University of Kaiserslautern
Department of Mathematics

Edgeworth Expansions for Binomial Trees

Alona Bock

Datum der Disputation: 22.08.2014

Vom Fachbereich Mathematik der Technischen Universität Kaiserslautern zur Verleihung des akademischen Grades Doktor der Naturwissenschaften (Doctor rerum naturalium, Dr. rer. nat.) genehmigte Dissertation.

1. Gutachter: Prof. Dr. Ralf Korn
2. Gutachter: Prof. Dr. Thomas Gerstner

D 386

Acknowledgements

First of all, I would like to thank my supervisor, Prof. Dr. Ralf Korn, for his advice and motivation, for his continuous support and understanding throughout my PhD studies, and for giving me this opportunity in the first place. This thesis would have been impossible without him.

I would also like to express my gratitude to Prof. Dr. Thomas Gerstner, at the Goethe-University Frankfurt, for acting as a referee for this thesis.

I am very grateful to my colleagues in the Financial Mathematics group at the Kaiserslautern University of Technology for all the helpful discussions.

I would like to thank my family and friends, who have been an amazing support system. I am especially grateful to my wonderful fiancé, Wolfgang Bock, for his love, help and encouragement, and for being there for me every step of the way.

Finally, I would like to thank the Department of Mathematics and the Kaiserslautern University for the financial support.

Contents

Chapter 1

Introduction

The binomial approach goes back to W.F. Sharpe and J.C. Cox, S.A. Ross and M. Rubinstein. It was introduced in [CRR79] as an approximation to the Black-Scholes model, in the sense that the prices of vanilla options computed in the binomial model converge to the Black-Scholes formula. It has gained a lot of popularity since, and has become a widespread tool in option pricing. Due to its simplicity, the binomial model is widely used to demonstrate the principals of arbitrage opportunities and risk-neutral pricing. However, in this thesis we concentrate on the binomial approach as a purely numerical method. We do not consider it as discretization of the continuous-time market, but rather as an approximation of the underlying distribution. This way we are not restricted to the risk-neutral measure, which gives us more freedom in the construction of the binomial models.

The binomial approach is based on the concept of weak convergence. The discrete-time model $S^{(n)}$ is constructed so as to ensure convergence in distribution to the continuous process S. This means that the expectations calculated in the binomial tree can be used as approximations of the option prices in the continuous model. Note that since we do not require risk-neutral transition probabilities, the expectations calculated in the discrete setting are not prices themselves.

Convergence of binomial trees

The binomial method is easy to implement and can be adapted to options with different types of payout structures, including American options. This makes the approach very appealing. However, the problem is that in many cases, the convergence of the method is slow and highly irregular, and even a fine discretization does not guarantee accurate price approximations. What complicates things even further, is that the convergence behavior is highly dependent on the specific option considered. This makes it impossible to construct a binomial tree which would be optimal for all payoff functions.

The reasons for the irregularities in the convergence of lattice methods as well as ways of overcoming them have been a topic of research for quite some time. There are two main sources of error in the binomial method, the discretization error and the payoff error (see e.g. [DKEB95]). The discretization error naturally arises in any discretization scheme, since the values are only considered at certain discrete time points. The effect of this error can be reduced by taking a larger number of points. The payoff error comes from the structure of the payout function (in [DKEB95] this is referred to as the specification error). In the lattice methods the values of the underlying asset at each time point are fixed to certain lattice points. This means that the payoff function of digital or barrier options, for example, is calculated in the lattice not with a theoretical strike or barrier, but with a slightly adjusted one, which matches one of the neighboring nodes. Since the adjusted values vary with the number of time steps, the convergence becomes irregular. This effect is present also for a larger number of steps, however, the oscillations do become smaller (see section 2.2). Therefore, this suggests that by controlling the position of the barrier or the strike, we can get a better convergence pattern.

Various approaches have been offered for different types of options. For digital and vanilla options, for example, some of the possible solutions can be found in [T99], [CP07], [KM13], for barrier options see e.g. [BL94] and [R95], for American options see [Lei98] and [KM13]. Of course, these are just a few examples and more literature is available on the topic. However, even though different methods have been offered, not that many theoretical results are available. In [DD04] second order expansions are presented for the digital and vanilla options in the one-dimensional Black-Scholes model. This is an important result, since it allows a theoretical comparison of different existing models and gives a better understanding of the asymptotic behavior of binomial models. This information can then be used to construct advanced binomial trees with a superior order of convergence (see [CP07] and [KM13]). In the case of barrier options a first order expansion is available in [G99]. However, the existing expansions in the literature mostly deal with one-dimensional binomial trees. There are hardly any results on the asymptotic behavior of multinomial or multidimensional trees. Therefore, in this thesis we present a framework that allows to construct advanced lattice models in a more general setting. For this purpose we will apply Edgeworth expansions.

Edgeworth expansions

Edgeworth expansions (see e.g. [Edge1905]) have been introduced as a generalization of the central limit theorem. They allow to present the distribution function of a normalized sum as a series of the normal distribution function and its derivatives, therefore, giving a better approximation than the CLT.

We apply an altered version of the Edgeworth expansion for lattice triangular arrays to obtain the asymptotics of the various option prices in the binomial model. The advantage of the Edgeworth expansion approach is that it is valid under very general conditions and can be extended to higher dimensions as well as multinomial trees.

Outline of the thesis

Chapter 2 contains an introduction to binomial trees. Here we present the basic definitions of weak convergence and the fundamental theory behind the lattice approach. In the second part of the chapter we present examples of commonly used trees and discuss some of the practical issues.

Chapter 3 contains the main theory on Edgeworth expansions. In binomial models we usually deal with triangular arrays of lattice random vectors. In this case the available expansions for lattices are not directly applicable. Therefore, the purpose of this chapter is to present Edgeworth expansions, which are also valid for the binomial tree setting. Since Edgeworth expansions are based on a Taylor series of the corresponding characteristic functions, we start with some basic theory on Fourier transforms. We then present the heuristics of the method. The main theoretical results are presented in the final section of the chapter.

The last two chapters present the applications of Edgeworth expansions to lattice models. Chapter 4 contains the theory in the one-dimensional case. Here we are able to obtain a third order expansion for general binomial and trinomial trees. Then advanced models are constructed for various types of options. For digital and vanilla options a $O\left(\frac{1}{n^2}\right)$ order of convergence is achieved. We present a way to obtain asymptotics for the price of a barrier option and construct an advanced model with a $O\left(\frac{1}{n^{3/2}}\right)$ order of convergence. In conclusion we discuss the convergence behavior of the Greeks in the Black-Scholes model.

In Chapter 5 we present expansions for multidimensional binomial trees. Here we confirm theoretically the $O\left(\frac{1}{\sqrt{n}}\right)$ order of convergence of the distribution function of a multidimensional tree at maturity to the corresponding normal distribution. Then we obtain expansions of order $O\left(\frac{1}{\sqrt{n}^3}\right)$ for the standard 2D binomial models and construct advanced binomial trees for the two-asset digital and the two-asset correlation options with a $O\left(\frac{1}{\sqrt{n}^3}\right)$ order of convergence. We also present advanced binomial models for a multidimensional setting. Finally, we consider ways of improving the convergence behavior of the decoupling approach with the Cholesky decomposition (see [KM09]).

Chapter 2

Fundamentals of the Binomial Method

In the theory of option pricing one is usually concerned with evaluating expectations under the risk-neutral measure Q

$$E_Q\left(e^{-rT}g\left(S\left(t\right),t\in[0,T]\right)\right), \tag{2.1}$$

where g is the payoff function of the financial instrument with maturity T, and S is the underlying. However, in many cases this expectation cannot be calculated explicitly and numerical methods need to be applied to approximate the desired quantity. Monte Carlo simulations, numerical methods for PDEs and the lattice approach are the methods typically employed. In this thesis we will consider the latter approach, with the main focus on binomial trees.

2.1 Binomial trees in theory

We start with a one-dimensional setting and consider the Black-Scholes model, where the risk-neutral dynamics of the underlying stock S can be represented as

$$dS(t)=S\left(t\right)\left(rdt+\sigma dW\left(t\right)\right),\quad S\left(0\right)=s_0,$$

with the risk-free interest rate r, volatility $\sigma>0$ and Brownian motion W with respect to the risk-neutral measure Q.

The idea of the lattice approach is to construct a discrete-time process $S^{(n)}$, such that its linear interpolation converges weakly to the given process S. The expectations calculated in the discrete setting can then be used as approximations for (2.1). However, before we move on to the construction of $S^{(n)}$ we first present some basic definitions and properties of the concept of weak convergence. For a detailed coverage of the topic we refer to [B68].

2.1.1 Weak convergence

Definition 2.1.1. Let M be a metric space with the Borel σ-field $\mathcal{B}(M)$, and let P_n, $n \in \mathbb{N}$ and P be probability measures on $(M, \mathcal{B}(M))$. The sequence $\{P_n\}_n$ is said to *converge weakly* to P $(P_n \Rightarrow P)$, if

$$\lim_{n\to\infty} \int_M f(x)\, dP_n(x) = \int_M f(x)\, dP(x),$$

for every bounded, continuous real-valued function f on M.

The concept of weak convergence (convergence in distribution) for random elements is equivalent to the weak convergence of the corresponding distributions. In other words, given a sequence $X_n : (\Omega_n, \mathcal{F}_n, P_n) \to (M, \mathcal{B}(M))$ and $X : (\Omega, \mathcal{F}, P) \to (M, \mathcal{B}(M))$ we say that $\{X_n\}_n$ converges weakly to X $\left(X_n \xrightarrow{\mathcal{D}} X\right)$ iff

$$P_n X_n^{-1} \Rightarrow P X^{-1}, \quad n \to \infty,$$

where $P_n X_n^{-1}$, PX^{-1} are probability measures on $(M, \mathcal{B}(M))$ defined as

$$P_{(n)} X_{(n)}^{-1}(A) = P_{(n)}(X \in A), \quad A \in \mathcal{B}(M).$$

Note that the above definition only makes sense if all the X_n and X have the same range, the underlying probability spaces, on the other hand, may all differ.

Remark 2.1.2. If $M = \mathbb{R}^k$, then weak convergence is equivalent to convergence of the corresponding distribution functions, i.e.

$$P_n \Rightarrow P \quad \text{iff} \quad F_n(x) \to F(x), \quad \text{for each continuity point } x \text{ of } F.$$

However, if $M = \mathcal{C}([0,1])$, i.e we are considering weak convergence of stochastic processes, then convergence of the finite-dimensional distributions is not enough. In addition we have to make sure that the sequence $\{P_n\}$ is tight.

A very useful theorem in the theory of weak convergence is Donsker's theorem, often referred to as the functional central limit theorem (see e.g. [KK01], Theorem 4.12).

Theorem 2.1.3 (Donsker's theorem for triangular schemes)**.** *Let $\xi_{n_1}, \ldots, \xi_{n_{k_n}}$, $n \in \mathbb{N}$, $k_n \in \mathbb{N}$ be i.i.d. random variables with zero mean and positive variance $\sigma^2_{n_1} < \infty$, and let $S_{n_m} = \sum_{l=1}^m \xi_{n_l}$, $s^2_{n_m} = m\sigma^2_{n_1}$, $m = 1, \ldots, k_n$, and $s^2_n = s^2_{n_{k_n}}$. Define the process $X_n(t)$, $t \in [0,1]$ as*

$$X_n(0) = 0,$$
$$X_n\left(s^2_{n_m}/s^2_n\right) = \frac{1}{s_n} S_{n_m}, \quad m = 1, \ldots, k_n,$$

and X_n is linear on the intervals $\left[s^2_{n_{m-1}}/s^2_n, s^2_{n_m}/s^2_n\right]$. Then, if $k_n \to \infty$ and $s_n \to \infty$ as $n \to \infty$, X_n converges weakly to a Brownian motion.

Let us go back to the Black-Scholes model. If we construct a process $S^{(n)}$, that ensures weak convergence to S, then by the Definition 2.1.1

$$E_n\left(f\left(S^{(n)}\right)\right) \to E_Q\left(f\left(S\right)\right), \quad n \to \infty,$$

for all bounded, continuous real-valued functions f. Furthermore, as we will see below, the statement can be shown to hold for a more general set of functions, which includes most types of options used in practice.

2.1.2 The model

We now proceed with the construction of the approximating binomial tree. We divide the time horizon $[0, T]$ into n time steps of equal length Δt. For each step $k = 1, \dots, n$ we define

$$S_k^{(n)} = S_{k-1}^{(n)} e^{\alpha_n \Delta t + \sigma\sqrt{\Delta t}\sum_{i=1}^{k}\xi_i^{(n)}}, \quad S_0^{(n)} = s_0, \tag{2.2}$$

where $\{\alpha_n\}_n$ is a bounded sequence and for each $n \in \mathbb{N}$, $\xi_k^{(n)}$, $k = 1, \dots, n$ are i.i.d random variables such that

$$\xi_k^{(n)} = \begin{cases} 1, & \text{with probability } p_n, \\ -1, & \text{with probability } 1 - p_n. \end{cases}$$

Therefore, $\Omega_n = \{\omega = (\omega_1, \dots, \omega_n) |\ \omega_k = \pm 1,\ k = 1, \dots, n\}$ and P_n is the product of the one-step transition probabilities, i.e.

$$P_n\left(\omega\right) = p_n^{n_u(\omega)}\left(1 - p_n\right)^{n_d(\omega)},$$

where $n_u(\omega)$ is the number of up-jumps, i.e. the number of 1's in the sequence $(\omega_1, \dots, \omega_n)$ and $n_d(\omega) = n - n_u(\omega)$ is the number of down-jumps.

Remark 2.1.4. Similar to the binomial model we can also consider m-nomial trees. In this case, for each time step $k = 1, \dots, n$, the variables $\xi_k^{(n)}$ can attain m different values with appropriate probabilities.

Note that to completely determine the model (2.2) we still have to specify the exact form of the drift α_n and the probability p_n. The goal now is to choose these parameters in such a way that weak convergence is ensured.

Moment-matching conditions

Let $\mu(n)$ and $\sigma^2(n)$ be the mean and variance of the one-period log-returns in the model (2.2), i.e.

$$\begin{aligned} \mu\left(n\right) &= \frac{1}{\Delta t} E_n\left(\log\left(\frac{S_k^{(n)}}{S_{k-1}^{(n)}}\right) \middle| S_{k-1}^{(n)}\right) = \alpha_n + \frac{1}{\sqrt{\Delta t}}\sigma E_n\left(\xi_1^{(n)}\right), \\ \sigma^2(n) &= \frac{1}{\Delta t} Var_n\left(\log\left(\frac{S_k^{(n)}}{S_{k-1}^{(n)}}\right) \middle| S_{k-1}^{(n)}\right) = \sigma^2 Var_n\left(\xi_1^{(n)}\right). \end{aligned} \tag{2.3}$$

In order to guarantee weak convergence, $\mu(n)$ and $\sigma^2(n)$ need to match the corresponding log-returns of the continuous process S, at least asymptotically. Therefore, we assume

$$\begin{aligned} \mu(n) &= r - \frac{1}{2}\sigma^2 + o(1), \quad n \to \infty, \\ \sigma^2(n) &= \sigma^2 + o(1), \quad n \to \infty. \end{aligned} \tag{2.4}$$

Due to Donsker's Theorem 2.1.3 we have the following theorem.

Theorem 2.1.5. *Let $S^{(c,n)}$ be constructed from $S^{(n)}$ by linear interpolation, i.e.*

$$S^{(c,n)} = \exp\left(\log\left(S^{(n)}_{\lfloor t/\Delta t\rfloor}\right) + (t/\Delta t - \lfloor t/\Delta t\rfloor)\left(\log\left(S^{(n)}_{\lfloor t/\Delta t+1\rfloor}\right) - \log\left(S^{(n)}_{\lfloor t/\Delta t\rfloor}\right)\right)\right).$$

Then under the moment-matching conditions (2.4)

$$S^{(c,n)} \xrightarrow{\mathcal{D}} S, \quad n \to \infty.$$

Furthermore,

$$E_n\left(e^{-rT} g\left(S^{(c,n)}\right)\right) \to E\left(g(S)\right), \quad n \to \infty,$$

for all continuous almost everywhere payoff functionals g, such that $\left\{g\left(S^{(c,n)}\right)\right\}_n$ is uniformly integrable.

Note that we consider weak convergence of the continuous time process $S^{(c,n)}$ and not the binomial tree $S^{(n)}$ directly, since as mentioned before, the approximating sequence and the limiting process need to have the same image space. For a detailed proof of Theorem 2.1.5 we refer the reader to [M09], Proposition 4 and Proposition 17.

Let us take a closer look at the moment-matching conditions (2.4). What does this actually mean for the parameters of the binomial tree (2.2) α_n and p_n?

Proposition 2.1.6. *The conditions (2.4) are satisfied iff*

$$p_n = \frac{1}{2} + \frac{r - \frac{1}{2}\sigma^2 - \alpha_n}{2\sigma}\sqrt{\Delta t} + o\left(\frac{1}{\sqrt{n}}\right), \quad n \to \infty. \tag{2.5}$$

Proof. Since $E_n\left(\xi_1^{(n)}\right) = 2p_n - 1$ and $Var_n\left(\xi_1^{(n)}\right) = 4p_n(1-p_n)$, (2.4) is equivalent to

$$\begin{aligned} \alpha_n + \frac{1}{\sqrt{\Delta t}}\sigma(2p_n - 1) &= r - \frac{1}{2}\sigma^2 + o(1), \\ p_n(1-p_n) &= \frac{1}{4} + o(1). \end{aligned}$$

The statement of the proposition follows directly. Note that the condition on the variance is implied by the condition on the mean. □

Obviously, (2.5) leaves a degree of freedom in the choice of α_n and p_n, and there are infinitely many combinations that satisfy the above proposition. Depending on these parameters the binomial tree will have different convergence properties. In section 2.2 we present the most typical models used in practice and discuss their advantages and disadvantages.

The algorithm

We now use $E_n\left(g\left(S^{(n)}\right)\right)$ to approximate the price $E_Q\left(g\left(S\right)\right)$. Consider a path-independent European option, i.e. an option with payoff

$$g\left(S\left(T\right)\right),$$

which can only be exercised at maturity. To calculate the expectation in the binomial model a simple backward induction algorithm can be used.

Step 1: Tree Initialization

- Calculate the possible values of the stock at maturity

$$S_n^{(n)}\left(l\right) = s_0 e^{\alpha_n T+\sigma\sqrt{\Delta t}(2l-n)}, \quad l = 0, \dots n.$$

- Calculate the option values at maturity

$$V_n(l) = g\left(S_n^{(n)}\left(l\right)\right), \quad l = 0, \dots n.$$

Step 2: Backward Induction

- For each time step $i = n-1, \dots, 0$ calculate the value of the option as the weighted sum of the values of the successor nodes (figure 2.1)

$$V_i\left(l\right) = p_n V_{i+1}\left(l+1\right) + (1-p_n)V_{i+1}\left(l\right), \quad l = 0, \dots, i.$$

Step 3: Return $e^{-rT}V_0\left(0\right)$.

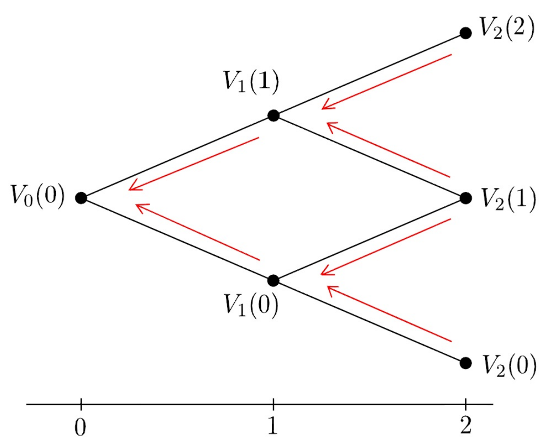


Figure 2.1: Backward induction for a 2-period binomial tree

The computational effort of the above algorithm is $O\left(n\right)$ for the tree initialization and $O\left(n^2\right)$ for the backward algorithm, i.e. $O\left(n^2\right)$ in total (see [M09], Proposition 22). For a d-dimensional asset pricing problem the computational effort would be $O\left(n^{d+1}\right)$ (see [Ku08], Lemma 3.2.1).

Remark 2.1.7. This algorithm can also be adjusted to incorporate path-dependent options. Depending on the exact payoff structure additional calculations are required in the backward induction step (see e.g. [M09]).

2.1.3 Binomial trees for American options

One of the main applications of the lattice approach is the valuation of American options. Unlike European options, American options can be exercised at any time before maturity. The optimal exercise time is unknown and is usually represented by a random stopping time τ^*. This makes calculating prices far more difficult. Closed-form solutions are hardly ever available, therefore, numerical methods have to be applied in most cases. The advantage of the binomial approach is that with a simple adjustment of the backward induction step, the algorithm described above can be applied to American options as well.

The price of an American contingent claim with payout function g is given by

$$\sup_{\tau \in \mathcal{T}_{0,T}} E\left(e^{-r\tau} g\left(S\left(\tau\right)\right)\right),$$

where $\mathcal{T}_{0,T}$ is the set of all stopping times with respect to the natural filtration of W with values in $[0,T]$. There exists an optimal stopping time τ^* such that the price can be represented as

$$E\left(e^{-r\tau^*} g\left(S\left(\tau^*\right)\right)\right).$$

The approximation of this expectation in the binomial model is given by

$$E_n\left(e^{-r\tau_n^*\Delta t} g\left(S_{\tau_n^*}^{(n)}\right)\right),$$

for an optimal stopping time τ_n^*. It has been shown (see [M09] Chapter 2) that a theorem similar to 2.1.5 also holds for American options.

In order to calculate the expectation in the binomial model we need to check at each node whether the option should be exercised or not, i.e. at every time step $i = n-1, \ldots, 0$ we have to compare the continuation value $E_n\left(e^{-r\Delta t} V_{i+1} \middle| S_i^{(n)}\right)$ and the early exercise value $g\left(S_i^{(n)}\right)$. The optimal stopping time τ_n^* is then defined as

$$\tau_n^* = \min\left\{ i = 0, \ldots, n \middle| V_i = g\left(S_i^{(n)}\right)\right\}.$$

The algorithm above can be modified as follows. The tree initialization remains the same.

Step 2: Backward Induction

- For each time step $i = n-1, \ldots, 0$

$$V_i\left(l\right) = \max\left\{e^{-r\Delta t}\left(p_n V_{i+1}\left(l+1\right) + (1-p_n) V_{i+1}\left(l\right)\right), g\left(S_i^{(n)}\left(l\right)\right)\right\}, \quad l = 0, \ldots i.$$

Step 3: Return $V_0(0)$.

For more details on American option pricing we refer to [KK01] and [M09].

2.2 Binomial trees in practice

We now present some of the standard binomial models and mention the issues that usually arise in practice.

2.2.1 One-dimensional trees

There are numerous different trees available in the literature. Even for the simple Black-Scholes setting there exist various examples. A comparison of some of the popular methods can be found in e.g. [J09]. Here we present only the two basic models.

The CRR tree

The binomial model was first introduced in [CRR79] by J. C. Cox, S. A. Ross and M. Rubinstein. They suggest to construct a tree, which is symmetric in the log-scale around $\log(s_0)$, i.e.

$$\alpha_n = 0.$$

The probability of an up-jump is then chosen as

$$p_n = \frac{1}{2} + \frac{r - \frac{1}{2}\sigma^2}{2\sigma}\sqrt{\Delta t}.$$

Note that the probability is well-defined only for n large enough. The moments are matched asymptotically

$$\mu(n) = r - \frac{1}{2}\sigma^2,$$
$$\sigma^2(n) = \sigma^2 - \left(r - \frac{1}{2}\sigma^2\right)^2 \Delta t.$$

The main advantage of the CRR tree is its symmetric construction. As we will see later on this allows to directly apply some useful results for random walks to the process $S^{(n)}$.

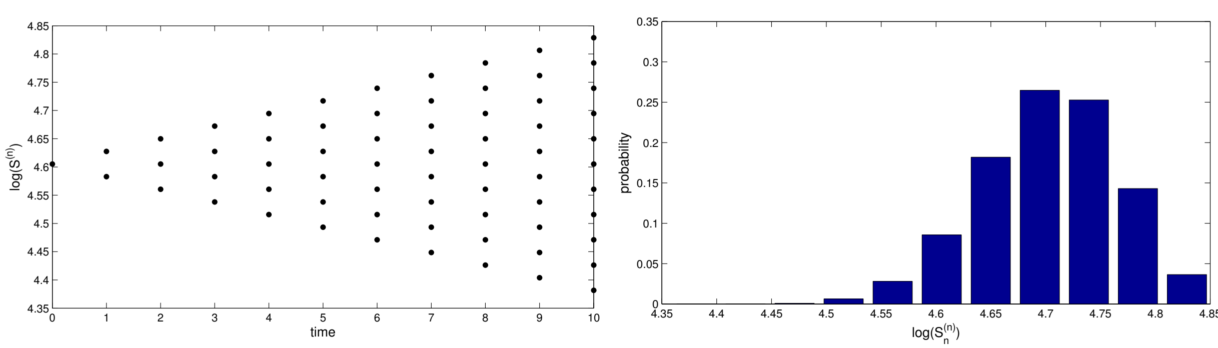


Figure 2.2: CRR tree with 10 periods

Indeed, as can be seen in the figure above the nodes are displayed symmetrically around $\log(s_0)$. To compensate this structure more probability is assigned to the upper nodes of the tree (right-hand plot).

The main disadvantage of the model is the oscillatory convergence behavior. Figure 2.3 shows the zig-zag convergence of the CRR tree often referred to as the saw-tooth effect. In addition, we can also see the so-called even-odd effect, the micro oscillations between even and odd n. If we only consider even n (green line) or only odd n (red line), these jumps are not present.

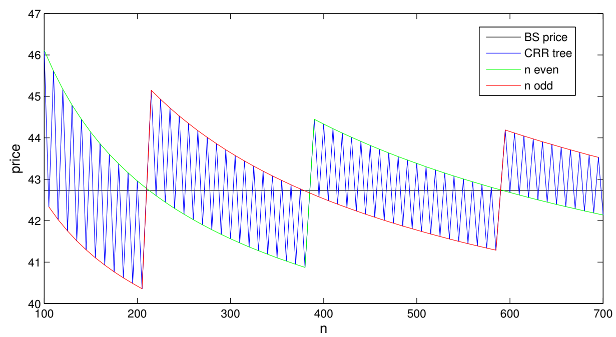


Figure 2.3: CRR tree for a digital put

The RB tree

In [RB79] R. J. Rendleman and B. J. Bartter introduce a model where the up- and down-jump have equal probability, i.e.

$$p_n = \frac{1}{2}.$$

The drift is then chosen as

$$\alpha_n = r - \frac{1}{2}\sigma^2.$$

Note that with this choice of parameters the moment-matching conditions (2.4) are satisfied exactly

$$\mu(n) = r - \frac{1}{2}\sigma^2,$$
$$\sigma^2(n) = \sigma^2.$$

Moreover, the equal probabilities prove to be an advantage in the implementation of the model, since only one multiplication per node is required in the backward induction step, which speeds up the algorithm.

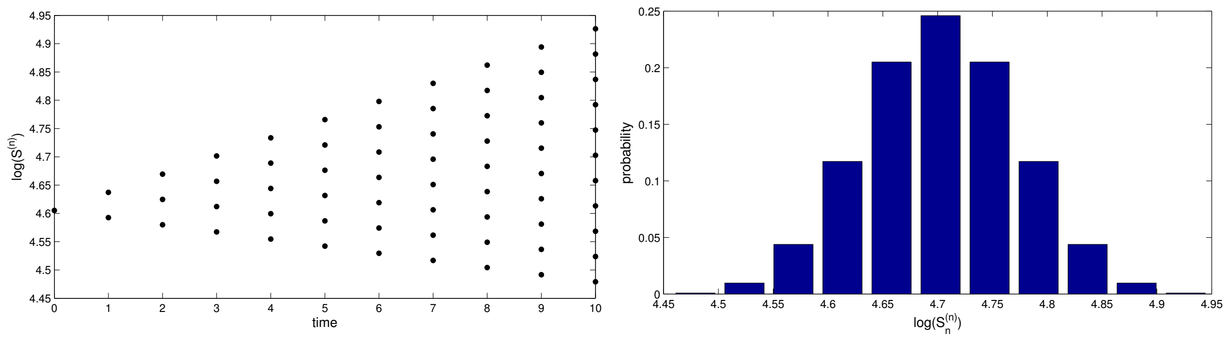


Figure 2.4: RB tree with 10 periods

Unlike the previous model the RB tree has a clear upward trend, but more evenly spread probabilities. The exact shape of the tree as well as the upward or downward trend, of course, depend on the parameters of the model.

The RB model also suffers from an oscillatory convergence pattern, although in many cases, the saw-tooth effect is less pronounced than for the CRR tree.

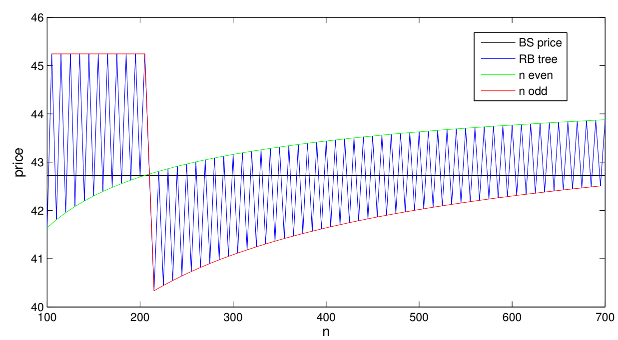


Figure 2.5: RB tree for a digital put

Remark 2.2.1. Note that in some references this model is also referred to as the Jarrow-Rudd (JR) tree (see e.g. [LR96] or [Ku08]).

2.2.2 Multidimensional trees

We now consider a general d-dimensional Black-Scholes model, where the dynamics of the d assets is given by

$$dS_i(t) = S_i(t)\left(rdt + \sigma_i dW_i(t)\right), \quad S_i(0) = s_{0,i}, \quad i = 1, \ldots, d, \tag{2.6}$$

where $W_i(t)$ and $W_j(t)$ have correlation $\rho_{i,j}$, $i, j = 1, \ldots, d$, $i < j$.

We construct the approximating d-dimensional n-period binomial tree $S^{(n)}$ as follows. For each time step $k = 1, \ldots, n$

$$S_k^{(n)} = \begin{pmatrix} S_{k,1}^{(n)} \\ \vdots \\ S_{k,d}^{(n)} \end{pmatrix} := \begin{pmatrix} S_{k-1,1}^{(n)} e^{\alpha_{n,1}\Delta t + \sigma_1 \sqrt{\Delta t} \xi_{k,1}^{(n)}} \\ \vdots \\ S_{k-1,d}^{(n)} e^{\alpha_{n,d}\Delta t + \sigma_d \sqrt{\Delta t} \xi_{k,d}^{(n)}} \end{pmatrix}, \quad S_0^{(n)} = s_0, \tag{2.7}$$

where $\{\alpha_{n,i}\}$, $i = 1, \ldots, d$ are bounded sequences and for each $n \in \mathbb{N}$, $\xi_k^{(n)}$ are i.i.d random vectors, each component taking on the values -1 or 1. In this case

$$\Omega_n = \left\{ \omega = (\omega_1, \ldots, \omega_n) | \; \omega_k \in \{-1, 1\}^d, \; k = 1, \ldots, n \right\}.$$

At each time step the probability of scenario ω_k is $P_{1,n}(\omega_k)$. The probability measure P_n is then given as the product of the one-step transition probabilities, i.e.

$$P_n(\omega) = \prod_{k=1}^{n} P_{1,n}(\omega_k).$$

As in the one-dimensional case, the drift sequences $\alpha_{n,i}$, $i = 1, \ldots, d$ and the one-step transition probabilities $P_{1,n}$ need to be chosen such that weak convergence is ensured. Note that in the multidimensional setting, $S^{(n)}$ has to match not only the mean and variance of the log-returns of each asset, but also the whole correlation structure. Therefore, the following moment-matching conditions need to be satisfied

$$\begin{aligned} \mu_i(n) &= \alpha_{n,i} + \frac{1}{\sqrt{\Delta t}} \sigma E_n\left(\xi_{1,i}^{(n)}\right) = r - \frac{1}{2}\sigma_i^2 + o(1), \quad i = 1, \ldots, d, \\ \sigma_i^2(n) &= \sigma_i^2 Var_n\left(\xi_{1,i}^{(n)}\right) = \sigma^2 + o(1), \quad i = 1, \ldots, d, \\ c_{i,j}(n) &= \frac{1}{\Delta t} Cov_n\left(\log\left(\frac{S_{k,i}^{(n)}}{S_{k-1,i}^{(n)}}\right), \log\left(\frac{S_{k,j}^{(n)}}{S_{k-1,j}^{(n)}}\right) \middle| S_{k-1,i}^{(n)}, S_{k-1,j}^{(n)} \right) \\ &= \sigma_i \sigma_j Cov_n\left(\xi_{1,i}^{(n)}, \xi_{1,j}^{(n)}\right) = \sigma_i \sigma_j \rho_{i,j} + o(1), \quad i = 1, \ldots, d, i < j. \end{aligned} \tag{2.8}$$

We now present examples of some well-known multidimensional binomial trees, which are generalizations of the one-dimensional trees discussed above.

The BEG model

P. P. Boyle, J. Evnine and S. Gibbs introduce in [BEG89] a multidimensional generalization of the CRR tree in the sense that all the one-dimensional projections coincide with the CRR tree. Therefore, the parameters of the model are given by

$$\alpha_{n,i} = 0, \quad i = 1, \ldots, d, \tag{2.9}$$

and

$$P_{1,n}(\omega_k) = \frac{1}{2^d}\left(1 + \sum_{i,j=1,i<j}^{d} \rho_{ij}\delta_{i,j}(\omega_k) + \sqrt{\Delta t}\sum_{i=1}^{d}\delta_i(\omega_k)\frac{r - \frac{1}{2}\sigma_i^2}{\sigma_i}\right), \tag{2.10}$$

where

$$\delta_{i,j}(\omega_k) = \begin{cases} 1 & \text{if } \omega_{k,i} = \omega_{k,j}, \\ -1 & \text{if } \omega_{k,i} \neq \omega_{k,j}, \end{cases} \tag{2.11}$$

and

$$\delta_i(\omega_k) = \begin{cases} 1 & \text{if } \omega_{k,i} = 1, \\ -1 & \text{if } \omega_{k,i} = -1. \end{cases}$$

For details see [BEG89] and [M09], Chapter 3. The BEG model inherits the properties of the CRR tree, and is also symmetric in the log-scale. Unfortunately, it also inherits the irregular convergence behavior, as can be see in the figure below.

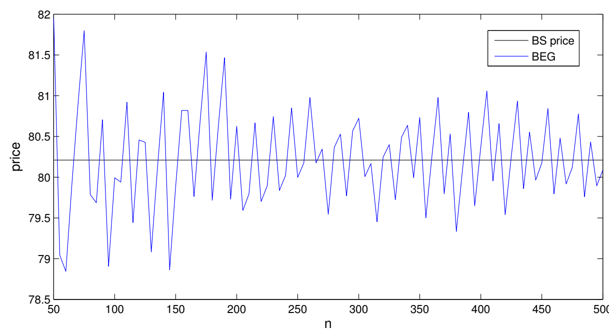


Figure 2.6: The BEG tree for a two-asset digital put

The m-dimensional RB model

Analogously, R. Korn and S. Müller introduce in [KM09] a model, where each of the components coincides with the RB tree. In this case we have

$$\alpha_{n,i} = r - \frac{1}{2}\sigma_i^2, \quad i = 1, \ldots, d, \tag{2.12}$$

and

$$P_{1,n}(\omega_k) = \frac{1}{2^d}\left(1 + \sum_{i,j=1,i<j}^{d} \rho_{ij}\delta_{i,j}(\omega_k)\right), \tag{2.13}$$

with $\delta_{i,j}$ defined as in (2.11).

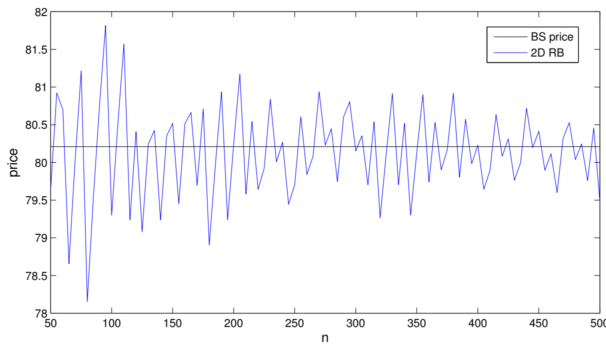


Figure 2.7: The 2D RB tree for a two-asset digital put

Remark 2.2.2. One of the drawbacks in both models presented above is that for $d \geq 3$ the transition probabilities $P_{1,n}$ might not be well-defined for certain parameter settings. As mentioned in [KM09], model parameters can be found for which the probabilities are negative for all $n \in \mathbb{N}$. As an alternative approach, which overcomes this problem we now present the decoupling method (see [M09] or [KM09]).

The decoupling approach

The multidimensional trees described above are constructed directly for the process S in (2.6). As a result, the nodes in the log-scale form a rectangular grid for each time-step which leads to the convergence behavior described above. In [KM09] a different method is proposed, where the continuous process S is first transformed to a process Y with independent components. An approximating tree is then constructed for Y. Then, by applying a back-transformation a binomial model for the original process S is obtained. This approach has the following advantages. First of all, the rectangular grid is destroyed, which leads to a better convergence pattern. Second, the correlation structure no longer enters the probabilities, which makes the decoupling approach applicable to any parameter setting.

We now present a more detailed description of the method. Let Σ be the covariance matrix of $\log(S)$, i.e.

$$\Sigma = \begin{pmatrix} \sigma_1^2 & \rho_{1,2}\sigma_1\sigma_2 & \cdots & \rho_{1,d}\sigma_1\sigma_2 \\ \rho_{1,2}\sigma_1\sigma_2 & & & \\ \vdots & & & \\ \rho_{1,d}\sigma_1\sigma_d & & \cdots & \sigma_d^2 \end{pmatrix}$$

Using the decomposition

$$\Sigma = GDG^T,$$

with $G \in \mathbb{R}^{m \times m}$ and $D \in \mathbb{R}^{m \times m}$ diagonal, construct the process $Y = G^{-1}\log(S)$ with independent components. The dynamics of Y is given by

$$dY_j\left(t\right) = \alpha_j dt + \sqrt{d_{jj}} d\bar{W}_j\left(t\right),$$

where $\alpha_i = \sum_{i=1}^{d} g_{ji}^{(-1)}(r - \frac{1}{2}\sigma_i^2)$, and $\bar{W}_j$, $j = 1, \ldots d$ are independent Brownian motions (see [M09], Proposition 37 or [KM09], Proposition 1). Now, construct the approximating tree as follows

$$Y_k^{(n)} = \begin{pmatrix} Y_{k-1,1}^{(n)} + \alpha_1 \Delta t + \sqrt{d_{11}}\sqrt{\Delta t}\xi_{k,1}^{(n)} \\ \vdots \\ Y_{k-1,d}^{(n)} + \alpha_d \Delta t + \sqrt{d_{dd}}\sqrt{\Delta t}\xi_{k,d}^{(n)} \end{pmatrix}, \quad Y_0^{(n)} = \log(s_0), \tag{2.14}$$

with

$$P_{1,n}(\omega_k) = \frac{1}{2^d}.$$

Note that $Y^{(n)}$ matches the mean and variance of the one-period returns of each component in Y, and, since all the components are independent,

$$Y^{(c,n)} \xrightarrow{\mathcal{D}} Y,$$

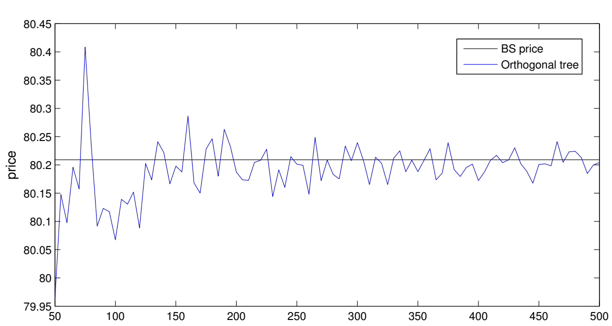
where $Y^{(c,n)}$ is obtained from $Y^{(n)}$ by linear interpolation. The tree for the original process S is now defined as

$$S_k^{(n)} = \exp\left(GY_k^{(n)}\right), \quad k = 0, \ldots, n. \tag{2.15}$$

Weak convergence for the S-tree is ensured due to the continuous mapping principle (see e.g. [B68], Theorem 5.1). For a detailed discussion of the approach see [KM09] or [M09].

Remark 2.2.3. The mentioned references consider two types of decompositions. In case of the spectral decomposition, where G is an orthogonal matrix and D contains the eigenvalues of Σ, the tree is referred to as the orthogonal tree. In case of the Cholesky decomposition $\Sigma = GG^T$, where G is a lower triangular matrix, the tree is called the Cholesky tree.

Compare the following convergence results for the orthogonal (left-hand side) and Cholesky (right-hand side) trees.



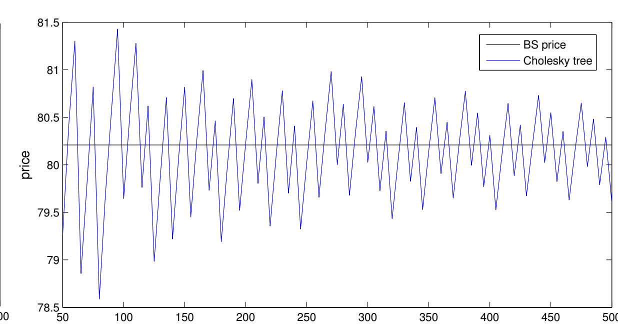


Figure 2.8: The decoupling method for a two-asset digital put

Note that for two-asset digital options the orthogonal tree performs much better than the models presented above, however, the Cholesky tree still has highly oscillatory convergence behavior. This is explained by the fact that with the Cholesky decomposition the rectangular grid is only partially destroyed (see e.g. [M09], pp. 143-146). However, the algorithm for the Cholesky tree is faster than the one for the orthogonal tree, due to the fewer calculations required for the back transformation (2.15) (see [M09], section 4.4).

Chapter 3

Edgeworth Expansions

In financial mathematics, statistics and probability theory problems often arise that deal with sums of random variables. The CLT plays an important role here in investigating the limiting distribution of such sums. However, in many cases we are interested in a more detailed analysis of the convergence behavior and a better approximation than simply the normal distribution function.

The idea of Edgeworth expansions goes back to P.L. Chebyshev (see e.g. [GK54]), F.Y. Edgeworth [Edge1905], C.V.L. Charlier [Ch1906], and others. They propose to expand the distribution of a sum of random variables in a series of the normal distribution function and its derivatives. These series and their convergence properties have been thoroughly studied by H. Cramer. In [BR76] R.N. Bhattacharya and R.R. Rao present numerous results on the topic, offer a generalization to higher dimensions and a detailed exposition for lattice random vectors, which will be the main focus of this chapter. We will be using the notation and following the proofs given in [BR76].

3.1 Introduction

The purpose of this thesis is to investigate the asymptotics of the general binomial model $S^{(n)}$ in (2.7). Our main concern is the distribution function of $S^{(n)}$ at maturity, i.e.

$$P\left(S_n^{(n)} \leq x\right) = P\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n} \xi_k^{(n)} \leq y_n\right),$$

with $y_{n,i} = \frac{\log\left(\frac{x_i}{s_{0,i}}\right) - \alpha_{n,i}T}{\sigma_i\sqrt{T}}$, $i = 1, \ldots, d$. Therefore, we are interested in the limit properties of a sum of triangular array vectors. The $\xi^{(n)}$ are lattice random vectors (see section 3.4.1). The results on Edgeworth expansions available for this class of distributions mainly deal with i.i.d. random vectors, and are not directly applicable here. So, the goal of this chapter is to justify the Edgeworth expansion also for the binomial tree setting.

Edgeworth expansions are based on a taylor series of the corresponding characteristic function. This series is then inverted to recover the necessary result for the distribution function itself. Therefore, we start with a brief introduction of the basic definitions and properties of moments and characteristic functions. Then, following the main literature on the topic, we give a heuristic explanation of the approach to get a better understanding of the intuition behind it. In the last section we present the main theorems.

3.2 The characteristic function: basic definitions and properties

We will follow the definitions in [BR76].

Definition 3.2.1. The *Fourier-Stieltjes transform* of a finite signed measure μ is the function $\psi_\mu : \mathbb{R}^d \to \mathbb{C}$ defined by

$$\psi_\mu(t) = \int_{\mathbb{R}^d} e^{i\langle t,x\rangle} \mu(dx), \quad t \in \mathbb{R}^d.$$

If μ has density f with respect to the Lebesgue measure, the $\psi_\mu = \psi_f$, where ψ_f is the Fourier transform of f. If μ is a probability measure, ψ_μ is usually referred to as the *characteristic function* of μ.

Remark 3.2.2. We will also use the notation ψ_X to denote the characteristic function of the probability measure corresponding to the random vector X.

Example 3.2.3. Consider the multidimensional normal distribution $\Phi_{m,\Sigma}$ with mean m and covariance matrix Σ. The density function is given by

$$\phi_{m,\Sigma}(x) = \frac{1}{(2\pi)^{d/2}\sqrt{\det\Sigma}} e^{-\frac{1}{2}\left\langle x,\Sigma^{-1}x\right\rangle}, \quad x \in \mathbb{R}^d.$$

In this case the characteristic function is equal to

$$\psi_{\Phi_{m,\Sigma}}(t) = e^{i\langle t,m\rangle - \frac{1}{2}\langle t,\Sigma t\rangle}, \quad t \in \mathbb{R}^d.$$

(See e.g. [BR76], Chapter 5.)

There exists a one-to-one correspondence between finite signed measures and Fourier-Stieltjes transforms (Uniqueness Theorem, see e.g. [BR76] Theorem 5.1 (i)). If μ is absolutely continuous with respect to the Lebesgue measure, then its density f can be recovered from ψ_μ using the following theorem (see [BR76], Theorems 4.1 and 5.1).

Theorem 3.2.4 (Fourier Inversion Formula). *i) If $\psi_\mu \in L^1(\mathbb{R}^d)$, then μ is absolutely continuous with respect to the Lebesgue measure and has a uniformly continuous, bounded density f,*

$$f(x) = \frac{1}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\langle t,x\rangle} \psi_\mu(t)\, dt, \quad x \in \mathbb{R}^k.$$

ii) Suppose $f \in L^1(\mathbb{R}^d)$. Let $x \in \mathbb{R}^d$ and $\alpha = (\alpha_1, \ldots, \alpha_d)$ be a nonnegative integer vector. If $x^\alpha f(x) \in L^1(\mathbb{R}^d)$, then $D^\alpha \psi_f$ exists and

$$x^\alpha f(x) = \frac{(-1)^{|\alpha|}}{(2\pi)^d} \int_{\mathbb{R}^d} e^{-i\langle t,x\rangle} D^\alpha \psi_f(t)\, dt.$$

In addition, weak convergence of probability measures on $\mathbb{R}^d$ can be characterized in terms of convergence of their characteristic functions.

Theorem 3.2.5.

i) Let $\{P_n\}_{n\geq 1}$, P be probability measures on $\mathbb{R}^d$. If $P_n \Rightarrow P$, then $\{\psi_{P_n}\}_{n\geq 1}$ converges pointwise to ψ_P.

ii) If $\{\psi_{P_n}\}_{n\geq 1}$ converges pointwise to a function ψ, which is continuous at zero, then there exists a probability measure P, such that $P_n \Rightarrow P$ and $\psi_P = \psi$.

(see e.g. [BR76] Theorem 5.2)

This suggests that the asymptotic expansions for distribution functions could be derived from the corresponding results for characteristic functions.

Let P be a probability measure on $\mathbb{R}^d$. To obtain an expansion for the characteristic function we will be using the following results.

Definition 3.2.6. (i) Let $\nu = (\nu_1, \ldots, \nu_k)$ be a nonnegative integer vector such that

$$\int_{\mathbb{R}^d} |x^\nu|\, P(dx) < \infty.$$

The *moment of order* ν of P is defined as

$$\mu_\nu = \int_{\mathbb{R}^d} x^\nu P(dx).$$

(ii) For $s \in \mathbb{R}$, $s \geq 0$ the *absolute moment of order* s of P is defined as

$$\rho_s = \int_{\mathbb{R}^d} \|x\|^s P(dx).$$

The following theorem characterizes the derivatives of ψ_P.

Theorem 3.2.7. *If $\rho_s < \infty$ for some positive integer s, then $D^\alpha \psi_P$ exists for every nonnegative integer vector $\nu \in \mathbb{R}^d$, $|\nu| \leq s$, and is equal to*

$$(D^\nu \psi_P)(t) = (i)^{|\nu|} \int_{\mathbb{R}^d} x^\nu e^{i\langle t,x\rangle} P(dx), \quad t \in \mathbb{R}^d.$$

Moreover,

$$i^{|\nu|} \mu_\nu = (D^\nu \psi_P)(0).$$

(see e.g. [BR76] Section 5).

