$ since there exists the finite limit

$$\begin{aligned}
\lim_{x \downarrow 0} g_0''(x; a, b) &= \lim_{x \downarrow 0} g_1'(x; a, b) + 4a^3 f^{(4)}(A) \int_{-1}^{1} \int_{0}^{s} \mathrm{d}u \, s \, \mathrm{d}s \\
&\quad + 8a^5 f^{(5)}(A) \int_{-1}^{1} \int_{0}^{s} \int_{0}^{u} \mathrm{d}v \, u \, \mathrm{d}u \, s \, \mathrm{d}s \\
&= 4a f^{(3)}(A) + \frac{16a^3 f^{(4)}(A)}{3} + \frac{16a^5 f^{(5)}(A)}{15}.
\end{aligned}$$

If, moreover, $f \in C^5_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$ satisfies estimates (7.0.1), then we have the corresponding estimate for g_0'':

$$\begin{aligned}
|g_0''(x, a, b)| &\leqslant 2|a| \int_{-1}^{1} |f^{(3)}(A + x + 2a\sqrt{x}s)| \, \mathrm{d}s \\
&\quad + 8|a|^3 \int_{-1}^{1} \int_{-s}^{s} |f^{(4)}(A + x + 2a\sqrt{x}u)| \, \mathrm{d}u \, |s| \, \mathrm{d}s \\
&\quad + 8|a|^5 \int_{-1}^{1} \int_{-s}^{s} \int_{-u}^{u} |f^{(5)}(A + x + 2a\sqrt{x}v)| \mathrm{d}v \, |u| \, \mathrm{d}u \, |s| \, \mathrm{d}s \\
&\leqslant 2|a| \int_{-1}^{1} C_3(1 + (A + x + 2|a|\sqrt{x}s)^{k_3}) \, \mathrm{d}s \\
&\quad + 8|a|^3 \int_{0}^{1} \int_{-s}^{s} C_4(1 + (A + x + 2|a|\sqrt{x}u)^{k_4}) \, \mathrm{d}u \, \mathrm{d}s \\
&\quad + 8|a|^5 \int_{0}^{1} \int_{-s}^{s} \int_{-u}^{u} C_5(1 + (A + x + 2|a|\sqrt{x}v)^{k_5}) \mathrm{d}v \, \mathrm{d}u \, \mathrm{d}s \\
&\leqslant C|a|(1 + A^k + x^k), \quad x \geq 0,
\end{aligned} \tag{7.0.4}$$

where the constants $C > 0$ and $k \in \mathbb{N}$ depend only on C_i and k_i, $i = 3,4,5$, and, in particular, on a good set of the function $f \in C^5_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$.

Now, for $l > 2$, we can proceed similarly. For $f \in C^{2l+1}_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$, denote

$$\begin{aligned} F^{0,q} &= F^{0,q}(x,a,b) := \int_{-1}^{1} f^{(q)}(A + x + 2a\sqrt{x}s)\,\mathrm{d}s, \\ F^{p,q} &= F^{p,q}(x,a,b) \\ &:= \int_{-1}^{1}\int_{0}^{s_1}\cdots\int_{0}^{s_p} f^{(q)}(A + x + 2a\sqrt{x}s_{p+1})\,\mathrm{d}s_{p+1}\ldots s_2\,\mathrm{d}s_2\,s_1\,\mathrm{d}s_1, \\ &\qquad\qquad p = 1,\ldots,l,\ q = l+1,\ldots,2l+1. \end{aligned}$$

Then, in addition to the first two derivatives

$$g_0'(x,a,b) = a(2F^{0,2} + 4a^2F^{1,3}) \quad \text{and} \quad g_0''(x,a,b) = a(2F^{0,3} + 8a^2F^{1,4} + 8a^4F^{2,5}),$$

we get

$$\begin{aligned} g_0'''(x,a,b) &= a(2F^{0,4} + 12a^2F^{1,5} + 24a^4F^{2,6} + 16a^6F^{3,7}), \\ &\cdots\cdots\cdots\cdots\cdots\cdots \\ g_0^{(l)}(x,a,b) &= a\sum_{j=0}^{l} c_{j,l}a^{2j}F^{j,l+j+1}(x;a,b), \end{aligned} \tag{7.0.5}$$