Using Theorem 3.2.7 and a Taylor expansion for complex-valued functions (see e.g. [BR76], Corollary 8.2) we get

$$\psi_P(t) = 1 + \sum_{|\nu|\leq s} \frac{\mu_\nu}{\nu!} (it)^\nu + o(\|t\|^s), \quad t \to 0, \tag{3.1}$$

if $\rho_s < \infty$ for some positive integer s. Therefore, we have an expansion for the characteristic function in terms of the moments of the distribution. However, in the theory of Fourier transforms it is usually more convenient to work with cumulants, also referred to as semi-invariants (e.g. [GK54], [P75]). Instead of (3.1) consider the Taylor expansion of the logarithm of the characteristic function.

$$\log \psi_P(t) = \sum_{|\nu|\leq s} \frac{\kappa_\nu}{\nu!} (it)^\nu + o(\|t\|^s), \quad t \to 0. \tag{3.2}$$

Here the coefficients κ_ν are the *cumulants of order* ν of P, and are given by

$$i^{|\nu|} \kappa_\nu = (D^\nu \log \psi_P)(0).$$

The cumulants κ_ν, $|\nu| \leq s$, as well as expansion (3.2) exist if $\rho_s < \infty$. Here $\log \psi_P(t)$ is the principal branch of the complex logarithm. Note that

$$\kappa_0 = 0, \quad \text{and} \quad \kappa_\nu = \mu_\nu, \ |\nu| = 1.$$

Example 3.2.8. Consider once again the normal distribution function $\Phi_{m,\Sigma}$. Since $\log \psi_{\Phi_{m,\Sigma}}(t)$ is quadratic in t all cumulants for $|\nu| > 2$ are equal to zero. For $|\nu| = 1, 2$ we have

$$\kappa_{e_i} = m_i, \quad i = 1, \ldots, d,$$

and

$$\kappa_{e_i+e_j} = \Sigma(i,j), \quad i,j = 1, \ldots, d$$

where e_i is the i-th unit vector (see [BR76], pp. 50-51).

Remark 3.2.9. The main reason why cumulants are usually the preferred choice is their additivity. Consider the sum $X = X_1 + \cdots + X_n$ of independent random vectors on $\mathbb{R}^d$, then

$$\psi_X(t) = \prod_{i=1}^{n} \psi_{X_i}(t),$$

and

$$\log \psi_X(t) = \sum_{i=1}^{n} \log \psi_{X_i}(t).$$

Therefore, if X_i, $i = 1, \dots, n$ have finite cumulants of order ν, $\kappa_\nu(X_i)$, then

$$\kappa_\nu(X) = \kappa_\nu(X_1) + \cdots + \kappa_\nu(X_n).$$

This is, obviously, not true for moments of higher orders.

The formal identity

$$\log\left(1 + \sum_{|\nu|\geq 1} \frac{\mu_\nu}{\nu!}(it)^\nu\right) = \sum_{|\nu|\geq 1} \frac{\kappa_\nu}{\nu!}(it)^\nu,$$

allows to uniquely express the cumulants in terms of the moments and vice versa. As a result we have

$$\begin{aligned} m_\nu &= \sum \frac{1}{q!} \frac{\nu!}{\nu_1! \cdots \nu_q!} \prod_{p=1}^{q} \kappa_{\nu_p} \quad \text{and} \\ \kappa_\nu &= \sum \frac{(-1)^{q-1}}{q} \frac{\nu!}{\nu_1! \cdots \nu_q!} \prod_{p=1}^{q} m_{\nu_p}, \end{aligned} \tag{3.3}$$

where the summation is over all $q-$tuples of nonnegative integer vectors ν_p, $|\nu_p| > 0$, such that

$$\sum_{p=1}^{q} \nu_p = \nu, \quad q = 1, \dots, |\nu|,$$

(see e.g. [Sh84], Theorem 12.8).

Example 3.2.10. In the one-dimensional case (3.3) gives

$$\begin{aligned} \mu_2 &= \kappa_2 + \kappa_1^2, \\ \mu_3 &= \kappa_3 + 3\kappa_2\kappa_1 + \kappa_1^3, \\ \mu_4 &= \kappa_4 + 4\kappa_3\kappa_1 + 3\kappa_2^2 + 6\kappa_2\kappa_1^2 + \kappa_1^4, \end{aligned}$$

and

$$\begin{aligned} \kappa_2 &= \mu_2 - \mu_1^2, \\ \kappa_3 &= \mu_3 - 3\mu_2\mu_1 + 2\mu_1^3, \\ \kappa_4 &= \mu_4 - 4\mu_3\mu_1 - 3\mu_2^2 + 12\mu_2\mu_1^2 - 6\mu_1^4, \end{aligned}$$

For ease of reference we state the following well-known properties of norms and moments.

Lemma 3.2.11. *For any nonnegative integer vector* $\nu \in \mathbb{R}^d$ *and integers* $0 \le m_1 \le m \le m_2$

(i) $|x^\nu| \le \|x\|^{|\nu|}, \quad x \in \mathbb{R}^d$,

(ii) $\|x\|^m \le \|x\|^{m_1} + \|x\|^{m_2}, \quad x \in \mathbb{R}^d$.

For a random vector X *in* $\mathbb{R}^d$ *with* $\rho_s < \infty$, $|\nu| \le s$

(iii) $|m_\nu| \le E|X^\nu| \le \rho_{|\nu|}$,

(iv) $|\kappa_\nu| \le c\rho_{|\nu|}$, *where the constant* c *depends only on* ν.

(Cf. [BR76], Lemma 6.3 and (9.13))

3.3 Heuristic considerations

Consider the following problem setting. Let $X_1, \ldots, X_n : (\Omega, \mathcal{F}, P) \to \mathbb{R}^d$ be i.i.d. random vectors with $EX_1 = 0$, $Cov(X_1) = V$, where V is a positive-definite matrix, and $\rho_s < \infty$ for some $s \ge 3$. We are interested in the asymptotics of the distribution function F_n,

$$F_n(x) := P(S_n \le x), \quad x \in \mathbb{R}^d,$$

where

$$S_n = \frac{1}{\sqrt{n}}(X_1 + \cdots + X_n).$$

By the properties of characteristic functions (see e.g. [BR76], Theorem 5.1)

$$\psi_{S_n}(t) = \left(\psi_{\frac{1}{\sqrt{n}}X_1}(t)\right)^n = \left(\psi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n.$$

Then by equation (3.2) and taking into account that $\kappa_\nu = \mu_\nu = 0$, for $|\nu| = 1$ we have

$$\begin{aligned}
\log \psi_{S_n}(t) &= n \log \psi_{X_1}\left(\frac{t}{\sqrt{n}}\right) \\
&= \sum_{r=1}^{s} \sum_{|\nu|=r} \frac{\kappa_\nu}{\nu!}(it)^\nu n^{-(r-2)/2} + n \cdot o\left(\left\|\frac{t}{\sqrt{n}}\right\|^s\right) \\
&= -\frac{1}{2}\langle t, Vt\rangle + \sum_{r=1}^{s-2} \sum_{|\nu|=r+2} \frac{\kappa_\nu}{\nu!}(it)^\nu n^{-r/2} + n \cdot o\left(\left\|\frac{t}{\sqrt{n}}\right\|^s\right), \quad \frac{t}{\sqrt{n}} \to 0.
\end{aligned}$$

If we now fix t we get

$$\psi_{S_n}(t) = e^{-\frac{1}{2}\langle t,Vt\rangle} \exp\left(\sum_{r=1}^{s-2} \sum_{|\nu|=r+2} \frac{\kappa_\nu}{\nu!}(it)^\nu n^{-r/2} + o\left(n^{-(s-2)/2}\right)\right), \quad n \to \infty. \tag{3.4}$$

Now define the functions $\tilde{P}_r(z, \{\kappa_\nu\})$ from the formal identity

$$1 + \sum_{r=1}^{+\infty} \tilde{P}_r(z, \{\kappa_\nu\}) u^r = \exp\left\{\sum_{r=1}^{+\infty} \sum_{|\nu|=r+2} \frac{\kappa_{n,\nu} z^\nu}{\nu!} u^r\right\}.$$

Lemma 3.3.1. *The functions $\tilde{P}_r$ are polynomials of degree $3r$ with coefficients that depend only on the cumulants κ_ν of order $|\nu| \le r+2$. They can be calculated as*

$$\tilde{P}_r(z, \{\kappa_\nu\}) = \sum_{m=1}^{r} \frac{1}{m!} \sum_{j_1,\ldots,j_m} \left(\sum_{\nu_1,\ldots,\nu_m} \frac{\kappa_{\nu_1} \cdots \kappa_{\nu_m}}{\nu_1! \cdots \nu_m!} z^{\nu_1+\cdots+\nu_m}\right), \quad r \ge 1,\ z \in \mathbb{R}^d,$$

where $\sum_{j_1,\ldots,j_m}$ is the summation over all m-tuples of positive integers $(j_1,\ldots,j_m)$ satisfying $\sum_{i=1}^m j_i = r$, and $\sum_{\nu_1,\ldots,\nu_m}$ is the summation over all m-tuples of nonnegative integral vectors $(\nu_1,\ldots,\nu_m)$ s.t. $|\nu_i| = j_i + 2$, $i = 1,\ldots,m$.

(See [BR76], Chapter 7).

Example 3.3.2. The first two polynomials are

$$\tilde{P}_1(z, \{\kappa_\nu\}) = \sum_{|\nu|=3} \frac{\kappa_\nu}{\nu!} z^\nu,$$
$$\tilde{P}_2(z, \{\kappa_\nu\}) = \sum_{|\nu|=4} \frac{\kappa_\nu}{\nu!} z^\nu + \frac{1}{2!} \sum_{|\nu_1|=|\nu_2|=3} \frac{\kappa_{\nu_1}\kappa_{\nu_2}}{\nu_1!\nu_2!} z^{\nu_1+\nu_2}.$$

We can now write (3.4) as

$$\psi_{S_n}(t) = e^{-\frac{1}{2}\langle t,Vt\rangle} + \sum_{r=1}^{s-2} n^{-r/2} \tilde{P}_r(it, \{\kappa_\nu\}) e^{-\frac{1}{2}\langle t,Vt\rangle} + o\left(n^{-(s-2)/2}\right), \quad n \to \infty, \tag{3.5}$$

and we have an expansion of the characteristic function of a sum of i.i.d. random vectors in terms of $\tilde{P}_r(it, \{\kappa_\nu\}) e^{-\frac{1}{2}\langle t,Vt\rangle}$. The question now is, how can we use (3.5) to get the corresponding asymptotics for the distribution function?

From example 3.2.3 we know that the first term $e^{-\frac{1}{2}\langle t,Vt\rangle}$ is the characteristic function of the normal distribution $\Phi_{0,V}$. Now define $P_r(-\phi_{0,V}, \{\kappa_\nu\})$ as the function whose Fourier transform is $\tilde{P}_r(it, \{\kappa_\nu\}) e^{-\frac{1}{2}\langle t,Vt\rangle}$, i.e.

$$\int_{\mathbb{R}^d} e^{i\langle t,x\rangle} P_r(-\phi_{0,V}, \{\kappa_\nu\})(x)\, dx = \tilde{P}_r(it, \{\kappa_\nu\}) e^{-\frac{1}{2}\langle t,Vt\rangle}. \tag{3.6}$$

By taking the ν-th derivative with respect to x on both sides of the Fourier inversion formula in Theorem 3.2.4 *i*) we get

$$\int_{\mathbb{R}^d} e^{i\langle t,x\rangle} D^\nu \phi_{0,V}(x)\, dx = (-it)^\nu \psi_{\Phi_{0,V}}(t) = (-it)^\nu e^{-\frac{1}{2}\langle t,Vt\rangle}.$$

Therefore, (3.6) will be satisfied if we set $P_r(-\phi_{0,V}, \{\kappa_\nu\})$ equal to $\tilde{P}_r(it, \{\kappa_\nu\})$ after substituting $(-1)^{|\nu|} D^\nu \phi_{0,V}$ for each power $(it)^\nu$, i.e.

$$P_r(-\phi_{0,V}, \{\kappa_\nu\}) = \tilde{P}_r(-D, \{\kappa_\nu\})\, \phi_{0,V}.$$

As a result, using Lemma 3.3.1 we get

Lemma 3.3.3. *$P_r(-\phi_{0,V}, \{\kappa_\nu\})$ is a polynomial multiple of $\phi_{0,V}$ and can be written as*

$$P_r(-\phi_{0,V}, \{\kappa_\nu\}) = \sum_{m=1}^{r} \frac{1}{m!} \sum_{j_1,\dots,j_m} \left(\sum_{\nu_1,\dots,\nu_m} \frac{\kappa_{\nu_1}\cdots\kappa_{\nu_m}}{\nu_1!\cdots\nu_m!} (-1)^{r+2m} D^{\nu_1+\cdots+\nu_m} \phi_{0,V} \right),$$

where the summation is as in Lemma 3.3.1.

Example 3.3.4. If $d = 1$, then $V = \sigma^2$ and

$$P_1(-\phi_{0,\sigma^2}, \{\kappa_\nu\})(x) = \frac{\kappa_3}{6\sigma^3} \left(\frac{x^3}{\sigma^3} - \frac{3x}{\sigma} \right) \phi_{0,\sigma^2}(x).$$

Finally, define $P_r(-\Phi_{0,V_n}, \kappa_{n,\nu})$ as the finite signed measure on $\mathbb{R}^d$ whose density is $P_r(-\phi_{0,V_n}, \kappa_{n,\nu})$. By the Lebesgue dominated convergence theorem we have

$$P_r(-\Phi_{0,V}, \{\kappa_\nu\}) = \tilde{P}_r(-D, \{\kappa_\nu\})\, \Phi_{0,V}.$$

For a detailed discussion see [BR76], Chapter 7.

We can now state the desired expansion for the distribution function, known as the *Edgeworth expansion*

$$F_n(x) = \Phi_{0,V}(x) + \sum_{r=1}^{s-2} n^{-r/2} P_r(-\Phi_{0,V}, \{\kappa_\nu\})(x) + o\left(n^{-(s-2)/2}\right), \quad n \to \infty. \tag{3.7}$$

For a more detailed argumentation see e.g. [BR76], Theorem 20.1 or [GK54], Chapter 45 for the one-dimensional case. Note that expansion (3.7) is only valid under an additional assumption on the distribution of the vectors X_i, $i = 1, \dots, n$, the so-called *Cramér condition*

$$\limsup_{\|t\|\to\infty} |\psi_{X_1}(t)| < 1, \tag{3.8}$$

which is satisfied, for example, for absolutely continuous distributions. However, for lattice distributions (3.8) does not hold, and, therefore, the Edgeworth expansion in the form (3.7) is not valid, and additional terms are required.

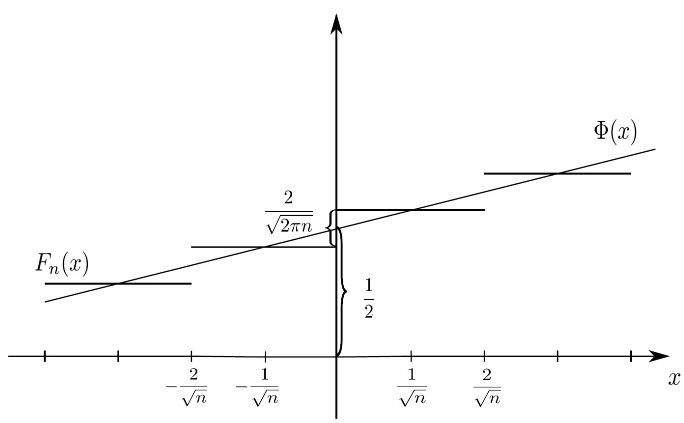
Let us go back to the context of binomial trees. Assume the one-dimensional RB model. In this case

$$S_n = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} \xi_i,$$

where the variables ξ_i are i.i.d. and have two possible values $+1$ and -1 with probabilities $\frac{1}{2}$. $F_n(x)$ is then purely discontinuous with jumps at points

$$x_m = \frac{m}{\sqrt{n}}, \quad \text{where } m = \begin{cases} 0, \pm 2, \ldots, \pm n, & n \text{ even}, \\ \pm 1, \pm 3, \ldots, \pm n, & \text{if } n \text{ odd}. \end{cases}$$

We now compare the relative behaviour of the functions $F_n(x)$ and $\Phi(x)$ in a neighborhood of a discontinuity point x_m. By the theorem of de Moivre-Laplace, the jump of F_n at x_m is asymptotically equal to $\frac{2}{\sqrt{2\pi n}} e^{-\frac{1}{2}x_m^2}$. Let n be even and consider the interval $\left(-\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}\right)$ around x_0. Here the function $\Phi(x)$ behaves like $\frac{1}{2} + \frac{x}{\sqrt{2\pi}}$, up to terms of order $o\left(\frac{1}{\sqrt{n}}\right)$. In addition, $\frac{1}{2}\left(F_n(0-) + F_n(0+)\right) = \Phi(0) = \frac{1}{2}$. Then, in the vicinity of x_0 up to order $o\left(\frac{1}{\sqrt{n}}\right)$ we have the following graph (see [E45])



Therefore, in the interval $\left(-\frac{1}{\sqrt{n}}, \frac{1}{\sqrt{n}}\right)$

$$F_n(x) - \Phi(x) = -\frac{2}{\sqrt{2\pi n}} S\left(\frac{x\sqrt{n}}{2}\right),$$

where

$$S(x) = x - \lfloor x \rfloor - \frac{1}{2}, \tag{3.9}$$

and $\lfloor x \rfloor$ denotes the integer part of x. For an expansion at other points the change of the slope of $\Phi(x)$ needs to be taken into account (see [GK54], Chapter 8.43). This suggests, that in addition to the right-hand side of (3.7) the asymptotics for the RB tree should contain terms of the form $-\frac{2}{\sqrt{2\pi n}}S\left(\frac{x\sqrt{n}}{2}\right)e^{-\frac{1}{2}x^2}$. The exact expansion for lattice random vectors is derived in the next section.

Remark 3.3.5. Note that the Edgeworth expansion for absolutely continuous distributions (3.7) is an expansion in powers of $n^{-1/2}$ in the classical sense, i.e. with constant coefficients. In the case of lattice distributions, on the other hand, we have expansion with bounded coefficients (see [DD04] or [M09]).

3.4 Edgeworth expansions for lattice triangular arrays

3.4.1 Lattice random vectors

We first consider lattice distributions in more detail.

Definition 3.4.1. The discrete subgroup L of $\mathbb{R}^d$ is called a *lattice* if there exist linearly independent vectors $h_1, \dots, h_d$ in L such that

$$L = \{m_1 h_1 + \cdots + m_d h_d |\ m_i \in \mathbb{Z},\ i = 1, \dots, d\}.$$

The set of vectors $\{h_1, \dots, h_d\}$ is the *basis* of L. The *volume* of the lattice is defined as

$$D(L) = |Det(h_1, \dots, h_d)|.$$

The volume is independent of the choice of the basis and is uniquely determined for each lattice.

Definition 3.4.2. Let $(\Omega, \mathcal{F}, P)$ be a probability space. The distribution of a random vector X on $\mathbb{R}^d$ is said to be a *lattice distribution* if there exists a $x_0 \in \mathbb{R}^d$ and a lattice L such that

$$P(X \in x_0 + L) = 1. \tag{3.10}$$

It is clear that one can find various vectors x_0 and lattices such that (3.10) holds. Just consider the possible representations for the values ± 1 of the random variables $\xi^{(n)}$ in the definition of the binomial tree. Therefore, lattice distributions are often characterized in terms of the unique minimal lattice.

Definition 3.4.3. The lattice L is called the *minimal lattice* of X in $\mathbb{R}^d$ if L satisfies (3.10) with some $x_0 \in \mathbb{R}^d$, and for every sublattice $L' \subset L$

$$P(X \in y_0 + L') < 1, \quad \forall y_0 \in \mathbb{R}^d.$$

Note that the minimal lattice has the maximal volume $D(L)$ out of all lattices satisfying (3.10).

Definition 3.4.4. A random vector X is called *degenerate* if there exists a hyperplane $H = \{x : \langle a, x\rangle = c\}$, $a \in \mathbb{R}^d$, $c \in R$, such that $P(X \in H) = 1$.

From now on we assume that X is a nondegenerate lattice random vector, and focus on the properties of the characteristic function of X

$$\psi_X(t) = \sum_{\alpha \in L} P(X = x_0 + \alpha)\, e^{i\langle t, x_0+\alpha\rangle}.$$

Note that $|\psi_X|$ is a periodic function, therefore, the Cramér condition (3.8) is indeed not satisfied. Consider the set L^* of periods of $|\psi_X|$. Let $\{h_1, \dots, h_d\}$ be a basis of L and $\{\hat{h}_1, \dots, \hat{h}_d\}$ its dual basis, i.e.

$$\left\langle h_i, \hat{h}_j \right\rangle = \delta_{i,j}, \quad i, j = 1, \dots, d,$$

where $\delta_{i,j}$ is Kronecker's delta. Then L^* is the lattice defined as

$$L^* = \{m_1 2\pi\hat{h}_1 + \cdots + m_d 2\pi\hat{h}_d |\, m_i \in \mathbb{Z},\ i = 1, \dots, d\}.$$

By the properties of a dual basis the volume of L^* is given by

$$D(L^*) = \frac{(2\pi)^d}{D(L)}.$$

Note that L^* can also be characterized in the following way (see e.g. [BR76], Lemma 21.6)

$$L^* = \{t : |\psi_X(t)| = 1\}.$$

We now introduce the *fundamental domain* $\mathcal{F}^*$ of L^*

$$\mathcal{F}^* = \{t_1\hat{h}_1 + \cdots + t_k\hat{h}_d \mid |t_i| < \pi\ \forall j\}.$$

$$vol\mathcal{F}^* = D(L^*).$$

The fundamental domain allows to partition the space $\mathbb{R}^d$ in the following way

$$\mathbb{R}^d = \bigcup_{\hat{\alpha} \in L^*} Cl(\mathcal{F}^* + \hat{\alpha}). \tag{3.11}$$

In addition $\mathcal{F}^* \cap L^* = \{0\}$, i.e. $\mathcal{F}^*$ doesn't contain any periods of $|\psi_X|$, other than 0, and, therefore,

$$|\psi_X(t)| < 1, \quad t \in \mathcal{F}^*, t \neq 0.$$

This is an important property that we will often make use of in the proofs below. Finally, we state the Fourier inversion formula for lattice random vectors.

Theorem 3.4.5 (Fourier inversion formula for lattice distributions)**.** *Let X be a nondegenerate lattice random vector in $\mathbb{R}^d$ with lattice L and $x_0 \in \mathbb{R}^d$, such that $P(X \in x_0 + L) = 1$. Then*

$$P(X = x_0 + \alpha) = \frac{D(L)}{(2\pi)^d} \int_{\mathcal{F}^*} e^{-i\langle t, x_0+\alpha\rangle} \psi_X(t)\, dt,$$

and

$$(x_0 + \alpha)^\nu P(X = x_0 + \alpha) = \frac{D(L)}{(2\pi)^d} (-i)^{|\nu|} \int_{\mathcal{F}^*} e^{-i\langle t, x_0+\alpha\rangle} D^\nu \psi_X(t)\, dt, \quad \alpha \in L.$$

For details see [BR76], Chapter 5.21.

3.4.2 Local expansions

We now derive Edgeworth expansions for the point masses of a sum of triangular array lattice random vectors. These results are used in section 3.4.3 to obtain expansions for the distribution function. Supplementary lemmas, that are used in the proofs below are presented in section 3.4.4.

Consider a triangular array of lattice random vectors $X_{n,1}, X_{n,2}, \ldots, X_{n,n}$ defined on the probability spaces $(\Omega_n, \mathcal{F}_n, P_n)$, with common minimal lattice $\mathbb{Z}^d$ s.t.

$$\begin{aligned} E(X_{n,1}) = \mu_n, \quad Cov(X_{n,1}) = V_n, \quad P\left(X_{n,1} \in \mathbb{Z}^d\right) = 1 \quad \text{and} \\ \rho_{n,s+1} = E\left\|X_{n,1} - \mu_n\right\|^{s+1} = O(1), \quad \text{for some integer } s \geq 2, \end{aligned} \tag{3.12}$$

where the sequence of positive-definite covariance matrices $\{V_n\}$ converges to a positive-definite limit matrix V. For each $n \in \mathbb{N}$ let S_n be the normalized sum

$$S_n = \frac{X_{n,1} + \cdots + X_{n,n} - n\mu_n}{\sqrt{n}}.$$

Note that S_n is also a lattice random vector with minimal lattice $L = n^{-1/2}\mathbb{Z}^d$ and

$$P\left(S_n \in -\sqrt{n}\mu_n + L\right) = 1.$$

Following the notation in [BR76] we define the values attained by S_n as $x_{\alpha,n} := \frac{1}{\sqrt{n}}(\alpha - n\mu_n)$, $n \in \mathbb{N}$, $\alpha \in \mathbb{Z}^d$. Set

$$\begin{aligned} p_n(x_{\alpha,n}) = P(X_{n,1} + \cdots + X_{n,n} = \alpha) = P(S_n = x_{\alpha,n}), \\ q_{n,s} = n^{-d/2} \sum_{r=0}^{s-2} n^{-r/2} P_r\left(-\phi_{0,V_n}, \{\kappa_{n,\nu}\}\right), \end{aligned}$$

where $\kappa_{n,\nu}$ is the $\nu - th$ cumulant of $X_{n,1}$.

Theorem 3.4.6. *Let $E(C) = \left\{t \in \mathbb{R}^d : \|t\| \leq C\right\}$, and $\mathcal{F}^*$ be a fundamental domain of $\left(\mathbb{Z}^d\right)^*$. Under conditions (3.12), if for all constants $C > 0$, s.t. $\mathcal{F}^* \setminus E(C)$ is non-empty, the characteristic functions $\psi_{X_{n,1}}$ satisfy the condition*

$$N_C := \sup\left\{\left|\psi_{X_{n,1}}(t)\right| : t \in \mathcal{F}^* \setminus E(C),\ n \in \mathbb{N}\right\} < 1, \tag{3.13}$$

then

$$\sup_{\alpha \in \mathbb{Z}^d}\left(1 + \|x_{\alpha,n}\|^s\right)\left|p_n(x_{\alpha,n}) - q_{n,s}(x_{\alpha,n})\right| = O\left(n^{-(d+s-1)/2}\right), \quad n \to \infty, \tag{3.14}$$

and

$$\sum_{\alpha \in \mathbb{Z}^d}\left|p_n(x_{\alpha,n}) - q_{n,s}(x_{\alpha,n})\right| = O\left(n^{-(s-1)/2}\right), \quad n \to \infty. \tag{3.15}$$

Proof of Theorem 3.4.6. We will follow the proof of Theorem 22.1 in [BR76].

For each $n \in \mathbb{N}$ set $Y_{i,n} := X_{i,n} - \mu_n$, $i = 1, \dots, n$. Then $\psi_{Y_{1,n}}(t) = \psi_{X_{1,n}}(t)\, e^{-i\langle t, \mu_n\rangle}$ and the characteristic function of S_n can be determined as

$$\psi_{S_n}(t) = \left(\psi_{Y_{1,n}}\left(\frac{t}{\sqrt{n}}\right)\right)^n.$$

Applying the inversion formulas in Theorem 3.4.5 we get

$$p_n(x_{\alpha,n}) = \frac{1}{(2\pi)^d n^{d/2}} \int_{\sqrt{n}\mathcal{F}^*} \psi_{S_n}(t)\, e^{-i\langle t, x_{\alpha,n}\rangle} dt,$$

and

$$x_{\alpha,n}^{\beta} p_n(x_{\alpha,n}) = \frac{1}{(2\pi)^d n^{d/2}} (-1)^{|\beta|} \int_{\sqrt{n}\mathcal{F}^*} e^{-i\langle t, x_{\alpha,n}\rangle} D^{\beta}\psi_{S_n}(t)\, dt, \tag{3.16}$$

where β is a non-negative integer vector with $|\beta| \leq s$. By Theorem 3.2.4 and the definition of the functions P_r (3.6) we have

$$x_{\alpha,n}^{\beta} q_{n,s}(x_{\alpha,n}) = \frac{1}{(2\pi)^d n^{d/2}} (-1)^{|\beta|} \int_{\mathbb{R}^d} e^{-i\langle t, x_{\alpha,n}\rangle} D^{\beta}\left(\sum_{r=0}^{s-2} n^{-r/2} \tilde{P}_r\left(it, \{\kappa_{n,\nu}\}\right) e^{-\frac{1}{2}\langle t, V_n t\rangle}\right) dt. \tag{3.17}$$

Since $\rho_{n,s+1} = O(1)$, by Lemma 3.4.11 *(ii)*, there exist a constant $\hat{c}_1$ and $n_0 \in \mathbb{N}$, such that

$$c_1 \Lambda_n^{-1/2} \lambda_n^{s/(2(s-2))} \rho_{n,s}^{-1/(s-2)} \geq \hat{c}_1, \quad \forall n \geq n_0.$$

Set

$$E(\hat{c}_1) = \left\{t \in \mathbb{R}^k : \|t\| \leq \hat{c}_1\right\}.$$

Note that for all $t \in \sqrt{n}E(\hat{c}_1)$ the assumptions of Lemma 3.4.10 hold for all $n \geq n_0$.

Now consider the difference $\left|x_{\alpha,n}^{\beta}\left(p_n\left(x_{\alpha,n}\right)-q_{n,s}\left(x_{\alpha,n}\right)\right)\right|$. Due to equations (3.16) and (3.17) we have

$$\begin{aligned}\left|x_{\alpha,n}^{\beta}\left(p_n\left(x_{\alpha,n}\right)-q_{n,s}\left(x_{\alpha,n}\right)\right)\right| &= \frac{1}{(2\pi)^d n^{d/2}}\Bigg|\int_{\sqrt{n}\mathcal{F}^*} e^{-i\langle t,x_{\alpha,n}\rangle} D^{\beta}\psi_{S_n}(t)\,dt \\ &\quad - \int_{\mathbb{R}^d} e^{-i\langle t,x_{\alpha,n}\rangle} D^{\beta}\left(\sum_{r=0}^{s-2} n^{-r/2}\tilde{P}_r\left(it,\{\kappa_{n,\nu}\}\right)e^{-\frac{1}{2}\langle t,V_n t\rangle}\right)dt\Bigg| \\ &\leq \frac{1}{(2\pi)^d n^{d/2}}\left(I_1+I_2+I_3\right),\end{aligned} \tag{3.18}$$

where

$$I_1 := \int_{\sqrt{n}E(\hat{c}_1)}\left|D^{\beta}\left[\psi_{S_n}(t) - e^{-\frac{1}{2}\langle t,V_n t\rangle}\sum_{r=0}^{s-2} n^{-r/2}\tilde{P}_r\left(it,\{\kappa_{n,\nu}\}\right)\right]\right|dt$$

$$I_2 := \int_{\sqrt{n}\mathcal{F}^*\setminus\sqrt{n}E(\hat{c}_1)}\left|D^{\beta}\psi_{S_n}(t)\right|dt$$

$$I_3 := \int_{\mathbb{R}^d\setminus\sqrt{n}E(\hat{c}_1)}\left|D^{\beta}\left[e^{-\frac{1}{2}\langle t,V_n t\rangle}\sum_{r=0}^{s-2} n^{-r/2}\tilde{P}_r\left(it,\{\kappa_{n,\nu}\}\right)\right]\right|dt$$

We will now estimate each of these integrals separately.

By Lemmas 3.4.10 and 3.4.11 (i), there exist constants $\hat{c}_2$ ans c_3, such that for all nonnegative integer vectors β, $0 \leq |\beta| \leq s$, for all $t \in \sqrt{n}E(\hat{c}_1)$

$$\begin{aligned}&\left|D^{\beta}\left[\psi_{S_n}(t) - e^{-\frac{1}{2}\langle t,V_n t\rangle}\sum_{r=0}^{s-2} n^{-r/2}\tilde{P}_r\left(it,\kappa_{n,\nu}\right)\right]\right| \\ &\leq \hat{c}_2 n^{-(s-1)/2}\left(\|t\|^{s+1-|\beta|}+\|t\|^{3(s-1)+|\beta|}\right)e^{-\frac{c_3\|t\|^2}{4}},\end{aligned} \tag{3.19}$$

Then

$$\begin{aligned}I_1 &\leq \hat{c}_2 n^{-(s-1)/2}\int_{\sqrt{n}E(\hat{c}_1)}\left(\|t\|^{s+1-|\beta|}+\|t\|^{3(s-1)+|\beta|}\right)e^{-\frac{c_3\|t\|^2}{4}}dt \\ &\leq c_{11} n^{-(s-1)/2}\int_{\mathbb{R}^d} e^{-\frac{c_3\|t\|^2}{8}}dt\end{aligned}$$

where $c_{11} = \hat{c}_2 \cdot \sup_{t\in\mathbb{R}^d} \left\{ \left(\|t\|^{s+1-|\beta|} + \|t\|^{3(s-1)+|\beta|} \right) e^{-\frac{c_3\|t\|^2}{8}} \right\}$. Therefore,

$$I_1 = O\left(n^{-(s-1)/2}\right), \quad n \to \infty. \tag{3.20}$$

Now let us consider I_3. By Lemma 3.4.12

$$\begin{aligned} I_3 &\leq c_5 \int_{\mathbb{R}^d \setminus \sqrt{n}E(\hat{c}_1)} e^{-c_3\|t\|^2/4} dt \\ &\leq c_5 \left(\hat{c}_1\sqrt{n}\right)^{-(s-1)} \int_{\mathbb{R}^d} \|t\|^{s-1} e^{-c_3\|t\|^2/4} dt, \end{aligned}$$

where the last step holds due to the Chebyshev-Markov inequality (see e.g. [Bau92], Lemma 20.1). Therefore,

$$I_3 = O\left(n^{-(s-1)/2}\right), \quad n \to \infty. \tag{3.21}$$

Finally, we estimate I_2. We assume that $\sqrt{n}E\left(\hat{c}_1\right) \subseteq \sqrt{n}\mathcal{F}^*$, otherwise, no further calculations are necessary. By Assumption (3.13)

$$\sup\left\{\left|\psi_{X_{1,n}}(t)\right| : t \in \mathcal{F}^* \setminus E\left(\hat{c}_1\right),\ n \in \mathbb{N}\right\} = N_{\hat{c}_1} < 1, \tag{3.22}$$

Since $\left|\psi_{Y_{n,1}}\right| = \left|\psi_{X_{n,1}}\right|$, by applying Lemma 3.4.13 for each $n \in \mathbb{N}$ we get

$$\begin{aligned} I_2 &\leq \rho_{n,|\beta|} n^{|\beta|/2} \int_{\sqrt{n}\mathcal{F}^* \setminus \sqrt{n}E(\hat{c}_1)} \left|\psi_{X_{1,n}}\left(\frac{t}{\sqrt{n}}\right)\right|^{n-|\beta|} dt \\ &= \rho_{n,|\beta|} n^{|\beta|/2} n^{d/2} \int_{\mathcal{F}^* \setminus E(\hat{c}_1)} \left|\psi_{X_{1,n}}(t)\right|^{n-|\beta|} dt \\ &\leq c_{12} N_{\hat{c}_1}^{n-|\beta|} n^{|\beta|/2+d/2}, \quad \forall n \geq n_0. \end{aligned} \tag{3.23}$$

Since $N_{\hat{c}_1} < 1$, $N_{\hat{c}_1}^{n-|\beta|}$ tends to zero faster than any power of $1/n$, as $n \to \infty$. Therefore,

$$I_2 = O\left(n^{-(s-1)/2}\right). \tag{3.24}$$

From equations (3.18), (3.20), (3.21) and (3.24) we get

$$\sup_{\alpha\in L} \left|x_{\alpha,n}^{\beta}\left(p_n\left(x_{\alpha,n}\right) - q_{n,s}\left(x_{\alpha,n}\right)\right)\right| = O\left(n^{-(d+s-1)/2}\right), \quad n \to \infty, \quad |\beta| \leq s. \tag{3.25}$$

By Jensen's inequality

$$\|x\|^s \leq d^{s/2-1}\left(|x_1|^s + \cdots + |x_d|^s\right) = d^{s/2-1} \sum_{\beta=se_i, 1\leq i\leq d} \left|x^{\beta}\right|,$$

where e_i is the i-th unit vector. Hence, by setting $\beta = (0, \ldots, 0)$, $\beta = (s, 0 \ldots, 0), \ldots, \beta = (0, 0 \ldots, s)$ in (3.25) we get the necessary relation (3.14).

We now consider equation (3.15).

$$\begin{aligned}
&\sum_{\alpha \in \mathbb{Z}^d} \left| p_n\left(x_{\alpha,n}\right) - q_{n,s}\left(x_{\alpha,n}\right) \right| \\
&\leq \sup_{\alpha \in \mathbb{Z}^d} \left[\left(1 + \|x_{\alpha,n}\|^s\right) \left| p_n\left(x_{\alpha,n}\right) - q_{n,s}\left(x_{\alpha,n}\right) \right| \right] \sum_{\alpha \in \mathbb{Z}^d} \frac{1}{\left(1 + \|x_{\alpha,n}\|^s\right)}.
\end{aligned}$$

By applying (3.14) we get

$$\begin{aligned}
n^{(s-1)/2} \sum_{\alpha \in \mathbb{Z}^d} \left| p_n\left(x_{\alpha,n}\right) - q_{n,s}\left(x_{\alpha,n}\right) \right| &\leq c_{13} n^{-d/2} \sum_{\alpha \in \mathbb{Z}^d} \frac{1}{\left(1 + \|x_{\alpha,n}\|^s\right)} \\
&= c_{13} n^{-d/2} \sum_{\alpha \in \mathbb{Z}^d - n\mu_n} \frac{1}{\left(1 + \left\| \frac{\alpha}{\sqrt{n}} \right\|^s\right)} \\
&\overset{(3.11)}{\leq} c_{13} n^{-d/2} \sup_{\|\xi\| \leq \frac{1}{\sqrt{n}} vol \mathcal{F}^*} \sum_{\alpha \in \mathbb{Z}^d} \frac{1}{\left(1 + \left\| \frac{\alpha}{\sqrt{n}} + \xi \right\|^s\right)} \\
&\to \int_{\mathbb{R}^d} \frac{1}{1 + x^s} dx, \quad n \to \infty,
\end{aligned}$$

where the last integral converges if $s \geq d+1$. Therefore, in this case, (3.15) follows directly from (3.14). The proof of (3.15) in the general case follows that of Bhattacharya and Rao pp. 233-236, dealing with the n-dependent parameters in (3.12) as presented above. □

Remark 3.4.7. The main difference from the i.i.d. case in the theorem above is the uniform condition (3.13). It is necessary to ensure that $\left|\psi_{X_{1,n}}\right|^n \to 0$, $n \to \infty$ on the fundamental domain. The question now is how restrictive this condition really is, and how applicable the theorem is in practice. In general, its verification is not straightforward. However, following the idea in [KorKu07], we can state a sufficient condition for it, which holds for multidimensional and multinomial trees.

Lemma 3.4.8. *Let $\xi_{n,1}, \ldots, \xi_{n,n} \in \mathbb{R}^d$, $n \in \mathbb{N}$ be a triangular array of lattice random vectors with a common minimal lattice L and support $S = \left\{ x \in \mathbb{R}^d \;\middle|\; p_{x,n} := P\left(\xi_{n,1} = x\right) > 0 \right\}$, $|S| = m$. If for each $x \in S$ there exists a constant $K_x > 0$ such that*

$$p_{x,n} \geq K_x, \quad n \in \mathbb{N},$$

then for all constants $C > 0$, s.t. $\mathcal{F}^* \setminus E(C)$ is non-empty

$$N_C := \sup \left\{ \left| \psi_{\xi_{n,1}}(t) \right| : t \in \mathcal{F}^* \setminus E(C),\ n \in \mathbb{N} \right\} < 1, \tag{3.26}$$

Here $\mathcal{F}^*$ is the fundamental domain of L^* and $E(C)$ is defined as in Theorem 3.4.6.

Proof. Since S has a finite number of elements, define $K := \min\{K_x, x \in S\}$. Then

$$\psi_{\xi_{n,1}}(t) = \sum_{x\in S} e^{i\langle t,x\rangle} p_{x,n} = \sum_{x\in S} e^{i\langle t,x\rangle}(p_{x,n} - K) + \sum_{x\in S} e^{i\langle t,x\rangle} K$$
$$= \sum_{x\in S} e^{i\langle t,x\rangle}(p_{x,n} - K) + Km \sum_{x\in S} e^{i\langle t,x\rangle}\frac{1}{m}.$$

Set

$$\psi(t) := \sum_{x\in S} e^{i\langle t,x\rangle}\frac{1}{m},$$

which is the characteristic function of a m-nomial random vector that has the same support S, but assigns an equal probability $\frac{1}{m}$ to each attainable value. Note that $\psi(t)$ is independent of n and for any constant $C > 0$, $\delta(C) := |\psi(t)| < 1$, for $t \in \mathcal{F}^* \setminus E(C)$. Then, since $p_{x,n} \geq K_x \geq K$, for all $x \in S$,

$$\left|\psi_{\xi_{n,1}}(t)\right| \leq \sum_{x\in S}(p_{x,n} - K) + Km|\psi(t)|.$$
$$= \sum_{x\in S} p_{x,n} + Km(\delta(C) - 1) = 1 - Km(1 - \delta(C)) := \varepsilon(C).$$

Since $\delta(C) < 1$, we have $\left|\psi_{\xi_{n,1}}(t)\right| \leq \varepsilon(C) < 1$ for all $n \in \mathbb{N}$ and $t \in \mathcal{F}^* \setminus E(C)$. Therefore, we have shown (3.26). □

3.4.3 Expansions for distribution functions

We now follow the proof of Theorem 23.1 in [BR76] to obtain an Edgeworth expansion for the distribution function of a sum of triangular array lattice random vectors. However, we will first need some additional notation.

Consider the sequence of functions S_j, $j \geq 0$, $S_j : \mathbb{R} \to \mathbb{R}$ defined by the Fourier series

$$S_j(x) = \begin{cases} (-1)^{j/2-1}\sum_{n=1}^{\infty}\frac{2\cos(2n\pi x)}{(2n\pi)^j}, & j \text{ even }, j > 0 \\ (-1)^{(j-1)/2}\sum_{n=1}^{\infty}\frac{2\sin(2n\pi x)}{(2n\pi)^j}, & j \text{ odd}, \end{cases}$$

for noninteger $x \in \mathbb{R}$. These functions are periodic with period 1. S_j is continuous for $j \geq 2$, right-continuous for $j = 1$ and $S'_{j+1}(x) = S_j(x)$, $j \geq 0$ for all noninteger values x. S_j can be determined by the j-th Bernoulli polynomial, for example

$$S_1(x) = x - \frac{1}{2},$$
$$S_2(x) = \frac{1}{2}\left(x^2 - x + \frac{1}{6}\right), \quad 0 \leq x < 1.$$

Due to periodicity, for all non-integer x

$$S_1(x) = x - \lfloor x \rfloor - \frac{1}{2}.$$

Note that this is exactly the value that appeared in equation (3.9) in the heuristic explanation.

In the following let $\alpha \in \mathbb{R}^d$ be a nonnegative integer vector and $x \in \mathbb{R}^d$. Define $S_\alpha(x) : \mathbb{R}^d \to \mathbb{R}$ as

$$S_\alpha(x) = S_{\alpha_1}(x_1) \cdots S_{\alpha_k}(x_d). \tag{3.27}$$

For more details see [BR76], Chapter A.4.

Theorem 3.4.9. *Under the conditions of Theorem 3.4.6, the distribution function of S_n satisfies*

$$\begin{aligned} \sup_{x \in \mathbb{R}^d} \Bigg| & P(S_n \le x) - \sum_{|\alpha| \le s-2} n^{-|\alpha|/2} (-1)^{|\alpha|} S_\alpha\left(n\mu_n + \sqrt{n}x\right) \left(D^\alpha \Phi_{0,V_n}\right)(x) \\ & - n^{-1/2} \sum_{|\alpha| \le s-3} n^{-|\alpha|/2} (-1)^{|\alpha|} S_\alpha\left(n\mu_n + \sqrt{n}x\right) \left(D^\alpha P_1\left(-\Phi_{0,V_n}, \{\kappa_{n,\nu}\}\right)\right)(x) \\ & - \cdots - n^{-(s-2)/2} P_{s-2}\left(-\Phi_{0,V_n}, \{\kappa_{n,\nu}\}\right)\Bigg| = O\left(n^{-(s-1)/2}\right), \end{aligned} \tag{3.28}$$

where the functions S_α are defined as in (3.27).

Proof. Let

$$P(S_n \le x) = \sum_{\alpha \in \mathbb{Z}^d :\, x_{\alpha,n} \le x} p_n(x_{\alpha,n})$$

and

$$Q_{n,s}(x) := \sum_{\alpha \in \mathbb{Z}^d :\, x_{\alpha,n} \le x} q_{n,s}(x_{\alpha,n})$$

Since by equation (3.15)

$$\begin{aligned} |P(S_n \le x) - Q_{n,s}(x)| &\le \sum_{\alpha \in \mathbb{Z}^d :\, x_{\alpha,n} \le x} |p_n(x_{\alpha,n}) - q_{n,s}(x_{\alpha,n})| \\ &\le \sum_{\alpha \in \mathbb{Z}^d} |p_n(x_{\alpha,n}) - q_{n,s}(x_{\alpha,n})| \\ &= O\left(n^{-(s-1)/2}\right), \end{aligned}$$

it remains to show (3.28) with $P(S_n \le x)$ replaced by $Q_{n,s}(x)$. We will apply the generalized Euler-Maclaurin summation formula (Theorem 3.4.14) to prove the latter statement. In Theorem 3.4.14 set $r = s-1$, $h = \frac{1}{\sqrt{n}}$, $v_n = -\sqrt{n}\mu_n$ and $f_n = \sum_{r=0}^{s-2} n^{-r/2} P_r\left(-\phi_{0,V_n}, \{\kappa_{n,\nu}\}\right)$. Due to the representation of $P_r\left(-\phi_{0,V_n}, \{\kappa_{n,\nu}\}\right)$ in Lemma 3.3.3, f_n is a Schwartz function (see (3.34)) for each $n \in \mathbb{N}$.

In this setting $F_n = \sum_{j=0}^{s-2} n^{-j/2} P_j\left(-\Phi_{0,V_n}, \{\kappa_{n,\nu}\}\right)$ and

$$\begin{aligned} Q_{n,s}(x) &= \sum_{\alpha \in \mathbb{Z}^d:\, x_{\alpha,n} \le x} q_{n,s}\left(\frac{\alpha - n\mu_n}{\sqrt{n}}\right) \\ &= h^d \sum_{\alpha \in \mathbb{Z}^d:\, h\alpha + v_n \le x} f\left(h\alpha + v_n\right). \end{aligned}$$

Then by Theorem 3.4.14, for every $m > d/2$

$$\begin{aligned} &\left| Q_{n,s}(x) - \sum_{j(\alpha)<s-1} (-1)^{|\alpha|} h^{|\alpha|} S_\alpha\left(\frac{x-v}{h}\right)\left(D^\alpha F_n\right)(x) \right| \\ &= \left| Q_{n,s}(x) - \sum_{j(\alpha)<s-1} (-1)^{|\alpha|} n^{-|\alpha|/2} S_\alpha\left(x\sqrt{n} + \mu_n n\right) D^\alpha \left(\sum_{j=0}^{s-2} n^{-j/2} P_j\left(-\Phi_{0,V_n}, \{\kappa_{n,\nu}\}\right)\right)(x) \right| \\ &\le c\left(s-1, m, d\right) \sum_{s-1 \le |\gamma| \le d(s-1)} n^{-|\gamma|/2} \sup_{x \in \mathbb{R}^d} \left(\left(1 + \|x\|^2\right)^{m/2} \left|D^\gamma f_n(x)\right|\right) \end{aligned} \tag{3.29}$$

Since the right-hand side of (3.29) is independent of x by taking the supremum on both sides we get

$$\begin{aligned} &\sup_{x \in \mathbb{R}^d} \left| Q_{n,s}(x) - \sum_{j(\alpha)<s-1} (-1)^{|\alpha|} n^{-|\alpha|/2} S_\alpha\left(x\sqrt{n} + \mu_n n\right) \right. \\ &\qquad \left. \times \left(D^\alpha \sum_{j=0}^{s-2} n^{-j/2} P_j\left(-\Phi_{0,V_n} : \kappa_{n,\nu}\right)\right)(x) \right| \\ &\qquad \le c\left(s-1, m, d\right) \sum_{s-1 \le |\gamma| \le d(s-1)} n^{-|\gamma|/2} \sup_{x \in \mathbb{R}^d} \left(\left(1 + \|x\|^2\right)^{m/2} \left|D^\gamma f_n(x)\right|\right) \\ &\qquad \le c\left(s-1, m, d\right) n^{-(s-1)/2} \sum_{s-1 \le |\gamma| \le d(s-1)} \sup_{x \in \mathbb{R}^d} \left(\left(1 + \|x\|^2\right)^{m/2} \left|D^\gamma f_n(x)\right|\right). \end{aligned} \tag{3.30}$$

As in the proof of Lemma 3.4.11, there exists a constant c_{14} s.t.

$$e^{-\frac{1}{2}\left\langle x, V_n^{-1} x\right\rangle} \le e^{-\frac{1}{2} c_{14} \|x\|^2}.$$

In addition $\kappa_{n,\nu} = O(1)$ for $|\nu| \le s$, $Det(V_n) = O(1)$ and, therefore, by Lemma 3.3.3

$$\sup_{x \in \mathbb{R}^d} \left(\left(1 + \|x\|^2\right)^{m/2} \left|D^\gamma f_n(x)\right|\right) = O(1), \quad n \to \infty,$$

for all γ s.t. $s-1 \leq |\gamma| \leq d(s-1)$. Relation (3.30) now becomes

$$\sup_{x\in\mathbb{R}^d} \left| Q_{n,s}(x) - \sum_{j(\alpha)<s-1} (-1)^{|\alpha|} n^{-|\alpha|/2} S_\alpha \left(x\sqrt{n} + \mu_n n\right) \right.$$
$$\left. \times D^\alpha \left(\sum_{j=0}^{s-2} n^{-j/2} P_j \left(-\Phi_{0,V_n}, \{\kappa_{n\nu}\}\right)\right)(x) \right|$$
$$= O\left(n^{-(s-1)/2}\right).$$

Since $|\alpha| > j(\alpha)$ for all $\alpha \in \mathbb{R}^d$ we get the necessary relation (3.28) by omitting from the above expansion all terms of order $n^{-j/2}$, $j \geq s-1$, and keeping in mind that the involved cumulants are bounded. □

3.4.4 Supplementary results

We now present some additional lemmas and theorems used in the proofs above.

Lemma 3.4.10. *Let P be a probability measure on $\mathbb{R}^d$ with zero mean, positive-definite covariance matrix V, and finite s-th absolute moment ρ_s for $s \geq 3$. Then there exist two positive constants c_1, c_2, s.t. for all $t \in \mathbb{R}^d$ satisfying*

$$\|t\| \leq c_1 n^{1/2} \Lambda^{-1/2} \lambda^{s/(2(s-2))} \rho_s^{-1/(s-2)},$$

where Λ and λ are the largest and smallest eigenvalue of V, one has

$$\left| D^\alpha \left[\psi_P^n \left(\frac{t}{\sqrt{n}}\right) - e^{-\frac{1}{2}\langle t,Vt\rangle} \sum_{r=0}^{s-3} n^{-r/2} \tilde{P}_r \left(it, \{\kappa_\nu\}\right)\right]\right|$$
$$\leq \frac{c_2 \Lambda^{|\alpha|/2} \lambda^{-s/2} \rho_s}{n^{(s-2)/2}} \left[\langle t, Vt\rangle^{(s-|\alpha|)/2} + \langle t, Vt\rangle^{(3(s-2)+|\alpha|)/2}\right] e^{-\frac{1}{4}\langle t,Vt\rangle}.$$

Proof. Let B be a symmetric positive-definite matrix, such that $B^2 = V^{-1}$. Define

$$\eta_s = \int_{\mathbb{R}^d} \|Bx\|^s P(dx).$$

Since $\|B\| = \left\|V^{-1/2}\right\| = \lambda^{-1/2}$, where λ is the smallest eigenvalue of V, and $\|Bx\| \leq \|B\| \cdot \|x\|$, we have

$$\eta_s \leq \lambda^{-s/2} \rho_s.$$

The statement of the lemma follows from [BR76], Theorem 9.10 and the remark on p. 83. The matrix norms above are the induced euclidean norms. □

Lemma 3.4.11. *Let $\{V_n\}$ be a sequence of matrices that converges to a positive-definite matrix V. Then the following statements hold.*

(i) There exist positive constants c_3, c_4 and $n_0 \in \mathbb{N}$, such that

$$c_3 \|t\|^2 \leq \langle t, V_n t\rangle \leq c_4 \|t\|^2, \quad \forall n \geq n_0, \forall t \in \mathbb{R}^k.$$

(ii) If Λ_n and λ_n are the smallest and largest eigenvalues of V_n, then

$$\Lambda_n, \lambda_n^{-1} = O(1), \quad n \to \infty.$$

Proof. (i) Since V is symmetric,

$$\lambda \|t\|^2 \leq \langle t, Vt\rangle \leq \Lambda \|t\|^2, \quad \forall t \in \mathbb{R}^d,$$

where $\lambda > 0$ and $\Lambda > 0$ are the smallest and largest eigenvalue of V.

$$\begin{aligned} \langle t, V_n t\rangle &= \langle t, V_n t\rangle - \langle t, Vt\rangle + \langle t, Vt\rangle \\ &= \sum_{i,j=1}^{k} t_i t_j \left(v_{ij}^{(n)} - v_{ij}\right) + \langle t, Vt\rangle \end{aligned} \tag{3.31}$$

Since V_n converges to V, $\forall \varepsilon > 0\ \exists n_\varepsilon$, s.t. $\forall n \geq n_\varepsilon$

$$-\varepsilon \leq v_{ij}^{(n)} - v_{ij} \leq \varepsilon,$$

and from (3.31) we have

$$(-\varepsilon + \lambda) \|t\|^2 \leq \langle t, V_n t\rangle \leq (\varepsilon + \Lambda) \|t\|^2.$$

(ii) Since $\|V_n\| = \Lambda_n$ and $\|V_n^{-1}\| = \frac{1}{\lambda_n}$, the sequences $\{\Lambda_n\}$ and $\left\{\frac{1}{\lambda_n}\right\}$ converge to Λ and $\frac{1}{\lambda}$, respectively.

□

Lemma 3.4.12. *Under the conditions (3.12), for any nonnegative integral vector β, $0 \leq |\beta| \leq s$, there exist a constant c_5 and $n_0 \in \mathbb{N}$, such that*

$$\left| D^\beta \left[e^{-\frac{1}{2}\langle t, V_n t\rangle} \sum_{r=0}^{s-2} n^{-r/2} \tilde{P}_r\left(it, \{\kappa_{n,\nu}\}\right)\right]\right| \leq c_5 e^{-c_3\|t\|^2/4}, \quad \forall n \geq n_0,$$

where c_3 is as in Lemma 3.4.11 (i).

Proof.

$$\left| D^\beta \left[e^{-\frac{1}{2}\langle t, V_n t\rangle} \sum_{r=0}^{s-2} n^{-r/2} \tilde{P}_r\left(it, \{\kappa_{n,\nu}\}\right)\right]\right| \leq \sum_{r=0}^{s-2} n^{-r/2} \left| D^\beta \left[\tilde{P}_r\left(it, \{\kappa_{n,\nu}\}\right) e^{-\frac{1}{2}\langle t, V_n t\rangle}\right]\right|. \tag{3.32}$$

By applying the product rule, the derivatives in the above equation can be represented as

$$D^{\beta}\left[\tilde{P}_r\left(it,\{\kappa_{n,\nu}\}\right)e^{-\frac{1}{2}\langle t,V_nt\rangle}\right]=\sum_{0\leq\alpha\leq\beta}c\left(\alpha\right)D^{\alpha}\tilde{P}_r\left(it,\{\kappa_{n,\nu}\}\right)D^{\beta-\alpha}e^{-\frac{1}{2}\langle t,V_nt\rangle}$$

By Lemmas 3.2.11 $(i),(ii)$ and 3.4.11 (i) we get

$$\left|D^{\beta-\alpha}e^{-\frac{1}{2}\langle t,V_nt\rangle}\right|\leq c_6\left(1+\|t\|^{|\beta-\alpha|}\right)e^{-c_3\|t\|^2/2}.$$

By Lemma 3.3.1, $\tilde{P}_r\left(it,\{\kappa_{n,\nu}\}\right)$ is a polynomial of degree $3r$. Since $\rho_{n,s+1}\in O(1)$, by Lemma 9.5 in [BR76] for $0\leq|\alpha|\leq 3r$, $r=0,\ldots,s-2$

$$\begin{aligned}\left|D^{\alpha}\tilde{P}_r\left(it,\{\kappa_{n,\nu}\}\right)\right|&\leq c_7\left(1+\rho_{n,2}^{r(s-3)/(s-2)}\right)\left(1+\|t\|^{3r-|\alpha|}\right)\rho_{n,s}^{r/(s-2)}\\&\leq c_8\left(1+\|t\|^{3r-|\alpha|}\right),\end{aligned}$$

starting from some $n\in\mathbb{N}$. Therefore,

$$\begin{aligned}\left|D^{\beta}\left[e^{-\frac{1}{2}\langle t,V_nt\rangle}\sum_{r=0}^{s-2}n^{-r/2}\tilde{P}_r\left(it,\{\kappa_{n,\nu}\}\right)\right]\right|&\leq c_9e^{-c_3\|t\|^2/2}\sum_{r=0}^{s-2}n^{-r/2}\\&\quad\times\sum_{0\leq\alpha\leq\min\{\beta,3r\}}\left(1+\|t\|^{|\beta-\alpha|}+\|t\|^{3r-|\alpha|}+\|t\|^{3r-|\alpha|+|\beta-\alpha|}\right)\\&\leq c_{10}e^{-c_3\|t\|^2/4},\quad\forall n\geq n_0,\end{aligned}$$

where the last inequality holds, since any polynomial multiplied by $e^{-c_3\|t\|^2/4}$ is bounded. □

Lemma 3.4.13. *Let $X_1,\ldots,X_n\in\mathbb{R}^d$ be i.i.d. random vectors with zero mean and finite s-th moment ρ_s, and $S_n=\frac{1}{\sqrt{n}}\sum_{i=1}^n X_i$. Then for any nonnegative integral vector $\beta\in\mathbb{R}^d$, $|\beta|\leq s$*

$$\left|D^{\beta}\psi_{S_n}\left(t\right)\right|\leq\rho_{|\beta|}n^{|\beta|/2}\left(\psi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right)^{n-|\beta|},\quad t\in\mathbb{R}^d.$$

Proof. By the Leibniz' formula for differentiation of a product of n functions $D^{\beta}\psi_{S_n}\left(t\right)=D^{\beta}\left(\psi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right)^n$ can be expressed as a sum of $n^{|\beta|}$ terms of the form

$$\left(\psi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right)^{n-r}\prod_{j=1}^{r}D^{\beta_j}\left(\psi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right),\tag{3.33}$$

where $1\leq r\leq|\beta|$, $\beta_1,\ldots,\beta_r$ are nonnegative integer vectors s.t. $|\beta_j|\geq 1$ for $1\leq j\leq r$, $\sum_{j=1}^r\beta_j=\beta$. By Theorem 3.2.7 and Lemma 3.2.11 (iii) each of the derivatives in the product above can be bounded by

$$\begin{aligned}\left|D^{\beta_j}\left(\psi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right)\right|&\leq n^{-|\beta_j|/2}E\left|X_1^{\beta_j}\right|\\&\leq n^{-|\beta_j|/2}\rho_{|\beta_j|}.\end{aligned}$$

As a result

$$\prod_{j=1}^{r}\left|D^{\beta_j}\left(\psi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right)\right| \leq n^{-(|\beta_1|+\cdots+|\beta_r|)/2}\rho_{|\beta_1|}\cdots\rho_{|\beta_r|}$$
$$\leq n^{-|\beta|/2}\rho_{|\beta|}^{(|\beta_1|+\cdots+|\beta_r|)/|\beta|} = n^{-|\beta|/2}\rho_{|\beta|},$$

where the last inequality holds since $\rho_s^{1/s}$ is a non-decreasing function (Lemma 6.2(ii), [BR76]). Therefore, taking into account that $\left|\psi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right| \leq 1$ and $n-r \geq n-|\beta|$, we get

$$\left|D^{\beta}\psi_{S_n}(t)\right| \leq n^{|\beta|}\left|\psi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right|^{n-r}\prod_{j=1}^{r}\left|D^{\beta_j}\left(\psi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right)\right|$$
$$\leq \rho_{n,|\beta|}n^{|\beta|/2}\left|\psi_{X_1}\left(\frac{t}{\sqrt{n}}\right)\right|^{n-|\beta|}.$$

□

The generalized Euler-Maclaurin summation formula

In the proof of the Edgeworth expansion for distribution functions we make use of the following version of the Euler-Maclaurin formula for an integral representation of multidimensional sums (see [BR76], Theorem A.4.3).

Consider the Schwartz space $\mathcal{S}$ on $\mathbb{R}^d$ of functions all of whose derivatives are rapidly decreasing, i.e $f \in \mathcal{S}$ iff f is infinitely differentiable and

$$\sup_{x\in\mathbb{R}^d}\left|x^{\beta}\left(D^{\alpha}f\right)(x)\right| < \infty \tag{3.34}$$

for all pairs of nonnegative integer vectors α, β.

Theorem 3.4.14 (Generalized Euler-Maclaurin Formula)**.** *Let $f \in \mathcal{S}$, $v \in \mathbb{R}^d$, and $h > 0$, and let r be a positive integer. Define*

$$\Lambda_r(x) = \sum_{j(\alpha)<r}(-1)^{|\alpha|}h^{|\alpha|}S_{\alpha}\left(\frac{x-v}{h}\right)\left(D^{\alpha}F\right)(x), \tag{3.35}$$

where F is defined as

$$F(x) = \int_{-\infty}^{x_1}\cdots\int_{-\infty}^{x_d}f(y)\,dy, \quad x \in \mathbb{R}^d,$$

and for any nonnegative integer vector $\alpha \in \mathbb{R}^d$

$$j(\alpha) = \sum_{\alpha_j\geq 2}(\alpha_j - 1), \quad j(\alpha) = 0, \text{ if } \alpha_j < 2\ \forall j.$$

For every $m > d/2$ *there exists a constant* $c(r, m, d)$ *such that for all Borel sets* A

$$\left| h^d \sum_{v+hn \in A} f(v + hn) - \int_A d\Lambda_r \right| \leq c(r, m, d) \sum_{r \leq |\gamma| \leq dr} h^{|\gamma|} \nu_m (D^\gamma f), \tag{3.36}$$

where ν_m *is*

$$\nu_m(\phi) = \sup \left\{ \left(1 + \|x\|^2\right)^{m/2} |\phi(x)| : x \in \mathbb{R}^d \right\}.$$

Chapter 4

Asymptotics of One-Dimensional Tree Models

As we have seen in Chapter 2, one of the main drawbacks of tree-based methods is their irregular convergence behavior. This issue has been addressed by numerous authors and different solutions have been offered (see, e.g. [BL94], [R95], [LR96], and [T99]). However, in order to better understand the source of the problem as well as to be able to compare existing methods theoretically, one has to consider the asymptotics of the discrete models ([DD04]). Moreover, this information can then be used to construct advanced tree models with faster or smoother convergence (see [CP07] and [KM13]).

In this chapter we focus on the one-dimensional setting. We offer a general method of constructing asymptotic expansions for lattices based on an appropriate Edgeworth expansion, and discuss ways of further improving the convergence behaviour for different types of options .

4.1 Distributional fit

Consider the one-dimensional Black-Scholes model

$$dS\left(t\right)=S\left(t\right)\left(rdt+\sigma dW\left(t\right)\right),\quad S(0)=s_0. \tag{4.1}$$

Recall (see section 2.1.2), that the approximating binomial tree $\left\{S_k^{(n)}\right\}$, $k=1,\ldots,n$ is constructed as follows

$$\begin{aligned} S_0^{(n)} &= s_0, \\ S_k^{(n)} &= S_{k-1}^{(n)} e^{\alpha_n \Delta t+\sigma\sqrt{\Delta t}\xi_k^{(n)}}, \end{aligned} \tag{4.2}$$

where $\{\alpha_n\}_n$ is a bounded sequence and for each $n\in\mathbb{N}$, $\xi_k^{(n)}$, $k=1,\ldots,n$ are i.i.d random variables taking on values 1 and -1 with probabilities p_n and $1-p_n$, respectively. To ensure weak convergence we assume (see (2.4))

$$\begin{aligned}\mu(n) &= r - \frac{1}{2}\sigma^2 + o(1), \\ \sigma^2(n) &= \sigma^2 + o(1).\end{aligned} \tag{4.3}$$

As mentioned previously, there are various choices for α_n and p_n, but which is the best one? It would be very convenient if we could optimize this problem for all types of options and construct a superior tree for every situation. Unfortunately, it is not that simple. As we will see below, the optimal α_n and p_n strongly depend on the type of option considered, and even though they are derived using the same idea, their exact form will differ.

However, before we consider specific options we would like to focus on the distribution of the stock price at maturity. Due to (4.3) we know that $S^{(n)}$ converges to S in distribution, we would now like to see how well it converges, i.e. we are interested in the discretization error

$$P\left(S_n^{(n)} \leq x\right) - Q\left(S(T) \leq x\right) = P\left(S_n^{(n)} \leq x\right) - \Phi\left(d_2\right), \tag{4.4}$$

where $d_2 := d_2(x)$, and

$$d_2(x) := \frac{\ln\left(\frac{x}{s_0}\right) - \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}. \tag{4.5}$$

F. Diener and M. Diener in [DD04] have already provided asymptotics for (4.4). However, their approach is based on an integral representation of binomial sums and has only been applied to one-dimensional binomial trees. In this section we present an alternative approach that makes use of an Edgeworth expansion for lattice triangular arrays. The advantage of this approach is that it can be easily generalized to both multinomial and multidimensional trees.

Consider the discrete-time stock price at maturity

$$S_n^{(n)} = s_0 e^{\alpha_n T + \sigma\sqrt{T}\frac{1}{\sqrt{n}}\sum_{k=1}^n \xi_k^{(n)}}.$$

Then

$$\begin{aligned}P\left(S_n^{(n)} \leq x\right) &= P\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n\left(\xi_k^{(n)} - E\xi_k^{(n)}\right) \leq \frac{\ln\left(\frac{x}{s_0}\right) - \mu(n)T}{\sigma\sqrt{T}}\right), \\ &= P\left(\frac{1}{\sqrt{n}}\sum_{k=1}^n\left(\eta_k^{(n)} - E\eta_k^{(n)}\right) \leq \frac{\ln\left(\frac{x}{s_0}\right) - \mu(n)T}{2\sigma\sqrt{T}}\right),\end{aligned}$$

where

$$\eta_k^{(n)} = \frac{\xi_k^{(n)} + 1}{2}, \quad k = 1, \ldots, n. \tag{4.6}$$

We can now apply the Edgeworth expansion for lattice triangular arrays (Theorem 3.4.9) to the variables $\eta_k^{(n)}$ to get an asymptotic expansion of the terminal distribution of the stock price.

Remark 4.1.1. The change of variables from $\xi_k^{(n)}$ to $\eta_k^{(n)}$ is necessary because of the minimal lattice condition in Theorem 3.4.9. The variables $\xi_k^{(n)}$ take values ± 1, therefore their minimal lattice is $2\mathbb{Z}$. However, after the transformation (4.6), for each $n \in \mathbb{N}$, $k = 1, \dots, n$, $\eta_k^{(n)}$ has a minimal lattice $\mathbb{Z}$ and $P\left(\eta_k^{(n)} \in \mathbb{Z}\right) = 1$. Therefore, Theorem 3.4.9 can be applied to $\eta_k^{(n)}$, but not directly to $\xi_k^{(n)}$. This transformation is not the only possibility, the resulting expansion, however, will remain the same.

By Theorem 3.4.9 the following asymptotic expansion holds.

Corollary 4.1.2. *Let $\mu_n = E\eta_1^{(n)}$, $\sigma_n^2 = Var\,\eta_1^{(n)}$ and $\kappa_{n,\nu}$ be the ν-th cumulant of $\eta_1^{(n)}$. The process $S^{(n)}$ defined in (4.2) satisfies*

$$\begin{aligned}
&P\left(S_n^{(n)} \leq x\right) \\
&= \Phi_{0,\sigma_n^2}(y_n) - \frac{1}{\sqrt{n}}\phi_{0,\sigma_n^2}(y_n)\left(S_1(a_n) + \frac{\kappa_{n,3}}{6\sigma_n^2}\left(\frac{y_n^2}{\sigma_n^2} - 1\right)\right) \\
&+ \frac{1}{n}\phi_{0,\sigma_n^2}(y_n)\frac{1}{\sigma_n}\left(-\frac{y_n}{\sigma_n}S_2(a_n) + \left(S_1(a_n)\frac{\kappa_{n,3}}{3!} + \frac{\kappa_{n,4}}{4!}\right)\frac{1}{\sigma_n^2}\left(\frac{3y_n}{\sigma_n} - \frac{y_n^3}{\sigma_n^3}\right)\right) \\
&- \frac{1}{n^{3/2}}\phi_{0,\sigma_n^2}(y_n)\frac{1}{\sigma_n^2}\left(S_3(a_n)\left(\frac{y_n^2}{\sigma_n^2} - 1\right) + \left(S_2(a_n)\frac{\kappa_{n,3}}{3!} + S_1(a_n)\frac{\kappa_{n,4}}{4!}\right)\frac{1}{\sigma_n^2}\left(\frac{y_n^4}{\sigma_n^4} - \frac{6y_n^2}{\sigma_n^2} + 3\right)\right) \\
&+ O\left(\frac{1}{n^2}\right),
\end{aligned} \tag{4.7}$$

where

$$y_n = \frac{\ln\left(\frac{x}{s_0}\right) - \mu(n)T}{2\sigma\sqrt{T}}, \quad a_n = n\mu_n + y_n\sqrt{n}, \tag{4.8}$$

for $\mu(n)$ is defined as in (2.3), and

$$\begin{aligned}
&S_1(z) = \{z\} - \frac{1}{2}, \quad S_2(z) = \frac{1}{2}\left(\{z\}^2 - \{z\} + \frac{1}{6}\right), \\
&S_3(z) = \frac{1}{6}\left(\{z\}^3 - \frac{3}{2}\{z\}^2 + \frac{1}{2}\{z\}\right),
\end{aligned} \tag{4.9}$$

with $\{z\}$ denoting the fractional part of z.

Proof. By proposition 2.1.6

$$p_n = \frac{1}{2} + O\left(\frac{1}{\sqrt{n}}\right), \quad n \to \infty, \tag{4.10}$$

therefore, the assumptions (3.12) are satisfied for any s and starting from some $n \in \mathbb{N}$, Lemma 3.4.8 is applicable, so the uniform condition (3.13) also holds. Furthermore, by Lemma 3.3.3, in the one-dimensional case

$$\begin{aligned}P_1\left(-\Phi_{0,\sigma_n^2},\kappa_{n,\nu}\right)(y_n) &= -\frac{\kappa_{n,3}}{3!}D^3\Phi_{0,\sigma_n^2}(y_n)\\ &= \frac{\kappa_{n,3}}{6\sigma_n^2}\left(1-\frac{y_n^2}{\sigma_n^2}\right)\phi_{0,\sigma_n^2}(y_n),\end{aligned}$$

$$\begin{aligned}P_2\left(-\Phi_{0,\sigma_n^2},\kappa_{n,\nu}\right)(y_n) &= \frac{\kappa_{n,4}}{4!}D^4\Phi_{0,\sigma_n^2}(y_n)+\frac{\kappa_{n,3}^2}{2!(3!)^2}D^6\Phi_{0,\sigma_n^2}(y_n)\\ &= \left(\frac{\kappa_{n,4}}{4!\sigma_n^3}\left(\frac{3y_n}{\sigma_n}-\frac{y_n^3}{\sigma_n^3}\right)-\frac{\kappa_{n,3}^2}{2!(3!)^2\sigma_n^5}\left(\frac{y_n^5}{\sigma_n^5}-\frac{10y_n^3}{\sigma_n^3}+\frac{15y_n}{\sigma_n}\right)\right)\phi_{0,\sigma_n^2}(y_n),\end{aligned}$$

and

$$\begin{aligned}P_3\left(-\Phi_{0,\sigma_n^2},\kappa_{n,\nu}\right)(y_n) &= \frac{\kappa_{n,5}}{5!}D^5\Phi_{0,\sigma_n^2}(y_n)+\frac{\kappa_{n,4}\kappa_{n,3}}{3!4!}D^7\Phi_{0,\sigma_n^2}(y_n)+\frac{\kappa_{n,3}^3}{(3!)^4}D^9\Phi_{0,\sigma_n^2}(y_n)\\ &= O\left(\frac{1}{\sqrt{n}}\right),\end{aligned}$$

since due to (4.10), $\kappa_{n,3},\kappa_{n,5} = O\left(\frac{1}{\sqrt{n}}\right)$, $n \to \infty$. Therefore, by applying Theorem 3.4.9 with $s = 5$ we get the statement of the Corollary. □

4.1.1 Asymptotics of the Normal distribution

Finally, in order to get an expansion for the discretization error $P\left(S_n^{(n)} \le x\right) - \Phi(d_2)$ we need to consider the asymptotics of the terms $\Phi_{0,\sigma_n^2}(y_n)$ and $\phi_{0,\sigma_n^2}(y_n)$.

For a fixed $z \in \mathbb{R}$ and sequences $\varepsilon_{1,n}$, $\varepsilon_{2,n}$ consider the function $\Phi(z\varepsilon_{1,n}+\varepsilon_{2,n}\varepsilon_{1,n})$. We know (see e.g. [P75], Theorem I.4.13) that

$$\Phi(z\varepsilon_{1,n}+\varepsilon_{2,n}\varepsilon_{1,n}) \to \Phi(z), \quad n\to\infty, \quad \text{iff}$$
$$\varepsilon_{1,n}\to 1 \quad \text{and} \quad \varepsilon_{2,n}\to 0, \quad n\to\infty. \tag{4.11}$$

The above condition is sufficient to ensure convergence, however if we are interested in the asymptotic behavior of $\Phi(z\varepsilon_{1,n}+\varepsilon_{2,n}\varepsilon_{1,n})$ we need stronger assumptions on the errors $\varepsilon_{1,n}$ and $\varepsilon_{2,n}$. Set

$$\varepsilon_n := z(\varepsilon_{1,n}-1)+\varepsilon_{2,n}\varepsilon_{1,n}.$$

Then $z\varepsilon_{1,n}+\varepsilon_{2,n}\varepsilon_{1,n} = z+\varepsilon_n$ and due to (4.11), $\varepsilon_n \to 0$, $n\to\infty$. Consider the first order Taylor expansion of the function $\Phi(z+\varepsilon_n)$ at point z

$$\Phi(z+\varepsilon_n) = \Phi(z)+\phi(z)\varepsilon_n+O\left(\varepsilon_n^2\right).$$

As we see, the order of convergence of $\Phi(z+\varepsilon_n)$ to $\Phi(z)$ coincides with that of the error ε_n, which in turn is determined by the minimum of the rates of convergence of $\varepsilon_{1,n}$ and $\varepsilon_{2,n}$ to 1 and 0, respectively. Since, as can be seen from Corollary 4.1.2, we are interested in an asymptotic expansion in orders of $1/\sqrt{n}$ we will require the following expansions of the errors $\varepsilon_{i,n}$, $i=1,2$

$$\varepsilon_{i,n} = e_{0,n}^{(i)} + e_{1,n}^{(i)}\frac{1}{\sqrt{n}} + e_{2,n}^{(i)}\frac{1}{n} + e_{3,n}^{(i)}\frac{1}{n^{3/2}} + O\left(\frac{1}{n^2}\right), \quad i=1,2, \tag{4.12}$$

where $e_{0,n}^{(1)}=1$ and $e_{0,n}^{(2)}=0$. Then

$$\varepsilon_n = e_{1,n}\frac{1}{\sqrt{n}} + e_{2,n}\frac{1}{n} + e_{3,n}\frac{1}{n^{3/2}} + O\left(\frac{1}{n^2}\right),$$

with

$$\begin{aligned}
e_{1,n} &= ze_{1,n}^{(1)} + e_{1,n}^{(2)},\\
e_{2,n} &= ze_{2,n}^{(1)} + e_{2,n}^{(2)} + e_{1,n}^{(1)}e_{1,n}^{(2)},\\
e_{3,n} &= ze_{3,n}^{(1)} + e_{3,n}^{(2)} + e_{1,n}^{(1)}e_{2,n}^{(2)} + e_{2,n}^{(1)}e_{1,n}^{(2)}.
\end{aligned}$$

If we again apply Taylor's theorem, we get the following 3/2-order expansions

$$\begin{aligned}
\Phi(z+\varepsilon_n) = \Phi(z) &+ \frac{1}{\sqrt{n}}\phi(z)\,e_{1,n} + \frac{1}{n}\phi(z)\left(e_{2,n} - \frac{1}{2}ze_{1,n}^2\right)\\
&+ \frac{1}{n^{3/2}}\phi(z)\left(e_{3,n} - ze_{1,n}e_{2,n} + \frac{1}{6}(z^2-1)e_{1,n}^3\right) + O\left(\frac{1}{n^2}\right),
\end{aligned} \tag{4.13}$$

$$\begin{aligned}
\phi(z+\varepsilon_n) = \phi(z) &- \frac{1}{\sqrt{n}}\phi(z)\,ze_{1,n} + \frac{1}{n}\phi(z)\left(-ze_{2,n} + \frac{1}{2}(z^2-1)e_{1,n}^2\right)\\
&+ \frac{1}{n^{3/2}}\phi(z)\left(-ze_{3,n} + (z^2-1)e_{1,n}e_{2,n} + \frac{1}{6}(3z-z^3)e_{1,n}^3\right) + O\left(\frac{1}{n^2}\right).
\end{aligned} \tag{4.14}$$

Remark 4.1.3. Note that at least a $O\left(n^{-\beta}\right)$ order expansion for the error is required for a $O\left(n^{-\beta}\right)$ expansion of $\Phi(z+\varepsilon_n)$.

Now let's go back to expansion (4.7). The first term is

$$\Phi_{0,\sigma_n^2}(y_n) = \Phi\left(\frac{y_n}{\sigma_n}\right) = \Phi\left(d_2\frac{1}{2\sigma_n} + \frac{\left(r-\frac{1}{2}\sigma^2-\mu(n)\right)\sqrt{T}}{\sigma}\frac{1}{2\sigma_n}\right).$$

Therefore, if

$$\frac{1}{2\sigma_n} \quad \text{and} \quad r-\frac{1}{2}\sigma^2-\mu(n) \quad \text{satisfy} \quad (4.12), \tag{4.15}$$

then by substituting (4.13) and (4.14) into (4.7) we get an asymptotic expansion of $P\left(S_n^{(n)} \le x\right)$ around $\Phi(d_2)$ up to order $O\left(\frac{1}{n^2}\right)$.

Remark 4.1.4. Note that assumption (4.15) is a stronger version of the moment-matching conditions (4.3).

Remark 4.1.5. We would like to point out once again that the order of convergence of $\Phi_{0,\sigma_n^2}(y_n)$ to $\Phi(d_2)$ is determined by the minimum of the rates of convergence of $\mu(n)$ to $r - \frac{1}{2}\sigma^2$ and $\frac{1}{2\sigma_n}$ to 1. If the moments are matched exactly, as in the case of the RB tree, then $\Phi_{0,\sigma_n^2}(y_n)$ coincides with $\Phi(d_2)$.

Remark 4.1.6. We will not present the exact expansion of $P\left(S_n^{(n)} \leq x\right)$ around $\Phi(d_2)$ in the general case. The expansion will be calculated separately for each specific model considered in the following sections by applying Corollary 4.1.2 and the expansions of the terms on the right-hand side of equation (4.7).

4.1.2 Expansions for trinomial trees

Even though in this work we will be mainly focused on binomial tree models, we would like to show the application of Edgeworth expansions to trinomial trees and to compare the obtained results with the ones presented above.

Consider the following discrete-time model $\left\{S_k^{(n)}\right\}$

$$\begin{aligned} S_0^{(n)} &= s_0, \\ S_k^{(n)} &= S_{k-1}^{(n)} e^{\alpha_n \Delta t + \lambda\sigma\sqrt{\Delta t}\xi_k^{(n)}},\ k = 1, \ldots, n, \end{aligned} \tag{4.16}$$

where $\lambda > 1$ and $\{\alpha_n\}_n$ is a bounded sequence. However, at each time step we now allow jumps to three different values, i.e. for each $n \in \mathbb{N}$, $\xi_k^{(n)}$, $k = 1, \ldots, n$ are i.i.d random variables s.t.

$$\xi_k^{(n)} = \begin{cases} 1, & \text{with probability } p_n^{(u)}, \\ 0, & \text{with probability } 1 - p_n^{(u)} - p_n^{(d)}, \\ -1, & \text{with probability } p_n^{(d)}. \end{cases}$$

The parameters α_n and p_n are again chosen such that the moment-matching conditions (4.3) are satisfied. The distribution of the terminal stock price is calculated as before, i.e.

$$P\left(S_n^{(n)} \leq x\right) = P\left(\frac{1}{\sqrt{n}} \sum_{k=1}^{n} \left(\xi_k^{(n)} - E\xi_k^{(n)}\right) \leq \frac{\ln\left(\frac{x}{s_0}\right) - \mu(n)T}{\lambda\sigma\sqrt{T}}\right),$$

however, in this case the minimal lattice of $\xi_k^{(n)}$, $k = 1, \ldots, n$ is $\mathbb{Z}$ and $P\left(\xi_k^{(n)} \in \mathbb{Z}\right) = 1$. Therefore, unlike in the binomial model, Theorem 3.4.9 is applicable directly to the variables $\xi_k^{(n)}$ and a transformation such as (4.6) is not necessary. As a result we can obtain the same expansion as in Corollary 4.1.2 with $y_n = \frac{\ln\left(\frac{x}{s_0}\right) - \mu(n)T}{\lambda\sigma\sqrt{T}}$, $\mu_n = E\xi_k^{(n)}$, $\sigma_n^2 = Var\xi_k^{(n)}$ and $\kappa_{n,\nu}$ be the ν-th cumulant of $\xi_1^{(n)}$.

This seemingly small difference, actually has a large effect on the convergence pattern, and explains the absence of the even-odd effect (see section 2.2) in trinomial trees. Consider the following table

Model	Binomial	Trinomial
Minimal lattice	$2\mathbb{Z}$	$\mathbb{Z}$
Leading error term	$S_1\left(\sqrt{n}\frac{\ln\left(\frac{x}{s_0}\right)-\alpha_n T}{2\sigma\sqrt{T}}+\frac{n}{2}\right)$	$S_1\left(\sqrt{n}\frac{\ln\left(\frac{x}{s_0}\right)-\alpha_n T}{\lambda\sigma\sqrt{T}}\right)$

Note the additional $n/2$ term in the binomial tree, which results in a jump of the value of S_1 for even and odd n. Since this term is not present in the trinomial expansion the even-odd effect does not occur in this case. Compare the convergence behavior of binomial (left-hand side) and trinomial (right-hand side) trees.

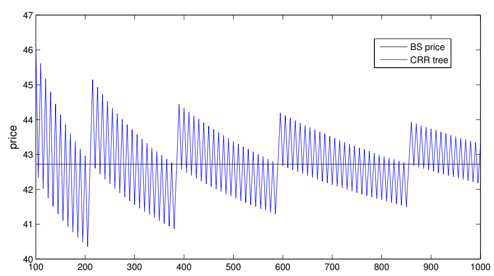


Figure 4.1: Binomial tree

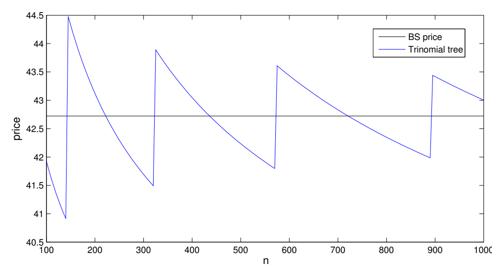


Figure 4.2: Trinomial tree with $\lambda = \sqrt{3/2}$

4.2 Improving the convergence behavior

We will now see how Corollary 4.1.2 can be used to influence the convergence pattern of one-dimensional binomial trees.

4.2.1 Existing methods in literature

As mentioned in [DD04], the irregularities in the convergence behavior of tree-methods can be explained by the periodic, n-dependent term $S_1(a_n)$ in (4.7). Since $S_1(a_n) = \{a_n\} - \frac{1}{2}$ does not have a limit as n goes to infinity and oscillates between $-\frac{1}{2}$ and $\frac{1}{2}$, even large values of n do not guarantee accurate results. As a solution, various methods have been offered to improve convergence by controlling the S_1 function and with that the leading error term. However, basically they pursue one of the following goals.

The first one, see e.g. [T99], is to achieve smooth convergence behavior, so that extrapolation methods can be applied to increase the order of convergence. The second one, see e.g. [CP07], is to construct the tree such that the leading error term becomes zero, thus increasing the order of convergence directly.

We now present the above methods in a little more detail. Consider the CRR tree model

$$S_k^{(n)} = S_{k-1}^{(n)} e^{\sigma\sqrt{\Delta t}\xi_k^{(n)}},\ k = 1, \ldots, n, \tag{4.17}$$

with risk-neutral transition probabilities p_n.

Fix the point $x \in \mathbb{R}$. The idea now is to tilt the tree, i.e. add a drift α_n for each $n \in \mathbb{N}$, in such a way that the point x always falls onto a fixed position between two neighboring nodes. As we will see shortly, this will also fix the function $S_1(a_n)$ and ensure that the leading error term remains constant in n. To be more precise, we want to construct a tilted tree $\left\{S_{k,\alpha_n}^{(n)}\right\}$, such that

$$S_{k,\alpha_n}^{(n)} = S_{k-1,\alpha_n}^{(n)} e^{\alpha_n \Delta t + \sigma\sqrt{\Delta t}\xi_k^{(n)}},\ k = 1, \ldots, n,$$

with risk-neutral transition probabilities p_{n,α_n} and an appropriate drift α_n.

For any drift α_n, the process $S_{n,\alpha_n}^{(n)}$ takes on the following values

$$S_{n,\alpha_n}(l) := s_0 e^{\alpha_n T + 2l\sigma\sqrt{\Delta t} - \sigma\sqrt{T}\sqrt{n}},$$

where $l \in \{0, \ldots, n\}$ denotes the possible values of $\sum_{k=1}^n \eta_k^{(n)}$, and can be interpreted as the number of up-jumps in the model. For each $n \in \mathbb{N}$ we now want to determine the two nodes, closest to the value x. For this we consider the following equation in $a \in \mathbb{R}$

$$s_0 e^{\alpha_n T + 2a\sigma\sqrt{\Delta t} - \sigma\sqrt{T}\sqrt{n}} = x. \tag{4.18}$$

Equation (4.18) is solved by

$$a_{n,\alpha_n} = \frac{\log(x/s_0) - \alpha_n T}{2\sigma\sqrt{\Delta t}} + \frac{n}{2}. \tag{4.19}$$

We now set $l_{n,\alpha_n} := \lfloor a_{n,\alpha_n} \rfloor$, i.e. l_{n,α_n} is the number of up-jumps, such that $S_{n,\alpha_n}(l_{n,\alpha_n}) \leq x < S_{n,\alpha_n}(l_{n,\alpha_n} + 1)$. The position of x with respect to the two neighboring nodes in the log-scale can be determined as

$$c_{n,\alpha_n} := \frac{\log(x) - \log(S_{n,\alpha_n}(l_{n,\alpha_n}))}{\log(S_{n,\alpha_n}(l_{n,\alpha_n} + 1)) - \log(S_{n,\alpha_n}(l_{n,\alpha_n}))} = a_{n,\alpha_n} - l_{n,\alpha_n} = \{a_{n,\alpha_n}\}, \tag{4.20}$$

(compare [M09], Section 2.2.3). Clearly, $0 \leq c_{n,\alpha_n} < 1$ for every α_n.

Let us go back to our problem. Our goal is to slightly shift the original CRR tree (4.17), so that c_{n,α_n}, i.e. the relative position of x, becomes constant in n. In other words, the new drift α_n has to satisfy the equation

$$a_{n,\alpha_n} = l_{n,0} + c, \tag{4.21}$$

for a fixed c, $0 \le c < 1$. Note that we require $l_{n,\alpha_n} = l_{n,0}$, so that the tree is not shifted too much and the neighboring nodes have the same number of up-jumps in both the original and tilted trees. By substituting (4.19) into (4.21) we get the following solution

$$\alpha_n^* = \frac{2\sigma\sqrt{\Delta t}\left(\{a_{n,0}\} - c\right)}{T}. \tag{4.22}$$

The above expression is bounded, therefore, $S_{n,\alpha_n^*}^{(n)}$ satisfies the moment-matching conditions, and by Corollary 4.1.2

$$P\left(S_{n,\alpha_n^*}^{(n)} \le x\right) = \Phi\left(d_2\right) - \frac{2}{\sqrt{n}}\phi\left(d_2\right) S_1\left(a_{n,\alpha_n^*}\right) + O\left(\frac{1}{n}\right). \tag{4.23}$$

The same expansion can be obtained by the methods presented in [DD04] (see e.g. [KM13] or [M09]). Note that

$$S_1\left(a_{n,\alpha_n^*}\right) = \left\{a_{n,\alpha_n^*}\right\} - \frac{1}{2} = c - \frac{1}{2}. \tag{4.24}$$

Therefore, as mentioned before, by fixing the position of the x to c, we also determine the term S_1.

Chang-Palmer (CP) approach. If x is fixed to the geometric average of the neighboring nodes $S_{n,\alpha_n^*}\left(l_{n,0}\right)$ and $S_{n,\alpha_n^*}\left(l_{n,0}+1\right)$, i.e. $c = \frac{1}{2}$, we have $S_1\left(a_{n,\alpha_n^*}\right) = 0$ and expansion (4.23) becomes

$$P\left(S_{n,\alpha_n^*}^{(n)} \le x\right) = \Phi\left(d_2\right) + O\left(\frac{1}{n}\right),$$

In this case $S_{n,\alpha_n^*}^{(n)}$ has a higher order of convergence (see [CP07] or [M09]).

Tian approach. If x is fixed to the lower node $S_{n,\alpha_n^*}\left(l_{n,0}\right)$, i.e. $c = 0$, then $S_1\left(a_{n,\alpha_n^*}\right) = -\frac{1}{2}$ and

$$P\left(S_{n,\alpha_n^*}^{(n)} \le x\right) = \Phi\left(d_2\right) + \frac{1}{\sqrt{n}}\phi\left(d_2\right) + O\left(\frac{1}{n}\right).$$

In this case the convergence is of the same order, but with a constant leading coefficient. This means that extrapolation can be applied to obtain the order $O\left(\frac{1}{n}\right)$, i.e. we consider the approximation

$$\hat{F}_n(x) = \frac{\sqrt{2}F_n(x) - F_{n/2}(x)}{\sqrt{2}-1} = \Phi\left(d_2\right) + O\left(\frac{1}{n}\right).$$

(Cf. [T99] or [M09]).

Remark 4.2.1. In the Tian approach c is chosen equal to 0. However, if we take any other value, $0 \leq c < 1$, $c \neq \frac{1}{2}$, the procedure will remain the same. The leading coefficient will be different, but the method still applies.

The methods described above concentrate only on the leading error term allowing to increase the order of convergence up to $O\left(\frac{1}{n}\right)$. However, if we also incorporate the subsequent terms in expansion (4.23), it is possible to further improve the convergence behavior.

Optimal Drift (OD) model. Instead of the traditional CRR model (4.17) consider the following construction

$$S_k^{(n)} = S_{k-1}^{(n)} e^{\alpha\Delta t + \sigma\sqrt{\Delta t}\xi_k^{(n)}},\ k = 1, \dots, n, \tag{4.25}$$

with a constant drift α and risk-neutral transition probabilities p_n. The additional parameter α gives the necessary flexibility to increase the order of convergence to $o\left(\frac{1}{n}\right)$ for most parameter settings used in practice. The model above satisfies

$$P\left(S_{n,\alpha_n^*}^{(n)} \leq x\right) = \Phi\left(d_2\right) - \frac{2}{\sqrt{n}}\phi\left(d_2\right) S_1\left(a_{n,\alpha_n^*}\right) + \frac{1}{n}\phi\left(d_2\right) f(\alpha_n^*) + O\left(\frac{1}{n^{3/2}}\right), \tag{4.26}$$

where $f(\alpha_n^*)$ is a bounded function (cf. expansion (4.36) or [M09], Proposition 9.). We now proceed as follows.

First, following the CP approach, a new drift α_n^* is introduced, so that the leading error term becomes zero, i.e.

$$\alpha_n^* = \frac{2\sigma\sqrt{\Delta t}\left(\{a_{n,0}\} - \frac{1}{2}\right)}{T} + \alpha.$$

Then, we choose the parameter α so that the second error term in expansion (4.26) is minimized. With the above choice of α_n^*, the term f becomes a quadratic function of the drift α. Therefore, there are two possibilities now.

- If the parabola f intersects the α-axis, then α is chosen from the equation $f = 0$. In this case the second error term is also zero and the order of convergence becomes $O\left(\frac{1}{n^{3/2}}\right)$.
- If f doesn't intersect the α-axis, then α is chosen as the vertex of the parabola. In this case the order of convergence remains $O\left(\frac{1}{n}\right)$, but it is optimized in the sense that the leading error term will have the minimal value possible for models of type (4.25).

Note, however, that in most situations of interest, the OD model will have the improved rate $O\left(\frac{1}{n^{3/2}}\right)$. For details see [KM13] or [M09], Proposition 14.

Remark 4.2.2. The mentioned references give the improved order $o\left(\frac{1}{n}\right)$, however, with Corollary 4.1.2 this can be stated more precisely as $O\left(\frac{1}{n^{3/2}}\right)$.

4.2.2 The Rendleman-Bartter model

As we have seen, most of the methods available in literature of improving convergence of one-dimensional binomial trees are based on the CRR tree, while very little attention has been given to the RB tree. However, the RB tree has some very appealing properties, such as a less pronounced saw-tooth effect compared to the CRR model and a faster backward induction algorithm due to the constant transition probabilities (see section 2.2). Therefore, in this section we investigate if it is possible to apply methods similar to that of Tian and Chang and Palmer to improve the convergence of the RB tree.

The model is defined as follows

$$S_k^{(n)} = S_{k-1}^{(n)} e^{\left(r-\frac{1}{2}\sigma^2\right)\Delta t + \sigma\sqrt{\Delta t}\xi_k}, \quad k = 1, \ldots, n, \tag{4.27}$$

with $p_n = \frac{1}{2}$.

Let $x \in \mathbb{R}$. We now want to tilt the tree such that the relative position of x with respect to the neighboring nodes is fixed, i.e. we consider the model

$$S_{k,\gamma_n}^{(n)} = S_{k-1,\gamma_n}^{(n)} e^{\left(r-\frac{1}{2}\sigma^2\right)\Delta t + \gamma_n\sqrt{\Delta t}^3 + \sigma\sqrt{\Delta t}\xi_k}, \quad k = 1, \ldots, n, \tag{4.28}$$

with transition probabilities 1/2 and a bounded sequence $\{\gamma_n\}$. Following the procedure from the previous section we fix the constant position c, $0 \leq c < 1$, and by the same arguments as for (4.22) we find

$$\gamma_n^* = \frac{2\sigma\left(\{a_{n,0}\} - c\right)}{T},$$

where

$$a_{n,\gamma_n} := \frac{\ln\left(\frac{x}{s_0}\right) - \left(r - \frac{1}{2}\sigma^2\right)T - \gamma_n T\sqrt{\Delta t}}{2\sigma\sqrt{\Delta t}} + n/2.$$

However, by Corollary 4.1.2 and expansion (4.13), $S_{n,\gamma_n^*}^{(n)}$ satisfies

$$P\left(S_{n,\gamma_n^*}^{(n)} \leq x\right) = \Phi\left(d_2(x)\right) - \frac{1}{\sqrt{n}}\phi\left(d_2\right)\left(2S_1\left(a_{n,\gamma_n^*}\right) + \frac{\gamma_n^* T}{\sigma}\right) + O\left(\frac{1}{n}\right). \tag{4.29}$$

Therefore, in this case the leading error term of the tilted tree will not be constant. Indeed,

$$\begin{aligned}
2S_1\left(a_{n,\gamma_n^*}\right) + \frac{\gamma_n^* T}{\sigma} &= 2\left\{a_{n,\gamma_n^*}\right\} - 1 + 2\left(\{a_{n,0}\} - c\right) \\
&= 2\left\{\lfloor a_{n,0}\rfloor + c\right\} - 1 + 2\left\{a_{n,0}\right\} - 2c \\
&= 2\left\{a_{n,0}\right\} - 1 \\
&= 2S_1\left(a_{n,0}\right),
\end{aligned}$$

which coincides with the coefficient in the expansion of the original RB tree (4.27). This means that fixing the position of x by tilting the tree will lead to the same convergence behavior. Compare the following results for a cash-or-nothing option with $s_0 = 100$, $T = 1$, $r = 0.1$, $\sigma = 0.25$, strike $= 100$ and payout $= 100$.

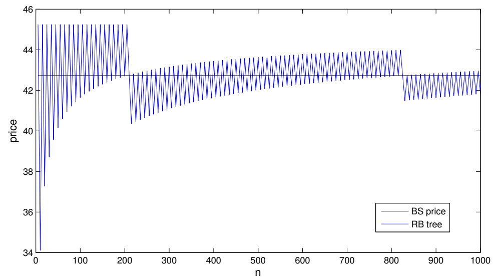


Figure 4.3: Standard RB tree

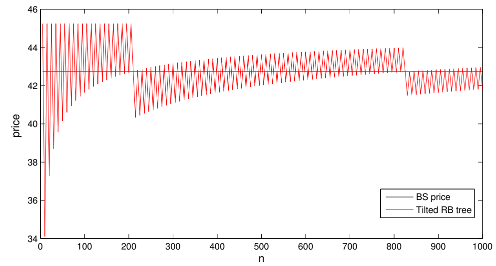


Figure 4.4: Tilted RB tree

Let us look at the problem from a slightly different angle. Consider the expansion (4.29). In order to improve convergence we need to find, for each $n \in \mathbb{N}$, a γ_n^*, such that

$$\begin{aligned} & 2S_1\left(a_{n,\gamma_n^*}\right) + \frac{\gamma_n^* T}{\sigma} = 0 \\ \Leftrightarrow \quad & 2\left\{a_{n,0} - \frac{\gamma_n^* T}{2\sigma}\right\} - 1 + \frac{\gamma_n^* T}{\sigma} = 0 \\ \Leftrightarrow \quad & a_{n,0} - \frac{1}{2} = \left\lfloor a_{n,0} - \frac{\gamma_n^* T}{2\sigma} \right\rfloor \end{aligned} \tag{4.30}$$

The above equation can be solved if $a_{n,0} - \frac{1}{2} = m$, for some $m \in \mathbb{Z}$. In this case we can choose γ_n^* to be any value such that $\frac{\gamma_n^* T}{2\sigma} = \left(-\frac{1}{2}, \frac{1}{2}\right]$. However, it is clear, that $a_{n,0} - \frac{1}{2}$ does not attain integer values for all n, therefore, there is no sequence $\{\gamma_n^*\}$, such that the leading error term in (4.29) vanishes for all $n \in \mathbb{N}$. Once again, this confirms that the convergence behavior of the RB tree cannot be improved by simply changing the drift.