where $c_{j,l}$, $0 \le j \le l$, are some constants. Note that, as before, in the right-hand side of Eq. (7.0.5), there are no negative powers of x, so that g_0 is l times continuously differentiable on the whole half-line $\overline{\mathbb{R}}_+$, provided that $f \in C^{2l+1}_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$. Moreover, as before, from (7.0.5) we get the following estimates for $g_0^{(r)}$:

$$|g_0^{(r)}(x,a,b)| \leqslant C|a|(1 + A^k + x^k), \qquad x \geqslant 0,\ r = 0,1,\ldots,l,$$

where the constants $C > 0$ and $k \in \mathbb{N}$ depend only on C_i and k_i, $i = 0,\ldots,2l+1$, that is, only on a good set of the function $f \in C^{2l+1}_{\mathrm{pol}}(\overline{\mathbb{R}}_+)$. □

Appendix: Proof of equality (5.3.16)

Proof. Let us start by taking $l = 1$. Then

$$\begin{aligned}\partial_x a(x) x \partial_x f(x) &= a'(x) x \partial_x f(x) + a(x)\partial_x f(x) + a(x) x \partial_x^2 f(x) \\ &= a(x)\partial_x f(x) + x \sum_{i=0}^{1} \binom{1}{i} a^{(1-i)}(x)\partial_x^{i+1} f(x).\end{aligned}$$

When $l = 2$, we have

$$\begin{aligned}\partial_x^2 a(x) x \partial_x f(x) &= (2a'(x) + xa''(x))\partial_x f(x) + (2a(x) + 2xa'(x))\partial_x^2 f(x) \\ &\quad + a(x) x \partial_x^3 f(x) \\ &= 2\sum_{i=0}^{1} \binom{1}{i} a^{(1-i)}(x)\partial_x^{i+1} f(x) \\ &\quad + x \sum_{i=0}^{2} \binom{2}{i} a^{(2-i)}(x)\partial_x^{i+1} f(x).\end{aligned}$$

Let us assume that for $l = k-1$, we have

$$\begin{aligned}\partial_x^{k-1} a(x) x \partial_x f(x) &= (k-1)\sum_{i=0}^{k-2} \binom{k-2}{i} a^{(k-2-i)}(x)\partial_x^{i+1} f(x) \\ &\quad + x \sum_{i=0}^{k-1} \binom{k-1}{i} a^{(k-1-i)}(x)\partial_x^{i+1} f(x).\end{aligned}$$

Then

$$\begin{aligned}\partial_x^{k} a(x) x \partial_x f(x) &= \partial_x \partial_x^{k-1} a(x) x \partial_x f(x) \\ &= (k-1)\sum_{i=0}^{k-2} \binom{k-2}{i} a^{(k-1-i)}(x)\partial_x^{i+1} f(x) \\ &\quad + (k-1)\sum_{i=0}^{k-2} \binom{k-2}{i} a^{(k-2-i)}(x)\partial_x^{i+2} f(x) \\ &\quad + \sum_{i=0}^{k-1} \binom{k-1}{i} a^{(k-1-i)}(x)\partial_x^{i+1} f(x) \\ &\quad + x \sum_{i=0}^{k-1} \binom{k-1}{i} a^{(k-i)}(x)\partial_x^{i+1} f(x) \\ &\quad + x \sum_{i=0}^{k-1} \binom{k-1}{i} a^{(k-1-i)}(x)\partial_x^{i+2} f(x)\end{aligned}$$