Remark 4.2.3. Note that the above results suggest, that any changes to the drift have to be balanced out by an appropriate change in the probabilities.

4.2.3 The 3/2 - Optimal model

We now consider a general setting that includes both the RB and CRR tree and show how convergence can be improved in this wider class of models.

Note that the problem of optimizing the convergence of the CRR tree to a certain order has already been addressed by different authors. The aforementioned OD model of R. Korn and S. Müller [KM13] allows to improve convergence up to order $O\left(\frac{1}{n^{3/2}}\right)$ and the method introduced in [L12] for vanilla options can also be adjusted to further improve the distributional fit, as well. However, these approaches are restricted to risk-neutral probabilities and involve solving quadratic equations, which rules out certain model parameters. We now present a slightly different approach, that involves only linear equations and is, therefore, applicable to any parameter setting.

Consider the following model

$$S_k^{(n)} = S_{k-1}^{(n)} e^{\alpha_n \Delta t + \sigma\sqrt{\Delta t}\xi_k^{(n)}}, k = 1, \dots, n, \tag{4.31}$$

where

$$\xi_k^{(n)} = \begin{cases} 1, & \text{with probability } p_n \\ -1, & \text{with probability } 1 - p_n \end{cases}.$$

and

$$\begin{aligned} p_n &= \frac{1}{2} + c_{1,n}\frac{1}{\sqrt{n}} + c_{2,n}\frac{1}{n} + c_{3,n}\frac{1}{\sqrt{n}^3} + c_{4,n}\frac{1}{n^2} + O\left(\frac{1}{n^{5/2}}\right), \\ \alpha_n &= k_{0,n} + k_{1,n}\frac{1}{\sqrt{n}} + k_{2,n}\frac{1}{n} + k_{3,n}\frac{1}{\sqrt{n}^3} + O\left(\frac{1}{n^2}\right), \end{aligned} \tag{4.32}$$

with bounded $c_{i,n}$, $k_{i,n}$ and

$$k_{0,n} + \frac{2\sigma c_{1,n}}{\sqrt{T}} = r - \frac{1}{2}\sigma^2. \tag{4.33}$$

Note that (4.32)-(4.33) are necessary and sufficient for the moment-matching condition (4.15) to be satisfied (see also proposition 2.1.6). From Corollary 4.1.2 we have the following proposition.

Proposition 4.2.4. *With an appropriate choice of parameters $c_{i,n}$, $i = 0, \dots, 4$ and $k_{j,n}$ $j = 0, \dots, 3$ the binomial process $S^{(n)}$ in (4.31) satisfies*

$$P\left(S_n^{(n)} \leq x\right) = \Phi(d_2) + O\left(\frac{1}{n^2}\right)$$

Proof. We assume the notation of Corollary 4.1.2. Given (4.32) we get the following dynamics

$$\mu(n) = r - \frac{1}{2}\sigma^2 + m_1\frac{\sigma}{\sqrt{T}}\frac{1}{\sqrt{n}} + m_2\frac{\sigma}{\sqrt{T}}\frac{1}{n} + m_3\frac{\sigma}{\sqrt{T}}\frac{1}{\sqrt{n}^3} + O\left(\frac{1}{n^2}\right),$$

with

$$m_i = \left(k_{i,n} + \frac{2\sigma c_{i+1,n}}{\sqrt{T}}\right)\frac{\sqrt{T}}{\sigma}, \quad i = 1, 2, 3, \tag{4.34}$$

and

$$\sigma_n^2 = p_n - (p_n)^2 = \frac{1}{4} - c_{1,n}^2\frac{1}{n} - 2c_{1,n}c_{2,n}\frac{1}{\sqrt{n}^3} + O\left(\frac{1}{n^2}\right).$$

Therefore, by the Binomial Series Theorem (Taylor series at 0 for $(1 + x)^\alpha$, $\alpha \in \mathbb{R}$)

$$\frac{1}{2\sigma_n} = 1 + 2c_{1,n}^2\frac{1}{n} + 4c_{1,n}c_{2,n}\frac{1}{\sqrt{n}^3} + O\left(\frac{1}{n^2}\right),$$

and as a result, for y_n as in (4.8)

$$\frac{y_n}{\sigma_n} = d_2 - m_1 \frac{1}{\sqrt{n}} + \left(2d_2 c_{1,n}^2 - m_2\right) \frac{1}{n} + \left(4d_2 c_{1,n} c_{2,n} - m_3 - 2c_{1,n}^2 m_1\right) \frac{1}{\sqrt{n}^3} + O\left(\frac{1}{n^2}\right). \quad (4.35)$$

Then, by (4.13) and (4.14)

$$\begin{aligned}\Phi_{0,\sigma_n^2}(y_n) = {} & \Phi(d_2) - \frac{1}{\sqrt{n}}\phi(d_2) m_1 + \frac{1}{n}\phi(d_2)\left(2d_2 c_{1,n}^2 - m_2 - \frac{1}{2} d_2 m_1^2\right) \\ & + \frac{1}{\sqrt{n}^3}\phi(d_2)\left(4d_2 c_{1,n} c_{2,n} - m_3 - d_2 m_1 m_2 + \left(d_2^2 - 1\right) m_1 \left(2c_1^2 - \frac{1}{6} m_1^2\right)\right) + O\left(\frac{1}{n^2}\right),\end{aligned}$$

$$\phi_{0,\sigma_n^2}(y_n) = \phi(d_2) + \frac{1}{\sqrt{n}}\phi(d_2) d_2 m_1 + \frac{1}{n}\phi(d_2)\left(m_2 d_2 - 2d_2^2 c_{1,n}^2 + \frac{1}{2}\left(d_2^2 - 1\right) m_1^2\right) + O\left(\frac{1}{\sqrt{n}^3}\right).$$

Let $\mu_{n,\nu}$, $\nu \in \mathbb{Z}$ denote the ν-th moment of $\eta_1^{(n)}$. Then by (3.3) the cumulants can be represented as

$$\begin{aligned}\kappa_{n,3} &= \mu_{n,3} - 3\mu_{n,1}\mu_{n,2} + 2\mu_{n,1}^3 \\ &= -\frac{1}{2} c_{1,n} \frac{1}{\sqrt{n}} - \frac{1}{2} c_{2,n} \frac{1}{n} + O\left(\frac{1}{\sqrt{n}^3}\right), \\ \kappa_{n,4} &= \mu_{n,4} - 4\mu_{n,3}\mu_{n,1} - 3\mu_{n,2}^2 + 12\mu_{n,2}\mu_{n,1}^2 - 6\mu_{n,1}^4 \\ &= -\frac{1}{8} + O\left(\frac{1}{n}\right).\end{aligned}$$

Substituting the above expansions into (4.7) we get

$$\begin{aligned}P\left(S_n^{(n)} \le x\right) = {} & \Phi(d_2) - \frac{1}{\sqrt{n}}\phi(d_2) f_1(\alpha_n, p_n) + \frac{1}{n}\phi(d_2) f_2(\alpha_n, p_n) \\ & + \frac{1}{\sqrt{n}^3}\phi(d_2) f_3(\alpha_n, p_n) + O\left(\frac{1}{n^2}\right),\end{aligned} \quad (4.36)$$

where

$$\begin{aligned}f_1(\alpha_n, p_n) &:= m_1 + 2S_1(a_n) \\ f_2(\alpha_n, p_n) &:= -m_2 + 2d_2 c_{1,n}^2 + \frac{2}{3} c_{1,n}\left(d_2^2 - 1\right) + \frac{d_2^3 - d_2}{12} - \frac{d_2}{2}\left(m_1 + 2S_1(a_n)\right)^2 \\ f_3(\alpha_n, p_n) &:= -m_3 + 4d_2 c_{1,n} c_{2,n} + \frac{2}{3} c_{2,n}\left(d_2^2 - 1\right) - 8S_3(a_n)\left(d_2^2 - 1\right) \\ & \quad - \frac{1}{6} m_1 \left(\left(d_2^2 - 1\right) m_1^2 + \left(d_2^2 - 1\right) m_1 S_1(a_n) + 4S_2(a_n)\left(d_2^2 - 1\right)\right) \\ & \quad - \left(m_1 + 2S_1(a_n)\right)\left(2c_{1,n}^2\left(1 - d_2^2\right) + d_2 m_2 + 2c_{1,n} d_2 - \frac{2}{3} c_{1,n} d_2^3 - \frac{1}{12}\left(d_2^4 - 6d_2^2 + 3\right)\right).\end{aligned} \quad (4.37)$$

The goal now is to choose the coefficients $c_{i,n}$, $k_{i,n}$ so that $f_i = 0$, $i = 1, 2, 3$. Clearly, there are various ways of doing that, since we have more variables than equations.

In [M09] risk-neutral probabilities are considered, therefore, the coefficients $c_{i,n}$ are completely determined by the drift, and they can't be chosen freely. In this case only the coefficients of the drift $\alpha_n = k_{0,n} + k_{1,n}\frac{1}{\sqrt{n}}$ are chosen. First, the equation $f_2(\alpha_n, p_n) = 0$ is solved, which is quadratic in $k_{0,n}$. Then the obtained value is substituted into $f_1(\alpha_n, p_n)$, and $k_{1,n}$ is chosen such that $f_1(\alpha_n, p_n) = 0$. As a result, the order $O\left(\frac{1}{\sqrt{n}^3}\right)$ can be obtained. If we now follow [L12] and add an additional coefficient $k_{3,n}$, we can also solve the equation $f_3(\alpha_n, p_n) = 0$. Note that with this approach the equation $f_1(\alpha_n, p_n) = 0$ has to be reevaluated with each additional coefficient, and since $k_{i,n}$, $i \geq 3$ depend on $k_{1,n}$, this might involve complicated calculations.

Instead, we propose a slightly different approach, where the coefficients of the probabilities are chosen instead of the drift. This way we are able to avoid quadratic equations, and hence we are able to increase the order of convergence for any parameter setting.

Set $k_{2,n} = k_{3,n} = 0$, however the same method holds for any values independent of n. Now choose $c_{2,n}$ such that $f_1(\alpha_n, p_n) = 0$ is satisfied, i.e.

$$
\begin{aligned}
& 2S_1(a_n) + \left(k_{1,n} + \frac{2\sigma c_{2,n}}{\sqrt{T}}\right)\frac{\sqrt{T}}{\sigma} = 0 \\
\Leftrightarrow \quad & \frac{\ln\left(\frac{x}{s_0}\right) - k_{0,n}T}{2\sigma\sqrt{\Delta t}} + \frac{n}{2} - \frac{1}{2} + c_{2,n} = \left\lfloor \frac{\ln\left(\frac{x}{s_0}\right) - k_{0,n}T - k_{1,n}\frac{T}{\sqrt{n}}}{2\sigma\sqrt{\Delta t}} + \frac{n}{2} \right\rfloor
\end{aligned}
$$

The above equation is solved by

$$
c_{2,n} = \frac{1}{2} - \left\{ \frac{\ln\left(\frac{x}{s_0}\right) - k_{0,n}T}{2\sigma\sqrt{\Delta t}} + \frac{n}{2} \right\}, \tag{4.38}
$$

and $k_{1,n}$ taking any of the following values

$$
-\frac{2\sigma c_{2,n}}{\sqrt{T}} - \frac{\sigma}{\sqrt{T}} < k_{1,n} \leq -\frac{2\sigma c_{2,n}}{\sqrt{T}} + \frac{\sigma}{\sqrt{T}}. \tag{4.39}
$$

The other terms in (4.36) become

$$
\begin{aligned}
f_2(\alpha_n, p_n) &= -2c_{3,n} + 2d_2 c_{1,n}^2 + \frac{2}{3}c_{1,n}\left(d_2^2 - 1\right) + \frac{d_2^3 - d_2}{12}, \\
f_3(\alpha_n, p_n) &= -2c_{4,n} + 4d_2 c_{1,n}c_{2,n} + \frac{2}{3}c_{2,n}\left(d_2^2 - 1\right), \\
&\quad - 8S_3(a_n)\left(d_2^2 - 1\right) - \frac{1}{6}\left(d_2^2 - 1\right)\left(m_1^3 - m_1\right),
\end{aligned}
$$

where we have used $S_2(a_n) = \frac{1}{2}\left(S_1^2(a_n) - \frac{1}{12}\right)$ (see(4.9)). Therefore, if we set

$$c_{3,n} = d_2 c_{1,n}^2 + \frac{1}{3} c_{1,n}\left(d_2^2 - 1\right) + \frac{d_2^3 - d_2}{24}, \tag{4.40}$$

and

$$\begin{aligned} c_{4,n} = 2d_2 c_{1,n} c_{2,n} &+ \frac{1}{3} c_{2,n}\left(d_2^2 - 1\right) - 4S_3(a_n)\left(d_2^2 - 1\right) \\ &- \frac{1}{12}\left(d_2^2 - 1\right)\left(m_1^3 - m_1\right) \end{aligned} \tag{4.41}$$

all chosen coefficients are bounded, and we get the statement of the proposition. □

Remark 4.2.5. Note that the name 3/2-Optimal model refers to the optimized convergence up to and including order 3/2, in the sense that the corresponding coefficients are set to zero. In the later sections we will use the name α-Optimal model for an optimized convergence of order α.

Remark 4.2.6. Note that the coefficient $k_{1,n}$ can be chosen freely, as long as (4.39) is satisfied. Numerical experiments suggest that the value of $k_{1,n}$ does not influence the $O\left(\frac{1}{n^2}\right)$ convergence behaviour. In this work we will usually choose $k_{1,n} = -\frac{2\sigma c_{2,n}}{\sqrt{T}}$, since in this case $m_1 = 0$, $S_3(a_n) = 0$ and, therefore, the values assigned to $c_{3,n}$ and $c_{4,n}$ have a simpler form and require less computations.

Remark 4.2.7. Other than (4.33), there are also no restrictions on $k_{0,n}$ and $c_{1,n}$. If we set $k_{0,n} = 0$, for example, we are in the CRR setting, if $c_{1,n} = 0$, then we have the RB tree extension. Either way, the order of convergence is $O\left(\frac{1}{n^2}\right)$, however, $k_{0,n}$ does influence the exact convergence pattern. The optimal choice of $k_{0,n}$ is still an open question.

Consider the following plots with the convergence behavior of the 3/2-Optimal RB-based and CRR-based trees. To verify the order of convergence we will use a log-log plot (right-hand side), i.e. we plot $\log(|F_n - \Phi(d_2)|)$ against $\log(n)$. Since by Proposition 4.2.4

$$|F_n(x) - \Phi(d_2(x))| = c(n)\frac{1}{n^2},$$

for $c(n) \in O(1)$, then

$$\log\left(|F_n(x) - \Phi(d_2(x))|\right) = -2\log(n) + \log(c(n)),$$

and we should see the -2 slope on the graph. Of course, since $c(n)$ is not constant, the convergence is not smooth, and the graph is not a straight line, but the general trend is present.

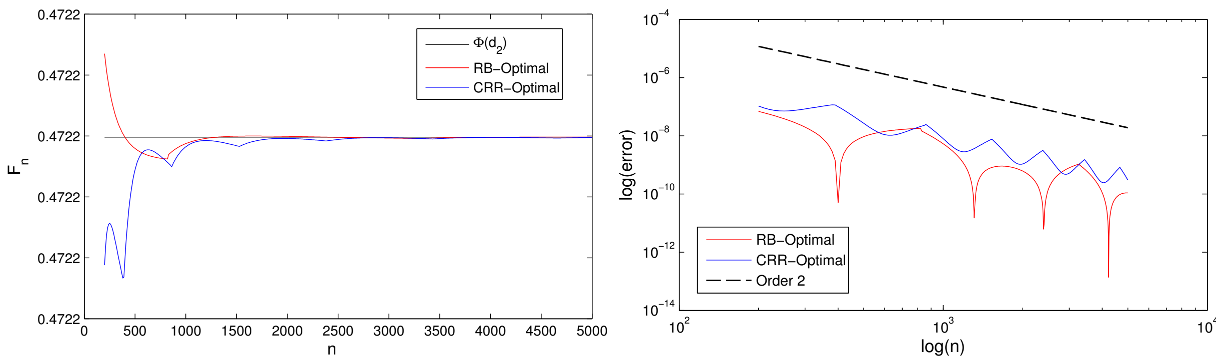


Figure 4.5: $s_0 = 95$, $x = 100$, $r = 0.1$, $\sigma = 0.25$, $T = 1$.

Remark 4.2.8. Since all absolute moments of $\eta_k^{(n)}$ are bounded, we can apply Theorem 3.4.9 to retrieve subsequent terms in the asymptotic expansion (4.7). We are then able to further increase the order of convergence, by adding more terms to the probability. With the approach described above all equations will be linear, and unlike the method in [L12], the previous coefficients will remain unaltered.

CRR vs RB We would now like to take a closer look at the distributional fit for different choices of the coefficient $k_{0,n}$. Figure 4.5 suggests that the RB-based model gives a slightly better approximation. However, is this the case for all options? To get a better idea of the general accuracy of the models we will use two types of errors, the root mean squared (RMS) error and the root mean squared relative (RMSR) error.

First, we randomly generate a sample of m parameter vectors $\pi = (s_0, x, r, \sigma, T)$, following the procedure described in [BD96], but allowing a slightly wider range for the parameters.

- The initial asset price s_0 is fixed to 100,
- the value x is uniformly distributed between 50 and 150,
- the riskless interest rate r is uniformly distributed between 0 and 0.2,
- the volatility σ is uniformly distributed between 0.1 and 0.8,
- the maturity T is chosen uniformly between 0 and 1 years with probability 0.75 and between 1 and 5 years with probability 0.25.

Note that s_0 remains fixed and we only vary x, since we are actually only interested in the ratio $\frac{x}{s_0}$ and not the values separately. Parameter vectors, for which $\Phi^\pi(d_2(x)) \leq 10^{-6}$ are then excluded from the sample to ensure a reliable relative error estimate.

For each $n \in \mathbb{N}$ and every parameter vector π, let $\varepsilon^\pi_{abs}(n)$ and $\varepsilon^\pi_{rel}(n)$ denote the absolute and relative error, respectively, i.e

$$\varepsilon^\pi_{abs}(n) = \left|F^\pi_n(x) - \Phi^\pi(d_2(x))\right|,$$

and

$$\varepsilon^{\pi}_{rel}(n) = \frac{\varepsilon^{\pi}_{abs}(n)}{\Phi^{\pi}(d_2(x))},$$

where F^{π}_{n} and Φ^{π} refer to the distribution functions of the discrete- and continuous-time models, corresponding to the parameters π. Then

$$RMS(n) = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(\varepsilon^{\pi_i}_{abs}(n)\right)^2},$$

and

$$RMSR(n) = \sqrt{\frac{1}{m}\sum_{i=1}^{m}\left(\varepsilon^{\pi_i}_{rel}(n)\right)^2}.$$

Due to Proposition 4.2.4, both errors are of order $O\left(\frac{1}{n^2}\right)$. Consider the following convergence behavior of the errors, taken over a sample of $m = 1000$ parameter vectors, 995 of which are included. For both errors a maximum of $n = 1000$ time steps is considered.

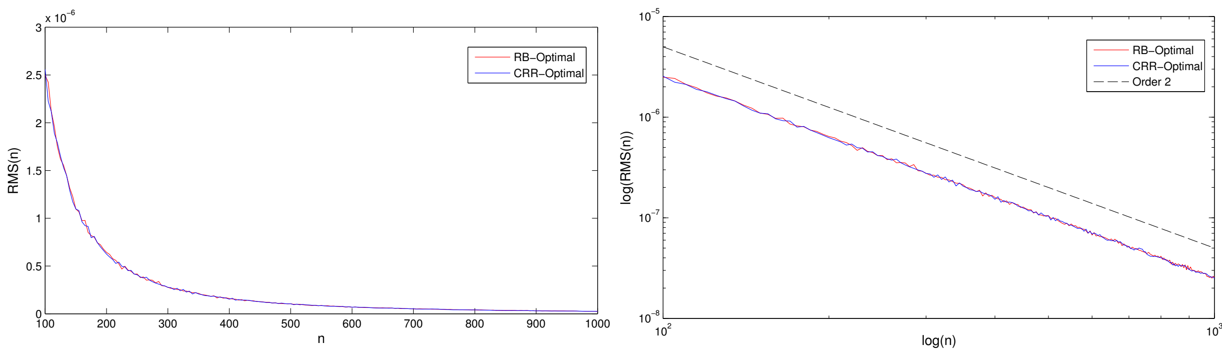


Figure 4.6: $RMS(n)$ error

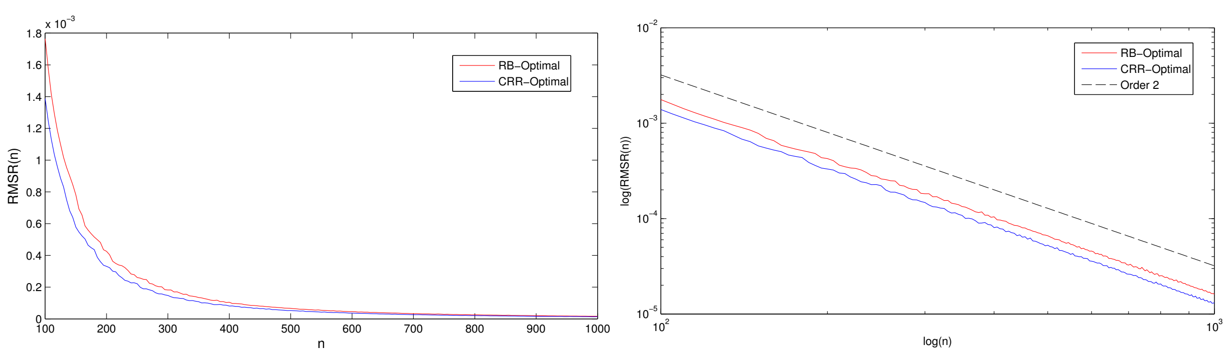


Figure 4.7: $RMSR(n)$ error.

Note that the absolute error has a very similar convergence pattern for both models (see figure Figure 4.6), however, figure 4.7 suggests, that the CRR-Optimal tree delivers better results for the relative error. A smaller relative error implies that the CRR-Optimal model deals with small values of $\Phi(d_2)$ better, then the RB model. Let us a take closer look at the difference in the absolute and relative errors for the RB- and CRR-based trees for different values of d_2.

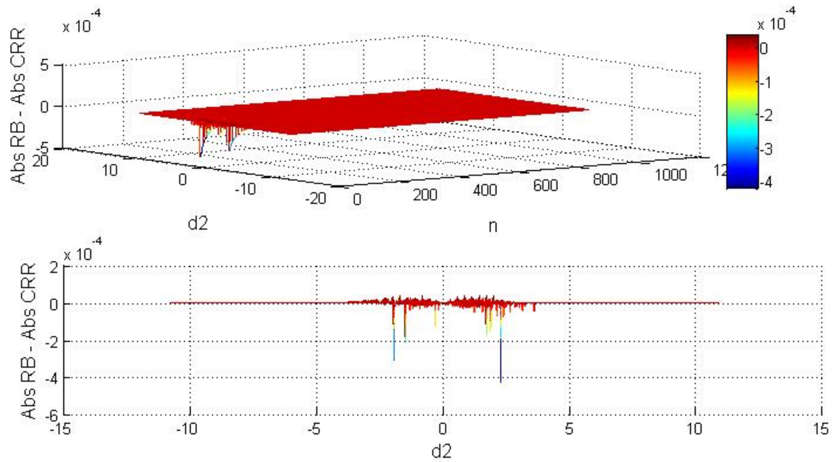


Figure 4.8: Difference in absolute error

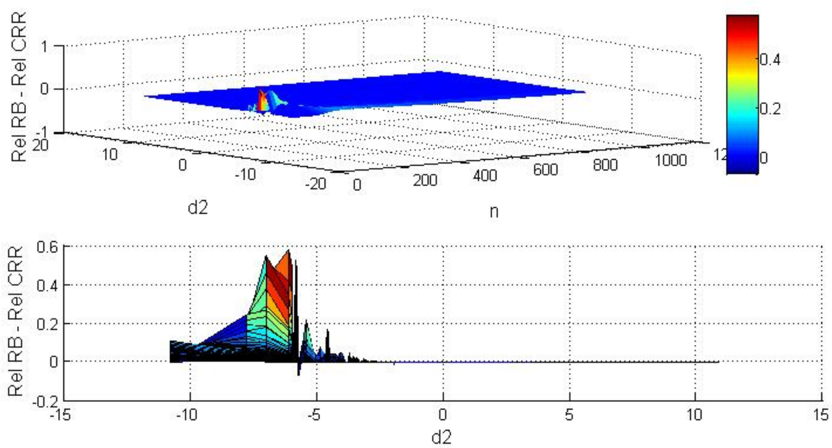


Figure 4.9: Difference in relative error

As can be seen above, depending on the exact value of d_2, either the RB - or CRR-Optimal model delivers a smaller absolute error, the trend being slightly better for RB-based tree. For deep in-the-money or deep out-of-the-money values the difference is minor. The relative error, on the other hand, seems to be better for the CRR-based tree only for very small values of d_2, where the absolute error is insignificant.

4.3 Expansions for option prices

We will now see how the above results can be applied to improve the convergence behaviour for specific types of options.

4.3.1 Digital options

An option that pays out a fixed amount G if the value of the underlying asset at maturity T is below the strike K, is called a digital (cash-or-nothing) put. The payout function of such options is

$$G\mathbb{1}_{\{S(T)\leq K\}},$$

and the price is given by

$$V = Ge^{-rT}\Phi\left(d_2\left(K\right)\right).$$

Despite this simple structure and easy closed-form solution, approximation methods for digital options exhibit extremely irregular convergence behavior, due to the discontinuities in the payout function. Therefore, methods of improving this convergence are of great interest.

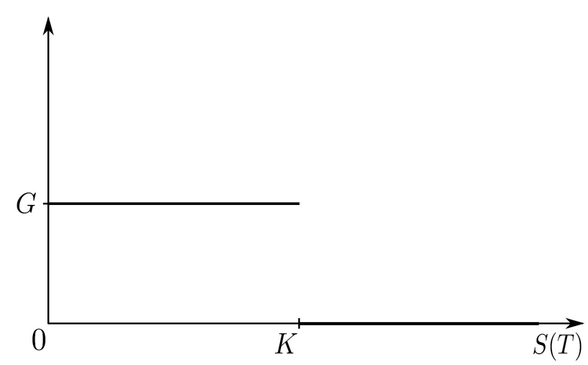


Figure 4.10: Payout function of a digital put

Proposition 4.3.1. *Let* $V^{(n)}$ *be the price of the cash-or-nothing put option in the discrete model* (4.31)*. If* $k_{0,n}$*,* $k_{1,n}$ *and* $c_{i,n}$*,* $i = 1, \ldots, 4$ *satisfy* (4.33), (4.38)-(4.41)*, then*

$$V^{(n)} = V + O\left(\frac{1}{n^2}\right), \quad n \to \infty$$

Proof. Since

$$V^{(n)} = Ge^{-rT}P(S_n^{(n)} \leq K),$$

the statement is a simple consequence of Proposition 4.2.4, where the strike K is taken as the point x. □

The algorithm

The coefficients of the drift α_n and probability p_n need to be calculated once for each $n \in \mathbb{N}$. Therefore, the algorithm presented in section 2.1.2 can be easily adjusted to incorporate the 3/2 - Optimal model, by adding an extra step in the beginning. As a result, for a given $n \in \mathbb{N}$ we have

Step 1: Calculate $k_{0,n}$, $k_{1,n}$ and $c_{i,n}$, $i = 1, \ldots, 4$.

Step 2: Tree Initialization (as in Chapter 1).

Step 3: Backward Induction (as in Chapter 1).

Step 4: Return $e^{-rT}V_0(0)$.

Numerical results

Compare the following convergence results for three different cash-or-nothing options.

Parameters	n	RB tree	OD model	RB 3/2-Optimal	BS value
$T = 1$	100	41.6411049	42.7287634	42.7233441	42.7233237
$K = 100$	200	42.6925603	42.7241952	42.7233299	
$\sigma = 0.25$	500	43.6283356	42.7230029	42.7233229	
	1000	41.8227457	42.7229920	42.7233232	
	2000	42.8222727	42.7233872	42.7233238	
	4000	42.3911188	42.7232924	42.7233237	
$T = 1$	100	0.3002757	0.3384204	0.3442454	0.3446987
$K = 80$	200	0.3946223	0.3415591	0.3445932	
$\sigma = 0.1$	500	0.3737622	0.3434433	0.3446817	
	1000	0.3239081	0.3440710	0.3446940	
	2000	0.3518514	0.3443847	0.3446977	
	4000	0.3406525	0.3445417	0.3446984	
$T = 3$	100	7.1617825	6.5563626	6.5649992	6.5650794
$K = 100$	200	6.6289911	6.5609228	6.5651036	
$\sigma = 0.1$	500	6.1329162	6.5634040	6.5650608	
	1000	6.4402779	6.5642490	6.5650794	
	2000	6.3914729	6.5646637	6.5650788	
	4000	6.6285242	6.5648714	6.5650794	

Figure 4.11: For all options: $s_0 = 95$, $r = 0.1$ and G= 100.

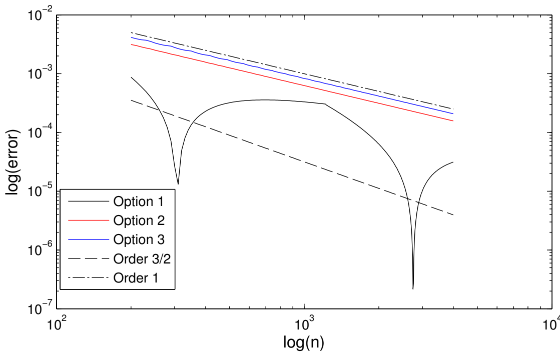


Figure 4.12: OD model

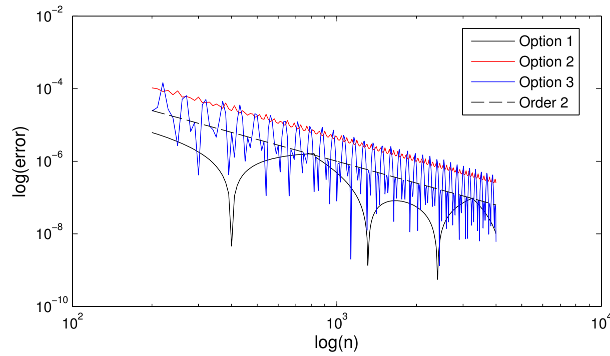


Figure 4.13: 3/2 - Optimal model

Note that the 3/2-Optimal model retains an order of convergence 2 even for the out-of-the-money options. The OD, on the other hand, has a smooth order 1 convergence (Options 2 and 3) instead of 3/2.

Computational effort The above results compare the convergence patterns of different methods, however, in order to be able to accurately compare the efficiency of these methods we have to consider their computational complexity as well. Standard d-dimensional binomial schemes have a total computational effort (see section 2.1.2)

$$\mathcal{C}(n) = O\left(n^{d+1}\right).$$

If the approximation error ε of a method is of order $\alpha > 0$, i.e. $\varepsilon = O\left(n^{-\alpha}\right)$, then the computational effort in terms of the error (computational complexity) can be written as

$$\mathcal{C}(\varepsilon) = O\left(\varepsilon^{-\frac{d+1}{\alpha}}\right),$$

which means that to decrease the error by 0.1 one would need to increase the computational effort by $10^{\frac{d+1}{\alpha}}$ (=10000 in a standard 1-dimensional binomial scheme).

The 1-dimensional 3/2-Optimal tree involves additional calculations of the coefficients $k_{j,n}$, $j = 0, 1$, $c_{i,n}$, $i = 1, \ldots, 4$ for each n, however, this contributes only $O(1)$ to the total computational effort. Therefore, this method also has $\mathcal{C}(n) = O\left(n^2\right)$, and we have

	CRR, RB tree	OD model	3/2-Optimal
ε	$O\left(n^{-\frac{1}{2}}\right)$	$O\left(n^{-\frac{3}{2}}\right)\ (O(n^{-1}))$	$O(n^{-2})$
$\mathcal{C}(\varepsilon)$	$O(\varepsilon^{-4})$	$O\left(\varepsilon^{-\frac{4}{3}}\right)(O(\varepsilon^{-2}))$	$O(\varepsilon^{-1})$

We now compare the computational complexity of the RB, OD and the 3/2-Optimal models. We will use the RMSR error and the same sampling procedure as described in section 4.2.3. We will only include those parameter vectors, for which the OD model has order of convergence $O\left(\frac{1}{\sqrt{n^3}}\right)$, i.e. we will have to make sure that the discriminant of the corresponding quadratic equation is greater than or equal to zero. Out of a sample of 2000 simulated options 1656 have been included.

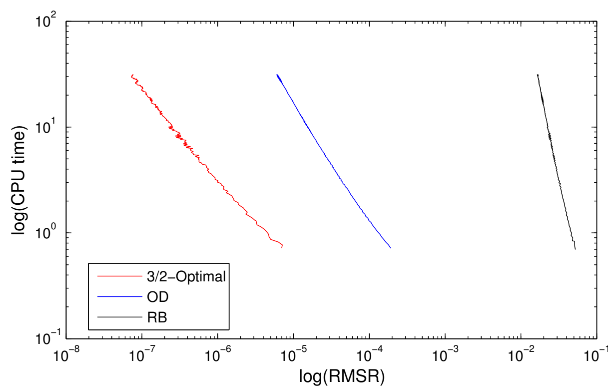


Figure 4.14: Computational complexity for a maximum of $n = 1000$ time steps.

Remark 4.3.2. In this section we have only considered digital puts, however, the above results can be easily adjusted to deal with cash-or-nothing call options as well, i.e. options that pay a fixed amount if the underlying at maturity is above the strike.

4.3.2 Plain vanilla options

A European *put* gives the holder the right to *sell* a share of the underlying stock at maturity T for a prespecified strike price $K \geq 0$. The payout of such an option will be

$$(K - S(T))^+.$$

A European *call* gives the holder the right to *buy* a share of the underlying stock at maturity T for a prespecified strike price $K \geq 0$, i.e. the payout in this case will be

$$(S(T) - K)^+.$$

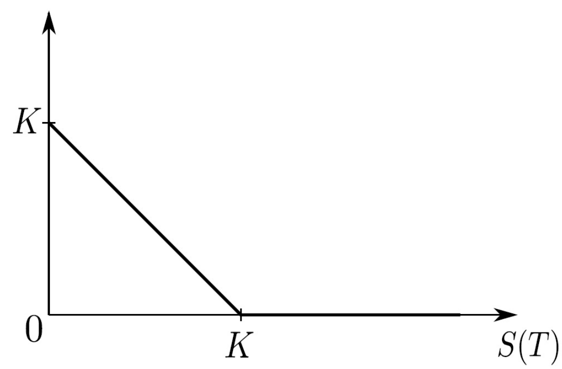


Figure 4.15: Payout function of a vanilla put

The prices of the plain vanilla options are given by the well-known Black-Scholes formula (see e.g. [KK01] Corollary 3.9)

$$V = \delta\left(Ke^{-rT}\Phi\left(\delta d_2\left(K\right)\right) - s_0\Phi\left(\delta d_1\left(K\right)\right)\right),$$

where $\delta = 1$ for put options and -1 for calls, d_2 is defined as in (4.5) and

$$d_1 := d_1\left(K\right) := d_2\left(K\right) - \sigma\sqrt{T}. \tag{4.42}$$

Note that our notation is slightly different from the conventional one in order to be in line with the previous results. We will focus on put options, however the same results can also be obtained for calls.

The continuous-time formula is based on an appropriate change of measure. We want to use the same idea also in the discrete-time case. For this we will need the following theorem.

Theorem 4.3.3 (Discrete change of measure)**.** *Consider two probability measures P and $\tilde{P}$ on a finite sample space Ω, such that $P\left(\omega\right) > 0$ and $\tilde{P}\left(\omega\right) > 0$ $\forall\omega \in \Omega$. Let Z be the Radon-Nikodým derivative of $\tilde{P}$ with respect to P, i.e.*

$$Z\left(\omega\right) = \frac{\tilde{P}\left(\omega\right)}{P\left(\omega\right)}, \quad \forall\omega \in \Omega.$$

Then

(i) $P\left(Z > 0\right) = 1$*;*

(ii) $E\left(Z\right) = 1$*;*

(iii) for any variable Y,

$$\tilde{E}(Y) = E(ZY). \tag{4.43}$$

(Cf. [S04], Theorem 3.1.1.).

Consider the general binomial model (4.31), where the drift and probabilities satisfy (4.32). The price of a European put in this model can be written as

$$\begin{aligned} V^{(n)} &= E_n\left(e^{-rT}\left(K - S_n^{(n)}\right)\mathbb{1}_{\left\{S_n^{(n)} \le K\right\}}\right) \\ &= e^{-rT} K P_n\left(S_n^{(n)} \le K\right) - s_0 E_n\left(e^{(\alpha_n - r)T + \sigma\sqrt{T}\sum_{k=1}^n \xi_k^{(n)}}\mathbb{1}_{\left\{S_n^{(n)} \le K\right\}}\right). \end{aligned} \tag{4.44}$$

The expansion of the first term can be obtained directly from Proposition 4.2.4. To get the dynamics of the second term we want to apply Theorem 4.3.3. If we set

$$Z = e^{(\alpha_n - r)T + \sigma\sqrt{T}\sum_{k=1}^n \xi_k^{(n)}},$$

then by equation (4.43) we could write

$$E_n\left(e^{(\alpha_n - r)T + \sigma\sqrt{T}\sum_{k=1}^n \xi_k^{(n)}}\mathbb{1}_{\left\{S_n^{(n)} \le K\right\}}\right) = \tilde{E}_n\left(\mathbb{1}_{\left\{S_n^{(n)} \le K\right\}}\right) = \tilde{P}_n\left(S_n^{(n)} \le K\right),$$

for the equivalent measure $\tilde{P}_n$, and Proposition 4.2.4 is again applicable. However, this approach only works if $E(Z) = 1$, otherwise $\tilde{P}_n$ will not be a well-defined probability measure. Therefore, it can only be applied in the case of a risk-neutral probability measure P_n, since this is the unique measure for which the discounted stock price is a martingale, i.e. $E(Z) = 1$. Note that this is exactly the procedure that appears in literature (see [CRR79] or [M09] section 2.5.2.).

In order to handle our more general setting we will need a slight modification. Let

$$\lambda_{n,k} = e^{(\alpha_n - r)\Delta t + \sigma\sqrt{\Delta t}\xi_k^{(n)}}, \quad k = 1, \ldots, n.$$

and

$$M_n = E(\lambda_{n,1}). \tag{4.45}$$

We can now define new transition probabilities for $\xi_k^{(n)}$, $k = 1, \ldots, n$. Let the new probabilities of an up-jump $\tilde{p}_n$ and a down-jump $\tilde{q}_n$ be

$$\tilde{p}_n = \frac{e^{(\alpha_n - r)\Delta t + \sigma\sqrt{\Delta t}} p_n}{M_n}, \tag{4.46}$$

and

$$\tilde{q}_n = \frac{e^{(\alpha_n - r)\Delta t - \sigma\sqrt{\Delta t}}(1 - p_n)}{M_n}. \tag{4.47}$$

Note that $\tilde{p}_n, \tilde{q}_n$ are well-defined, starting from some $n \in \mathbb{N}$, and $\tilde{p}_n + \tilde{q}_n = 1$. The new probability measure $\tilde{P}_n$ is now defined as

$$\begin{aligned}\tilde{P}_n(\omega_1, \ldots, \omega_n) &= \tilde{p}_n^{n_u(\omega_1,\ldots,\omega_n)} \cdot \tilde{q}_n^{n_d(\omega_1,\ldots,\omega_n)} \\ &= \left(\frac{e^{(\alpha_n - r)\Delta t + \sigma\sqrt{\Delta t}} p_n}{M_n}\right)^{n_u(\omega_1,\ldots,\omega_n)} \left(\frac{e^{(\alpha_n - r)\Delta t - \sigma\sqrt{\Delta t}}(1-p_n)}{M_n}\right)^{n_d(\omega_1,\ldots,\omega_n)} \\ &= \prod_{k=1}^{n} \frac{\lambda_{n,k}(\omega_k)}{M_n} P_n(\omega_1, \ldots, \omega_n),\end{aligned} \tag{4.48}$$

where $\omega_k \in \{-1, 1\}$, $k = 1, \ldots n$, and as defined in section 2.1.2, $n_u(\omega_1, \ldots, \omega_n)$ and $n_d(\omega_1, \ldots, \omega_n)$ are the number of 1's and -1's in the sequence $(\omega_1, \ldots, \omega_n)$. Therefore, the Radon-Nikodým derivative of $\tilde{P}_n$ with respect to P_n is given by

$$\Lambda_n = \prod_{k=1}^{n} \frac{\lambda_{n,k}}{M_n}.$$

All conditions of Theorem 4.3.3 are satisfied and by equation (4.43) the second term in (4.44) can be written as

$$\begin{aligned}s_0 E_n\left(e^{(\alpha_n - r)T + \sigma\sqrt{T}\sum_{k=1}^{n}\xi_k^{(n)}} \mathbb{1}_{\left\{S_n^{(n)} \le K\right\}}\right) &= s_0 M_n^n E_n\left(\Lambda_n \mathbb{1}_{\left\{S_n^{(n)} \le K\right\}}\right) \\ &= s_0 M_n^n \tilde{P}_n\left(S_n^{(n)} \le K\right).\end{aligned}$$

As a result

$$V^{(n)} = e^{-rT} K P_n\left(S_n^{(n)} \le K\right) - s_0 M_n^n \tilde{P}_n\left(S_n^{(n)} \le K\right). \tag{4.49}$$

Proposition 4.3.4. *With an appropriate choice of parameters $k_{i,n}$, $i = 1, 3$ and $c_{i,n}$, $i = 1, 4$, the price of a European put option in the binomial model (4.31) satisfies*

$$V^{(n)} = V + O\left(\frac{1}{n^2}\right), \quad n \to \infty.$$

Proof. By the proof of Proposition 4.2.4 (equation (4.36))

$$\begin{aligned}e^{-rT} K P_n\left(S_n^{(n)} \le K\right) = e^{-rT} K \Phi(d_2) - \frac{1}{\sqrt{n}} e^{-rT} K \phi(d_2) f_1(\alpha_n, p_n) + \frac{1}{n} e^{-rT} K \phi(d_2) f_2(\alpha_n, p_n) \\ + \frac{1}{\sqrt{n}^3} e^{-rT} K \phi(d_2) f_3(\alpha_n, p_n) + O\left(\frac{1}{n^2}\right),\end{aligned} \tag{4.50}$$

where the functions f_i, $i = 1, 3$ are defined as in (4.37). By the Taylor theorem

$$e^{(\alpha_n - r)\Delta T \pm \sigma\sqrt{\Delta T}} = 1 \pm \frac{1}{\sqrt{n}}\sigma\sqrt{T} - \frac{1}{n}2\sigma c_{1,n}\sqrt{T} + \frac{1}{\sqrt{n}^3}\left(k_{1,n}T \mp 2\sigma^2 c_{1,n}T \mp \frac{1}{3}\sigma^3\sqrt{T}^3\right)$$
$$+\frac{1}{n^2}\left(k_{2,n}T \pm \sigma k_{1,n}\sqrt{T}^3 + 2\sigma^2 c_{1,n}^2 T - \frac{1}{12}\sigma^4 T^2\right)$$
$$+\frac{1}{\sqrt{n}^5}\left(k_{3,n}T \pm \sigma k_{2,n}\sqrt{T}^3 - 2\sigma k_{1,n}c_{1,n}\sqrt{T}^3 \pm \frac{1}{2}\sigma\sqrt{T}^3\left(2\sigma c_{1,n} + \frac{1}{3}\sigma^2\sqrt{T}\right)^2 \pm \frac{1}{48}\sigma^5\sqrt{T}^5\right)$$
$$+O\left(\frac{1}{n^3}\right), \tag{4.51}$$

therefore, M_n in (4.45) satisfies

$$M_n = 1 + \frac{1}{\sqrt{n}^3}\sigma\sqrt{T}m_1 + \frac{1}{n^2}\left(\sigma\sqrt{T}m_2 - 2\sigma^2 T\left(c_{1,n} + \frac{1}{6}\sigma\sqrt{T}\right)^2 - \frac{1}{36}\sigma^4 T^2\right)$$
$$+ \frac{1}{\sqrt{n}^5}\left(\sigma\sqrt{T}m_3 - 4\sigma^2 c_{1,n}c_{2,n}T - \frac{2}{3}c_{2,n}\sigma^3\sqrt{T}^3\right) \tag{4.52}$$
$$+ O\left(\frac{1}{n^3}\right),$$

with m_i, $i = 1, \ldots, 3$ as in (4.34). Then from (4.46) - (4.47)

$$\tilde{p}_n = \frac{1}{2} + \tilde{c}_{1,n}\frac{1}{\sqrt{n}} + \tilde{c}_{2,n}\frac{1}{n} + \tilde{c}_{3,n}\frac{1}{\sqrt{n}^3} + \tilde{c}_{4,n}\frac{1}{n^2} + O\left(\frac{1}{n^{5/2}}\right),$$
$$\tilde{q}_n = 1 - \tilde{p}_n,$$

where

$$\tilde{c}_{1,n} = c_{1,n} + \frac{1}{2}\sigma\sqrt{T},$$
$$\tilde{c}_{2,n} = c_{2,n},$$
$$\tilde{c}_{3,n} = c_{3,n} - 2\sigma\sqrt{T}\left(c_{1,n} + \frac{1}{4}\sigma\sqrt{T}\right)^2 - \frac{1}{24}\sigma^3\sqrt{T}^3,$$
$$\tilde{c}_{4,n} = c_{4,n} - \sigma^2 c_{2,n}T - 4\sigma c_{1,n}c_{2,n}\sqrt{T}.$$

Note that under the new measure

$$\tilde{\mu}(n) = k_{0,n} + \frac{2\sigma\tilde{c}_{1,n}}{\sqrt{T}} + O\left(\frac{1}{\sqrt{n}}\right)$$
$$\overset{(4.33)}{=} r + \frac{1}{2}\sigma^2 + O\left(\frac{1}{\sqrt{n}}\right),$$

and

$$\tilde{\sigma}^2(n) = \sigma^2 + O\left(\frac{1}{\sqrt{n}}\right),$$

for $\tilde{\mu}(n)$ and $\tilde{\sigma}^2(n)$ as in (2.3). Therefore, $\tilde{P}_n\left(S_n^{(n)} \leq K\right)$ converges to $\Phi(d_1)$, and again, following the proof of Proposition 4.2.4 we get the following expansion

$$\begin{aligned} s_0\tilde{P}_n\left(S_n^{(n)} \leq K\right) = s_0\Phi(d_1) - \frac{1}{\sqrt{n}}s_0\phi(d_1)\tilde{f}_1(\alpha_n, \tilde{p}_n) + \frac{1}{n}s_0\phi(d_1)\tilde{f}_2(\alpha_n, \tilde{p}_n) \\ + \frac{1}{\sqrt{n}^3}s_0\phi(d_1)\tilde{f}_3(\alpha_n, \tilde{p}_n) + O\left(\frac{1}{n^2}\right), \end{aligned} \tag{4.53}$$

where the coefficients $\tilde{f}_i$, $i = 1, \ldots, 3$ are defined as in (4.37) with d_1 substituted for d_2 and $\tilde{c}_{i,n}$ substituted for $c_{i,n}$, $i = 1, \ldots, 4$. To get the asymptotics of M_n^n consider the binomial formula

$$\begin{aligned} (1+x)^n &= \sum_k^n x^k \binom{n}{k} \\ &= 1 + nx + \frac{n(n-1)}{2}x^2 + \frac{n(n-1)(n-2)}{6}x^3 + \sum_{k=4}^n x^k \binom{n}{k}. \end{aligned} \tag{4.54}$$

If $x = O\left(n^{-\frac{3}{2}}\right)$, then $\sum_{k=4}^n x^k \binom{n}{k} = O\left(\frac{1}{n^2}\right)$. Indeed,

$$\begin{aligned} n^2 \cdot \sum_{k=4}^n \left(O\left(n^{-\frac{3}{2}}\right)\right)^k \binom{n}{k} &= n^2 \cdot \sum_{k=0}^{n-4} \left(O\left(n^{-\frac{3}{2}}\right)\right)^{k+4} \binom{n}{k+4} \\ &\leq n^2 \cdot \frac{n^4}{4!}\left(O\left(n^{-\frac{3}{2}}\right)\right)^4 \sum_{k=0}^{n-4} \left(O\left(n^{-\frac{3}{2}}\right)\right)^k \binom{n-4}{k} \\ &\leq C\left(1 + O\left(n^{-\frac{3}{2}}\right)\right)^{n-4}. \end{aligned}$$

The last expression is convergent, and, therefore, bounded, and we have the necessary result. If we now substitute (4.52) into (4.54) we get

$$M_n^n = 1 + a_{1,n}\frac{1}{\sqrt{n}} + a_{2,n}\frac{1}{n} + a_{3,n}\frac{1}{\sqrt{n}^3} + O\left(\frac{1}{n^2}\right), \tag{4.55}$$

where

$$\begin{aligned}
a_{1,n} &= \sigma\sqrt{T}m_1, \\
a_{2,n} &= \sigma\sqrt{T}m_2 - 2\sigma^2 T\left(c_{1,n} + \frac{1}{6}\sigma\sqrt{T}\right)^2 - \frac{1}{36}\sigma^4 T^2 + \frac{1}{2}\sigma^2 T m_1^2, \\
a_{3,n} &= \sigma\sqrt{T}m_3 - 4\sigma^2 c_{1,n}c_{2,n}T - \frac{2}{3}c_{2,n}\sigma^3\sqrt{T}^3 \\
&\quad + \sigma\sqrt{T}m_1\left(\sigma\sqrt{T}m_2 - 2\sigma^2 T\left(c_{1,n} + \frac{1}{6}\sigma\sqrt{T}\right)^2 - \frac{1}{36}\sigma^4 T^2\right) + \frac{1}{6}m_1^3\sigma^3\sqrt{T}^3.
\end{aligned} \tag{4.56}$$

Note that the term $S_1(a_n)$ depends only on the coefficients of the drift α_n and not the probabilities p_n or $\tilde{p}_n$, therefore, $f_1 = \tilde{f}_1$, and by substituting (4.50), (4.53) and (4.55) into (4.49), we get the following dynamics

$$\begin{aligned}
V^{(n)} &= V - \frac{1}{\sqrt{n}}s_0 a_{1,n}\Phi(d_1) + \frac{1}{n}s_0\left(\phi(d_1)\left(f_2 - \tilde{f}_2 + a_{1,n}\tilde{f}_1\right) - a_{2,n}\Phi(d_1)\right) \\
&\quad + \frac{1}{\sqrt{n}^3}s_0\left(\phi(d_1)\left(f_3 - \tilde{f}_3 - a_{1,n}\tilde{f}_2 + a_{2,n}\tilde{f}_1\right) - a_{3,n}\Phi(d_1)\right) + O\left(\frac{1}{n^2}\right),
\end{aligned} \tag{4.57}$$

where we have used the well-known formula $s_0\phi(d_1) = Ke^{-rT}\phi(d_2)$. Note that for risk-neutral transition probabilities $M_n = 1$, i.e. $a_{i,n} = 0$ for all i, so the coefficient of the leading error term in (4.57) becomes zero. However, this is not true for the general case. Therefore, not every binomial tree automatically delivers an $O\left(\frac{1}{n}\right)$ order of convergence for vanilla options. We now choose the coefficients $k_{i,n}$ and $c_{i,n}$, such that the first three error terms in (4.57) become zero.

First, as in Proposition 4.2.4 we set $k_{2,n} = k_{3,n} = 0$ and $c_{2,n} = \frac{1}{2} - \left\{\frac{\ln\left(\frac{x}{s_0}\right) - k_{0,n}T}{2\sigma\sqrt{\Delta t}} + \frac{n}{2}\right\}$. The latter ensures that $f_1 = 0$. Note that, unlike the digital option case, this is not necessary here, i.e. f_1 could take on any value, but the above choice simplifies the calculations. Next we set

$$k_{1,n} = \frac{2\sigma c_{2,n}}{\sqrt{T}},$$

so that $m_1 = a_{1,n} = 0$. In this case, taking into account that $d_2 - d_1 = \sigma\sqrt{T}$, we get

$$f_2 - \tilde{f}_2 = \sigma\sqrt{T}\left(-2c_{1,n}^2 - c_{1,n}\left(\sigma\sqrt{T} + \frac{d_2 + d_1}{3}\right) - \frac{1}{8}\sigma^2 T + \frac{1}{4} - \frac{d_2^2}{8} + \frac{d_1^2}{24}\right),$$

$$f_3 - \tilde{f}_3 = -2\sigma\sqrt{T}c_{2,n}\left(\frac{2}{3}\sigma\sqrt{T} + 2c_{1,n} + \frac{1}{3}d_1\right).$$

If we now set

$$c_{3,n} = \frac{1}{2\sigma\sqrt{T}}\left(\frac{\phi(d_1)}{\Phi(d_1)}(f_2 - \tilde{f}_2) + 2\sigma^2 T\left(c_{1,n} + \frac{1}{6}\sigma\sqrt{T}\right)^2 + \frac{1}{36}\sigma^4 T^2\right),$$

$$c_{4,n} = \frac{1}{2\sigma\sqrt{T}} \left(\frac{\phi(d_1)}{\Phi(d_1)} \left(f_3 - \tilde{f}_3 \right) + 4\sigma^2 c_{1,n} c_{2,n} T + \frac{2}{3} c_{2,n} \sigma^3 \sqrt{T}^3 \right),$$

we will get the statement of the proposition. □

Remark 4.3.5. Note that, as in Proposition 4.2.4, the parameters $k_{0,n}$ and $c_{1,n}$ can be chosen freely, as long as condition (4.33) is satisfied.

We now compare the RMS and RMSR errors for $k_{0,n} = 0$ (CRR setting) and $c_{1,n} = 0$ (RB setting). We employ the same sampling procedure as in the previous section. Out of a sample of 2000 generated options, 1966 were included. A maximum step of $n = 1000$ has been used.

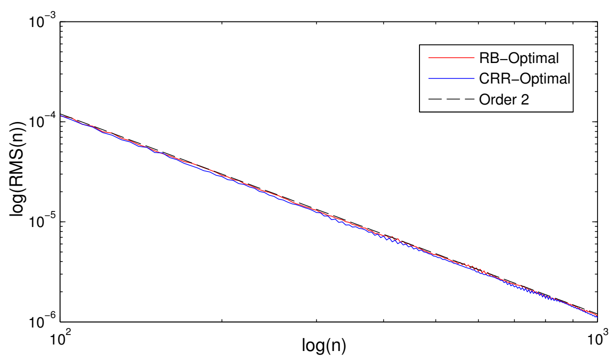


Figure 4.16: RMS(n) error

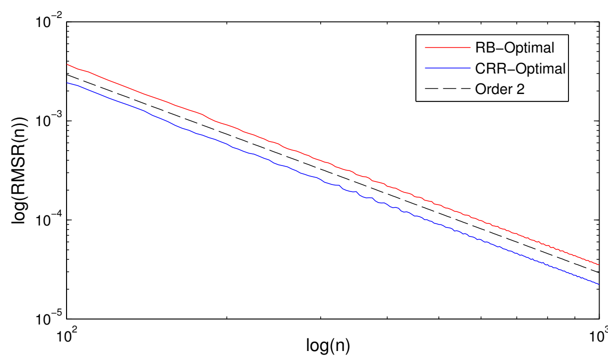


Figure 4.17: RMSR(n) error

Note that, the CRR-Optimal tree has a lower relative error for vanilla options as well, which suggests that it performs better than the RB based model for small option prices. The convergence pattern of the absolute error is practically the same for both methods.

Remark 4.3.6. For the standard RB and CRR models the expected value M_n is not equal to 1, therefore, conventional methods cannot be used to show the order of convergence. However, with the method above we see that, since, in both cases, $k_{1,n} = c_{2,n} = 0$, we have $m_1 = 0$ and the coefficient of the leading error term in (4.57) disappears. Therefore, we can prove the $O\left(\frac{1}{n}\right)$ order of convergence of the standard binomial schemes.

Remark 4.3.7 (Advanced methods in literature)*.* The Tian, CP and OD methods are also applicable for vanilla options.

The Tian and CP approaches are based on fixing the value of $S_1(a_n)$. However, for vanilla options, the CP approach doesn't directly increase the order of convergence. Both methods result in a constant leading error coefficient and, therefore, have smooth convergence of order $O\left(\frac{1}{n}\right)$. Extrapolation can then be applied to get a $O\left(\frac{1}{\sqrt{n}^3}\right)$ convergence.

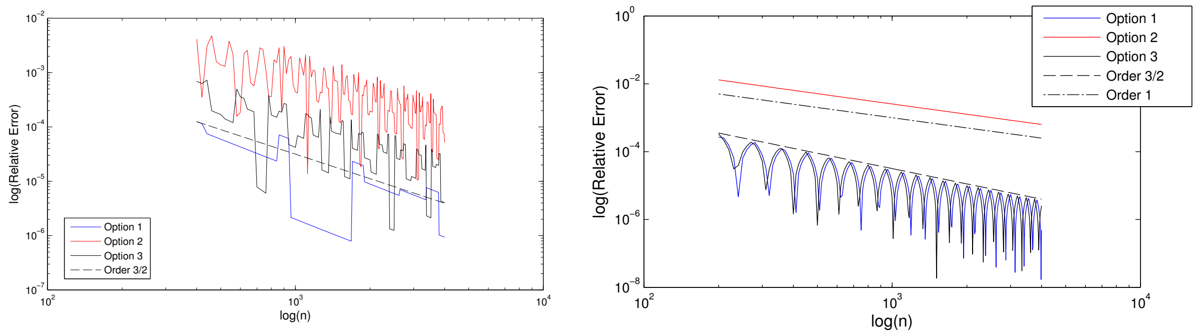
The OD method for vanilla options follows the same procedure as described in the previous section. The transition probabilities are assumed to be risk-neutral. $S_1(a_n)$ is fixed to 0, and $k_{0,n}$ $(c_{1,n})$ is chosen to solve the quadratic equation $f_2 - \tilde{f}_2 = 0$. As in the case with digital options this method delivers an order of convergence $O\left(\frac{1}{\sqrt{n}^3}\right)$ if the discriminant of the mentioned quadratic equation is greater than or equal to zero, otherwise the order remains $O\left(\frac{1}{n}\right)$. This, of course, puts some restrictions on the parameter setting, but as mentioned in [M09], p. 100, these restrictions are very weak. For more details on these methods see [KM13], [M09], [CP07] and [T99].

Numerical results

We now compare the convergence behavior of the Tian method with extrapolation, the OD method and the 3/2-Optimal method for different types of options.

Parameters	n	Tian with Extrapolation	OD model	CRR 3/2-Optimal	BS value
$T = 1$	100	7.4069566	7.4008690	7.4046753	7.4049833
$K = 105$	200	7.4018715	7.4069738	7.4049790	
$\sigma = 0.25$	500	7.4045424	7.4048572	7.4049777	
$r = 0.1$	1000	7.4049982	7.4048921	7.4049809	
	2000	7.4050493	7.4050257	7.4049833	
	4000	7.4049764	7.4049870	7.4049832	
$T = 0.2$	100	0.00012923	0.00013346	0.00013565	0.00013712
$K = 72$	200	0.00013937	0.00013533	0.00013680	
$\sigma = 0.2$	500	0.00013731	0.00013642	0.00013706	
$r = 0.05$	1000	0.00013718	0.00013677	0.00013710	
	2000	0.00013715	0.00013695	0.00013711	
	4000	0.00013713	0.00013703	0.00013712	
$T = 3$	100	5.0995900	5.0834292	5.0861481	5.0858511
$K = 130$	200	5.0877658	5.0841486	5.0859119	
$\sigma = 0.1$	500	5.0849887	5.0858476	5.0858542	
$r = 0.1$	1000	5.0856365	5.0858300	5.0858518	
	2000	5.0857252	5.0859012	5.0858515	
	4000	5.0858305	5.0858640	5.0858512	

Figure 4.18: For all options: $s_0 = 100$.



Tian model with extrapolation

OD model

Figure 4.19: Convergence behavior for even n

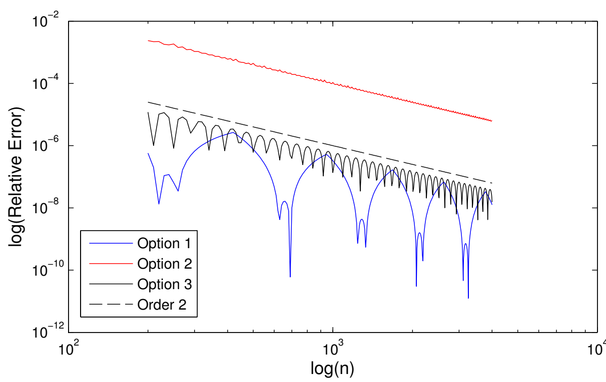


Figure 4.20: 3/2-Optimal model, n even

Note that for the out-of-the-money Option 2 the OD model delivers an order $O\left(\frac{1}{n}\right)$, while both the Tian approach with extrapolation and the 3/2-Optimal model have order $O\left(\frac{1}{n^{3/2}}\right)$, the latter having a smoother convergence. For Option 1, the 3/2-Optimal model delivers an error smaller than 10^{-6} for as few as 200 steps.

Computational effort The computational effort of the Tian (CP) tree is $\mathcal{C}(n) = O(n^2)$. If we additionally perform extrapolation

$$\hat{V}^{(n)} = 2V^{(n)} - V^{(n/2)},$$

then the computational effort becomes $\mathcal{C}(n) = O(n^2) + O((n/2)^2) = O(n^2)$, however, the leading error constant is larger.

As in the case of digital options, the computational effort of the 3/2-Optimal model is $\mathcal{C}(n) = O(n^2)$, since the additional calculations contribute only $O(1)$ to the effort.

As a result, for plain vanilla options we have

	CRR, RB	Tian, CP (Extrapolation)	OD model	3/2 - Optimal
ε	$O\left(n^{-1}\right)$	$O\left(n^{-\frac{3}{2}}\right)$	$O\left(n^{-\frac{3}{2}}\right)\ \left(O\left(n^{-1}\right)\right)$	$O\left(n^{-2}\right)$
$\mathcal{C}\left(\varepsilon\right)$	$O\left(\varepsilon^{-2}\right)$	$O\left(\varepsilon^{-\frac{4}{3}}\right)$	$O\left(\varepsilon^{-\frac{4}{3}}\right)\ \left(O\left(\varepsilon^{-2}\right)\right)$	$O\left(\varepsilon^{-1}\right)$

We now compare the computational complexity of the Tian model with extrapolation, the OD and the 3/2-Optimal models. As before, we will use the RMSR error and will only include those parameter vectors, for which the OD model has order of convergence $O\left(n^{-\frac{3}{2}}\right)$. Out of a sample of 2000 simulated options 1928 have been included.

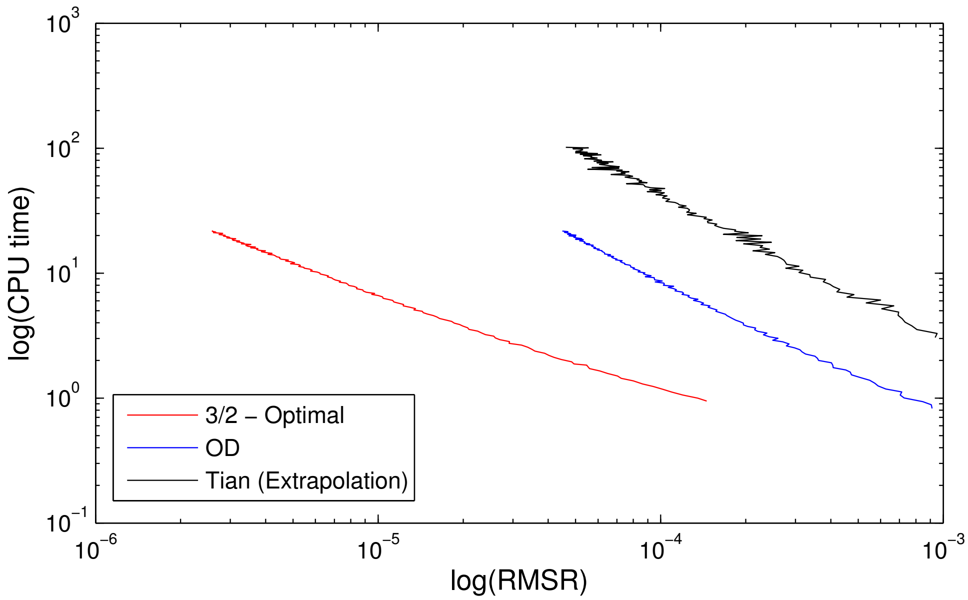


Figure 4.21: Computational complexity for a maximum of $n = 750$ time steps.

Note that the OD model and the Tian model with extrapolation do have the same slope, but due to the additional computations the latter requires a larger effort to obtain the same precision.

4.3.3 Barrier options

So far we have considered options whose payout depends only on the stock price at maturity. To conclude this section we would now like to show how to derive asymptotic expansions and construct advanced binomial trees for path-dependent options, namely barrier options.

A barrier put (call) option has the same payoff at maturity as the European put (call), however, whether or not the owner of the option will receive this payoff depends on the behavior of the whole path of the stock. One distinguishes between knock-in and knock-out options.

- A *knock-in* put (call) pays out $(K - S(T))^+$ $\left((S(T) - K)^+\right)$ if the stock price reaches a certain level (barrier) B before maturity T.
- A *knock-out* put (call) pays out $(K - S(T))^+$ $\left((S(T) - K)^+\right)$ if the stock price doesn't cross the given barrier B before maturity T.

Knock-in(out) options are further subdivided into *up-and-in(out)* and *down-and-in(out)* options, depending on the position of the initial stock price relative to the barrier and whether the barrier is hit from below (up-and-in(out)) or from above (down-and-in(out)). The following in-out-parity relations hold for put (call) options with the same maturity, strike and barrier

$$\begin{aligned} V_{P(C)} &= V^{ui}_{P(C)} + V^{uo}_{P(C)}, \\ V_{P(C)} &= V^{di}_{P(C)} + V^{do}_{P(C)}, \end{aligned} \tag{4.58}$$

where $V_{P(C)}$ denotes the price of the corresponding vanilla put (call).

We will focus on the price of an up-and-in put. Prices of the other barrier options can be obtained in a similar manner or using the relations (4.58). The payout of an up-and-in put is given by

$$(K - S(T))^+ \, \mathbb{1}_{\left\{S(t)\geq B, \text{ for some } t\in[0,T]\right\}}.$$

We assume $s_0 < B$, since otherwise the barrier option becomes a plain vanilla option. Let $K < B$, then the price of an up-and-in put is given by (see, e.g. [H06])

$$V_P^{ui} = Ke^{-rT}\left(\frac{B}{s_0}\right)^{\mu}\Phi\left(d_3\right) - s_0\left(\frac{B}{s_0}\right)^{\mu+2}\Phi\left(d_4\right),$$

where $\mu = \frac{2r}{\sigma^2} - 1$ and

$$\begin{aligned} d_3 &= \frac{\log\left(\frac{Ks_0}{B^2}\right) - \left(r - \frac{1}{2}\sigma^2\right)T}{\sigma\sqrt{T}}, \\ d_4 &= d_3 - \sigma\sqrt{T}. \end{aligned}$$

Lattice methods for barrier options have a very irregular convergence behaviour due to the position of the barrier. This phenomenon as well as possible solutions have been studied by various authors, see for example [BL94], [DKEB95], [R95], etc. However, not that many results are available about the approximation error. A first order asymptotic expansion for binomial trees has been obtained in [G99]. We would now like to apply the Edgeworth expansion to get second-order asymptotics for binomial.

Binomial trees for barrier options

Consider the model

$$S_k^{(n)} = S_{k-1}^{(n)} e^{\sigma\sqrt{\Delta t}\xi_k^{(n)}}, \quad k = 1\dots,n, \tag{4.59}$$

with the probability of an up-jump

$$p_n = \frac{1}{2} + c_{1,n}\frac{1}{\sqrt{n}} + c_{2,n}\frac{1}{n} + c_{3,n}\frac{1}{\sqrt{n}^3}, \tag{4.60}$$

where

$$c_{1,n} = \frac{r - \frac{1}{2}\sigma^2}{2\sigma}\sqrt{T},$$

and $c_{2,n}, c_{3,n}$ are bounded. The price of the up-and-in put in the above model is given by

$$V_n^{ui} = E\left(e^{-rT}\left(K - S_n^{(n)}\right)\mathbb{1}_{\left\{\max_{1\le k\le n} S_k^{(n)}\ge B,\, S_n^{(n)}\le K\right\}}\right).$$

Let

$$a_n^K = \frac{\log\left(\frac{K}{s_0}\right)}{\sigma\sqrt{\Delta t}}, \quad a_n^B = \frac{\log\left(\frac{B}{s_0}\right)}{\sigma\sqrt{\Delta t}}, \quad l_n^B = \lceil a_n^B \rceil .$$

Then $\left\{-a_n^B\right\}$ is the overshoot of the barrier in the discrete model.

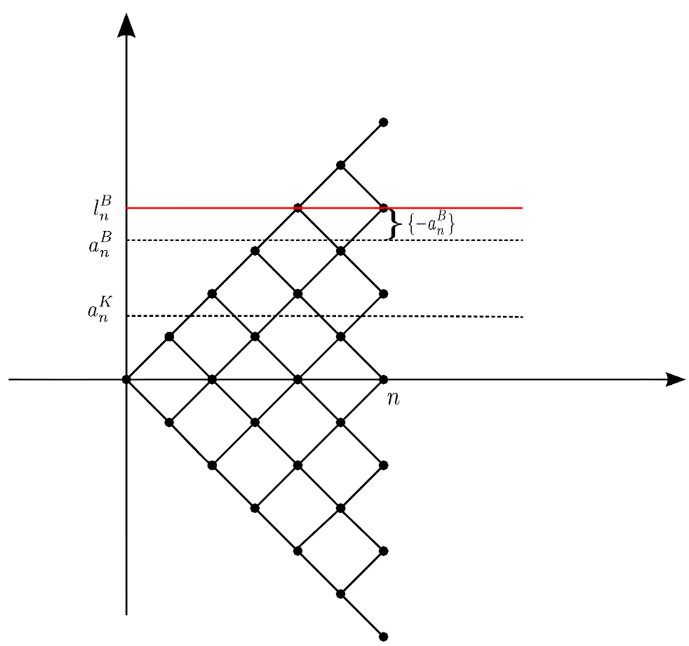


Figure 4.22: Dynamics of $\sum_{i=1}^n \xi_i^{(n)}$

Therefore,

$$\begin{aligned}
V_n^{ui} &= \sum_{x \le a_n^K} E\left(e^{-rT}\left(K - s_0 e^{\sigma\sqrt{\Delta t}\sum_{i=1}^n \xi_i^{(n)}}\right)\mathbb{1}_{\left\{\max_{1\le k\le n}\sum_{i=1}^k \xi_i^{(n)} \ge a_n^B,\ \sum_{i=1}^n \xi_i^{(n)} = x\right\}}\right) \\
&= e^{-rT} K \sum_{x \le a_n^K} P\left(\max_{1\le k\le n}\sum_{i=1}^k \xi_i^{(n)} \ge l_n^B,\ \sum_{i=1}^n \xi_i^{(n)} = x\right) \\
&- s_0 \sum_{x \le a_n^K} e^{-rT + \sigma\sqrt{\Delta t}\cdot x} P\left(\max_{1\le k\le n}\sum_{i=1}^k \xi_i^{(n)} \ge l_n^B,\ \sum_{i=1}^n \xi_i^{(n)} = x\right)
\end{aligned} \tag{4.61}$$

In order to calculate the probabilities in (4.61) we will need the following lemma that makes use of the reflection principle for a simple random walk.

Lemma 4.3.8. *Let $S_n := X_1 + \cdots + X_n$, where X_i are i.i.d, $X_i = \pm 1$ with probabilities p and q. Then*

$$P\left(\max_{1\le k\le n} S_k \ge b,\ S_n = x\right) = \begin{cases} \left(\frac{p}{q}\right)^b P\left(S_n = x - 2b\right), & \text{if } x < b \\ \\ P\left(S_n = x\right), & \text{if } x \ge b \end{cases}. \tag{4.62}$$

Proof. Let $x \ge b$. In this case $S_n = x$ implies $\max_{1\le k\le n} S_k \ge b$ and the statement of the lemma trivially holds.

Now let $x < b$. Following [Konst09] we introduce the notation

$$\{(0,0) \to (n,x)\} := \{\text{all the paths that start at } (0,0) \text{ and end up at } (n,x)\}.$$

Every path from $(0,0)$ to (n,x) has $(n+x)/2$ up-moves and $(n-x)/2$ down-moves, and the probability of each of these paths is $p^{(n+x)/2}q^{(n-x)/2}$, if $|x| \le n$, $n + x$ is even, and 0, otherwise. We assume the latter conditions on n and x, since otherwise, both sides of (4.62) are zero. Then

$$P\left(\max_{1\le k\le n} S_k \ge b,\ S_n = x\right) = \#\left\{(0,0) \to (n,x) \middle|\, \text{cross } b\right\} p^{(n+x)/2} q^{(n-x)/2}.$$

By the reflection principle

$$\#\left\{(0,0) \to (n,x) \middle|\, \text{cross } b\right\} = \#\left\{(0,0) \to (n, 2b - x)\right\}.$$

Indeed, every path that crosses b and ends up in (n,x) can be reflected starting from the first hitting point of b to create a path that ends up in $(n, 2b - x)$ and vice versa (see figure 4.23). Therefore,

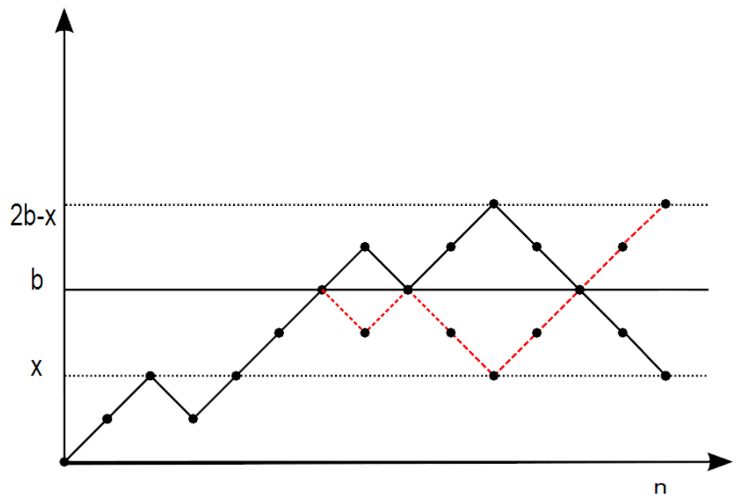


Figure 4.23: Reflection principle for a binomial random walk

$$
\begin{aligned}
P\left(\max_{1\leq k\leq n} S_k \geq b,\ S_n = x\right) &= \left(\begin{array}{c} n \\ \frac{n+2b-x}{2} \end{array}\right) p^{(n+x)/2} q^{(n-x)/2} \\
&= \left(\frac{p}{q}\right)^b \left(\begin{array}{c} n \\ \frac{n+x-2b}{2} \end{array}\right) p^{(n+x)/2-b} q^{(n-x)/2+b} \\
&= \left(\frac{p}{q}\right)^b P\left(S_n = x - 2b\right),
\end{aligned}
$$

for $x \geq -n + 2b$, otherwise both sides are zero. $\square$

Substituting lemma 4.3.8 into (4.61) we get

$$
\begin{aligned}
& e^{-rT} K \sum_{x\leq a_n^K} P\left(\max_{1\leq k\leq n} \sum_{i=1}^{k} \xi_i^{(n)} \geq l_n^B,\ \sum_{i=1}^{n} \xi_i^{(n)} = x\right) \\
&= e^{-rT} K \left(\frac{p}{q}\right)^{l_n^B} \sum_{x\leq a_n^K} P\left(\sum_{i=1}^{n} \xi_i^{(n)} = x - 2l_n^B\right) \\
&= e^{-rT} K \left(\frac{p}{q}\right)^{l_n^B} P\left(\sum_{i=1}^{n} \xi_i^{(n)} \leq a_n^K - 2l_n^B\right),
\end{aligned}
$$

and with the same change of measure as in (4.48)

$$
\begin{aligned}
s_0 \sum_{x \le a_n^K} e^{-rT+\sigma\sqrt{\Delta t}\cdot x} P\left(\max_{1\le k\le n} \sum_{i=1}^{k} \xi_i^{(n)} \ge l_n^B,\ \sum_{i=1}^{n} \xi_i^{(n)} = x\right) \\
= s_0 \left(\frac{p}{q}\right)^{l_n^B} \sum_{x \le a_n^K} e^{-rT+\sigma\sqrt{\Delta t}\cdot x} P\left(\sum_{i=1}^{n} \xi_i^{(n)} = x - 2l_n^B\right) \\
= s_0 \left(\frac{p}{q}\right)^{l_n^B} e^{2l_n^B \sigma\sqrt{\Delta t}} M_n^n \tilde{P}\left(\sum_{i=1}^{n} \xi_i^{(n)} \le a_n^K - 2l_n^B\right).
\end{aligned}
$$

As a result

$$
\begin{aligned}
V_n^{ui} = e^{-rT} K \left(\frac{p}{q}\right)^{l_n^B} P\left(\sum_{i=1}^{n} \xi_i^{(n)} \le a_n^K - 2l_n^B\right) \\
- s_0 \left(\frac{p}{q}\right)^{l_n^B} e^{2l_n^B \sigma\sqrt{\Delta t}} M_n^n \tilde{P}\left(\sum_{i=1}^{n} \xi_i^{(n)} \le a_n^K - 2l_n^B\right).
\end{aligned}
\tag{4.63}
$$

Proposition 4.3.9. *With an appropriate choice of coefficients $c_{2,n}$ and $c_{3,n}$, the binomial model in (4.59) satisfies*

$$
V_n^{ui} = V^{ui} + O\left(\frac{1}{\sqrt{n}^3}\right).
$$

Proof.

$$
P\left(\sum_{i=1}^{n} \xi_i^{(n)} \le a_n^K - 2l_n^B\right) = P\left(\frac{1}{\sqrt{n}} \sum_{i=1}^{n} \left(\eta_i^{(n)} - \mu_n\right) \le y_n\right),
$$

where $\eta_i^{(n)}$ is defined as in (4.6), $\mu_n = E\left(\eta_i^{(n)}\right)$ and

$$
y_n = \frac{\log\left(\frac{K s_o}{B^2}\right) - \mu(n)\, T}{2\sigma\sqrt{T}} - \frac{\left\{-a_n^B\right\}}{\sqrt{n}}.
$$

From (4.60) we get

$$
\mu(n) = r - \frac{1}{2}\sigma^2 + \frac{2c_{2,n}\sigma}{\sqrt{T}} \frac{1}{\sqrt{n}} + \frac{2c_{3,n}\sigma}{\sqrt{T}} \frac{1}{n},
$$

and following the proof of Proposition 4.2.4 we get

$$
\frac{y_n}{\sigma_n} = d_3 - \frac{2}{\sqrt{n}}\left(\left\{-a_n^B\right\} + c_{2,n}\right) + \frac{2}{n}\left(2d_3 c_{1,n}^2 - 2c_{3,n}\right) + O\left(\frac{1}{\sqrt{n}^3}\right),
$$

and

$$P\left(\sum_{i=1}^{n}\xi_i^{(n)} \le a_n^K - 2l_n^B\right) = \Phi\left(d_3\right) - \frac{2}{\sqrt{n}}\phi(d_3)\left(\left\{-a_n^B\right\} + c_{2,n} + S_1\left(a_n\right)\right)$$
$$+\frac{2}{n}\phi(d_3)\left(-c_{3,n} + d_3 c_{1,n}^2 + \frac{1}{3}c_{1,n}(d_3^2-1) + \frac{d_3^3-d_3}{24} - d_3\left(\left\{-a_n^B\right\} + c_{2,n} + S_1\left(a_n\right)\right)^2\right)$$
$$+O\left(\frac{1}{\sqrt{n}^3}\right), \tag{4.64}$$

where $a_n = \frac{\log\left(\frac{Ks_0}{B^2}\right)}{2\sigma\sqrt{\Delta t}} + \frac{n}{2} - \left\{-a_n^B\right\}$.

Similarly

$$\tilde{\mu}\left(n\right) = r + \frac{1}{2}\sigma^2 + \frac{2\tilde{c}_{2,n}\sigma}{\sqrt{T}}\frac{1}{\sqrt{n}} + \frac{2\tilde{c}_{3,n}\sigma}{\sqrt{T}}\frac{1}{n},$$

and

$$\tilde{P}\left(\sum_{i=1}^{n}\xi_i^{(n)} \le a_n^K - 2l_n^B\right) = \Phi\left(d_4\right) - \frac{2}{\sqrt{n}}\phi(d_4)\left(\left\{-a_n^B\right\} + c_{2,n} + S_1\left(a_n\right)\right)$$
$$+\frac{2}{n}\phi(d_4)\left(-\tilde{c}_{3,n} + d_4 \tilde{c}_{1,n}^2 + \frac{1}{3}\tilde{c}_{1,n}(d_4^2-1) + \frac{d_4^3-d_4}{24} - d_4\left(\left\{-a_n^B\right\} + c_{2,n} + S_1\left(a_n\right)\right)^2\right)$$
$$+O\left(\frac{1}{\sqrt{n}^3}\right). \tag{4.65}$$

$$\begin{aligned} e^{2l_n^B\sigma\sqrt{\Delta t}} &= \left(\frac{B}{s_0}\right)^2 e^{2\left\{-a_n^B\right\}\sigma\sqrt{\Delta t}} \\ &= \left(\frac{B}{s_0}\right)^2\left(1 + \frac{1}{\sqrt{n}}2\sigma\sqrt{T}\left\{-a_n^B\right\} + \frac{1}{n}2\sigma^2 T\left\{-a_n^B\right\}^2\right) + O\left(\frac{1}{\sqrt{n}^3}\right). \end{aligned} \tag{4.66}$$

$$\begin{aligned} \left(\frac{p}{q}\right)^{l_n^B} &= \left(\frac{B}{s_0}\right)^{\mu}\left(1 + \frac{4}{\sqrt{n}}\left(c_{1,n}\left\{-a_n^B\right\} + c_{2,n}\frac{\log\frac{B}{s_0}}{\sigma\sqrt{T}}\right)\right) \\ &+ \frac{4}{n}\left(c_{2,n}\left\{-a_n^B\right\} - \frac{\log\frac{B}{s_0}}{\sigma\sqrt{T}}\left(\frac{2}{3}c_{1,n}^3 - c_{3,n}\right) + 2\left(c_{1,n}\left\{-a_n^B\right\} + c_{2,n}\frac{\log\frac{B}{s_0}}{\sigma\sqrt{T}}\right)^2\right) \\ &+ O\left(\frac{1}{\sqrt{n}^3}\right). \end{aligned} \tag{4.67}$$

The above result was obtained with the help of the built-in function TAYLOR in MATLAB. Substituting (4.64), (4.65), (4.66), (4.67) and the expansions from the previous section (4.55) and (4.57) into (4.63), and taking into account $s_0\phi(d_4) = Ke^{-rT}\phi(d_3)$, we get

$$V_n^{ui} = V^{ui} - \frac{1}{\sqrt{n}}\left(c_{2,n}g_{1,n} + \left\{-a_n^B\right\}g_{2,n}\right) + \frac{1}{n}\left(g_{3,n} - c_{3,n}g_{1,n}\right) + O\left(\frac{1}{\sqrt{n}^3}\right), \tag{4.68}$$

where

$$\begin{aligned} g_{1,n} &= 2\sigma\sqrt{T}s_0\left(\frac{B}{s_0}\right)^{\mu+2}\Phi\left(d_4\right) - 4V^{ui}\log\left(\frac{B}{s_0}\right)/(\sigma\sqrt{T}), \\ g_{2,n} &= 2\sigma\sqrt{T}s_0\left(\frac{B}{s_0}\right)^{\mu+2}\Phi\left(d_4\right) - 4V^{ui}c_{1,n}, \\ g_{3,n} &= \sigma\sqrt{T}s_0\left(\frac{B}{s_0}\right)^{\mu+2}\left(\phi\left(d_4\right)h_{1,n} - \Phi\left(d_4\right)h_{2,n}\right) + 4V^{ui}h_{3,n}, \end{aligned}$$

and

$$\begin{aligned} h_{1,n} &= -2c_{1,n}^2 - c_{1,n}\left(\sigma\sqrt{T} + \frac{d_3+d_4}{3}\right) - \frac{1}{8}\sigma^2 T + \frac{1}{4} - \frac{d_3^2}{8} + \frac{d_4^2}{24} - 2\left(\left\{-a_n^B\right\} + c_{2,n} + S_1\left(a_n\right)\right)^2 \\ &\quad + 4\left(\left\{-a_n^B\right\} + c_{2,n}\right)\left(\left\{-a_n^B\right\} + c_{2,n} + S_1\left(a_n\right)\right), \\ h_{2,n} &= 2\sigma\sqrt{T}\left(\left\{-a_n^B\right\} + c_{2,n}\right)^2 - 2\sigma\sqrt{T}(c_{1,n} + \frac{1}{6}\sigma\sqrt{T})^2 - \frac{1}{36}\sigma^3\sqrt{T}^3 \\ &\quad + 8\left(\left\{-a_n^B\right\} + c_{2,n}\right)\left(c_{1,n}\left\{-a_n^B\right\} + c_{2,n}\frac{\log\frac{B}{s_0}}{\sigma\sqrt{T}}\right), \\ h_{3,n} &= c_{2,n}\left\{-a_n^B\right\} - \frac{2\log\left(\frac{B}{s_0}\right)}{3\sigma\sqrt{T}}c_{1,n}^3 + 2\left(c_{1,n}\left\{-a_n^B\right\} + c_{2,n}\frac{\log\frac{B}{s_0}}{\sigma\sqrt{T}}\right)^2. \end{aligned}$$

Therefore, if $g_{1,n} \neq 0$, then by setting

$$\begin{aligned} c_{2,n} &= -\frac{\left\{-a_n^B\right\}g_{2,n}}{g_{1,n}}, \\ c_{3,n} &= \frac{g_{3,n}}{g_{1,n}}, \end{aligned}$$

we get the statement of the proposition. □

Remark 4.3.10. Consider the leading error coefficient in (4.68). Note that $g_{1,n}$, $g_{2,n}$ are constant, therefore the oscillatory convergence behavior is due to the overshoot of the barrier $\left\{-a_n^B\right\}$. The relative position of the strike with respect to the two neighboring nodes at maturity does not enter into the expression. The strike is only present starting from the $O\left(\frac{1}{n}\right)$ coefficient in the term $S_1(a_n)$ together with the barrier in quadratic form. Therefore, the position of the barrier will have a much stronger effect on the convergence pattern than the position of the strike.

Numerical results

We now consider the convergence pattern of a specific barrier option.

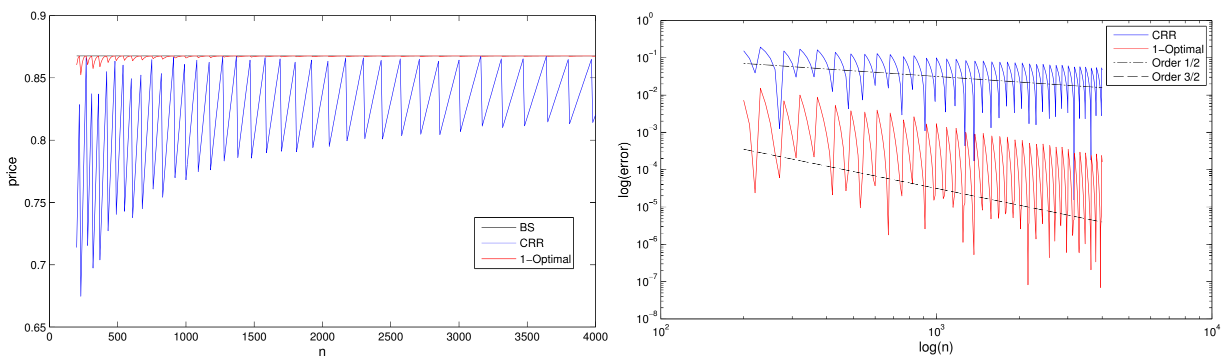


Figure 4.24: Up-and-in barrier put: $T = 1$, $r = 0.1$, $\sigma = 0.25$, $s_0 = 100$, $K = 105$, $B = 120$

Remark 4.3.11. Note that in the CRR tree $c_{2,n} = 0$, therefore, the leading error coefficient in (4.68) becomes $\left\{-a_n^B\right\} g_{2,n}$, where $0 < \left\{-a_n^B\right\} \leq 1$. Therefore, the binomial tree will either overestimate or underestimate the BS price for all n, depending on the sign of $g_{2,n}$, as can be seen in the above figures.

CRR vs RB We would now like to compare the convergence behaviour of the CRR and the RB models for barrier options. As we have seen in the previous chapters the CRR and RB trees for digital and plain vanilla options have a similar convergence behavior and no clear preferences can be stated. However, this is not the case for barrier options. Expansions for the RB tree cannot be obtained with the method described above, since the reflection principle is directly applicable only to the CRR tree, however, numerical results show that the RB tree has a much smoother convergence compared to the CRR tree.

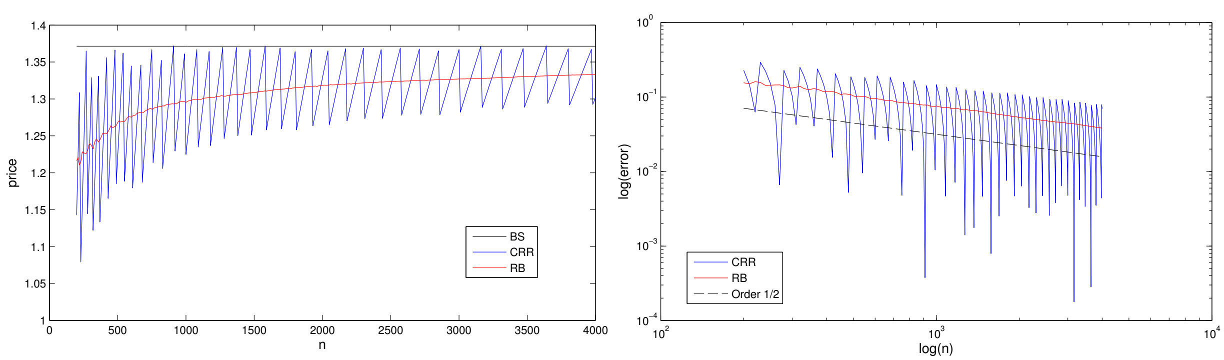


Figure 4.25: Barrier option: $T = 1$, $r = 0.1$, $\sigma = 0.25$, $s_0 = 100$, $K = 110$, $B = 120$

This can be explained by the fact that, since the CRR tree is symmetric around 0 in the log-scale, the overshoot of the barrier $\left\{-a_n^B\right\}$ is the same for each time step. Therefore, with an increase of n, a whole row of nodes becomes out-of-the-money. The RB tree, on the other hand, is tilted, therefore, this effect is not that pronounced.

Due to the smooth convergence pattern we can apply extrapolation to the RB tree to increase the order of convergence. As a result we get

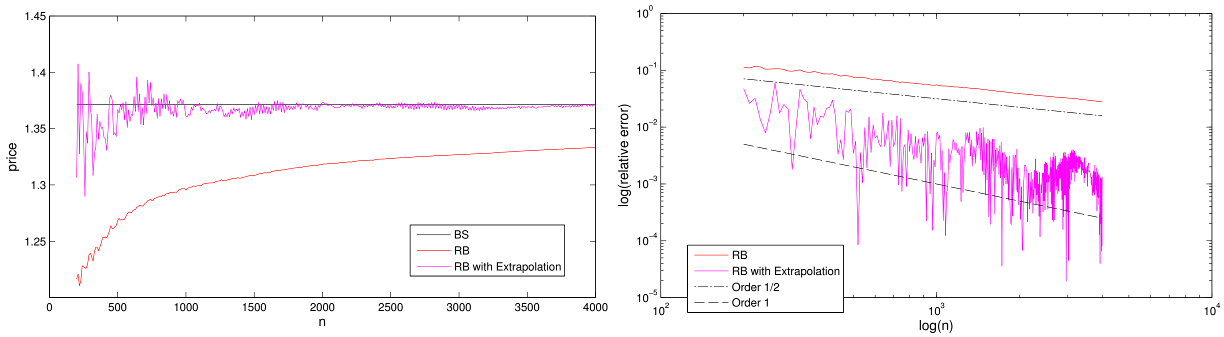


Figure 4.26: Barrier option: $T = 1$, $r = 0.1$, $\sigma = 0.25$, $s_0 = 100$, $K = 110$, $B = 120$

Parameters	n	CRR tree	RB Extrapolation	1-Optimal	BS value
$T = 1$	100	1.0370950	1.2728844	1.3071811	1.3714613
$s_0 = 100$	200	1.1428755	1.3064705	1.3528020	
$K = 110,\ B = 120$	500	1.2210427	1.3667218	1.3668495	
$\sigma = 0.25$	1000	1.2248525	1.3608690	1.3671287	
$r = 0.1$	2000	1.3285299	1.3731534	1.3713814	
	4000	1.3018025	1.3696310	1.3710375	

4.4 Approximating the Greeks

In practice numerical methods are used not only to calculate the option price, but also the sensitivity of the price with respect to the change in the underlying parameters or variables. These sensitivities are usually referred to as the Greeks, and are used for hedging. The most common Greeks are

- Delta: $\Delta = \frac{\partial V}{\partial s_0}$;
- Gamma: $\Gamma = \frac{\partial \Delta}{\partial s_0}$;
- Theta. $\Theta = \frac{\partial V}{\partial t}$;
- Rho: $\rho = \frac{\partial V}{\partial r}$;
- Vega: $\nu = \frac{\partial V}{\partial \sigma}$.

In this thesis we will consider the Δ, Γ and Θ for European options in the Black-Scholes model.

The delta represents the change of the option price with changes in the underlying spot price. It indicates the number of units of the underlying one needs to hold to hedge a short option position against small movements in the market. For European call (put) options the delta is given by

$$\Delta = \Phi(-d_1) \quad (\Delta = -\Phi(d_1)),$$

where d_1 is defined as in (4.42). Therefore, the delta is positive for call options, and negative for puts. Note that $|\Delta| \leq 1$, approaching zero for out-of-the-money options and 1 for in-the-money options.

Gamma is a second-order derivative and is also an important tool in hedging. It measures the sensitivity of an option's delta to changes in the price of the underlying. For both call and put options the gamma can be calculated as

$$\Gamma = \frac{\phi\left(-d_1\right)}{s_0 \sigma \sqrt{T}}.$$

Note that at-the-money options have the largest gamma.

Theta is the change in the price with the passage of time. It can be calculated as

$$\Theta = -\frac{\sigma s_0 \phi\left(-d_1\right)}{2\sqrt{T}} + \delta r K e^{-rT} \Phi\left(\delta d_2\right),$$

with d_2 as in (4.5) and $\delta = 1$ for put options and -1 for calls.

We will now discuss how the above values can be approximated using the binomial model. There are various ways of calculating the above Greeks with the lattice approach. The first option is a numerical differentiation approach, which requires the construction of additional trees with shifted initial values to calculate the necessary difference. The drawback of this method is that it is very time-consuming, and it often leads to inaccurate approximations of the gamma (see e.g. [PV94]). Another approach is to read off the delta and gamma directly from the tree by using the values of the option and the stock after the first and second steps to approximate the derivatives (see e.g. [Hull06]). Let $S_{i,j}$ denote the stock price at time $i\Delta t$ with j up-jumps and $i-j$ down-jumps. From Figure 4.27 we get

$$\begin{aligned} \Delta^{(n)} &= \frac{V^{(n)}\left(S_{1,1}\right) - V^{(n)}\left(S_{1,0}\right)}{S_{1,1} - S_{1,0}}, \\ \Gamma^{(n)} &= \left(\frac{V^{(n)}\left(S_{2,2}\right) - V^{(n)}\left(S_{2,1}\right)}{S_{2,2} - S_{2,1}} - \frac{V^{(n)}\left(S_{2,1}\right) - V^{(n)}\left(S_{2,0}\right)}{S_{2,1} - S_{2,0}}\right) \frac{2}{\left(S_{2,2} - S_{2,0}\right)}, \end{aligned} \tag{4.69}$$

where $V^{(n)}\left(S_{i,j}\right)$ denotes the price of the option calculated at node $S_{i,j}$.

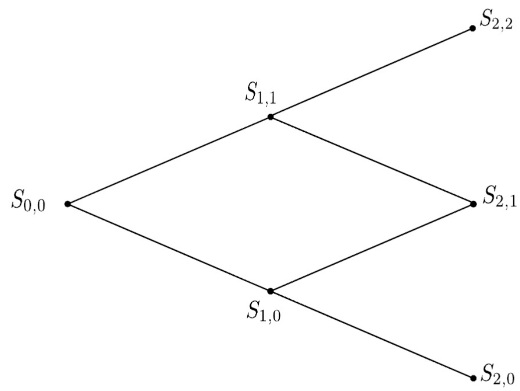


Figure 4.27

Theta can then be approximated using the Black-Scholes PDE as

$$\Theta^{(n)} = rV^{(n)} - rs_0\Delta^{(n)} - \frac{1}{2}\sigma^2 s_0^2 \Gamma^{(n)}. \tag{4.70}$$

The problem with this method is that the values of $S_{i,j}$ in (4.69) are taken at time Δt and $2\Delta t$, although we are supposed to consider change in the stock price without changing the time. However, this can be neglected by taking a large enough n. Nevertheless we will consider a different approach, the extended binomial tree method introduced in [PV94]. The idea of the method is to extend the tree two periods before the starting date as in Figure 4.28.