$$
\begin{aligned}
&= (k-1)a^{(k-1)}(x)\partial_x f(x) \\
&\quad + (k-1)\sum_{i=1}^{k-2}\binom{k-2}{i}a^{(k-1-i)}(x)\partial_x^{i+1}f(x) \\
&\quad + (k-1)\sum_{i=1}^{k-2}\binom{k-2}{i-1}a^{(k-1-i)}(x)\partial_x^{i+1}f(x) \\
&\quad + (k-1)a(x)\partial_x^k f(x) \\
&\quad + a^{(k-1)}(x)\partial_x f(x) + a(x)\partial_x^k f(x) \\
&\quad + \sum_{i=1}^{k-2}\binom{k-1}{i}a^{(k-1-i)}(x)\partial_x^{i+1}f(x) \\
&\quad + xa^{(k)}(x)\partial_x f(x) \\
&\quad + x\sum_{i=1}^{k-1}\binom{k-1}{i}a^{(k-i)}(x)\partial_x^{i+1}f(x) \\
&\quad + x\sum_{i=1}^{k-1}\binom{k-1}{i-1}a^{(k-i)}(x)\partial_x^{i+1}f(x) \\
&\quad + xa(x)\partial_x^{k+1}f(x) \\
&= k\sum_{i=0}^{k-1}\binom{k-1}{i}a^{(k-1-i)}(x)\partial_x^{i+1}f(x) \\
&\quad + x\sum_{i=0}^{k}\binom{k}{i}a^{(k-i)}(x)\partial_x^{i+1}f(x),
\end{aligned}
$$

because $\binom{k-1}{i} + \binom{k-1}{i-1} = \binom{k}{i}$. $\square$

Bibliography

[1] A. Ahdida and A. Alfonsi, *A mean-reverting SDE on correlation matrices*, Stochastic Processes and their Applications, Elsevier **123** (2013), no. 4, 1472–1520.

[2] A. Alfonsi, *On the discretization schemes for the CIR (and Bessel squared) processes*, Monte Carlo Methods and Applications **11** (2005), no. 4, 355–384.

[3] ———, *High order discretization schemes for the CIR process: Application to affine term structure and Heston models*, Mathematics of Computation, American Mathematical Society **79** (2010), no. 269, 209–237.

[4] ———, *Affine Diffusions and Related Processes: Simulation, Theory and Applications*, 1 ed., Bocconi & Springer Series, vol. 6, Springer, 2015.

[5] P.E. Arrouy, S. Mehalla, B. Lapeyre, and A. Boumezoued, *Jacobi stochastic volatility factor for the Libor market model*, Working Papers hal-02468583, HAL, 2020.

[6] G. Bernis and S. Scotti, *Alternative to beta coefficients in the context of diffusions*, Quantitative Finance **17** (2017), 275–288.

[7] F. Black and M. Scholes, *The pricing of options and corporate liabilities*, Journal of Political Economy **81** (1973), no. 3, 637–654.

[8] K.C. Chan, G.A. Karolyi, F.A. Longstaff, and A.B. Sanders, *An empirical investigation of alternative models of the short-term interest rate*, Journal of Finance (1992), no. 47, 1209–1227.

[9] S. Cheng, *Differentiation under the integral sign with weak derivatives*, Tech. report, 2006.

[10] J.C. Cox, *Notes on option pricing I: Constant elasticity of variance diffusions*, Working Paper, Stanford University, 1975 Reprinted in *The Journal of Portfolio Management* (1996), no. 23, 15–17.

[11] J.C. Cox, J.E. Ingersoll, and S.A. Ross, *A theory of the term structure of interest rates*, Econometrica (1985), no. 53, 385–407.

[12] F. Delbaen and H. Shirakawa, *An interest rate model with upper and lower bounds*, Asia-Pacific Financial Markets **9** (2002), 191–209.

[13] C.L. Epstein and R. Mazzeo, *Wright–Fisher diffusion in one dimension*, SIAM J. Math. Anal. **42** (2010), 568–608.

[14] S.N. Ethier and T.G. Kurtz, *Markov Processes. Characterization and Convergence*, 2 ed., John Wiley & Sons, Inc., Hoboken. New Jerse, 2005.

[15] N. Gabrielli, *Affine processes from the perspective of path space valued Lévy processes*, Ph.D Thesis, ETH Zurich (2014).

[16] C. Gourieroux and J. Jasiak, *Multivariate Jacobi process with application to smooth transitions*, Journal of Econometrics **131** (2006), 475–505.