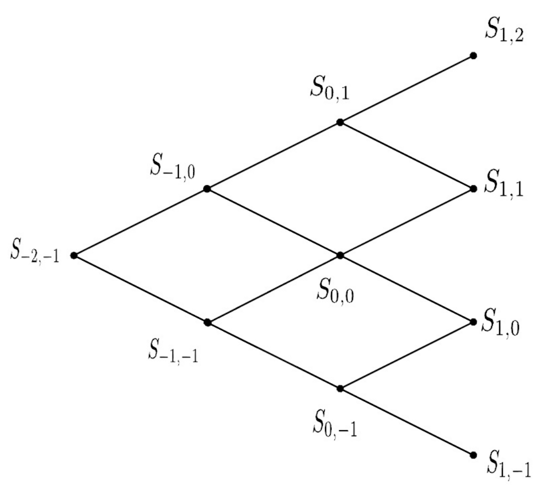


Figure 4.28

The delta and gamma can then be calculated from the option and stock values at time zero.

$$\begin{aligned}\Delta^{(n)} &= \frac{V^{(n)}\left(S_{0,1}\right) - V^{(n)}\left(S_{0,-1}\right)}{S_{0,1} - S_{0,-1}},\\ \Gamma^{(n)} &= \left(\frac{V^{(n)}\left(S_{0,1}\right) - V^{(n)}\left(S_{0,0}\right)}{S_{0,1} - S_{0,0}} - \frac{V^{(n)}\left(S_{0,0}\right) - V^{(n)}\left(S_{0,-1}\right)}{S_{0,0} - S_{0,-1}}\right)\frac{2}{\left(S_{0,1} - S_{0,-1}\right)}.\end{aligned} \tag{4.71}$$

Substituting these values into (4.70) we get an approximation for Θ.

The convergence behavior of $\Delta^{(n)}$ and $\Gamma^{(n)}$ in (4.71) has already been studied in [CHLS11] for various types of trees. They confirm the $O\left(\frac{1}{n}\right)$ order of convergence for $\Delta^{(n)}$ theoretically and show that the relative position of the strike with respect to the neighboring nodes affects the convergence pattern of the greeks as well. In this section we apply the Edgeworth expansion to not only prove the $O\left(\frac{1}{n}\right)$ rate of convergence for $\Delta^{(n)}$, but also to get the leading error coefficient. We consider only the European put, however, similar results can also be obtained for the call option.

Proposition 4.4.1. *Consider the general binomial model (4.31) with drift and probabilities as in (4.32) and* $k_{1,n} = \frac{2\sigma c_{2,n}}{\sqrt{T}}$. *Then* $\Delta^{(n)}$ *in (4.71) for a European put satisfies*

$$\Delta^{(n)} = \Delta + \left(\phi\left(d_1\right)\delta_n - \Phi\left(d_1\right)a_{2,n}\right)\frac{1}{n} + O\left(\frac{1}{n^{3/2}}\right), \quad n \to \infty, \tag{4.72}$$

where

$$\delta_n = 2d_2 - \tilde{f}_2\left(0\right) + \frac{1}{4\sigma\sqrt{T}}\left(4a_{2,n} + f_3\left(2\right) - f_3\left(-2\right) - \left(\tilde{f}_3\left(2\right) - \tilde{f}_3\left(-2\right)\right)\right),$$

with $a_{2,n}$ *defined as in* (4.56) *and* $\tilde{f}_2\left(0\right) := \tilde{f}_2\left(\alpha_n, p_n, m_1 = 0\right)$, $f_3\left(\pm 2\right) := f_3\left(\alpha_n, p_n, m_1 = \pm 2\right)$, $\tilde{f}_3\left(\pm 2\right) := \tilde{f}_3\left(\alpha_n, p_n, m_1 = \pm 2\right)$, f_3 *defined as in (4.37) and* $\tilde{f}_2$, $\tilde{f}_3$ *as in* (4.53).

Proof. Extending the model (4.31) as in Figure 4.28 we get $S_{0,0} = s_0$, $S_{0,1} = s_0 e^{2\sigma\sqrt{\Delta t}}$ and $S_{0,-1} = s_0 e^{-2\sigma\sqrt{\Delta t}}$. Following (4.49) we have

$$V^{(n)}\left(s_0 e^{\pm 2\sigma\sqrt{\Delta t}}\right) = e^{-rT} K P_n\left(S_n^{(n)} \le \frac{K}{e^{\pm 2\sigma\sqrt{\Delta t}}}\right) - s_0 e^{\pm 2\sigma\sqrt{\Delta t}} M_n^n \tilde{P}_n\left(S_n^{(n)} \le \frac{K}{e^{\pm 2\sigma\sqrt{\Delta t}}}\right). \tag{4.73}$$

As usual, setting $\eta_k^{(n)} = \frac{\xi_k^{(n)}+1}{2}$, $k = 1, \ldots, n$, $\mu_n = E\,\eta_1^{(n)}$ and $\sigma_n^2 = Var\,\eta_1^{(n)}$, the first term above becomes

$$P_n\left(S_n^{(n)} \le \frac{K}{e^{\pm 2\sigma\sqrt{\Delta t}}}\right) = P\left(\frac{1}{\sqrt{n}}\sum_{k=1}^{n}\left(\eta_k^{(n)} - \mu_n\right) \le y_n^{\pm}\right),$$

where $y_n^{\pm} = \frac{\ln\left(\frac{x}{s_0}\right) - \mu(n)T}{2\sigma\sqrt{T}} \mp \frac{1}{\sqrt{n}}$. Then

$$\frac{y_n^{\pm}}{\sigma_n} = d_2 \mp 2\frac{1}{\sqrt{n}} + \left(2d_2 c_{1,n}^2 - m_2\right)\frac{1}{n} + \left(4d_2 c_{1,n}c_{2,n} - m_3 \mp 2c_{1,n}^2 2\right)\frac{1}{\sqrt{n}^3} + O\left(\frac{1}{n^2}\right),$$

with m_2, m_3 as in (4.34). Note that the above expansion coincides with (4.35) if m_1 is substituted by ± 2, therefore, following (4.36) we get

$$e^{-rT}KP_n\left(S_n^{(n)} \leq \frac{K}{e^{\pm 2\sigma\sqrt{\Delta t}}}\right) = e^{-rT}K\Phi\left(d_2\right) - \frac{1}{\sqrt{n}}e^{-rT}K\phi(d_2)f_1\left(\pm 2\right) + \frac{1}{n}e^{-rT}K\phi(d_2)f_2\left(\pm 2\right)$$
$$+ \frac{1}{\sqrt{n}^3}e^{-rT}K\phi(d_2)f_3\left(\pm 2\right) + O\left(\frac{1}{n^2}\right).$$

Analogously, following (4.53) we get an expansion for the second term in (4.73). Substituting these expressions into (4.71) and taking into account the Taylor series for $e^{\pm 2\sigma\sqrt{\Delta t}}$, we get the statement of the proposition. □

Consider the 3/2-Optimal model from Proposition 4.3.4. Note that despite the $O\left(\frac{1}{n^2}\right)$ order of convergence of the Black-Scholes price approximations in the 3/2-Optimal model, the delta, gamma and theta converge with order $O\left(\frac{1}{n}\right)$, which is typical for binomial trees. However, the convergence is smooth, therefore, following [CHLS11] we can use extrapolation to improve the order of convergence, which delivers results comparable to those of other advanced tree methods.

Remark 4.4.2. Note that the additional degrees of freedom available in multinomial trees could be used to improve the order of convergence for the greeks. This is left for future research.

Numerical results

Compare the convergence behavior of the delta, gamma and theta in the extended CRR and 3/2-Optimal trees for a European put with $T = 1$, $r = 0.1$, $\sigma = 0.25$, Strike $= 110$ and $s_0 = 100$.

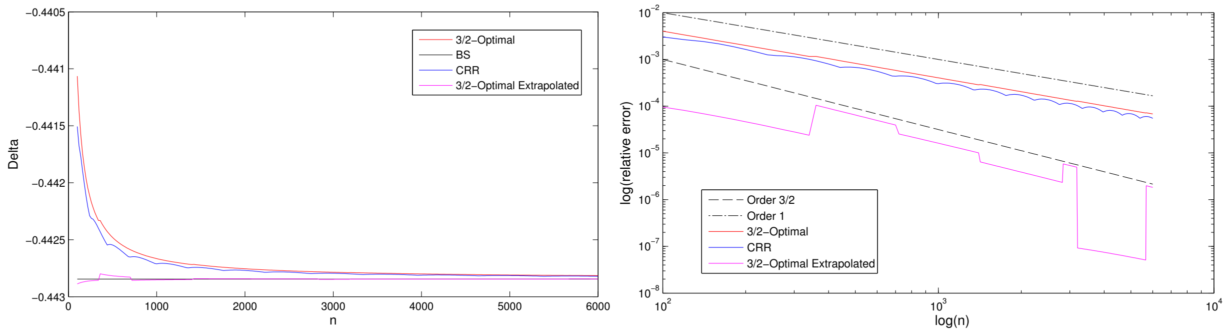


Figure 4.29: Delta, n even

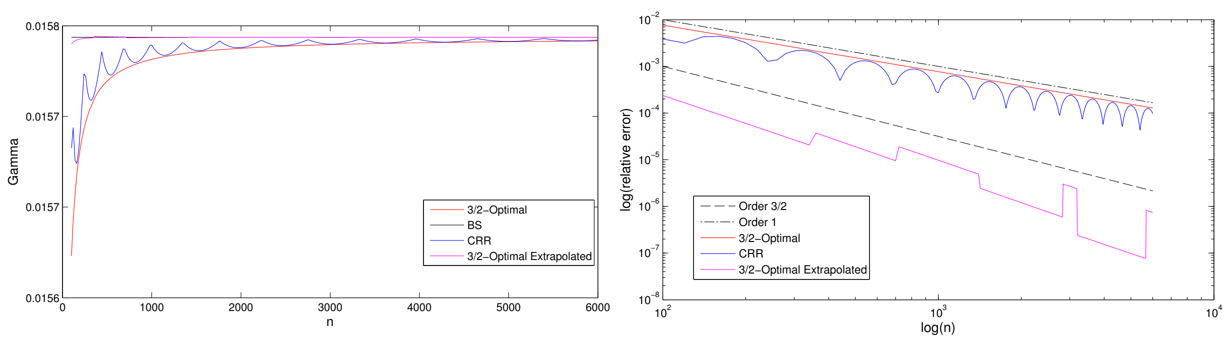


Figure 4.30: Gamma, n even

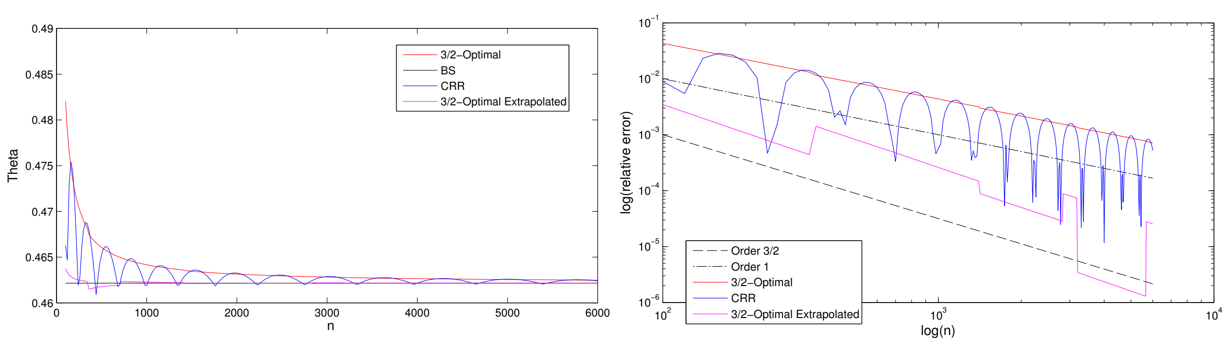


Figure 4.31: Theta, n even

Consider the approximation results for a European put with $T = 1$, $r = 0.1$, $\sigma = 0.25$, Strike $= 100$ and different initial values.

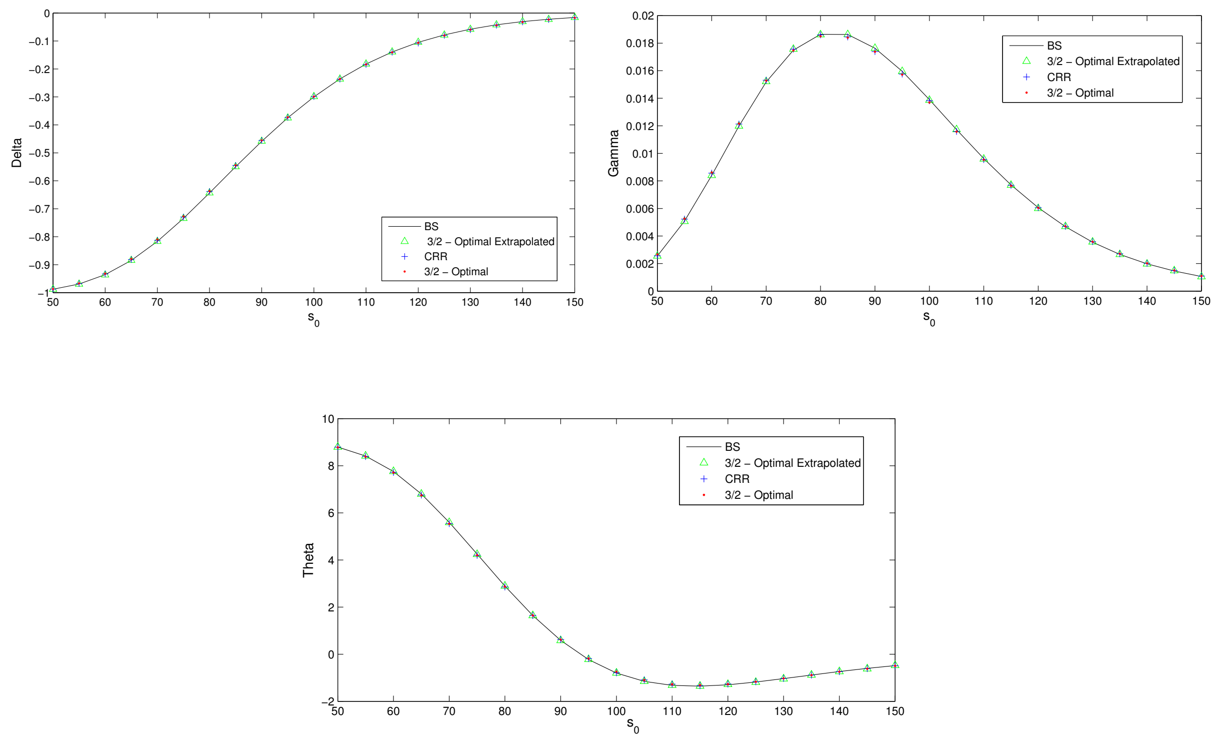


Figure 4.32: Delta, gamma and theta for $n = 50$

4.5 Conclusion

In this chapter we have considered applications of the Edgeworth expansions to one-dimensional tree models in the Black-Scholes setting. We have seen how expansions can be obtained for digital, vanilla and barrier options, and how these results can be used to improve convergence behavior.

Once again, we consider the binomial method as a purely numerical approach, and do not restrict ourselves to the equivalent martingale measure in the binomial setting. This gives us more freedom in the construction of the advanced trees, as it allows to choose the probabilities as well as the drift of the tree. Nevertheless, the expansions obtained in the proofs of the propositions above hold for a very general setting, and can also be applied to trees under the risk-neutral measure.

Chapter 5

Asymptotics of Multidimensional Tree Models

Unlike the one-dimensional case, very little theory is available for the multidimensional setting. Therefore, the goal of this chapter is to provide theoretical results on the asymptotic behavior of multidimensional trees, that would allow to compare existing methods and to develop new ones with superior convergence properties.

We apply Edgeworth expansions to obtain the distributional fit of multi-asset binomial trees. As an example, we present asymptotic expansions of order $O\left(\frac{1}{n^{3/2}}\right)$ for the 2D BEG and the 2D-RB models mentioned in section 2.2.2. We then show how advanced trees can be constructed in the multidimensional case and present models with a $O\left(\frac{1}{n^{3/2}}\right)$ order of convergence for the two-asset digital and two-asset correlation options. In conclusion we consider the asymptotics of the decoupling approach.

We will start with a two-dimensional model and consider further generalizations later on.

5.1 Two-dimensional expansions

We assume a two-dimensional Black-Scholes model:

$$dS_i\left(t\right) = S_i\left(t\right)\left(rdt + \sigma_i dW_i\left(t\right)\right), \quad S_i\left(0\right) = s_{0,i} \quad i = 1, 2,$$

where $W_1(t)$ and $W_2(t)$ are Brownian motions with correlation ρ. As the approximating binomial model we consider the discrete-time process $S_k^{(n)}, k = 1, \ldots, n$ such that

$$\begin{aligned} S_0^{(n)} &= \begin{pmatrix} s_{0,1} \\ s_{0,2} \end{pmatrix}, \\ S_k^{(n)} &= \begin{pmatrix} S_{k-1,1}^{(n)} e^{\alpha_{n,1}\Delta t + \sigma_1\sqrt{\Delta t}\xi_{k,1}^{(n)}} \\ S_{k-1,2}^{(n)} e^{\alpha_{n,2}\Delta t + \sigma_2\sqrt{\Delta t}\xi_{k,2}^{(n)}} \end{pmatrix}, \end{aligned} \tag{5.1}$$

where $\alpha_{n,i}$, $i = 1, 2$, are bounded drifts and for each $n \in \mathbb{N}$, $\xi_k^{(n)}$, $k = 1, \ldots, n$ are i.i.d. random vectors such that

$$\xi_k^{(n)} = \left(\xi_{k,1}^{(n)}, \xi_{k,2}^{(n)}\right) = \begin{cases} (1,1), & \text{with probability } p_n^{(uu)}, \\ (1,-1), & \text{with probability } p_n^{(ud)}, \\ (-1,1), & \text{with probability } p_n^{(du)}, \\ (-1,-1), & \text{with probability } p_n^{(dd)}. \end{cases}$$

To ensure weak convergence we assume that the moment matching conditions (2.8) are satisfied, i.e.

$$\begin{aligned} \mu_i(n) &= r - \frac{1}{2}\sigma_i^2 + o(1), \quad i = 1, 2, \\ \sigma_i^2(n) &= \sigma_i^2 + o(1), \quad i = 1, 2, \\ c(n) &= \rho\sigma_1\sigma_2 + o(1). \end{aligned} \tag{5.2}$$

5.1.1 Distributional Fit

We now focus on the distributional fit of the process $S^{(n)}$ in (5.1) at maturity. As in the one-dimensional case, we are interested in the error

$$P\left(S_n^{(n)} \le x\right) - Q\left(S(T) \le x\right) = P\left(S_n^{(n)} \le x\right) - \Phi_{0,V}\left(d_{2,1}, d_{2,2}\right),$$

where

$$d_{2,i} := \frac{\ln\left(\frac{x_i}{s_{0,i}}\right) - \left(r - \frac{1}{2}\sigma_i^2\right)T}{\sigma_i\sqrt{T}}, \quad i = 1, 2,$$

$$V = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}.$$

At maturity the discrete-time stock price satisfies

$$S_n^{(n)} = \begin{pmatrix} s_{0,1}e^{\alpha_{n,1}T + \sigma_1\sqrt{\Delta t}\sum_{k=1}^n \xi_{k,1}^{(n)}} \\ s_{0,2}e^{\alpha_{n,2}T + \sigma_2\sqrt{\Delta t}\sum_{k=1}^n \xi_{k,2}^{(n)}} \end{pmatrix}.$$

Set $\eta_{j,i}^{(n)} = \left(\xi_{j,i}^{(n)} + 1\right)/2$, so that $\eta_j^{(n)}$ has minimal lattice $\mathbb{Z}^2$ with $P\left(\eta_j^{(n)} \in \mathbb{Z}^2\right) = 1$, for all $j = 1, \ldots, n$. Then

$$\begin{aligned} P\left(S_n^{(n)} \le x\right) &= P\left(S_{n,1}^{(n)} \le x_1, S_{n,2}^{(n)} \le x_2\right) \\ &= P\left(\frac{1}{\sqrt{n}}\left(\sum_{j=1}^n \eta_{j,1}^{(n)} - nE\eta_{1,1}^{(n)}\right) \le y_{n,1}, \frac{1}{\sqrt{n}}\left(\sum_{j=1}^n \eta_{j,2}^{(n)} - nE\eta_{1,2}^{(n)}\right) \le y_{n,2}\right). \end{aligned}$$

where

$$y_{n,i} = \frac{\ln\left(\frac{x_i}{s_{0,i}}\right) - \mu_i(n)T}{2\sigma_i\sqrt{T}}, \quad i = 1, 2. \tag{5.3}$$

We have the following corollary to Theorem 3.4.9.

Corollary 5.1.1. *Let* $\mu_n = E\left(\eta_1^{(n)}\right)$, V_n *be the covariance matrix of* $\eta_1^{(n)}$, *i.e.*

$$V_n = \begin{pmatrix} \sigma_{n,1}^2 & \rho_n \sigma_{n,1}\sigma_{n,2} \\ \rho_n \sigma_{n,1}\sigma_{n,2} & \sigma_{n,2}^2 \end{pmatrix}. \tag{5.4}$$

and $\kappa_{n,\nu}$ *be the* ν*-th cumulant of* $\eta_1^{(n)}$. *The binomial process* $S^{(n)}$ *in (5.1) satisfies*

$$\begin{aligned}
&P\left(S_n^{(n)} \le x\right) = \Phi_{0,V_n}(y_n) \\
&- \frac{1}{\sqrt{n}} \left(S_1(a_{n,1}) D_1 \Phi_{0,V_n}(y_n) + S_1(a_{n,2}) D_2 \Phi_{0,V_n}(y_n) - P_1(-\Phi_{0,V_n}, \{\kappa_{n,\nu}\})(y_n)\right) \\
&+ \frac{1}{n} \left[S_1(a_{n,1}) S_1(a_{n,2}) D_1 D_2 \Phi_{0,V_n}(y_n) + S_2(a_{n,1}) D_1^2 \Phi_{0,V_n}(y_n)\right. \\
&+ S_2(a_{n,2}) D_2^2 \Phi_{0,V_n}(y_n) - S_1(a_{n,1}) D_1 P_1(-\Phi_{0,V_n}, \{\kappa_{n,\nu}\})(y_n) \\
&\left.-S_1(a_{n,2}) D_2 P_1(-\Phi_{0,V_n}, \{\kappa_{n,\nu}\})(y_n) + P_2(-\Phi_{0,V_n}, \{\kappa_{n,\nu}\})(y_n)\right] + O\left(\frac{1}{\sqrt{n}^3}\right),
\end{aligned} \tag{5.5}$$

where y_n *is defined in* (5.3), $a_n = n\mu_n + y_n\sqrt{n}$ *and*

$$P_1(-\Phi_{0,V_n}, \{\kappa_{n,\nu}\})(u) = -\sum_{|\nu|=3} \frac{\kappa_{n,\nu}}{\nu!} D^\nu \Phi_{0,V_n}(u),$$

$$\begin{aligned}
P_2(-\Phi_{0,V_n}, \{\kappa_{n,\nu}\})(u) = &\sum_{|\nu|=4} \frac{\kappa_{n,\nu}}{\nu!} D^\nu \Phi_{0,V_n}(u) \\
&+ \frac{1}{2!} \sum_{\nu_1,\nu_2,\, |\nu_i|=3} \frac{\kappa_{n,\nu_1}\kappa_{n,\nu_2}}{\nu_1!\nu_2!} D^{\nu_1+\nu_2} \Phi_{0,V_n}(u),
\end{aligned}$$

Proof. The moment-matching conditions (5.2) are equivalent to

$$\begin{aligned}
p_n^{(uu)} &= \frac{1}{4}(1+\rho) + o(1), \quad p_n^{(ud)} = \frac{1}{4}(1-\rho) + o(1) \\
p_n^{(du)} &= \frac{1}{4}(1-\rho) + o(1), \quad p_n^{(dd)} = \frac{1}{4}(1+\rho) + o(1)
\end{aligned}$$

Therefore, Lemma 3.4.8 is applicable and the conditions of Theorem 3.4.9 are satisfied for all s. Taking $s = 4$ we have the statement of the corollary. The expressions for $P_i(-\Phi_{0,V_n}, \{\kappa_{n,\nu}\})$ are obtained from Lemma 3.3.3. □

The cumulants $\kappa_{n,\nu}$ can be calculated by using their representation (3.3) in terms of moments. For $|\nu| = 3$ the possible $\kappa_{n,\nu}$ are

$$\begin{aligned}
\kappa_{n,(3,0)} &= \mu_{n,(3,0)} - 3\mu_{n(1,0)}\mu_{n(2,0)} + 2\mu_{n(1,0)}^3, \\
\kappa_{n,(0,3)} &= \mu_{n,(0,3)} - 3\mu_{n(0,1)}\mu_{n(0,2)} + 2\mu_{n(0,1)}^3, \\
\kappa_{n,(2,1)} &= \mu_{n,(2,1)} - \mu_{n,(0,1)}\mu_{n,(2,0)} - 2\mu_{n,(1,1)}\mu_{n,(1,0)} + 2\mu_{n,(1,0)}^2\mu_{n,(0,1)}, \\
\kappa_{n,(1,2)} &= \mu_{n,(1,2)} - \mu_{n,(1,0)}\mu_{n,(0,2)} - 2\mu_{n,(1,1)}\mu_{n,(0,1)} + 2\mu_{n,(0,1)}^2\mu_{n,(1,0)},
\end{aligned}$$

and for $|\nu| = 4$

$$\begin{aligned}
\kappa_{n,(4,0)} &= \mu_{n,(4,0)} - 4\mu_{n,(3,0)}\mu_{n,(1,0)} - 3\mu_{n,(2,0)}^2 + 12\mu_{n,(2,0)}\mu_{n,(1,0)}^2 - 6\mu_{n,(1,0)}^4, \\
\kappa_{n,(0,4)} &= \mu_{n,(0,4)} - 4\mu_{n,(0,3)}\mu_{n,(0,1)} - 3\mu_{n,(0,2)}^2 + 12\mu_{n,(0,2)}\mu_{n,(0,1)}^2 - 6\mu_{n,(0,1)}^4, \\
\kappa_{n,(3,1)} &= \mu_{n,(3,1)} - 3\mu_{n,(2,1)}\mu_{n,(1,0)} - \mu_{n,(3,0)}\mu_{n,(0,1)} - 3\mu_{n,(2,0)}\mu_{n,(1,1)} \\
&\quad + 6\mu_{n,(1,1)}\mu_{n,(1,0)}^2 + 6\mu_{n,(2,0)}\mu_{n,(1,0)}\mu_{n,(0,1)} - 6\mu_{n,(1,0)}^3\mu_{n,(0,1)}, \\
\kappa_{n,(1,3)} &= \mu_{n,(1,3)} - 3\mu_{n,(1,2)}\mu_{n,(0,1)} - \mu_{n,(0,3)}\mu_{n,(1,0)} - 3\mu_{n,(0,2)}\mu_{n,(1,1)} \\
&\quad + 6\mu_{n,(1,1)}\mu_{n,(0,1)}^2 + 6\mu_{n,(0,2)}\mu_{n,(0,1)}\mu_{n,(1,0)} - 6\mu_{n,(0,1)}^3\mu_{n,(1,0)}, \\
\kappa_{n,(2,2)} &= \mu_{n,(2,2)} - 2\mu_{n,(1,2)}\mu_{n,(1,0)} - 2\mu_{n,(2,1)}\mu_{n,(0,1)} - \mu_{n,(2,0)}\mu_{n,(0,2)} - 2\mu_{n,(1,1)}^2 \\
&\quad + 2\mu_{n,(2,0)}\mu_{n,(0,1)}^2 + 2\mu_{n,(0,2)}\mu_{n,(1,0)}^2 + 8\mu_{n,(1,1)}\mu_{n,(1,0)}\mu_{n,(0,1)} - 6\mu_{n,(1,0)}^2\mu_{n,(0,1)}^2
\end{aligned}$$

Asymptotics of the Normal distribution and its derivatives

In order to get the expansion of $P\left(S_n^{(n)} \le x\right)$ around $\Phi_{0,V}(d_2)$ we need to consider the asymptotics of $\Phi_{0,V_n}(y_n)$ and its derivatives.

We assume that the drifts and transition probabilities in the model (5.1) are chosen such that the following expansions hold

$$\begin{aligned}
\mu_i(n) &= r - \frac{1}{2}\sigma_i^2 + a_{i,1}^{(n)}\frac{1}{\sqrt{n}} + a_{i,2}^{(n)}\frac{1}{n} + O\left(n^{-3/2}\right), \\
\sigma_{n,i}^2 &= \frac{1}{4} + b_{i,1}^{(n)}\frac{1}{\sqrt{n}} + b_{i,2}^{(n)}\frac{1}{n} + O\left(n^{-3/2}\right) \\
\rho_n &= \rho + c_1^{(n)}\frac{1}{\sqrt{n}} + c_2^{(n)}\frac{1}{n} + O\left(n^{-3/2}\right),
\end{aligned} \tag{5.6}$$

where $a_{i,1(2)}^{(n)}$, $b_{i,1(2)}^{(n)}$ and $c_{1(2)}^{(n)}$ are bounded in n for $i = 1, 2$. Note that this is a stronger version of the moment-matching conditions (5.2). From (5.4) the inverse of the covariance matrix is

$$V_n^{-1} = \frac{1}{1-\rho_n^2}\begin{pmatrix} \frac{1}{\sigma_{n,1}^2} & -\frac{\rho_n}{\sigma_{n,1}\sigma_{n,2}} \\ -\frac{\rho_n}{\sigma_{n,1}\sigma_{n,2}} & \frac{1}{\sigma_{n,2}^2} \end{pmatrix},$$

and, therefore,

$$\begin{aligned}
\phi_{0,V_n}(u) &= \frac{1}{2\pi}\frac{1}{\sigma_{n,1}\sigma_{n,2}\sqrt{1-\rho_n^2}}e^{-\frac{1}{2}\left\langle u, V_n^{-1}u\right\rangle} \\
&= \frac{1}{2\pi}\frac{1}{\sigma_{n,1}\sigma_{n,2}\sqrt{1-\rho_n^2}}e^{-\frac{1}{2\left(1-\rho_n^2\right)}\left(\frac{u_1^2}{\sigma_{n,1}^2} - \frac{2\rho_n u_1 u_2}{\sigma_{n,1}\sigma_{n,2}} + \frac{u_2^2}{\sigma_{n,2}^2}\right)}.
\end{aligned}$$

Let us first consider the derivatives of the bivariate normal distribution.

$$\begin{aligned} D_1\Phi_{0,V_n}(y_1,y_2) &= \frac{\partial}{\partial y_1}\int_{-\infty}^{y_1}\int_{-\infty}^{y_2}\frac{1}{2\pi}\frac{1}{\sigma_{n,1}\sigma_{n,2}\sqrt{1-\rho_n^2}}e^{-\frac{1}{2\sigma_{n,1}^2}u_1^2}e^{-\frac{1}{2(1-\rho_n^2)}\left(\frac{u_2}{\sigma_{n,2}}-\rho_n\frac{u_1}{\sigma_{n,1}}\right)^2}du_2du_1 \\ &= \frac{1}{2\pi}\frac{1}{\sigma_{n,1}\sigma_{n,2}\sqrt{1-\rho_n^2}}e^{-\frac{1}{2\sigma_{n,1}^2}y_1^2}\int_{-\infty}^{y_2}e^{-\frac{1}{2(1-\rho_n^2)}\left(\frac{u_2}{\sigma_{n,2}}-\rho_n\frac{y_1}{\sigma_{n,1}}\right)^2}du_2 \\ &= \frac{1}{\sigma_{n,1}}\phi\left(\frac{y_1}{\sigma_{n,1}}\right)\Phi\left(\frac{\frac{y_2}{\sigma_{n,2}}-\rho_n\frac{y_1}{\sigma_{n,1}}}{\sqrt{1-\rho_n^2}}\right), \quad y\in\mathbb{R}^2. \end{aligned}$$

We can now use the one-dimensional results (4.13) and (4.14) to get the expansions of the standard normal distribution and density functions. Due to (5.6) we have

$$\begin{aligned} \frac{y_{n,i}}{\sigma_{n,i}} &= \frac{d_{2,i}}{2\sigma_{n,i}}+\frac{\left(r-\frac{1}{2}\sigma_i^2-\mu_i(n)\right)T}{\sigma_i\sqrt{T}}\frac{1}{2\sigma_{n,i}} \\ &= d_{2,i}+h_{i,1}^{(n)}\frac{1}{\sqrt{n}}+h_{i,2}^{(n)}\frac{1}{n}+O\left(n^{-3/2}\right)\ , i=1,2, \end{aligned} \tag{5.7}$$

where

$$\begin{aligned} h_{i,1}^{(n)} &:= -2d_{2,i}b_{i,1}^{(n)}-a_{i,1}^{(n)}\frac{\sqrt{T}}{\sigma_i} \\ h_{i,2}^{(n)} &:= 2d_{2,i}\left(3\left(b_{i,1}^{(n)}\right)^2-b_{i,2}^{(n)}\right)+\left(2b_{i,1}^{(n)}a_{i,1}^{(n)}-a_{i,2}^{(n)}\right)\frac{\sqrt{T}}{\sigma_i}. \end{aligned}$$

Also, by applying the Binomial Series Theorem to (5.6) we get

$$\frac{1}{\sqrt{1-\rho_n^2}} = \frac{1}{\sqrt{1-\rho^2}}+\tilde{c}_1^{(n)}\frac{1}{\sqrt{n}}+\tilde{c}_2^{(n)}\frac{1}{n}+O\left(n^{-3/2}\right), \tag{5.8}$$

where

$$\begin{aligned} \tilde{c}_1^{(n)} &= \frac{\rho}{\sqrt{1-\rho^2}^3}c_1^{(n)}, \\ \tilde{c}_2^{(n)} &= \frac{1}{2}\left(\frac{\left(c_1^{(n)}\right)^2+2\rho c_2^{(n)}}{\sqrt{1-\rho^2}^3}+\frac{3\rho^2\left(c_1^{(n)}\right)^2}{\sqrt{1-\rho^2}^5}\right), \end{aligned}$$

and, therefore,

$$\frac{\frac{y_{n,i}}{\sigma_{n,i}}-\rho_n\frac{y_{n,j}}{\sigma_{n,j}}}{\sqrt{1-\rho_n^2}} = \frac{d_{2,i}-\rho d_{2,j}}{\sqrt{1-\rho^2}}+k_{i,1}^{(n)}\frac{1}{\sqrt{n}}+k_{i,2}^{(n)}\frac{1}{n}+O(n^{-3/2}), \tag{5.9}$$

with

$$k_{i,1}^{(n)} := \frac{h_{i,1} - \rho h_{j,1} - d_{2,j}c_1}{\sqrt{1-\rho^2}} + \tilde{c}_1 \left(d_{2,i} - \rho d_{2,j}\right)$$
$$k_{i,2}^{(n)} := \frac{h_{i,2} - \rho h_{j,2} - h_{j,1}c_1 - d_{2,j}c_2}{\sqrt{1-\rho^2}} + \tilde{c}_1 \left(h_{i,1} - \rho h_{j,1} - d_{2,j}c_1\right) + \tilde{c}_2 \left(d_{2,i} - \rho d_{2,j}\right)$$

We now substitute (5.7) and (5.9) into (4.14) and (4.13) to get

$$D_i \Phi_{0,V_n}\left(y_{n,1}, y_{n,2}\right) = 2\phi\left(d_i\right) \Phi\left(\frac{d_j - \rho d_i}{\sqrt{1-\rho^2}}\right) + \frac{1}{\sqrt{n}} l_{i,1}^{(n)} + \frac{1}{n} l_{i,2}^{(n)} + O(n^{-3/2}), \tag{5.10}$$

with

$$l_{i,1}^{(n)} := 2k_{j,1}\left(n\right) \phi\left(d_i\right) \phi\left(\frac{d_{2,j} - \rho d_{2,i}}{\sqrt{1-\rho^2}}\right) - \left(2h_{i,1}d_{2,i} + 4b_{i,1}\right) \phi\left(d_i\right) \Phi\left(\frac{d_{2,j} - \rho d_{2,i}}{\sqrt{1-\rho^2}}\right)$$
$$l_{i,2}^{(n)} := 2\left(k_{j,2} - \frac{1}{2}\frac{d_{2,j} - \rho d_{2,i}}{\sqrt{1-\rho^2}} k_{j,1}^2 - h_{i,1}d_i k_{j,1} - 2b_{i,1}k_{j,1}\right) \phi\left(d_i\right) \phi\left(\frac{d_{2,j} - \rho d_{2,i}}{\sqrt{1-\rho^2}}\right)$$
$$+ \left(\left(d_{2,i}^2 - 1\right) h_{i,1}^2 - 2d_{2,i}h_{i,2} - 4b_{i,1}h_{i,1}d_{2,i} + 4\left(3b_{i,1}^2 - b_{i,2}\right)\right) \phi\left(d_i\right) \Phi\left(\frac{d_{2,j} - \rho d_{2,i}}{\sqrt{1-\rho^2}}\right).$$

In a similar manner one can obtain asymptotic expansions for the following derivatives. Set $\alpha_i := \alpha_i\left(y_{n,i}\right) = y_{n,i}/\sigma_{n,i}$, $\beta_{i,j} := \beta_{i,j}\left(y_{n,i}, y_{n,j}\right) = \left(\frac{y_{n,i}}{\sigma_{n,i}} - \rho_n \frac{y_{n,j}}{\sigma_{n,j}}\right) / \sqrt{1-\rho_n^2}$, $i, j = 1, 2$.

$$\begin{aligned} D_1 D_2 \Phi_{0,V_n}\left(y_{n,1}, y_{n,2}\right) &= \phi_{0,V_n}\left(y_{n,1}, y_{n,2}\right) \\ &= \frac{1}{\sigma_{n,1}\sigma_{n,2}\sqrt{1-\rho_n^2}} \phi\left(\alpha_1\right) \phi\left(\beta_{2,1}\right) \\ &= \frac{1}{\sigma_{n,1}\sigma_{n,2}\sqrt{1-\rho_n^2}} \phi\left(\alpha_2\right) \phi\left(\beta_{1,2}\right) \end{aligned}$$

$$\begin{aligned} D_1^2 \Phi_{0,V_n}\left(y_{n,1}, y_{n,2}\right) &= \frac{\partial^2}{\partial y_{n,1}^2} \Phi_{0,V_n}\left(y_{n,1}, y_{n,2}\right) \\ &= -\frac{y_1}{\sigma_{n,1}^3} \phi\left(\alpha_1\right) \Phi\left(\beta_{2,1}\right) - \frac{\rho_n}{\sigma_{n,1}^2 \sqrt{1-\rho_n^2}} \phi\left(\alpha_1\right) \phi\left(\beta_{2,1}\right) \\ &= -\frac{1}{\sigma_{n,1}^2} \phi\left(\alpha_1\right) \left(\frac{y_{n,1}}{\sigma_{n,1}} \Phi\left(\beta_{2,1}\right) + \frac{\rho_n}{\sqrt{1-\rho_n^2}} \phi\left(\beta_{2,1}\right)\right) \end{aligned}$$

$$D_2^2 \Phi_{0,V_n}\left(y_{n,1}, y_{n,2}\right) = -\frac{1}{\sigma_{n,2}^2} \phi\left(\alpha_2\right) \left(\frac{y_{n,2}}{\sigma_{n,2}} \Phi\left(\beta_{1,2}\right) + \frac{\rho_n}{\sqrt{1-\rho_n^2}} \phi\left(\beta_{1,2}\right)\right)$$

Finally, we consider $\Phi_{0,V_n}(y_n)$.

$$\Phi_{0,V_n}(y_{n,1},y_{n,2}) = \int_{-\infty}^{y_{n,1}}\int_{-\infty}^{y_{n,2}} \frac{1}{2\pi}\frac{1}{\sigma_{n,1}\sigma_{n,2}\sqrt{1-\rho_n^2}} e^{-\frac{1}{2(1-\rho_n^2)}\left(\frac{u_1^2}{\sigma_{n,1}^2}-\frac{2\rho_n u_1 u_2}{\sigma_{n,1}\sigma_{n,2}}+\frac{u_2^2}{\sigma_{n,2}^2}\right)} du_2 du_1$$
$$= \frac{1}{2\pi}\frac{1}{\sqrt{1-\rho_n^2}} \int_{-\infty}^{y_{n,1}/\sigma_{n,1}}\int_{-\infty}^{y_{n,2}/\sigma_{n,2}} e^{-\frac{1}{2(1-\rho_n^2)}\left(v_1^2-2\rho_n v_1 v_2+v_2^2\right)} dv_2 dv_1$$

Set $\hat{h}_{n,i} = h_{i,1}^{(n)}\frac{1}{\sqrt{n}} + h_{i,2}^{(n)}\frac{1}{n} + O\left(n^{-3/2}\right)$, then by (5.7)

$$\Phi_{0,V_n}(y_{n,1},y_{n,2}) = \Phi_{0,V_{\rho_n}}\left(d_{2,1}+\hat{h}_{n,1}, d_{2,2}+\hat{h}_{n,2}\right), \quad V_{\rho_n} = \begin{pmatrix} 1 & \rho_n \\ \rho_n & 1 \end{pmatrix}.$$

Therefore, by applying Taylor's theorem for multivariate functions at the point $(\rho, d_{2,1}, d_{2,2})$ and using the relation $\frac{\partial \Phi_{0,V}(x_1,x_2)}{\partial \rho} = \phi_{0,V}(x_1,x_2)$ (see e.g. [Pl54]) we get

$$\begin{aligned}
&\Phi_{0,V_n}(y_{n,1},y_{n,2}) = \sum_{|\alpha|\le 2} \frac{D^\alpha \Phi_{0,V}(d_{2,1},d_{2,2})}{\alpha!}\left(\rho_n-\rho, \hat{h}_{n,1}, \hat{h}_{n,2}\right)^\alpha + O\left(n^{-3/2}\right) \\
&= \Phi_{0,V}(d_{2,1},d_{2,2}) + \frac{1}{\sqrt{n}}\left(\phi_{0,V}(d_{2,1},d_{2,2})\, c_1^{(n)} + D_1\Phi_{0,V}(d_{2,1},d_{2,2})\, h_{1,1}^{(n)} + D_2\Phi_{0,V}(d_{2,1},d_{2,2})\, h_{2,1}^{(n)}\right) \\
&+ \frac{1}{n}\left(\phi_{0,V}(d_{2,1},d_{2,2})\, c_2^{(n)} + \frac{1}{2}\frac{\partial \phi_{0,V}(d_{2,1},d_{2,2})}{\partial \rho}\left(c_1^{(n)}\right)^2 + D_1 D_2 \Phi_{0,V}(d_{2,1},d_{2,2})\, h_{1,1}^{(n)} h_{2,1}^{(n)}\right. \\
&+ D_1\Phi_{0,V}(d_{2,1},d_{2,2})\, h_{1,2}^{(n)} + \frac{1}{2}D_1^2\Phi_{0,V}(d_{2,1},d_{2,2})\left(h_{1,1}^{(n)}\right)^2 + D_1\phi_{0,V}(d_{2,1},d_{2,2})\, h_{1,1}^{(n)} c_1^{(n)} \\
&\left.+ D_2\phi_{0,V}(d_{2,1},d_{2,2})\, h_{2,1}^{(n)} c_1^{(n)} + D_2\Phi_{0,V}(d_{2,1},d_{2,2})\, h_{2,2}^{(n)} + \frac{1}{2}D_2^2\Phi_{0,V}(d_{2,1},d_{2,2})\left(h_{2,1}^{(n)}\right)^2\right) \\
&+ O\left(n^{-3/2}\right).
\end{aligned} \tag{5.11}$$

Existing methods in literature

As an example of the application of corollary 5.1.1 we present the expansions for the 2D-RB and the BEG models.

Proposition 5.1.2 (Dynamics of the 2D-RB model)**.** *The 2D-RB process (2.12)-(2.13) satisfies*

$$\begin{aligned}
P\left(S_n^{(n)} \le x\right) &= \Phi_{0,V}(d_{2,1},d_{2,2}) \\
&- \frac{1}{\sqrt{n}}\left(2S_1(a_{n,1})\,\phi(d_{2,1})\,\Phi\left(\frac{d_{2,2}-\rho d_{2,1}}{\sqrt{1-\rho^2}}\right) + 2S_1(a_{n,2})\,\phi(d_{2,2})\,\Phi\left(\frac{d_{2,1}-\rho d_{2,2}}{\sqrt{1-\rho^2}}\right)\right) \\
&+ \frac{1}{n} f\left(d_{2,1}, d_{2,2}, S_1(a_{n,1}), S_1(a_{n,2})\right) + O\left(\frac{1}{\sqrt{n}^3}\right),
\end{aligned}$$

where

$$a_{n,i} = \frac{\ln\left(\frac{x_i}{s_{0,i}}\right) - \left(r - \frac{1}{2}\sigma_i^2\right)T}{2\sigma_i\sqrt{\Delta t}} + \frac{n}{2},$$

and

$$\begin{aligned} f\left(d_{2,1}, d_{2,2}, S_1\left(a_{n,1}\right), S_1\left(a_{n,2}\right)\right) = {} & \phi\left(d_{2,1}\right)\Phi\left(\frac{d_{2,2} - \rho d_{2,1}}{\sqrt{1-\rho^2}}\right)\left(\frac{1}{12}\left(d_{2,1}^3 - d_{2,1}\right) - 2S_1^2\left(a_{n,1}\right)d_{2,1}\right) \\ & + \phi\left(d_{2,2}\right)\Phi\left(\frac{d_{2,1} - \rho d_{2,2}}{\sqrt{1-\rho^2}}\right)\left(\frac{1}{12}\left(d_{2,2}^3 - d_{2,2}\right) - 2S_1^2\left(a_{n,2}\right)d_{2,2}\right) \\ & + \phi_{0,V}(d_{2,1}, d_{2,2})\left(4S_1\left(a_{n,1}\right)S_1\left(a_{n,2}\right) - 2\rho S_1^2\left(a_{n,1}\right) - 2\rho S_1^2\left(a_{n,2}\right)\right. \\ & \left. + \frac{\rho}{2} - \frac{\rho(1+\rho^2)(d_{2,1}^2 + d_{2,2}^2)}{12(1-\rho^2)} + \frac{\rho^2 d_{2,1}d_{2,2}}{3(1-\rho^2)}\right). \end{aligned} \tag{5.12}$$

Proof. We assume the notation of Corollary 5.1.1. The moments (5.2) are matched exactly by the 2D-RB model,

$$\begin{aligned} \mu_n = \left(\frac{1}{2}, \frac{1}{2}\right), \quad & V_n = \frac{1}{4}\begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}, \\ \kappa_{n,\nu} &= 0, \quad |\nu| = 3, \\ \kappa_{n,(4,0)} = \kappa_{n,(0,4)} &= -\frac{1}{8}, \\ \kappa_{n,(3,1)} = \kappa_{n,(1,3)} &= -\frac{1}{8}\rho, \\ \kappa_{n,(2,2)} &= -\frac{1}{8}\rho^2, \end{aligned} \tag{5.13}$$

and

$$\begin{aligned} D^{(2,2)}\Phi_{0,V_n}\left(y_{n,1}, y_{n,2}\right) &= \phi_{0,V}(d_1, d_2)\frac{16\left(\left(1+\rho^2\right)d_1 d_2 - \rho\left(d_1^2 + d_2^2\right) + \rho\left(1-\rho^2\right)\right)}{\left(1-\rho^2\right)^2}, \\ D^{(3,1)}\Phi_{0,V_n}\left(y_{n,1}, y_{n,2}\right) &= -\phi_{0,V}(d_1, d_2)\frac{16\left(1 - \rho^2 - d_1^2 - \rho^2 d_2^2 + 2\rho d_1 d_2\right)}{\left(1-\rho^2\right)^2}, \\ D^{(4,0)}\Phi_{0,V_n}\left(y_{n,1}, y_{n,2}\right) &= -16\rho\phi_{0,V}(d_1, d_2)\left(3d_1^2 - \frac{3\rho d_1(d_2 - \rho d_1)}{1-\rho^2} - 3 - \frac{\rho^2}{1-\rho^2} + \frac{\rho^2\left(d_2 - \rho d_1\right)^2}{\left(1-\rho^2\right)^2}\right) \\ & \quad - 16\phi\left(d_1\right)\Phi\left(\frac{d_2 - \rho d_2}{\sqrt{1-\rho^2}}\right)\left(d_1^3 - 3d_1\right). \end{aligned} \tag{5.14}$$

The expressions for $D^{(1,3)}$ and $D^{(0,4)}$ are completely symmetric. Substituting these values into (5.5) we get the statement of the proposition. □

Proposition 5.1.3 (Dynamics of the BEG model)**.** *The BEG process (2.9)-(2.10) satisfies*

$$\begin{aligned} P\left(S_n^{(n)} \le x\right) &= \Phi_{0,V}\left(d_{2,1}, d_{2,2}\right) \\ &\quad - \frac{1}{\sqrt{n}}\left(2S_1\left(a_{n,1}\right)\phi\left(d_{2,1}\right)\Phi\left(\frac{d_{2,2}-\rho d_{2,1}}{\sqrt{1-\rho^2}}\right) + 2S_1\left(a_{n,2}\right)\phi\left(d_{2,2}\right)\Phi\left(\frac{d_{2,1}-\rho d_{2,2}}{\sqrt{1-\rho^2}}\right)\right) \\ &\quad + \frac{1}{n}g\left(d_{2,1}, d_{2,2}, S_1\left(a_{n,1}\right), S_1\left(a_{n,2}\right)\right) + O\left(\frac{1}{\sqrt{n}^3}\right), \end{aligned}$$

where

$$a_{n,i} = \frac{\ln\left(\frac{x_i}{s_{0,i}}\right)}{2\sigma_i\sqrt{\Delta t}} + \frac{n}{2},$$

and

$$\begin{aligned} & g\left(d_{2,1}, d_{2,2}, S_1\left(\alpha_1\right), S_1\left(\alpha_2\right)\right) \\ & = \phi\left(d_{2,1}\right)\Phi\left(\frac{d_{2,2}-\rho d_{2,1}}{\sqrt{1-\rho^2}}\right)\left(h_{1,2}^{(n)} + \frac{r-\frac{1}{2}\sigma_1^2}{3\sigma_1}(d_{2,1}^2-1) + \frac{1}{12}\left(d_{2,1}^3-d_{2,1}\right) - 2S_1^2\left(a_{n,1}\right)d_{2,1}\right) \\ & + \phi\left(d_{2,2}\right)\Phi\left(\frac{d_{2,1}-\rho d_{2,2}}{\sqrt{1-\rho^2}}\right)\left(h_{2,2}^{(n)} + \frac{r-\frac{1}{2}\sigma_2^2}{3\sigma_2}(d_{2,2}^2-1) + \frac{1}{12}\left(d_{2,2}^3-d_{2,2}\right) - 2S_1^2\left(a_{n,2}\right)d_{2,2}\right) \\ & + \phi_{0,V}(d_{2,1}, d_{2,2})\left(c_2^{(n)} + 4S_1\left(a_{n,1}\right)S_1\left(a_{n,2}\right) - 2\rho S_1^2\left(a_{n,1}\right) - 2\rho S_1^2\left(a_{n,2}\right) + \frac{\rho}{2} + \frac{\rho^2 d_{2,1}d_{2,2}}{3(1-\rho^2)}\right. \\ & \left. -\frac{\rho(1+\rho^2)(d_{2,1}^2+d_{2,2}^2)}{12(1-\rho^2)} + \frac{\rho\left(r-\frac{1}{2}\sigma_1^2\right)\left(2\rho d_{2,2}-d_{2,1}\left(1+\rho^2\right)\right)}{3\sigma_1(1-\rho^2)} + \frac{\rho\left(r-\frac{1}{2}\sigma_2^2\right)\left(2\rho d_{2,1}-d_{2,2}\left(1+\rho^2\right)\right)}{3\sigma_2(1-\rho^2)}\right). \end{aligned}$$

with

$$\begin{aligned} c_2^{(n)} &= \rho\frac{\left(r-\frac{1}{2}\sigma_1^2\right)^2}{2\sigma_1^2}T + \rho\frac{\left(r-\frac{1}{2}\sigma_2^2\right)^2}{2\sigma_2^2}T - \frac{\left(r-\frac{1}{2}\sigma_1^2\right)\left(r-\frac{1}{2}\sigma_2^2\right)}{\sigma_1\sigma_2}T, \\ h_{i,2}^{(n)} &= \frac{d_i\left(r-\frac{1}{2}\sigma_i^2\right)^2T}{2\sigma_i^2}, \quad i=1,2. \end{aligned} \tag{5.15}$$

Proof. We assume the notation of Corollary 5.1.1. For the BEG model the coefficients in (5.6) are

$$\begin{aligned} & a_{i,j}^{(n)} = 0, \quad i=1,2,\ j=1,2, \\ & b_{i,1}^{(n)} = 0, \quad b_{i,2}^{(n)} = -\frac{\left(r-\frac{1}{2}\sigma_i^2\right)^2}{4\sigma_i^2}T, \quad i=1,2, \\ & c_1^{(n)} = 0, \quad \tilde{c}_1^{(n)} = 0, \quad h_{i,1}^{(n)} = 0, \end{aligned}$$

and $c_2^{(n)}$, $h_{i,2}^{(n)}$, $i=1,2$ are defined as in equation (5.15). Substituting these values into (5.11) we get

$$\begin{aligned}\Phi_{0,V_n}(y_{n,1},y_{n,2}) &= \Phi_{0,V}(d_{2,1},d_{2,2})\\ &+\frac{1}{n}\left(\phi(d_{2,1})\,\Phi\left(\frac{d_{2,2}-\rho d_{2,1}}{\sqrt{1-\rho^2}}\right)h_{1,2}^{(n)}+\phi(d_{2,2})\,\Phi\left(\frac{d_{2,1}-\rho d_{2,2}}{\sqrt{1-\rho^2}}\right)h_{2,2}^{(n)}\right.\\ &\left.+\phi_{0,V}(d_{2,1},d_{2,2})\,c_2^{(n)}\right)+O\left(n^{-3/2}\right)\end{aligned}$$

The cumulants for $|\nu|=3$ are

$$\begin{aligned}\kappa_{n,(3,0)} &= -\frac{1}{4}\frac{\left(r-\frac{1}{2}\sigma_1^2\right)\sqrt{\Delta t}}{\sigma_1}+O\left(\frac{1}{n}\right),\\ \kappa_{n,(2,1)} &= -\frac{1}{4}\rho\frac{\left(r-\frac{1}{2}\sigma_1^2\right)\sqrt{\Delta t}}{\sigma_1}+O\left(\frac{1}{n}\right),\end{aligned}$$

and $\kappa_{n,(0,3)}$, $\kappa_{n,(1,2)}$ are completely analogous.

$$\begin{aligned}D^{(3,0)}\Phi_{0,V_n}(y_{n,1},y_{n,2}) &= 8\phi(d_{2,1})\,\Phi\left(\frac{d_{2,2}-\rho d_{2,1}}{\sqrt{1-\rho^2}}\right)\left(d_{2,1}^2-1\right)\\ &\quad+8\rho\phi_{0,V}(d_{2,1},d_{2,2})\left(2d_{2,1}-\frac{\rho}{1-\rho^2}(d_{2,2}-\rho d_{2,1})\right)+O\left(\frac{1}{\sqrt{n}}\right)\\ D^{(2,1)}\Phi_{0,V_n}(y_{n,1},y_{n,2}) &= -8\phi_{0,V}(d_{2,1},d_{2,2})\frac{d_{2,1}-\rho d_{2,2}}{1-\rho^2}+O\left(\frac{1}{\sqrt{n}}\right),\end{aligned} \tag{5.16}$$

and $D^{(0,3)}$, $D^{(1,2)}$ are symmetric. For $|\nu|=4$ the cumulants are as in (5.13) up to order $O\left(\frac{1}{\sqrt{n}}\right)$, and the derivatives $D^{\nu}\Phi_{0,V_n}$ are as in (5.14) up to order $O\left(\frac{1}{\sqrt{n}}\right)$. Substituting the obtained expansions into equation (5.5) we get the statement of the proposition. □

5.1.2 Improving the convergence behavior

As in the one-dimensional case we now apply Corollary 5.1.1 to construct an advanced 2D tree with an $O\left(n^{-3/2}\right)$ order of convergence. We will refer to this tree as the 2D 1-Optimal tree.

Consider the following binomial model

$$S_k^{(n)}=\begin{pmatrix} S_{k-1,1}^{(n)}e^{\alpha_{n,1}\Delta t+\sigma_1\sqrt{\Delta t}\xi_{k,1}^{(n)}}\\ S_{k-1,2}^{(n)}e^{\alpha_{n,2}\Delta t+\sigma_2\sqrt{\Delta t}\xi_{k,2}^{(n)}}\end{pmatrix},\ k=1,\dots,n, \tag{5.17}$$

where

$$\alpha_{n,i}=r-\frac{1}{2}\sigma_i^2-\frac{c_{i,n}\sigma_i}{\sqrt{T}}\frac{1}{\sqrt{n}},\quad i=1,2 \tag{5.18}$$

and

$$\begin{aligned} p_n^{(uu)} &= \frac{1}{4}(1+\rho) + \frac{c_{1,n}+c_{2,n}+c_{3,n}}{4}\frac{1}{n} + \frac{c_{4,n}+c_{5,n}}{4}\frac{1}{\sqrt{n}^3}, \\ p_n^{(ud)} &= \frac{1}{4}(1-\rho) + \frac{c_{1,n}-c_{2,n}-c_{3,n}}{4}\frac{1}{n} + \frac{c_{4,n}-c_{5,n}}{4}\frac{1}{\sqrt{n}^3}, \\ p_n^{(du)} &= \frac{1}{4}(1-\rho) - \frac{c_{1,n}-c_{2,n}+c_{3,n}}{4}\frac{1}{n} - \frac{c_{4,n}-c_{5,n}}{4}\frac{1}{\sqrt{n}^3}, \\ p_n^{(dd)} &= \frac{1}{4}(1+\rho) - \frac{c_{1,n}+c_{2,n}-c_{3,n}}{4}\frac{1}{n} - \frac{c_{4,n}+c_{5,n}}{4}\frac{1}{\sqrt{n}^3}, \end{aligned} \tag{5.19}$$

for bounded sequences $\{c_{i,n}\}_n$, $i = 1, \dots, 5$. These transition probabilities are well-defined, i.e. they sum up to 1 and lie between 0 and 1 for a sufficiently large n. Moreover, with the above choice of parameters (5.6) becomes

$$\mu_i(n) = r - \frac{1}{2}\sigma_i^2 + \frac{\sigma_i c_{i+3,n}}{\sqrt{T}}\frac{1}{n}, \quad i = 1, 2,$$

$$\sigma_{n,i}^2 = \frac{1}{4} + O\left(\frac{1}{n^2}\right), \quad \rho_n = \rho + c_{3,n}\frac{1}{n} + O\left(\frac{1}{n^2}\right).$$

Therefore, the moment-matching conditions are satisfied and $S^{(n)}$ converges weakly to the continuous-time process S.

Proposition 5.1.4. *With the appropriate choice of the coefficients $c_{i,n}$, $i = 1, \dots, 5$, $S^{(n)}$ in (5.17) satisfies*

$$P\left(S_{n,1}^{(n)} \le x_1,\ S_{n,2}^{(n)} \le x_2\right) = \Phi_{0,V}(d_{2,1}, d_{2,2}) + O\left(\frac{1}{\sqrt{n}^3}\right).$$

Proof. In the above model, with y_n as in (5.3),

$$\frac{y_{n,i}}{\sigma_{n,i}} = d_{2,i} - c_{i+3,n}\frac{1}{n} + O\left(\frac{1}{n^2}\right), \quad i = 1, 2,$$

and by (5.11) we get

$$\Phi_{0,V_n}(y_{n,1}, y_{n,2}) = \Phi_{0,V}(d_{2,1}, d_{2,2}) + \frac{1}{n}\left(\phi_{0,V}(d_{2,1}, d_{2,2})c_{3,n} - \phi(d_{2,1})\Phi\left(\frac{d_{2,2}-\rho d_{2,1}}{\sqrt{1-\rho^2}}\right)c_{4,n}\right.$$
$$\left. - \phi(d_{2,2})\Phi\left(\frac{d_{2,1}-\rho d_{2,2}}{\sqrt{1-\rho^2}}\right)c_{5,n}\right) + O\left(\frac{1}{n^2}\right).$$

Also, since

$$\frac{\frac{y_{n,i}}{\sigma_{n,i}} - \rho_n\frac{y_{n,j}}{\sigma_{n,j}}}{\sqrt{1-\rho_n^2}} = \frac{d_{2,i}-\rho d_{2,j}}{\sqrt{1-\rho^2}} + O\left(\frac{1}{n}\right),$$

by (5.10) we have

$$D_i\Phi_{0,V_n}\left(y_{n,1},y_{n,2}\right)=2\phi\left(d_{2,i}\right)\Phi\left(\frac{d_{2,j}-\rho d_{2,i}}{\sqrt{1-\rho^2}}\right)+O\left(\frac{1}{n}\right),\quad i,j=1,2,\ i\neq j.$$

$$D_i^2\Phi_{0,V_n}\left(y_{n,1},y_{n,2}\right)=-4d_{2,i}\phi\left(d_{2,i}\right)\Phi\left(\frac{d_{2,j}-\rho d_{2,i}}{\sqrt{1-\rho^2}}\right)-4\rho\phi_{0,V}\left(d_{2,1},d_{2,2}\right)+O\left(\frac{1}{n}\right),\quad i,j=1,2,\ i\neq j.$$

Since $\kappa_{n,\nu}=O\left(\frac{1}{n}\right)$, $|\nu|=3$,

$$P_1\left(-\Phi_{0,V_n},\{\kappa_{n,\nu}\}\right)(y_n)=O\left(\frac{1}{n}\right)$$
$$P_2\left(-\Phi_{0,V_n},\{\kappa_{n,\nu}\}\right)(y_n)=\sum_{|\nu|=4}\frac{\kappa_{n,\nu}}{\nu!}D^\nu\Phi_{0,V_n}\left(y_n\right)+O\left(\frac{1}{n}\right).$$

For $|\nu|=4$ the cumulants are as in (5.13) up to order $O\left(\frac{1}{n}\right)$, and the derivatives $D^\nu\Phi_{0,V_n}$ are as in (5.14) up to order $O\left(\frac{1}{n}\right)$. Substituting these results into (5.5) we get the following asymptotics

$$\begin{aligned}&P\left(S_{n,1}^{(n)}\leq x_1,\ S_{n,2}^{(n)}\leq x_2\right)=\Phi_{0,V}\left(d_{2,1},d_{2,2}\right)\\&-\frac{1}{\sqrt{n}}\left(2S_1\left(a_{n,1}\right)\phi\left(d_{2,1}\right)\Phi\left(\frac{d_{2,2}-\rho d_{2,1}}{\sqrt{1-\rho^2}}\right)+2S_1\left(a_{n,2}\right)\phi\left(d_{2,2}\right)\Phi\left(\frac{d_{2,1}-\rho d_{2,2}}{\sqrt{1-\rho^2}}\right)\right)\\&+\frac{1}{n}\left(\phi_{0,V}\left(d_{2,1},d_{2,2}\right)c_{3,n}-\phi\left(d_{2,1}\right)\Phi\left(\frac{d_{2,2}-\rho d_{2,1}}{\sqrt{1-\rho^2}}\right)c_{4,n}-\phi\left(d_{2,2}\right)\Phi\left(\frac{d_{2,1}-\rho d_{2,2}}{\sqrt{1-\rho^2}}\right)c_{5,n}\right.\\&\quad\left.+f\left(d_{2,1},d_{2,2},S_1\left(a_{n,1}\right),S_1\left(a_{n,2}\right)\right)\right)+O\left(\frac{1}{\sqrt{n}^3}\right)\end{aligned}\tag{5.20}$$

where

$$a_{n,i}=\frac{\ln\left(\frac{x_i}{s_{0,i}}\right)-\left(r-\frac{1}{2}\sigma_i^2\right)T+c_{i,n}\sigma_i\sqrt{\Delta t}}{2\sigma_i\sqrt{\Delta t}}+\frac{n}{2},\quad i=1,2,\tag{5.21}$$

and $f\left(d_1,d_2,S_1\left(a_{n,1}\right),S_1\left(a_{n,2}\right)\right)$ is defined as in (5.12). If we now set

$$c_{i,n}=1-2\left\{\frac{\ln\left(\frac{x_i}{s_{0,i}}\right)-\left(r-\frac{1}{2}\sigma_i^2\right)T}{2\sigma_i\sqrt{\Delta t}}+\frac{n}{2}\right\},\quad i=1,2,\tag{5.22}$$

so that $S_1(a_{n,i}) = 0$, $i = 1, 2$, and

$$\begin{aligned} c_{3,n} &= -\frac{\rho}{2} + \frac{\rho(1+\rho^2)(d_{2,1}^2 + d_{2,2}^2)}{12(1-\rho^2)} - \frac{\rho^2 d_{2,1} d_{2,2}}{3(1-\rho^2)}, \\ c_{4,n} &= \frac{1}{12}\left(d_{2,1}^3 - d_{2,1}\right), \\ c_{5,n} &= \frac{1}{12}\left(d_{2,2}^3 - d_{2,2}\right), \end{aligned} \tag{5.23}$$

then we have the statement of the proposition. □

Remark 5.1.5. Note that the choice of parameters in (5.18) and (5.19) is only one option. Other combinations of the coefficients $c_{i,n}$ are also possible. However, with the above choice we are able to avoid division by $\phi(d_{2,i})\,\Phi\left(\frac{d_{2,j}-\rho d_{2,i}}{\sqrt{1-\rho^2}}\right)$, which attain very small values for certain parameter settings, therefore, leading to unstable numerical results.

Remark 5.1.6. By choosing the drifts as in (5.18), we apply a generalized 2D version of the Chang-Palmer method. Indeed, with $c_{i,n}$, $i = 1, 2$ defined in (5.22), by applying (4.20) and (4.24) to each of the components we get that x_i coincides with the geometric average of the two neighboring nodes of $S_{n,i}^{(n)}$, or alternatively

$$\frac{\log\left(S_{n,i}^{(n)}(l_{n,i})\right) + \log\left(S_{n,i}^{(n)}(l_{n,i}+1)\right)}{2} = \log(x_i), \tag{5.24}$$

where, as usual, $l_{n,i}$ is the largest number of up-moves, s.t. $S_{n,i}^{(n)}(l_{n,i}) \leq x_i < S_{n,i}^{(n)}(l_{n,i}+1)$, $i = 1, 2$. Note that due to the rectangular grid, (5.24) holds for each component independent of the other one, which means that the relative position of x with respect to two rows of nodes will be fixed.

Remark 5.1.7. Here we have considered a generalization of the 2D-RB model, however, similar results can also be obtained for the BEG model.

We now compare the distributional fit of the 2D-RB model, the decoupling method presented in [KM09] and the 2D 1-Optimal tree. To get a better idea of the general convergence behavior of the methods we consider the RMSR error. We use the same sampling procedure as in the one-dimensional case, only now we simulate two σ and two x parameters and we additionally need the correlation coefficient ρ which we choose to be uniformly distributed between 0.1 and 0.9 with probability 0.75 and between -0.9 and -0.1 with probability 0.25. As in the one-dimensional case, we omit those parameters, for which $\Phi(d_{2,1}, d_{2,2}) \leq 10^{-6}$. Out of a sample of 200 simulated options 196 have been included.

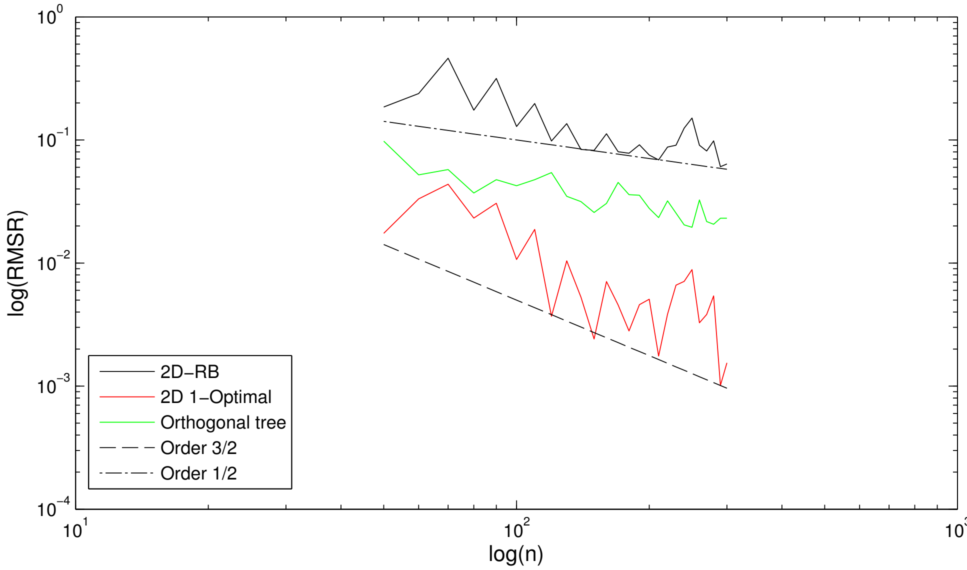


Figure 5.1: RMSR(n) error

In the figure above we can see the higher order of convergence of the 2D 1-Optimal model. The RMSR(n) error for both the 2D-RB and the decoupling approach has a slope of 1/2, whereas the 2D 1-Optimal model exhibits a 3/2 slope.

Remark 5.1.8. Note that, as in the one-dimensional case, the 2D 1-Optimal model requires an additional calculation of the coefficients $c_{i,n}$, $i = 1, \ldots, 5$ for each $n \in \mathbb{N}$. However, this only contributes $O(1)$ to the total computational effort $O(n^3)$.

5.1.3 Expansions for two-asset options

We now consider two types of two-asset options and apply the above results to obtain a better asymptotic behavior of their prices.

Two-asset digital options

A two-asset digital call (put) pays out a fixed amount G if the values of both underlyings are above (below) given strikes K_1, K_2. As usual, we will only consider the put option here, however, similar results can be easily obtained for the call as well. The payout function of the two-asset digital put is

$$G\mathbb{1}_{\{S_1(T)\leq K_1,\, S_2(T)\leq K_2\}},$$

and the price in the Black-Scholes model is

$$V = Ge^{-rT}\Phi_{0,V}\left(d_{2,1}, d_{2,2}\right),$$

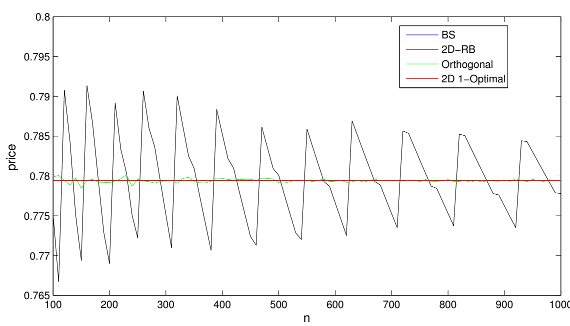
where $d_{2,i}$, $i = 1, 2$ and the covariance matrix V are defined as in subsection 5.1.1.

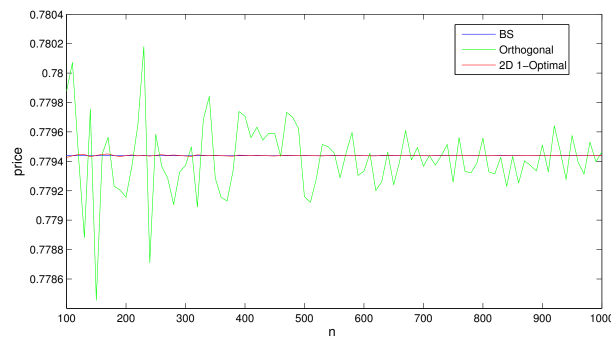
Proposition 5.1.9. *Let $V^{(n)}$ be the price of the two-asset digital put in the discrete model (5.17). If the coefficients $c_{i,n}$, $i = 1, \ldots, 5$ satisfy (5.22)-(5.23) then*

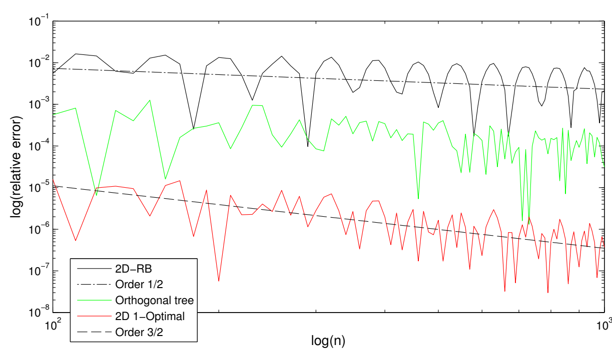
$$V^{(n)} = V + O\left(\frac{1}{\sqrt{n}^3}\right), \quad n \to \infty.$$

Proof. Substituting (K_1, K_2) for (x_1, x_2), the statement is a direct consequence of Proposition 5.1.4. $\square$

Compare the following convergence results for a two-asset digital put with $s_0 = [12, 12]$, $r = 0.1$, $x = [20, 17]$, $\sigma = [0.2, 0.25]$, $\rho = 0.5$, $T = 1$. The first two figures show the convergence pattern of the 2D-RB, the 2D 1-Optimal and the orthogonal trees, while the third one is a log-log plot and gives the rate of convergence.







Note that the 2D 1-Optimal model also has oscillatory convergence behavior, however, these oscillations are much smaller compared to the 2D-RB and the orthogonal trees.

The 2D 1-Optimal model matches the first 5 decimal points (see table below) with as few as 200 time steps, whereas the decoupling approach requires 900 steps to match the first four decimal points. The 2D-RB tree has highly irregular convergence and doesn't provide accurate enough results even with 1000 time steps.

Parameters	n	2D RB tree	Orthogonal tree	2D 1-Optimal	BS value
$s_0 = [12, 12]$	100	0.7751250	0.779877	0.7794261	0.7794385
$r = 0.1$	200	0.7690237	0.7791561	0.7794384	
$x = [20, 17]$	300	0.7751430	0.7793718	0.7794362	
$\sigma = [0.2, 0.25]$	500	0.7800901	0.7791601	0.7794398	
$\rho = 0.5$	700	0.7752795	0.7793661	0.7794380	
$T = 1$	900	0.7762390	0.7795109	0.7794384	
	1000	0.7777951	0.7794628	0.7794388	

Two-asset correlation options

To demonstrate how the change of measure technique presented in subsection 4.3.2 works in higher dimensions we now consider the two-asset correlation put option. This option has the payout of a plain vanilla put on one asset, if the other underlying is below a certain strike at maturity, i.e. the payout function is

$$(S_2 (T) - K_2)^+ \mathbb{1}_{\{S_1(T) \le K_1\}}.$$

The Black-Scholes price is

$$V = K_2 e^{-rT} \Phi_{0,V} (d_{2,1}, d_{2,2}) - s_{0,2} \Phi_{0,V} (d_{1,1}, d_{1,2}),$$

where $d_{2,i}$, $i = 1, 2$ are as in subsection 5.1.1 and

$$d_{1,1} = d_{2,1} - \rho\sigma_2\sqrt{T}, \quad d_{1,2} = d_{2,2} - \sigma_2\sqrt{T}.$$

(See [H06], Chapter 5). The price $V^{(n)}$ in the model (5.17) is

$$\begin{aligned} V^{(n)} &= E_n \left(e^{-rT} \left(K_2 - s_{0,2} e^{\alpha_{n,2}T + \sigma_2\sqrt{\Delta t}\sum_{i=1}^n \xi_{i,2}^{(n)}} \right) \mathbb{1}_{\left\{S_{n,1}^{(n)} \le K_1,\, S_{n,2}^{(n)} \le K_2\right\}} \right) \\ &= e^{-rT} K_2 P_n \left(S_{n,1}^{(n)} \le K_1,\ S_{n,2}^{(n)} \le K_2 \right) - s_{0,2} E_n \left(e^{(\alpha_{n,2}-r)T + \sigma_2\sqrt{\Delta t}\sum_{i=1}^n \xi_{i,2}^{(n)}} \mathbb{1}_{\left\{S_{n,1}^{(n)} \le K_1,\, S_{n,2}^{(n)} \le K_2\right\}} \right). \end{aligned}$$

The first term can be directly calculated as in Proposition 5.1.4, whereas for the second term we will apply a change of measure as in subsection 4.3.2. Let

$$\lambda_{n,i} = e^{(\alpha_{n,2}-r)\Delta t + \sigma_2\sqrt{\Delta t}\xi_{i,2}^{(n)}},$$

and $M_n = E(\lambda_{n,1})$. Consider the change of measure

$$\tilde{p}_n^{(uu)} = \frac{p_n^{(uu)} e^{(\alpha_{n,2}-r)\Delta t + \sigma_2\sqrt{\Delta t}}}{M_n},$$
$$\tilde{p}_n^{(ud)} = \frac{p_n^{(ud)} e^{(\alpha_{n,2}-r)\Delta t - \sigma_2\sqrt{\Delta t}}}{M_n},$$
$$\tilde{p}_n^{(du)} = \frac{p_n^{(du)} e^{(\alpha_{n,2}-r)\Delta t + \sigma_2\sqrt{\Delta t}}}{M_n},$$
$$\tilde{p}_n^{(dd)} = \frac{p_n^{(dd)} e^{(\alpha_{n,2}-r)\Delta t - \sigma_2\sqrt{\Delta t}}}{M_n}.$$

These probabilities are well-defined, and the Radon-Nikodým derivative of the new measure $\tilde{P}_n$ with respect to P_n is given by

$$\Lambda_n = \prod_{i=1}^{n} \frac{\lambda_{n,i}}{M_n} = \frac{e^{(\alpha_{n,2}-r)T + \sigma_2\sqrt{\Delta t}\sum_{i=1}^{n}\xi_{i,2}^{(n)}}}{M_n^n}.$$

Therefore, by Theorem 4.3.3 we have

$$V^{(n)} = e^{-rT} K_2 P_n\left(S_{n,1}^{(n)} \le K_1,\ S_{n,2}^{(n)} \le K_2\right) - s_{0,2} M_n^n \tilde{P}_n\left(S_{n,1}^{(n)} \le K_1,\ S_{n,2}^{(n)} \le K_2\right). \quad (5.25)$$

Proposition 5.1.10. *With the appropriate choice of the coefficients $c_{i,n}$, $i = 1, \dots, 5$ the price of a two-asset correlation put in the model (5.17) satisfies*

$$V^{(n)} = V + O\left(\frac{1}{\sqrt{n}^3}\right), \quad n \to \infty.$$

Proof. Set $c_{3,n} = c_{4,n} = 0$. By (5.20) we have

$$\begin{aligned}
&e^{-rT} K_2 P_n\left(S_{n,1}^{(n)} \le K_1,\ S_{n,2}^{(n)} \le K_2\right) = e^{-rT} K_2 \Phi_{0,V}(d_{2,1}, d_{2,2}) \\
&- \frac{1}{\sqrt{n}} e^{-rT} K_2 \left(2S_1(a_{n,1})\phi(d_{2,1})\Phi\left(\frac{d_{2,2} - \rho d_{2,1}}{\sqrt{1-\rho^2}}\right) + 2S_1(a_{n,2})\phi(d_{2,2})\Phi\left(\frac{d_{2,1} - \rho d_{2,2}}{\sqrt{1-\rho^2}}\right)\right) \\
&- \frac{1}{n} e^{-rT} K_2 \left(\phi(d_{2,2})\Phi\left(\frac{d_{2,1} - \rho d_{2,2}}{\sqrt{1-\rho^2}}\right) c_{5,n} - f(d_{2,1}, d_{2,2}, S_1(a_{n,1}), S_1(a_{n,2}))\right) + O\left(\frac{1}{\sqrt{n}^3}\right),
\end{aligned}$$

with $f(d_1, d_2, S_1(a_{n,1}), S_1(a_{n,2}))$ defined as in (5.12). We now consider the second term in (5.25). Since $p_n^{(\cdot u)} = \frac{1}{2} + \frac{1}{n}\frac{c_{2,n}}{2} + \frac{1}{\sqrt{n}^3}\frac{c_{5,n}}{2}$, as in (4.52) we get

$$M_n = 1 + \frac{1}{n^2}\left(\sqrt{T}\sigma_2 c_{5,n} - \frac{\sigma_2^4 T^2}{12}\right) + O\left(\frac{1}{\sqrt{n}^5}\right),$$

and, therefore,

$$M_n^n = 1 + \frac{1}{n}\left(\sqrt{T}\sigma_2 c_{5,n} - \frac{\sigma_2^4 T^2}{12}\right) + O\left(\frac{1}{\sqrt{n}^3}\right). \tag{5.26}$$

Using equation (4.51) the new probabilities are

$$\begin{aligned}
\tilde{p}_n^{(uu)} &= \frac{1}{4}(1+\rho) + \frac{1}{4\sqrt{n}}\sigma_2\sqrt{T}(1+\rho) + \frac{1}{4n}(c_{1,n}+c_{2,n}) \\
&+ \frac{1}{4\sqrt{n}^3}\left(c_{5,n} + \sigma_2\sqrt{T}(c_{1,n}-\rho c_{2,n}) - (1+\rho)\frac{1}{3}\sigma_2^3\sqrt{T}^3\right), \\
\tilde{p}_n^{(ud)} &= \frac{1}{4}(1-\rho) - \frac{1}{4\sqrt{n}}\sigma_2\sqrt{T}(1-\rho) + \frac{1}{4n}(c_{1,n}-c_{2,n}) \\
&- \frac{1}{4\sqrt{n}^3}\left(c_{5,n} + \sigma_2\sqrt{T}(c_{1,n}-\rho c_{2,n}) - (1-\rho)\frac{1}{3}\sigma_2^3\sqrt{T}^3\right), \\
\tilde{p}_n^{(du)} &= \frac{1}{4}(1-\rho) + \frac{1}{4\sqrt{n}}\sigma_2\sqrt{T}(1-\rho) - \frac{1}{4n}(c_{1,n}-c_{2,n}) \\
&+ \frac{1}{4\sqrt{n}^3}\left(c_{5,n} - \sigma_2\sqrt{T}(c_{1,n}-\rho c_{2,n}) - (1-\rho)\frac{1}{3}\sigma_2^3\sqrt{T}^3\right), \\
\tilde{p}_n^{(dd)} &= \frac{1}{4}(1+\rho) - \frac{1}{4\sqrt{n}}\sigma_2\sqrt{T}(1+\rho) - \frac{1}{4n}(c_{1,n}+c_{2,n}) \\
&- \frac{1}{4\sqrt{n}^3}\left(c_{5,n} - \sigma_2\sqrt{T}(c_{1,n}-\rho c_{2,n}) - (1+\rho)\frac{1}{3}\sigma_2^3\sqrt{T}^3\right).
\end{aligned}$$

Then (5.6) becomes

$$\begin{aligned}
\tilde{\mu}_1(n) &= r - \frac{1}{2}\sigma_1^2 + \rho\sigma_1\sigma_2 - \frac{1}{3n}\rho\sigma_1\sigma_2^3 T + O\left(\frac{1}{\sqrt{n}^3}\right), \\
\tilde{\mu}_2(n) &= r + \frac{1}{2}\sigma_2^2 + \frac{1}{n}\frac{\sigma_2}{\sqrt{T}}\left(c_{5,n} - \frac{1}{3}\sigma_2^3\sqrt{T}^3\right) + O\left(\frac{1}{\sqrt{n}^3}\right),
\end{aligned}$$

$$\begin{aligned}
\tilde{\sigma}_1^2(n) &= \frac{1}{4} - \frac{1}{4n}\rho^2\sigma_2^2 T + O\left(\frac{1}{\sqrt{n}^3}\right), \\
\tilde{\sigma}_2^2(n) &= \frac{1}{4} - \frac{1}{4n}\sigma_2^2 T + O\left(\frac{1}{\sqrt{n}^3}\right), \\
\tilde{\rho}_n &= \rho + \frac{1}{2n}\rho(\rho^2-1)\sigma_2^2 T + O\left(\frac{1}{\sqrt{n}^3}\right),
\end{aligned}$$

and for y_n as in (5.3)

$$\frac{\tilde{y}_{n,1}}{\tilde{\sigma}_{n,1}} = d_{1,1} + \frac{1}{n}\left(\frac{1}{3}\rho\sigma_2^3\sqrt{T}^3 + \frac{1}{2}d_{1,1}\rho^2\sigma_2^2 T\right) + O\left(\frac{1}{\sqrt{n}^3}\right),$$

$$\frac{\tilde{y}_{n,2}}{\tilde{\sigma}_{n,2}} = d_{1,2} - \frac{1}{n}\left(c_{5,n} - \frac{1}{3}\sigma_2^3\sqrt{T}^3 - \frac{1}{2}d_{1,2}\sigma_2^2 T\right) + O\left(\frac{1}{\sqrt{n}^3}\right).$$

Therefore, by (5.11) we get

$$\begin{aligned}
&\Phi_{0,\tilde{V}_n}\left(\tilde{y}_{n,1}, \tilde{y}_{n,2}\right) = \Phi_{0,V}\left(d_{1,1}, d_{1,2}\right)\\
&+\frac{1}{n}\left(\phi_{0,V}\left(d_{1,1}, d_{1,2}\right)\frac{1}{2}\rho(\rho^2-1)\sigma_2^2 T + \phi\left(d_{1,1}\right)\Phi\left(\frac{d_{1,2}-\rho d_{1,1}}{\sqrt{1-\rho^2}}\right)\left(\frac{1}{3}\rho\sigma_2^3\sqrt{T}^3 + \frac{1}{2}d_{1,1}\rho^2\sigma_2^2 T\right)\right.\\
&\left.-\phi\left(d_{1,2}\right)\Phi\left(\frac{d_{1,1}-\rho d_{1,2}}{\sqrt{1-\rho^2}}\right)\left(c_{5,n} - \frac{1}{3}\sigma_2^3\sqrt{T}^3 - \frac{1}{2}d_{1,2}\sigma_2^2 T\right)\right) + O\left(\frac{1}{\sqrt{n}^3}\right).
\end{aligned}$$

The cumulants for $|\nu| = 3$ are

$$\kappa_{n,(3,0)} = -\frac{1}{4\sqrt{n}}\rho\sigma_2\sqrt{T} + O\left(\frac{1}{n}\right),$$

$$\kappa_{n,(0,3)} = -\frac{1}{4\sqrt{n}}\sigma_2\sqrt{T} + O\left(\frac{1}{n}\right),$$

$$\kappa_{n,(2,1)} = -\frac{1}{4\sqrt{n}}\rho^2\sigma_2\sqrt{T} + O\left(\frac{1}{n}\right),$$

$$\kappa_{n,(2,1)} = -\frac{1}{4\sqrt{n}}\rho\sigma_2\sqrt{T} + O\left(\frac{1}{n}\right),$$

For $|\nu| = 4$ the cumulants are as in (5.13) up to order $O\left(\frac{1}{\sqrt{n}}\right)$. Using equations (5.14), (5.16), (5.26) and Corollary 5.1.1 we get the following asymptotics

$$\begin{aligned}
&s_{0,2}M_n^n\tilde{P}_n\left(S_{n,1}^{(n)} \le K_1,\ S_{n,2}^{(n)} \le K_2\right) = s_{0,2}\Phi_{0,V}\left(d_{1,1}, d_{1,2}\right)\\
&-\frac{2s_{0,2}}{\sqrt{n}}\left(S_1\left(a_{n,1}\right)\phi\left(d_{1,1}\right)\Phi\left(\frac{d_{1,2}-\rho d_{1,1}}{\sqrt{1-\rho^2}}\right) + S_1\left(a_{n,2}\right)\phi\left(d_{1,2}\right)\Phi\left(\frac{d_{1,1}-\rho d_{1,2}}{\sqrt{1-\rho^2}}\right)\right)\\
&-\frac{s_{0,2}}{n}\left(\phi\left(d_{1,2}\right)\Phi\left(\frac{d_{1,1}-\rho d_{1,2}}{\sqrt{1-\rho^2}}\right)c_{5,n} - \Phi_{0,V}\left(d_{1,1}, d_{1,2}\right)\sigma_2\sqrt{T}c_{5,n} - \tilde{f}\left(d_{1,1}, d_{1,2}, S_1\left(a_{n,1}\right), S_1\left(a_{n,2}\right)\right)\right)\\
&+O\left(\frac{1}{\sqrt{n}^3}\right),
\end{aligned} \tag{5.27}$$

where $a_{n,i}$, $i = 1, 2$ are defined as in (5.21) and

$$\tilde{f}\left(d_{1,1}, d_{1,2}, S_1\left(a_{n,1}\right), S_1\left(a_{n,2}\right)\right) =$$
$$\phi\left(d_{1,1}\right)\Phi\left(\frac{d_{1,2}-\rho d_{1,1}}{\sqrt{1-\rho^2}}\right)\left(\frac{d_{1,1}^3-d_{1,1}}{12} - 2S_1^2\left(a_{n,1}\right)d_{1,1} + \frac{1}{3}\rho\sigma_2^3\sqrt{T}^3 + \frac{1}{2}d_{1,1}\rho^2\sigma_2^2 T + \frac{1}{3}\rho\sigma_2\sqrt{T}(d_{1,1}^2-1)\right)$$
$$+\phi\left(d_{1,2}\right)\Phi\left(\frac{d_{1,1}-\rho d_{1,2}}{\sqrt{1-\rho^2}}\right)\left(\frac{d_{1,2}^3-d_{1,2}}{12} - 2S_1^2\left(a_{n,2}\right)d_{1,2} + \frac{1}{3}\sigma_2^3\sqrt{T}^3 + \frac{1}{2}d_{1,2}\sigma_2^2 T + \frac{1}{3}\sigma_2\sqrt{T}(d_{1,2}^2-1)\right)$$
$$+\phi_{0,V}(d_{1,1}, d_{1,2})\left(4S_1\left(a_{n,1}\right)S_1\left(a_{n,2}\right) - 2\rho S_1^2\left(a_{n,1}\right) - 2\rho S_1^2\left(a_{n,2}\right) + \frac{1}{2}\rho(\rho^2-1)\sigma_2^2 T + \frac{\rho^2 d_{1,1}d_{1,2}}{3(1-\rho^2)}\right.$$
$$\left.+\frac{\rho}{2} - \frac{\rho(1+\rho^2)(d_{1,1}^2+d_{1,2}^2)}{12(1-\rho^2)} + \frac{1}{3}\rho\sigma_2\sqrt{T}\left(\rho d_{1,1}-d_{1,2}\right)\right)$$
$$-\Phi_{0,V}\left(d_{1,1}, d_{1,2}\right)\frac{\sigma_2^4 T^2}{12}.$$

Substituting the above expansions into (5.25) and taking into account that

$$Ke^{-rT}\phi\left(d_{2,2}\right)\Phi\left(\frac{d_{2,1}-\rho d_{2,2}}{\sqrt{1-\rho^2}}\right) = s_{0,2}\phi\left(d_{1,2}\right)\Phi\left(\frac{d_{1,1}-\rho d_{1,2}}{\sqrt{1-\rho^2}}\right),$$

we get

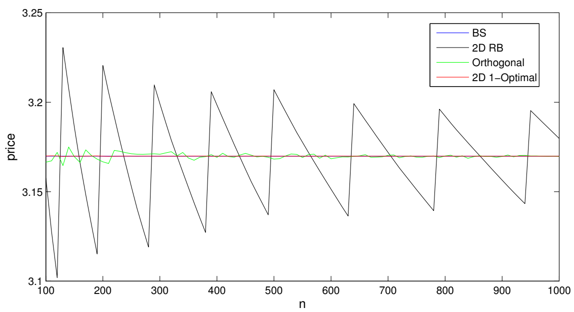
$$V^{(n)} = V - \frac{1}{\sqrt{n}}2S_1(a_{n,1})\left(K_2e^{-rT}\phi\left(d_{2,1}\right)\Phi\left(\frac{d_{2,2}-\rho d_{2,1}}{\sqrt{1-\rho^2}}\right) - s_{0,2}\phi\left(d_{1,1}\right)\Phi\left(\frac{d_{1,2}-\rho d_{1,1}}{\sqrt{1-\rho^2}}\right)\right)$$
$$+\frac{1}{n}\left(K_2e^{-rT}f\left(d_{2,1}, d_{2,2}, S_1\left(a_{n,1}\right), S_1\left(a_{n,2}\right)\right) - s_{0,2}\tilde{f}\left(d_{1,1}, d_{1,2}, S_1\left(a_{n,1}\right), S_1\left(a_{n,2}\right)\right)\right.$$
$$\left.-s_{0,2}\Phi_{0,V}\left(d_{1,1}, d_{1,2}\right)\sigma_2\sqrt{T}c_{5,n}\right) + O\left(\frac{1}{\sqrt{n}^3}\right).$$

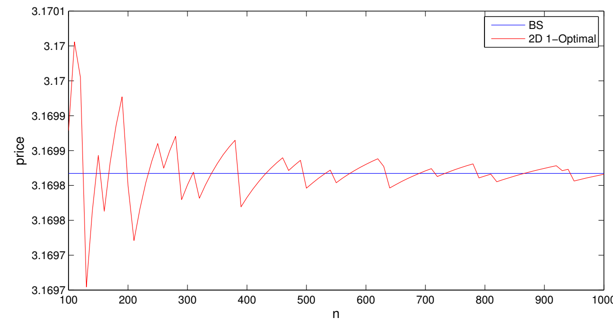
By setting $c_{i,n}$, $i = 1, 2$ as in (5.22) and

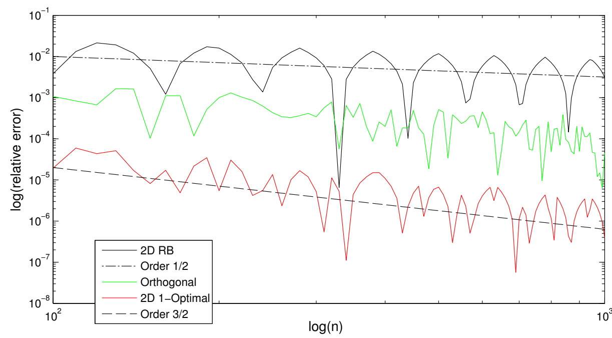
$$c_{5,n} = \frac{K_2e^{-rT}f\left(d_{2,1}, d_{2,2}, 0, 0\right) - s_{0,2}\tilde{f}\left(d_{1,1}, d_{1,2}, 0, 0\right)}{s_{0,2}\Phi_{0,V}\left(d_{1,1}, d_{1,2}\right)\sigma_2\sqrt{T}},$$

we have the statement of the proposition. □

Compare the numerical results of the 2D RB model, the Orthogonal tree and the 2D 1-Optimal model for a two-asset correlated option with $s_0 = [12, 12]$, $r = 0.1$, $K = [15, 17]$, $\sigma = [0.2, 0.25]$, $\rho = 0.5$, $T = 1$. The first two plots give the convergence pattern of the methods, while the third gives the order of convergence.







Note the y-axis scale in the right-hand side plot above. The 2D 1-Optimal model also has oscillatory behavior, but on a much smaller scale. From the log-log plot we can see, that the 2D 1-Optimal model indeed has a higher, $O\left(\frac{1}{\sqrt{n}^3}\right)$, order of convergence.

Parameters	n	2D RB tree	Orthogonal tree	2D 1-Optimal	BS value
$s_0 = [12, 12]$	100	3.1577028	3.1664811	3.1699286	3.1698672
$r = 0.1$	200	3.2205542	3.1667005	3.1698499	
$K = [15, 17]$	300	3.1991868	3.1709520	3.1698507	
$\sigma = [0.2, 0.25]$	500	3.2069667	3.1682395	3.1698460	
$\rho = 0.5$	700	3.1719940	3.1703365	3.1698709	
$T = 1$	900	3.1564813	3.1697090	3.1698745	
	1000	3.1797558	3.1697243	3.1698660	

5.2 Multidimensional generalizations

We now consider a general d-dimensional problem

$$dS_i(t) = S_i(t)\left(rdt + \sigma_i dW_i(t)\right), \quad i \in \{1, \ldots, d\}, \tag{5.28}$$

where the Brownian motions $W_i(t)$ and $W_j(t)$ have correlation $\rho_{i,j},\ i,j = 1,\dots,d,\ i<j$. The approximating binomial process is now

$$S_k^{(n)} = \begin{pmatrix} S_{k-1,1}^{(n)} e^{\alpha_{n,1}\Delta t + \sigma_1 \sqrt{\Delta t}\xi_{k,1}^{(n)}} \\ \vdots \\ S_{k-1,2}^{(n)} e^{\alpha_{n,2}\Delta t + \sigma_2 \sqrt{\Delta t}\xi_{k,d}^{(n)}} \end{pmatrix},\ k = 1,\dots,n, \tag{5.29}$$

where $\{\alpha_{n,i}\}$, $i = 1,\dots,d$ are bounded sequences, and for each $n \in \mathbb{N}$ $\xi_k^{(n)}$ are i.i.d random vectors s.t.

$$\xi_k^{(n)} \in \{\omega_k = (\omega_{k,1},\dots,\omega_{k,d}) |\, \omega_{k,i} = \pm 1,\ i = 1,\dots,d\},$$

with the one-step transition probabilities $P_{1,n}(\omega_k)$. To ensure weak convergence we assume the following moment-matching conditions (see (2.8))

$$\begin{aligned} \mu_i(n) &= r - \frac{1}{2}\sigma_i^2 + o(1), \quad i = 1,\dots,d, \\ \sigma_i^2(n) &= \sigma_i^2 + o(1), \quad i = 1,\dots,d, \\ c_{i,j}(n) &= \sigma_i \sigma_j \rho_{i,j}, \quad i = 1,\dots,d, i<j. \end{aligned} \tag{5.30}$$

To get the distributional fit in the multidimensional case we again apply the Edgeworth expansion Theorem 3.4.9.