[17] J.B.S. Haldane, *A mathematical theory of natural and artificial selection. Part II: The influence of partial self-fertilisation, inbreeding, assortative mating, and selective fertilisation on the composition of Mendelian populations, and on natural selection*, Biological Reviews **1** (1924).

[18] S. Heston, *A closed-form solution for options with stochastic volatility with applications to bond and currency options*, The Review of Financial Studies **6** (1993), no. 2, 327–343.

[19] D.J. Higham, X. Mao, and M. Stuart, *Strong convergence of Euler-type methods for nonlinear stochastic differential equations*, SIAM Journal on Numerical Analysis **40** (2003), no. 3, 1041—-1063.

[20] M. Jeanblanc, M. Yor, and M. Chesney, *Mathematical Methods for Financial Markets*, 1 ed., Springer, 2009.

[21] S. Karlin and H.M. Taylor, *A Second Course in Stochastic Processes*, 1 ed., Academic Press, New York, 1981.

[22] M. Kimura, *Solution of a process of random genetic drift with a continuous model*, Proceedings of the National Academy of Sciences of the United States of America **41** (1955), 144–150.

[23] ———, *Diffusion models in population genetics*, Journal of Applied Probability **1** (1964), 177–232.

[24] P.E. Kloeden and E. Platen, *Numerical Solution of Stochastic Differential Equations*, 1 ed., Springer, 1992.

[25] S.G. Krantz and H.R Parks, *A Primer of Real Analytic Functions*, 2 ed., Birkhäuser Boston, 2002.

[26] A. Lenkšas and V. Mackevičius, *Weak approximation of Heston model by discrete random variables*, Mathematics and Computers in Simulation **113** (2015), 1–15.

[27] ———, *A second-order weak approximation of Heston model by discrete random variables*, Lithuanian Mathematical Journal **55** (2015), 555–572.

[28] P. Lévy, *Processus Stochastiques et Mouvement Brownien*, 2 ed., Paris, Gauthier-Villars, 1965.

[29] G. Lileika and V. Mackevičius, *Weak approximation of CKLS and CEV processes by discrete random variables*, Lithuanian Mathematical Journal **60** (2020), no. 2, 208–224.

[30] ______, *Second-order weak approximations of CKLS and CEV processes by discrete random variables.*, Mathematics **9** (2021).

[31] J. Ma, *Pricing foreign equity options with stochastic correlation and volatility*, Annals of Economics and Finance **10** (2009), 303–327.

[32] V. Mackevičius, *Introduction to Stochastic Analysis: Integrals and Differential Equations*, 1 ed., Wiley, 2011.

[33] ______, *Weak approximation of CIR equation by discrete random variables*, Lith. Math. J. **51** (2011), no. 3, 385–401.

[34] G. Milstein and M. V. Tretyakov, *Stochastic Numerics for Mathematical Physics*, 1 ed., Scientific Computation, Springer-Verlag, 2004.

[35] E. Moro and H. Schurz, *Boundary preserving semianalytic numerical algorithms for stochastic differential equations*, SIAM Journal on Scientific Computing **29** (2007), no. 4, 1525–1549.

[36] S. Ninomiya and N. Victoir, *Weak approximation of stochastic differential equations and application to derivative pricing*, Applied Mathematical Finance **15** (2008), no. 2, 107–121.

[37] B. Øksendal, *Stochastic Differential Equations: An Introduction with Applications*, 5 ed., Springer, 1998.

[38] V. Paulauskas and A. Račkauskas, *Funkcinė Analizė, Kn. 1*, 1 ed., Vaistų žinios, 2007.

[39] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion*, 3 ed., Grundlehren der mathematischen Wissenschaften, Springer, Berlin, Heidelberg, 1999.

[40] O. Vašíček *An equilibrium characterization of the term structure*, Journal of Financial Economics **5** (1977), no. 2, 177–188.

[41] S. Wright et al., *Adaptation and selection*, Genetics, paleontology and evolution (1949), 365–389.

NOTES

NOTES

Vilnius University Press
9 Saulėtekio Ave., Building III, LT-10222 Vilnius
Email: info@leidykla.vu.lt, www.leidykla.vu.lt
bookshop.vu.lt, journals.vu.lt
Print run 20