Corollary 5.2.1. *For each $n \in \mathbb{N}$ let $\eta_i^{(n)} = \frac{\xi_i^{(n)}+1}{2}$, $i = 1,\dots,d$, $\mu_n = E\left(\eta_1^{(n)}\right)$ and $V_n = Cov\left(\eta_1^{(n)}\right)$. Then the process $S^{(n)}$ in (5.29) satisfies*

$$\begin{aligned} P\left(S_n^{(n)} \le x\right) = \Phi_{0,V_n}(y_n) &- \frac{1}{\sqrt{n}} \sum_{i=1}^{d} S_1(a_{n,i}) D_i \Phi_{0,V_n}(y_n) \\ &+ \frac{1}{\sqrt{n}} P_1\left(-\Phi_{0,V_n}, \{\kappa_{n,\nu}\}\right)(y_n) + O\left(\frac{1}{n}\right), \end{aligned} \tag{5.31}$$

where $y_{n,i} = \left(\ln\left(\frac{x_i}{s_{0,i}}\right) - \mu_i(n)T\right)/(2\sigma_i\sqrt{T})$, $i = 1,\dots,d$, $a_n = n\mu_n + y_n\sqrt{n}$, and $P_1\left(-\Phi_{0,V_n}, \{\kappa_{n,\nu}\}\right)(y_n)$ is defined as in Corollary 5.1.1.

Proof. Due to the moment-matching condition (5.30) the assumptions of Theorem 3.4.9 are satisfied for all s. Taking $s = 3$ we get the statement of the corollary. □

Remark 5.2.2. Let $V_n^{(i)}$ be the covariance matrix V_n, where the 1-st and i-th row and column are interchanged, and consider the following partition

$$V_n^{(i)} = \begin{pmatrix} \sigma_i^2 & V_{n,12}^{(i)} \\ V_{n,21}^{(i)} & V_{n,22}^{(i)} \end{pmatrix},$$

where $V_{n,22}^{(i)}$ is a $(d-1)\times(d-1)$ matrix and $V_{n,21}^{(i)} = \left(V_{n,12}^{(i)}\right)^T$. Further, let $y_n^{(i)}$ be obtained from y_n by interchanging the first and the i-th component, i.e.
$y_n^{(i)} := (y_{n,i}, y_{n,2}, \ldots, y_{n,i-1}, y_{n,1}, y_{n,i+1}, \ldots, y_{n,d})$. Set

$$\mu_n^{(i)} = \frac{y_{n,i}}{\sigma_i^2} V_{n,21}^{(i)}, \quad \Sigma_n^{(i)} = V_{n,22}^{(i)} - \frac{1}{\sigma_i^2} V_{n,21}^{(i)} V_{n,12}^{(i)}.$$

Then, the derivatives in the above expansion can be written as

$$D_i \Phi_{0,V_n}(y_n) = \frac{1}{\sigma_i} \phi\left(\frac{y_{n,i}}{\sigma_i}\right) \Phi_{\mu_n^{(i)},\Sigma_n^{(i)}}\left(y_{n,2}^{(i)}, \ldots, y_{n,d}^{(i)}\right).$$

Indeed, using the formulas for conditional and marginal densities of the normal distribution (see e.g. [PP12], Chapter 8.1), for any $x \in \mathbb{R}^d$ we have

$$\phi_{0,V_n}(x) = \phi_{0,\sigma_1^2}(x_1)\, \phi_{\mu_n^{(1)},\Sigma_n^{(1)}}(x_2, \ldots, x_d). \tag{5.32}$$

Therefore, for $i = 1$ the statement is proven. For $i \geq 2$ note that

$$\phi_{0,V_n}(x) = \phi_{0,V_n}\left(T_{1,i} x^{(i)}\right) = \phi_{0,V_n^{(i)}}\left(x_2^{(i)}, \ldots, x_d^{(i)}\right),$$

where $T_{1,i}$ is the $d \times d$ elementary matrix that changes rows 1 and i and $x^{(i)}$ is constructed as $y_n^{(i)}$, $i = 1, \ldots, d$. By applying the same formulas as in (5.32) we get the statement for all i.

We now construct advanced multi-dimensional trees with an $O\left(\frac{1}{n}\right)$ order of convergence, which we will refer to as the 1/2-Optimal tree. Following the two-dimensional case we choose the drifts to be

$$\alpha_i^{(n)} = r - \frac{1}{2}\sigma_i^2 - \frac{c_{i,n}\sigma_i}{\sqrt{T}} \frac{1}{\sqrt{n}}$$

with bounded $\{c_{i,n}\}_n$, $i = 1, \ldots, d$, and the transition probabilities equal to

$$P_{1,n}(\omega_k) = \frac{1}{2^d}\left(1 + \sum_{i,j=1,i<j}^{d} \delta_{i,j}(\omega_k)\, \rho_{i,j} + \frac{1}{n} \sum_{i=1}^{d} \delta_i(\omega_k)\, c_{i,n}\right), \tag{5.33}$$

where

$$\delta_{i,j}(\omega_k) = \begin{cases} 1 & \text{if } \omega_{k,i} = \omega_{k,j}, \\ -1 & \text{if } \omega_{k,i} \neq \omega_{k,j}, \end{cases}$$

and

$$\delta_i(\omega_k) = \begin{cases} 1 & \text{if } \omega_{k,i} = 1, \\ -1 & \text{if } \omega_{k,i} = -1. \end{cases}$$

for $\omega_k \in \{-1, 1\}^d$ (see also section 2.2.2).

Proposition 5.2.3. *If the drift coefficients $c_{i,n}$ are chosen as*

$$c_{i,n} = 1 - 2\left\{\frac{\ln\left(\frac{x_i}{s_{0,i}}\right) - \left(r - \frac{1}{2}\sigma_i^2\right)T}{2\sigma_i\sqrt{\Delta t}} + \frac{n}{2}\right\}, \quad i = 1, \ldots, d, \tag{5.34}$$

then the distribution function of $S_n^{(n)}$ in (5.29) satisfies

$$P\left(S_{n,1}^{(n)} \leq x_1, \ldots, S_{n,d}^{(n)} \leq x_d\right) = \Phi_{0,V}\left(d_{2,1}, \ldots, d_{2,d}\right) + O\left(\frac{1}{n}\right),$$

where $d_{2,i} = \frac{\ln\left(\frac{x_i}{s_{0,i}}\right) - \left(r - \frac{1}{2}\sigma_i^2\right)T}{\sigma_i\sqrt{T}}$, $i = 1, \ldots, d$, *and* $V = \begin{pmatrix} 1 & \rho_{1,2} & \cdots & \rho_{1,d} \\ \rho_{1,2} & & & \\ \vdots & & & \\ \rho_{1,d} & & \cdots & 1 \end{pmatrix}$.

Proof. If $c_{n,i}$ is chosen as in (5.34), then $S_1\left(a_{n,i}\right) = 0$, $i = 1, \ldots, d$ and we are only left with the $O\left(\frac{1}{\sqrt{n}}\right)$ term $P_1\left(-\Phi_{0,V_n}, \{\kappa_{n,\nu}\}\right)$ in (5.31). However, for $|\nu| = 3$ the possible cumulants of the d-dimensional process are

$$\begin{aligned}
\kappa_{n,(3,0\ldots,0)} &= \mu_{n,(3,0\ldots,0)} - 3\mu_{n,(1,0\ldots,0)}\mu_{n(2,0\ldots,0)} + 2\mu_{n,(1,0\ldots,0)}^3, \\
\kappa_{n,(2,1,\ldots,0)} &= \mu_{n,(2,1,\ldots,0)} - \mu_{n,(0,1,\ldots,0)}\mu_{n,(2,0,\ldots,0)} - 2\mu_{n,(1,1,\ldots,0)}\mu_{n,(1,0,\ldots,0)} \\
&\quad + 2\mu_{n,(1,0,\ldots,0)}^2\mu_{n,(0,1,\ldots,0)}, \\
\kappa_{n,(1,1,1,\ldots,0)} &= \mu_{n,(1,1,1,\ldots,0)} - \mu_{n,(1,0,0,\ldots,0)}\mu_{n,(0,1,1,\ldots,0)} - \mu_{n,(0,1,0,\ldots,0)}\mu_{n,(1,0,1,\ldots,0)} \\
&\quad - \mu_{n,(0,0,1,\ldots,0)}\mu_{n,(1,1,0,\ldots,0)} + 2\mu_{n,(1,0,0,\ldots,0)}\mu_{n,(0,1,0,\ldots,0)}\mu_{n,(0,0,1,\ldots,0)},
\end{aligned} \tag{5.35}$$

and completely symmetric for the other indices. Due to (5.33) we have $\forall r, s, q \in \mathbb{N}$

$$\begin{aligned}
\mu_{n,re_i} &= \frac{1}{2} + O\left(\frac{1}{n}\right), \\
\mu_{n,re_i+se_j} &= \frac{1}{4}\left(1 + \rho_{i,j}\right) + O\left(\frac{1}{n}\right), \\
\mu_{n,re_i+se_j+qe_k} &= \frac{1}{8}\left(1 + \rho_{i,j} + \rho_{i,k} + \rho_{j,k}\right) + O\left(\frac{1}{n}\right), \quad i, j, k = 1, \ldots, d,
\end{aligned}$$

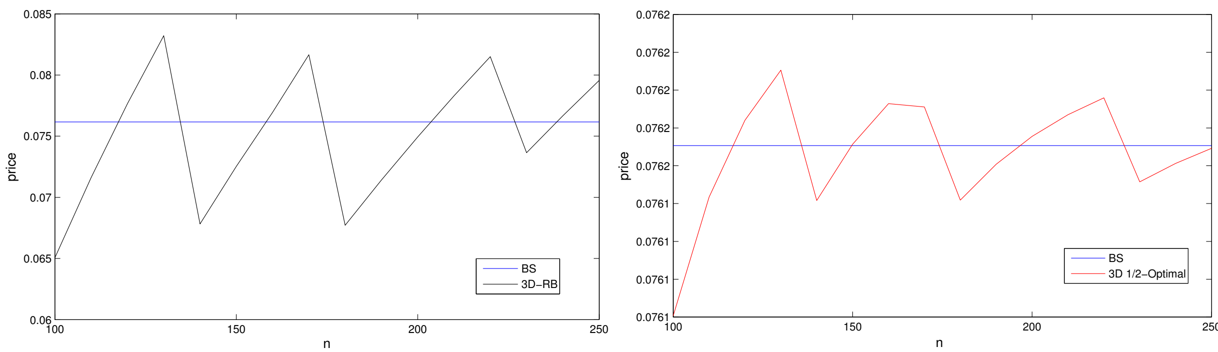
where e_i is the i-th unit vector. Substituting these expressions into (5.35) we get $\kappa_{n,\nu} = O\left(\frac{1}{n}\right)$, $|\nu| = 3$, therefore,

$$P_1\left(-\Phi_{0,V_n}, \{\kappa_{n,\nu}\}\right)\left(y_n\right) = O\left(\frac{1}{n}\right),$$

and we have the statement of the proposition. $\square$

Remark 5.2.4. Note that the above tree construction inherits the drawbacks of the underlying multidimensional RB model, i.e. for $d \geq 3$ the probabilities in (5.33) might not be well-defined for all parameter settings. (For details see e.g. [M09], Chapter 3).

To conclude this section we present numerical results for the 3D-RB tree and the 3D 1/2-Optimal tree introduced above. The first two plots show the convergence pattern, while the third one confirms the order of convergence. Once again, note that the oscillations of the 3D 1/2-Optimal tree are much smaller than in the case of the 3D-RB tree.



log-log plot

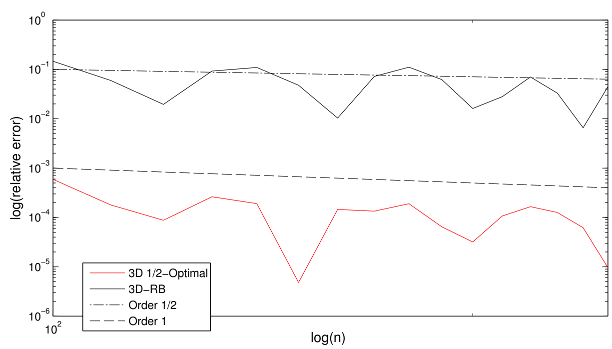


Figure 5.2: 3D distributional fit for $s_0 = [20, 16, 25]$, $r = 0.1$, $x = [17, 20, 25]$, $\sigma = [0.2, 0.25, 0.6]$, $\rho = [0.6, 0.2, 0.4]$, $T = 1$

Parameters	n	3D RB tree	3D 1/2-Optimal	BS value
$s_0 = [20, 16, 25]$	50	0.0695389	0.0760962	0.0761653
$r = 0.1$	100	0.0650581	0.0761203	
$x = [17, 20, 25]$	150	0.0725303	0.0761657	
$\sigma = [0.2, 0.25, 0.6]$	200	0.0749387	0.0761678	
$\rho = [0.6, 0.2, 0.4]$	250	0.0795593	0.0761646	
$T = 1$	300	0.0778009	0.0761620	

Note that the 3D 1/2-Optimal model exhibits a significant improvement and has a relative error smaller than 10^{-4} for 150 time steps.

5.3 Edgeworth expansions for decoupled processes

In this section we consider the decoupling method introduced in [KM09] (see section 2.2.2). Numerical experiments suggest that this approach has a better convergence behavior than the standard multidimensional BEG and RB models. We apply an Edgeworth expansion to obtain a theoretical justification of this result. We then discuss ways of further improving the convergence, following the methods described above.

Consider the two-dimensional Black-Scholes model

$$dS_i\left(t\right) = S_i\left(t\right)\left(rdt + \sigma_i dW_i\left(t\right)\right), \quad i = 1, 2, \tag{5.36}$$

where $W_1(t)$ and $W_2(t)$ have correlation ρ. Let $Y\left(t\right)$ be the decoupled process, i.e. $Y\left(t\right) = G^{-1}S\left(t\right)$, $t \in [0, T]$, where G is obtained from the Cholesky decomposition $\Sigma = GG^T$, i.e.

$$G = \begin{pmatrix} \sigma_1 & 0 \\ \rho\sigma_2 & \sqrt{1-\rho^2}\sigma_2 \end{pmatrix}.$$

Recall that in the decoupling approach the binomial tree $Y^{(n)}$ in (2.14) is first constructed to approximate the process Y with independent components. The tree $S^{(n)}$ for the original process S is then obtained by a back transformation (2.15), i.e. for $k = 1, \dots, n$, we have

$$\begin{aligned} S_0^{(n)} &= s_0, \\ S_k^{(n)} &= \exp\left(G\begin{pmatrix} Y_{k,1}^{(n)} \\ Y_{k,2}^{(n)} \end{pmatrix}\right) = \begin{pmatrix} S_{k-1,1}^{(n)} e^{\left(r-\frac{1}{2}\sigma_1^2\right)\Delta t + \sigma_1\sqrt{\Delta t}\xi_{k,1}} \\ S_{k-1,2}^{(n)} e^{\left(r-\frac{1}{2}\sigma_2^2\right)\Delta t + \sigma_2\sqrt{\Delta t}\left(\rho\xi_{k,1} + \sqrt{1-\rho^2}\xi_{k,2}\right)} \end{pmatrix}, \end{aligned} \tag{5.37}$$

where ξ_k are i.i.d. random vectors with $P\left(\xi_k = \omega\right) = \frac{1}{4}$, for all $\omega \in \left\{\left(\omega_1, \omega_2\right) \middle| \omega_i = \pm 1, i = 1, 2\right\}$.

5.3.1 Distributional fit

Now consider the distributional fit of the process $S^{(n)}$ at maturity.

$$\begin{aligned} P\left(S_n^{(n)} \le x\right) &= P\left(\ln S_n^{(n)} \le \ln x\right) = P\left(GY_n^{(n)} \le \ln x\right) \\ &= P\left(G\frac{1}{\sqrt{n}}\left(\sum_{k=1}^{n} \eta_k - nE\left(\eta_k\right)\right) \le y\right), \end{aligned} \tag{5.38}$$

where $\eta_{k,i} = \left(\xi_{k,i} + 1\right)/2$ and $y_i = \frac{\ln\left(\frac{x_i}{s_{0,i}}\right) - \left(r - \frac{1}{2}\sigma_i^2\right)T}{2\sqrt{T}}$, $i = 1, 2$.

Independent processes

We first assume the simplest case, when $\rho = 0$, i.e. S_1 ans S_2 are independent. In this case

$$\begin{aligned} P\left(S_n^{(n)} \le x\right) &= P(Y_{n,1}^{(n)} \le \frac{\log x_1}{\sigma_1}, Y_{n,2}^{(n)} \le \frac{\log x_2}{\sigma_2}) \\ &= P(Y_{n,1}^{(n)} \le \frac{\log x_1}{\sigma_1}) P(Y_{n,2}^{(n)} \le \frac{\log x_2}{\sigma_2}). \end{aligned} \tag{5.39}$$

We can now apply Theorem 3.4.9 to each of the components separately.

Proposition 5.3.1. *If $\rho = 0$, then the process $S^{(n)}$ in (5.37) satisfies*

$$P\left(S_n^{(n)} \le K\right) = \Phi_{0,I}\left(d_{2,1}, d_{2,2}\right) - \frac{2}{\sqrt{n}}\left(S_1\left(\frac{n}{2} + \frac{d_{2,1}}{2}\sqrt{n}\right)\phi\left(d_{2,1}\right)\Phi\left(d_{2,2}\right) + S_1\left(\frac{n}{2} + \frac{d_{2,2}}{2}\sqrt{n}\right)\phi\left(d_{2,2}\right)\Phi\left(d_{2,1}\right)\right) + O\left(\frac{1}{n}\right),$$

where, as usual, $d_{2,i} = \frac{\log(\frac{K_i}{s_{0,i}}) - (r - \frac{1}{2}\sigma_i^2)T}{\sigma_i\sqrt{T}}$, *and* $S_1(x) = \{x\} - \frac{1}{2}$.

Proof. By the 1-D Edgeworth expansion

$$P\left(Y_{n,i}^{(n)} \le \frac{\log x_i}{\sigma_i}\right) = \Phi\left(d_{2,i}\right) - \frac{2}{\sqrt{n}} S_1\left(\frac{n}{2} + \frac{d_{2,i}}{2}\sqrt{n}\right)\phi\left(d_{2,i}\right) + O\left(\frac{1}{n}\right).$$

Together with (5.39) this gives the necessary result. □

Correlated processes

We now assume correlated Brownian motions in (5.36). In this case a simple decomposition as in (5.39) is not possible and the 1-D results are not directly applicable. From (5.38) we have

$$P\left(S_n^{(n)} \le K\right) = P\left(\frac{1}{\sqrt{n}}\left(\sum_{k=1}^{n}\eta_k - nE\left(\eta_k\right)\right) \in A\right),$$

where $A = \left\{(y_1, y_2) \mid y_1 \le \frac{d_{2,1}}{2},\ \rho y_1 + \sqrt{1-\rho^2}\, y_2 \le \frac{d_{2,2}}{2}\right\}$.

Proposition 5.3.2. *The process $S^{(n)}$ in (5.37) satisfies*

$$P\left(S_n^{(n)} \le K\right) = \Phi_{0,V}\left(d_{2,1}, d_{2,2}\right) - \frac{2}{\sqrt{n}} S_1\left(\frac{n}{2} + \frac{d_{2,1}}{2}\sqrt{n}\right)\phi\left(d_{2,1}\right)\Phi\left(\frac{d_{2,2} - \rho d_{2,1}}{\sqrt{1-\rho^2}}\right) - \frac{2}{\sqrt{n}}\phi\left(d_{2,2}\right)\int_{-\infty}^{d_{2,1}}\left(\frac{\rho}{\sqrt{1-\rho^2}} S_1\left(\frac{n}{2} + \frac{x}{2}\sqrt{n}\right) + S_1\left(\frac{n}{2} + \frac{d_{2,2} - \rho x}{2\sqrt{1-\rho^2}}\sqrt{n}\right)\right)\phi\left(\frac{x - \rho d_{2,2}}{\sqrt{1-\rho^2}}\right) dx + O\left(\frac{1}{n}\right), \quad n \to \infty, \tag{5.40}$$

where $V = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$.

Proof. By applying the multidimensional Euler-Maclaurin formula (Theorem 3.4.14) with the Borel set A, and following the proof of Theorem 3.4.9 we get

$$P\left(S_n^{(n)} \leq K\right) = \Phi_{0,V}\left(d_1, d_2\right) - \frac{1}{\sqrt{n}}\left(\Lambda_1\left(A\right) + \Lambda_2\left(A\right)\right) + O\left(\frac{1}{n}\right),$$

where Λ_1, Λ_2 are signed measures with distribution functions defined as

$$\begin{aligned}\Lambda_1(x_1, x_2) &= S_1\left(\frac{n}{2} + \sqrt{n}x_1\right)\frac{\partial}{\partial x_1}\Phi_{0,E}(x_1, x_2)\\ &= 2S_1\left(\frac{n}{2} + \sqrt{n}x_1\right)\phi(2x_1)\Phi\left(2x_2\right),\end{aligned}$$

and analogously

$$\begin{aligned}\Lambda_2(x_1, x_2) &= S_1\left(\frac{n}{2} + \sqrt{n}x_2\right)\frac{\partial}{\partial x_2}\Phi_{0,E}(x_1, x_2)\\ &= 2S_1\left(\frac{n}{2} + \sqrt{n}x_2\right)\phi(2x_2)\Phi\left(2x_1\right),\end{aligned}$$

with $E = \begin{pmatrix} 1/4 & 0 \\ 0 & 1/4 \end{pmatrix}$. Note, that here we used that $P_1\left(-\phi_{0,V_n}, \{\kappa_\nu\}\right) = 0$, since all the ν-th cumulants, κ_ν, of η_k are equal to zero for $|\nu| = 3$. Applying integration by parts for Stieltjes integrals we get

$$\begin{aligned}\Lambda_1\left(A\right) &= 2\int_{-\infty}^{\frac{d_{2,1}}{2}}\int_{-\infty}^{\frac{\frac{d_{2,2}}{2} - \rho x_1}{\sqrt{1-\rho^2}}} dS_1\left(\frac{n}{2} + \sqrt{n}x_1\right)\phi(2x_1)\Phi\left(2x_2\right)\\ &= 2\int_{-\infty}^{d_{2,1}}\Phi\left(\frac{d_{2,2} - \rho x_1}{\sqrt{1-\rho^2}}\right) dS_1\left(\frac{n}{2} + \sqrt{n}\frac{x_1}{2}\right)\phi(x_1)\\ &= 2S_1\left(\frac{n}{2} + \sqrt{n}\frac{d_{2,1}}{2}\right)\phi(d_{2,1})\Phi\left(\frac{d_{2,2} - \rho d_{2,1}}{\sqrt{1-\rho^2}}\right)\\ &+ \frac{2\rho}{\sqrt{1-\rho^2}}\phi(d_{2,2})\int_{-\infty}^{d_{2,1}} S_1\left(\frac{n}{2} + \sqrt{n}\frac{x_1}{2}\right)\phi\left(\frac{x_1 - \rho d_{2,2}}{\sqrt{1-\rho^2}}\right)dx_1,\end{aligned}$$

and

$$\begin{aligned}\Lambda_2\left(A\right) &= 2\int_{-\infty}^{\frac{d_{2,1}}{2}}\int_{-\infty}^{\frac{\frac{d_{2,2}}{2} - \rho x_1}{\sqrt{1-\rho^2}}} dS_1\left(\frac{n}{2} + \sqrt{n}x_2\right)\phi(2x_2)\Phi\left(2x_1\right)\\ &= 2\phi(d_{2,2})\int_{-\infty}^{d_{2,1}} S_1\left(\frac{n}{2} + \sqrt{n}\frac{d_{2,2} - \rho x_1}{2\sqrt{1-\rho^2}}\right)\phi\left(\frac{x_1 - \rho d_{2,2}}{\sqrt{1-\rho^2}}\right)dx_1.\end{aligned}$$

The statement of the proposition follows directly. □

As mentioned in [KM09] (see also section 2.2.2), the reason why the decoupling approach has a better performance than the usual methods, is that the rectangular grid is destroyed (or partially destroyed, as with the Cholesky decomposition) due to the transformation G. This can also be seen in the above expansion. Unlike the 2D-RB (Proposition 5.1.2) or the BEG model (Proposition 5.1.3), the leading error term in the expansion for the decoupling method does not have a closed form. Instead we have an integral component over a sum of two periodic functions S_1, whose effect might cancel out, leading to a better convergence pattern.

Note that a similar expansion can also be shown to hold for the decoupling approach with a spectral decomposition. However, in this case both terms in the $\frac{1}{\sqrt{n}}$ coefficient will have an integral form. On one hand, this allows even smoother convergence, on the other hand, however, the absence of a closed formula prevents from further improving the convergence behavior with the methods described above.

5.3.2 Improving the convergence behavior

We will now demonstrate how the distributional fit of the Cholesky tree can be improved by controlling the position of the nodes at maturity. We will apply a two-dimensional version of the Tian approach, which together with extrapolation provides better convergence results.

Consider the 2-D $Y_n^{(n)}$ grid. The in-the-money points are $y \in A$, i.e.

$$\begin{aligned} y_1 &\leq \frac{\log K_1}{\sigma_1}, \\ \rho y_1 + \sqrt{1-\rho^2} y_2 &\leq \frac{\log K_2}{\sigma_2}. \end{aligned}$$

The idea of the Tian approach would be to tilt the tree in such a way, that the nodes of the tree coincide with the boundary of the in-the-money region A for every $n \in \mathbb{N}$. Since the region is not rectangular, we cannot simply match two rows of nodes. We can first match $\frac{\log K_1}{\sigma_1}$ with the $Y_{n,1}^{(n)}$ coordinate independent of $Y_{n,2}^{(n)}$, i.e. match a whole row. Then the second coordinate can be chosen to match a node with the second boundary, however this will depend on the value of $Y_{n,1}^{(n)}$. Based on numerical results we propose to match a node with $\left(\frac{\log K_1}{\sigma_1}, \frac{1}{\sqrt{1-\rho^2}} \left(\frac{\log K_2}{\sigma_2} - \rho \frac{\log K_1}{\sigma_1} \right) \right)$, which is equivalent to matching a node in the $S_n^{(n)}$ grid to (K_1, K_2). We now proceed in the usual way. We tilt the tree by introducing bounded drifts $\gamma_{n,i}$. For $k = 1, \ldots, n$ the tree is now constructed as

$$Y_{k,i}^{(n)} := Y_{k-1,i}^{(n)} + \left(\alpha_i + \gamma_{n,i} \sqrt{\Delta t} \right) \Delta t + \xi_{k,i}^{(n)} \sqrt{\Delta t}, \quad i = 1, 2, \tag{5.41}$$

where $\xi_{k,i}^{(n)}$ are independent variables with probabilities $p_{n,i} := P\left(\xi_{k,i}^{(n)} = 1 \right) = \frac{1}{2} - \frac{\gamma_{n,i}}{2} \Delta t$. The one-step transition probabilities of the vector $\xi_k^{(n)}$ are defined as the product of the corresponding one-dimensional probabilities. With this choice of parameters the moment-matching conditions (5.2) are satisfied and $S^{(n)}$ converges weakly to S.

The possible values for $Y^{(n)}$ at maturity are

$$Y_{n,i}^{(n)}(l_i) = Y_{0,i}^{(n)} + \left(\alpha_i + \gamma_{n,i}\sqrt{\Delta t}\right) T + 2l_i\sqrt{\Delta t} - n\sqrt{\Delta t},$$

where $l_i \in \{0, \ldots, n\}$ is the number of up-jumps in each of the dimensions. To determine the nodes closest to $G^{-1}\log(K)$ we consider the equations

$$\begin{pmatrix} Y_1(a_1) \\ Y_2(a_2) \end{pmatrix} = G^{-1}\log\begin{pmatrix} K_1 \\ K_2 \end{pmatrix},$$

which are solved by

$$a_{n,i}^{\gamma_n} = \frac{G_{i.}^{-1}\left(\log(K/s_0) - \left(r - \frac{1}{2}\sigma^2\right)T\right) - \gamma_{n,i}\sqrt{\Delta t}T}{2\sqrt{\Delta t}} + \frac{n}{2}, \quad i = 1, 2.$$

Set

$$l_{n,i}^{\gamma_n} = \left\lfloor a_{n,i}^{\gamma_n} \right\rfloor,$$

i.e. $l_{n,i}^{\gamma_n}$ is the number of up-jumps, s.t. $Y_i^{(n)}\left(l_{n,i}^{\gamma_n}\right) \leq G_{i.}^{-1}\log(K) < Y_i\left(l_{n,i}^{\gamma_n} + 1\right)$. Having found the four surrounding nodes, we choose γ_n to match $G^{-1}\log\begin{pmatrix} K_1 \\ K_2 \end{pmatrix}$ with the lower left node. Note that this is just one option, matching other nodes is also possible. Following section 4.2.1 we set

$$\gamma_{n,i} := \frac{2}{T}\left\{a_{n,i}^0\right\}. \tag{5.42}$$

In this case the first summand in the leading error coefficient in expansion (5.40) becomes constant, equal to $\frac{1}{\sqrt{n}}\phi(d_{2,1})\Phi\left(\frac{d_{2,2}-\rho d_{2,1}}{\sqrt{1-\rho^2}}\right)$. Note that the integral term is still present in the expansion, however, by correctly choosing the components, we can get a better convergence behavior. To be more precise, the described method is not symmetric, i.e. it provides a better fit for the first coordinate of $S_n^{(n)}$. Therefore, the question is, which of the two correlated assets to take as the first coordinate?

Consider the leading error term in expansion (5.40).

$$\left|2S_1\left(\frac{n}{2} + \frac{d_{2,1}}{2}\sqrt{n}\right)\phi(d_{2,1})\Phi\left(\frac{d_{2,2} - \rho d_{2,1}}{\sqrt{1-\rho^2}}\right)\right| < \phi(d_{2,1})\Phi\left(\frac{d_{2,2} - \rho d_{2,1}}{\sqrt{1-\rho^2}}\right)$$

$$\left|2\phi(d_{2,2})\int_{-\infty}^{d_{2,1}}\left(\frac{\rho}{\sqrt{1-\rho^2}}S_1\left(\frac{n}{2} + \frac{x}{2}\sqrt{n}\right) + S_1\left(\frac{n}{2} + \frac{d_{2,2} - \rho x}{2\sqrt{1-\rho^2}}\sqrt{n}\right)\right)\phi\left(\frac{x - \rho d_{2,2}}{\sqrt{1-\rho^2}}\right)dx\right|$$
$$< \left(\rho + \sqrt{1-\rho^2}\right)\phi(d_{2,2})\Phi\left(\frac{d_{2,1} - \rho d_{2,2}}{\sqrt{1-\rho^2}}\right)$$

Therefore, we propose the following heuristic approach. Take the asset i with the higher value of $\phi\left(d_{2,i}\right)\Phi\left(\frac{d_{2,j}-\rho d_{2,i}}{\sqrt{1-\rho^2}}\right)$ as the first component. This way more weight will be given to the first term and in many cases the second term can be neglected. Now extrapolation can be applied to speed up convergence. Note that by the above criterion, the asset, which is more out-of-the-money will be chosen as the first component.

The effectiveness of the method depends on the exact values of $\phi\left(d_{2,i}\right)\Phi\left(\frac{d_{2,j}-\rho d_{2,i}}{\sqrt{1-\rho^2}}\right)$ as well as the correlation coefficient ρ. Numerical experiments suggest, that the method performs better for at-the-money options.

Numerical results

We now compare the method proposed above with and without extrapolation, and the standard decoupling approach with both Cholesky and spectral decomposition. We consider three different options.

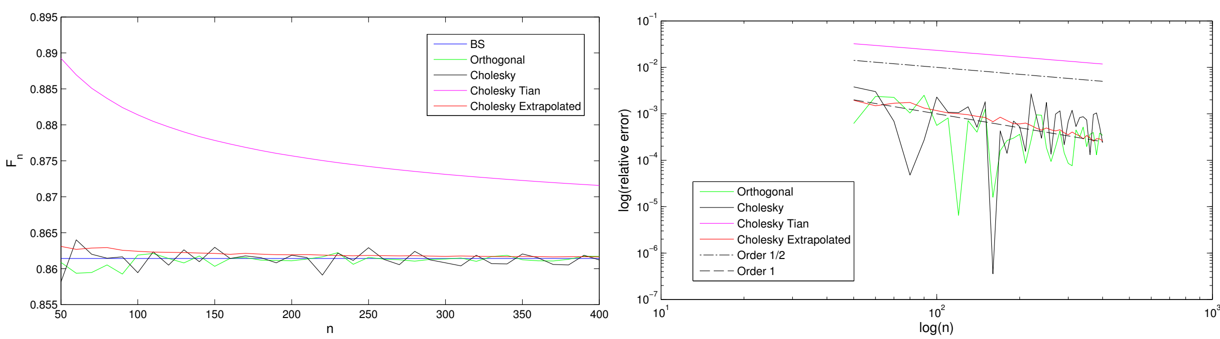


Figure 5.3: $s_0 = [12, 12]$, $r = 0.1$, $x = [20, 17]$, $\sigma = [0.2, 0.25]$, $\rho = 0.5$, $T = 1$

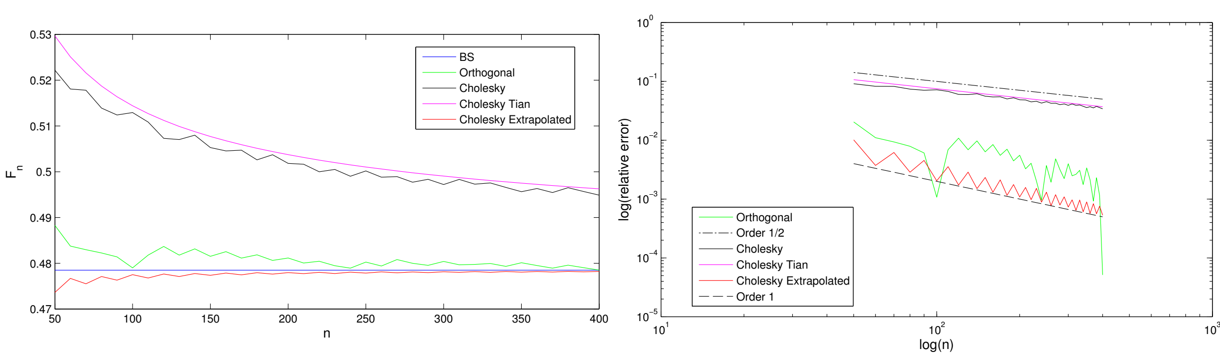


Figure 5.4: $s_0 = [12, 12]$, $r = 0.1$, $x = [13, 15]$, $\sigma = [0.2, 0.25]$, $\rho = 0.8$, $T = 1$

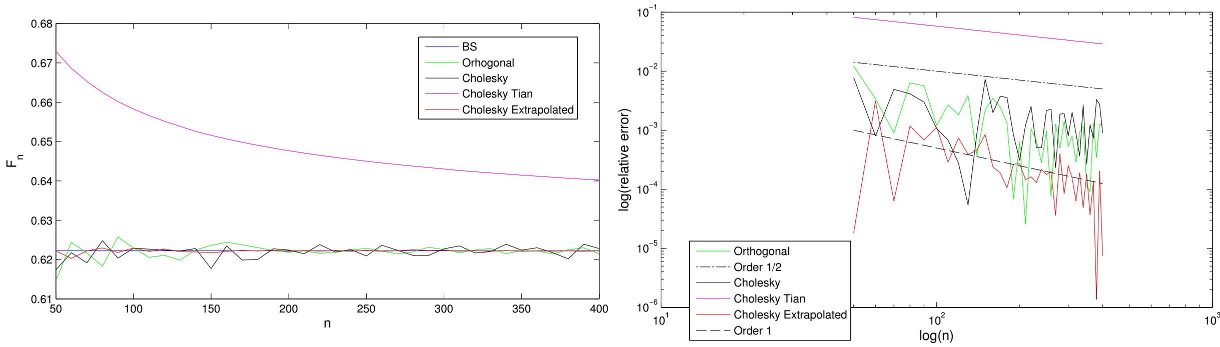


Figure 5.5: $s_0 = [12, 12]$, $r = 0.1$, $x = [18, 14]$, $\sigma = [0.2, 0.25]$, $\rho = 0.5$, $T = 1$

n	Cholesky tree	Orthogonal tree	Cholesky extrapolation	BS value
50	0.8581496	0.8608818	0.8630937	0.8614127
100	0.8594436	0.8618970	0.8624206	
150	0.8629707	0.8603272	0.8621292	
200	0.8618657	0.8611007	0.8619335	
300	0.8608183	0.8613391	0.8617172	
400	0.8612073	0.8617086	0.8616369	
50	0.5221727	0.4882740	0.4736266	0.4784877
100	0.5129303	0.4790018	0.4775078	
150	0.5052736	0.4815309	0.4773739	
200	0.5018282	0.4811512	0.4779660	
300	0.4972109	0.4803988	0.4781373	
400	0.4948763	0.4785123	0.4782329	
50	0.6173913	0.6146295	0.6222416	0.6222303
100	0.6228891	0.6229835	0.6229239	
150	0.6176984	0.6236023	0.6217028	
200	0.6224228	0.6218326	0.6220626	
300	0.6227290	0.6225625	0.6223870	
400	0.6227937	0.6214632	0.6222256	

Note that for at-the-money options (see figure 5.4) the method described above gives results which are even better than the orthogonal tree. For deep in-the-money options (figure 5.3), however, there is little improvement.

Remark 5.3.3. The computational effort of the standard decoupling approach is $O(n^3)$ (see [M09], Proposition 44). Due to the extrapolation procedure, the method described above also requires $O(n^3)$, but with a larger constant.

Choosing the correct asset as the first component does indeed influence the convergence behavior. Consider, for example, the option presented in figure 5.5. Here

$$\phi\left(d_{2,1}\right)\Phi\left(\frac{d_{2,2}-\rho d_{2,1}}{\sqrt{1-\rho^2}}\right)=0.03,$$
$$\phi\left(d_{2,2}\right)\Phi\left(\frac{d_{2,1}-\rho d_{2,2}}{\sqrt{1-\rho^2}}\right)=0.36,$$

therefore, the second asset needs to be taken as the first component. If this is disregarded, then we get convergence behavior as in figure 5.6.

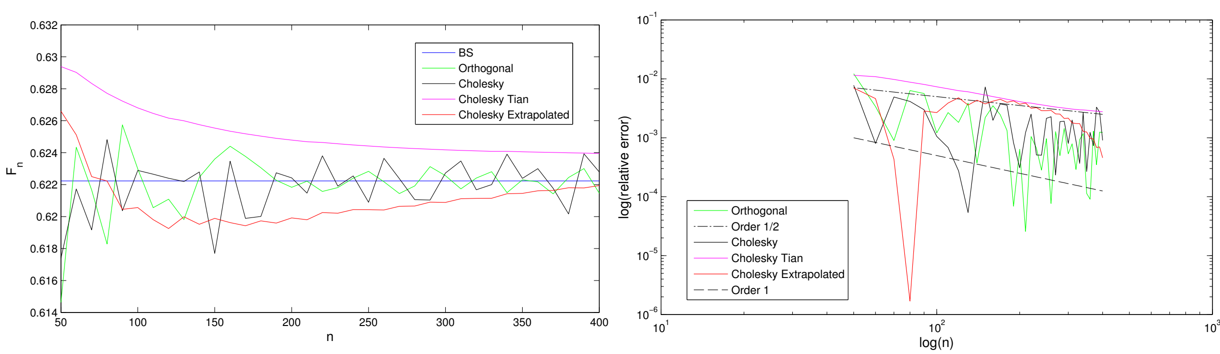


Figure 5.6: $s_0 = [12, 12]$, $r = 0.1$, $x = [18, 14]$, $\sigma = [0.2, 0.25]$, $\rho = 0.5$, $T = 1$

5.4 Conclusion

Edgeworth expansions provide a general method of constructing asymptotic expansions for the distribution functions of multidimensional trees. This allows to get a better understanding of the source of irregularity in the convergence of tree methods as well as to construct advanced models with a higher order of convergence. So far these results have been applied to European options in the Black-Scholes model, however, possible extensions to American-type payouts as well as other models could be the topic of further research.

Bibliography

[AK94] Amin, K. and Khanna, A. (1994). *Convergence of American Option Values from Discrete-to Continuous-Time Financial Models*. Mathematical Finance, 4:289-304.

[B68] Billingsley, P. (1968). *Convergence of Probability Measures*. John Wiley & Sons, New York, US.

[Bau92] Bauer, H. (1992). *Maß- und Integrationstheorie*. New York: Wiley.

[BD96] Broadie, M. and Detemple, J. (1996). *American Option Valuation: New Bounds, Approximations and a Comparison of Existing Methods*. Review of Financial Studies 9, 4:1221-1250.

[BEG89] Boyle, P.P., Evnine, J. and Gibbs, S. (1989). *Numerical Evaluation of Multivariate Contingent Claims*. Review of Financial Studies, 2:241-250.

[BL94] Boyle, P.P. and Lau, S.H. (1994). *Bumping Up Against the Barrier with the Binomial Method*. The Journal of Derivatives 1, 4:6-14.

[BR76] Bhattacharya, R.N. and Rao, R.R. (1976). *Normal Approximation and Asymptotic Expansions*. Berlin; New York: de Gruyter.

[BW90] Bhattacharya, R.N. and Waymire, E.C. (1990). *Stochastic Processes with Applications*. John Wiley & Sons.

[C45] Cramér, H. (1945). *Mathematical Methods of Statistics*. Princeton University Press.

[CJM92] Carr, P., Jarrow, R. and Myneni, R. (2007). *Alternative Characterizations of American Put Options*. Mathematical Finance 2, 2:87-106.

[Ch1906] Charlier, C.V.L. (1906). *Über das Fehlergesetz*. Ark. Mat. Astr. Fiz. 2, 8.

[CHLS11] Chung, S.-L., Hung, W., Lee, H.-H. and Shih, P.-T. (2011). *On the Rate of Convergence of Binomial Greeks*. The Journal of Futures Markets, 316:562-597.

[CP07] Chang, L.-B. and Palmer, K. (2007). *Smooth Convergence in the Binomial Model*. Finance and Stochastics, 11:91-105.

[CRR79] Cox, J.C., Ross, S.A. and Rubinstein, M. (1979). *Option Pricing: a Simplified Approach*. Journal of Financial Econometrics, 7:229-263.

[DD04] Diener, F. and Diener, M. (2004). *Asymptotics of the Price Oscillations of a European Call Option in a Tree Model*. Mathematical Finance, 14:271-293.

[DGS05] Denault, M., Gauthier, G. and Simonato, J.-G. (2005). *Improving Lattice Schemes Through Bias Reduction*. The Journal of Futures Markets 26, 6:733-757.

[DKEB95] Derman, E., Kani I., Ergener, D. and Bardhan, I. (1995). *Enhanced Numerical Methods for Options with Barriers*. Financial Analyst Journal 51, 2:8-17.

[E45] Esseen, C.G. (1945). *Fourier Analysis for Distribution Functions*. Acta Mathematica, 77:1-125.

[Edge1905] Edgeworth, F.Y. (1905). *The Law of Error*. Proc. Cambridge Philos. Soc. 20: 6-65.

[F68] Feller, P. (1968). *An Introduction to Probability Theory and its Applications, Volume 1*. John Wiley & Sons, Third edition.

[F71] Feller, P. (1971). *An Introduction to Probability Theory and its Applications, Volume 2*. John Wiley & Sons, Second edition.

[G99] Gobet, E. (1999). *Analysis of the Zigzag Convergence for Barrier Options with Binomial Trees*. Working paper 536, Université Paris 6 et Paris 7.

[GK54] Gnedenko, B.V. and Kolmogorov, A.N. (1954). *Limit Distributions for Sums of Independent Random Variables*. Addison-Wesley.

[GS98] Garcia-Soidan, P.H. (1998). *Edgeworth Expansions for Triangular Arrays*. Communications in Statistics-Theory and Methods, 27, 3:705-722.

[H06] Haug, E.G. (2006). *The Complete Guide to Option Pricing Formulas*. McGraw-Hill.

[Hull06] Hull, J. (2006). *Options, Futures, and Other Derivatives*. Upper Saddle River, NJ: Prentice Hall, 6th Edition.

[HW90] Hull, J. and White, A. (1990). *Valuing Derivative Securities Using the Explicit Finite Difference Method*. Journal of Financial and Quantitative Analysis 25, 1:87-100.

[HZ00] Heston, S. and Zhou, G. (2000). *On the Rate of Convergence of Discrete-Time Contingent Claims*. Mathematical Finance, 10:53-75.

[J09] Joshi, M.S. (2009). *The Convergence of Binomial Trees for Pricing the American Put*. The Journal of Risk, 87-108.

[K94] Kolassa, J.E. (1994). *Series Approximation Methods in Statistics*. New York : Springer-Verlag.

[KK01] Korn, R. and Korn, E. (2001). *Option Pricing and Portfolio Optimization: Modern Methods of Financial Mathematics*.

[KM09] Korn, R. and Müller, S. (2009). *The Decoupling Approach to Binomial Pricing of Multi-Asset Options*. The Journal of Computational Finance 12, 3:1-30.

[KM13] Korn, R. and Müller, S. (2013). *The Optimal-Drift Model: an Accelerated Binomial Scheme*. Finance Stoch. 17: 135-160.

[Konst09] Konstantopoulos, T. (1994). *Markov Chains and Random Walks*. Lecture notes.

[KorKu07] Korchinsky, I.A. and Kulik, A.M. (2007). *Local limit theorems for triangular array of random variables*. Theory of Stochastic Processes 13, 29, 3:48-54.

[KS98] Karatzas, I. and Shreve, S.E. (1998). *Brownian Motion ad Stochastic Calculus*. Springer, New York, US, second edition.

[Kr99] Krekel, M. (1999). *Bewertung von Exotischen Optionen mit diskreten Dividenden*. Diplomarbeit.

[Ku08] Kürten, C. (2008). *Dünngitter-Binomialbäume zur Bewertung von Multiasset-Optionen*. Diplomarbeit.

[L12] Leduc, G. (2012). *Arbitrarily Fast CRR Schemes*. MPRA.

[Lei98] Leisen, D.P.J. (1998). *Pricing the American Put Option: A Detailed Convergence Analysis for Binomial Models*. Journal of Economic Dynamics and Control, 22:1419-1444

[LR96] Leisen, D.P.J. and Reimer, M. (1996). *Binomial Models for Option Valuation-Examininig and Improving Convergence*. Applied Mathematical Finance, 3:319-346

[M09] Müller, S. (2009). *The Binomial Approach to Option Valuation*.PhD thesis.

[My92] Myneni, R. (1992). *The Pricing of the American Option*. The Annals of Applied Probability 2, 1:1-23.

[P75] Petrov, V.V. (1975). *Sums of Independent Random Variables.* Springer-Verlag, Berlin.

[Pl54] Plackett, R.L. (1954). *A Reduction Formula for Normal Multivariate Integrals.* Biometrica 41, 351-360.

[PP12] Petersen, K.B. and Pedersen M.S. (2012). *The Matrix Cookbook.* http://matrixcookbook.com.

[PR69] Prohorov, Y.V. and Rozanov Y.A. (1969). *Probability Theory.* Springer-Verlag, Berlin.

[PV94] Pelsser, A. and Vorst, T. (1994). *The Binomial Model and the Greeks.* The Journal of Derivatives 1, 3:45-49.

[R95] Ritchken, P. (1995). *On Pricing Barrier Options.* The Journal of Derivatives 3, 2:19-28.

[RB79] Rendleman, R.J. and Bartter, B.J. (1979). *Two-State Option Pricing.* Journal of Finance, 34:1093-1110.

[S04] Shreve, S.E. (2004). *Stochastic Calculus for Finance I: the Binomial Asset Pricing Model.* Springer.

[Sha79] Sharpe, W.F., (1979). *Investments.* Prentice-Hall, Endlewood Cliffs, US.

[Sh84] Shiryaev, A.N. (1984). *Probability.* Springer.

[T99] Tian, Y.S. (1999). *A Flexible Binomial Option Pricing Model.* Journal of Future Markets, 19:817-843.

[W07] Wilmott, P. (2007). *Paul Wilmott Introduces Quantitative Finance.* John Wiley & Sons, Second edition.

[Wil62] Wilks, S.S. (1962). *Mathematical Statistics.* John Wiley & Sons.

[WW02] Walsh, J.B. and Walsh O.D. (2002). *Embedding and the Convergence of the Binomial and Trinomial Tree Schemes.* Technical report, University of British Columbia, Department of Mathematics.

Curriculum Vitae

Name:	Alona Bock

Education:

09/2001-06/2004	Lyceum "Intellect", Kyiv, Ukraine
09/2004-06/2008	Bachelor of Science studies in Statistics, Mathematics Department, Taras Schevchenko University, Kyiv
10/2008-03/2011	Master of Science studies in Mathematics, University of Kaiserslautern
04/2011-04/2014	PhD studies in Financial Mathematics, University of Kaiserslautern
10/2010-03/2011	Research assistant; Fraunhofer ITWM, Kaiserslautern
04/2011-10/2012	Research assistant; Department of Mathematics, University of Kaiserslautern

Wissenschaftlicher Werdegang

Name:	Alona Bock

Ausbildung:

09/2001-06/2004	Lyceum "Intellect", Kiew, Ukraine
09/2004-06/2008	Studium zum Bachelor of Science in Statistik, Fachbereich Mathematik an der Taras Schevchenko Universität, Kiew
10/2008-03/2011	Studium zum Master of Science in Mathematik, der TU Kaiserslautern
04/2011-04/2014	Promotionsstudentin im Fachgebiet Finanzmathematik, TU Kaiserslautern
10/2010-03/2011	wiss. Hilfkraft am Fraunhofer ITWM, Kaiserslautern
04/2011-10/2012	wiss. Mitarbeiter am Fachbereich Mathematik der TU Kaiserslautern