DISCUSSION PAPER NO. B–309

BINOMIAL MODELS FOR OPTION VALUATION – EXAMINING AND IMPROVING CONVERGENCE

DIETMAR LEISEN AND MATTHIAS REIMER

Abstract. Binomial models, which rebuild the continuous setup in the limit, serve for approximative valuation of options, especially where formulas cannot be derived mathematically due to properties of the considered option type. Unfortunately, even with the valuation of European call options distorting irregularities occur calculating prices along iteration of tree refinements. For the first time, these convergence patterns in binomial option valuation models are examined and it is proved order of convergence one for the Cox–Ross–Rubinstein[79] model as well as for the tree parameter selections of Jarrow and Rudd[83], and Tian[93]. Then, we define new binomial models, where the calculated option prices converge smoothly to the Black–Scholes solution and we achieve order of convergence two with much smaller initial error. Notably, solely the formulas to determine the constant up– and down–factors change. Finally, all tree approaches are compared with respect to speed and accuracy calculating relative root–mean–squared error of approximative option values for a sample of randomly selected parameters across a set of refinements. This approach was used in Broadie and Detemple[94]. Approximation of American type options with the new models exhibits order of convergence one but smaller initial error than with previously existing binomial models.

1. Introduction

With arbitrage pricing theory, the present value of any derivative security is obtained by calculating the initial cost of a dynamically payoff–replicating portfolio with proportions of the underlying asset and savings. The usual assumption of idealized financial markets leads to a model where this underlying asset is characterized by a stochastic lognormal diffusion process. Consequently, in this setup the replicating portfolio is reconstructed continuously.

Alternatively, in a simplified approach, asset price changes are decomposed into a sequence of Bernoulli steps implying a time– and state–discrete replicating strategy. Due to the central limit theorem, if this binomial tree is defined correspondingly to the continuous framework, in particular with matching distribution parameters, this model coincides with the continuous model when refined infinitly.

For practical applications however, a binomial tree of only fixed length approximates the continuous set of trading occurences and asset prices always by covering but some finite range of asset prices with a discrete grid structure. With every application of binomial trees, one inevitably must examine the approximation quality by careful consideration of the approximation result with changing tree refinements.

Investigations reveal, that the acquired degree of precision in binomially computed option prices in comparison to a continuously calculated option price varies with the refinement of the binomial trees in a bumpy manner. The option prices unsymmetrically oscillate with changing amplitude around the Black–Scholes solution for a European call option.

When considering lattice approaches primarily as a means to design a limit distribution, we desire that an approximation method should have a convergence speed as fast as possible. Convergence speed is

Date. March 1995. This version: December 20, 1995.

JEL Classification. G13.

Key words and phrases. binomial model, option valuation, order of convergence, convergence pattern.

We are grateful to Hans Föllmer, Stefan Look, Dieter Sondermann a.o. for comments and assistance. Financial support from the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 and from the Graduiertenkolleg, both at Rheinische Friedrich-Wilhelms-Universität Bonn is gratefully acknowledged. Document typeset in LaTeX.

measured here by the degree of change with iterated refinement in absolute difference of binomial price and continuous solution price. Furthermore, the approximation should improve "smoothly", that is improvement with each increase of the refinement regardless of the given parameter constellation.

Our paper goes beyond the findings of the existing literature in several ways. First, for several existing lattice approaches the order of convergence is shown and proved. Astonishingly, convergence speed of binomially computed option prices has not been examined technically so far. We present here a theorem which characterises the order of convergence in terms of the by one reduced highest order contained in first, second, and third moments and a pseudomoment. This allows for a comparison of methods with well measurable and quantifiable criteria.

Second, the presentation of models with higher order of convergence and smooth convergence pattern follows. We exploit findings of the mathematical literature of normal approximations to the binomial function. The approach is general with respect to the choice of inverted normal approximations. In particular, there is the Peizer and Pratt[68] method which consists of a convolution of several methods. Constructing a sequence of binomial models and each time iterating the tree refinement, a stepwise approximation improvement is revealed. All the presented models, in particular the previously existing and the newly developed models, have the same computation speed given the same tree refinement, because only the formulas to calculate the tree parameters change. Importantly, lattice approaches establish much more than a vehicle to achieve a certain limit distribution of future asset prices. Here, the arbitrage relationships which imply the replicating portfolio can be characterized clearly, whereas this theoretical construction is somewhat concealed in the continuous setup. Notably, these neat properties are retained entirely in the newly established models with improved approximation properties. From the theoretical point of view all the considered models can be used interchangeably. Consequently, we have shown how the applicability of lattice approaches can be improved tremendously.

Finally, we give some numerical examples to underline the strength of the new approximation. A method recently presented by Broadie and Detemple[94] follows, where option prices are computed for a large sample of random parameters and then the relative standard deviation to the true solution is calculated and compared to computation time with increasing refinement. Graphically it is shown that previous models stay behind drastically with respect to accuracy. On top, using the same sample of parameters it is shown that our tree models perform better than all previous lattice approaches when computing American type option prices. Here is the specific attraction of fast performing models, because binomial models can approximate prices of those derivative securities for which explicit formulas cannot be derived in the continuous setup.

A binomial option pricing model was first developed simultaneously by Cox, Ross, and Rubinstein[79] (CRR) and Rendleman and Bartter[79]. CRR presented the fundamental economic principles of option pricing by arbitrage considerations in the most simplest manner. By application of a central limit theorem they proved that their model merges into the Black and Scholes model when the time steps between successive trading instances approach zero. In their proof the binomial functions are transformed to standard normal functions repectively. This serves as a starting point for our construction of improved models. Additionally, the model was used to evaluate American type options and options on assets with continuous dividend payments.

In the meantime, innumerable contributions to lattice approaches have been published. We must excuse that not all of them can be mentioned here.

Jarrow and Rudd[83] (JR) constructed a binomial model where the first two moments of the discrete and continuous time return processes match. Furthermore, a probability measure equal to one half results.

Boyle[88] constructed a trinomial lattice, which is fixed up to some arbitrary parameter λ, which is determined heuristically. Although this model lacks a universal solution, he realizes indeed that there are potentialities to improve lattice approaches by an ingenious choice of parameters.

Omberg[88] deduced a whole family of lattice trees using the technique of Gauss–Hermite quadrature as solution to the backward recursive integration problem. Unfortunately, trees with four or more vertices do not recombine properly and interpolation methods must be applied to keep a tracktable grid of asset prices. Notably, he recognizes that even with a 20th order Gauss–Hermite jump process the location of the exercise price within the tree structure may cause trouble.

Tian[93] proposed binomial and trinomial models where the model parameters are derived as unique solutions to equation systems, established from sufficient conditions to acquire weak convergence due to the Lindeberg theorem, supplemented to use remaining degrees of freedom to equalize further moments of the continuous and discrete asset–distributions. Unfortunately, this interesting contribution lacks to support the ideas by mathematical arguments.

Numerous adjustments have been introduced to apply lattice approaches to various types of options. There is the broad field of exotic options. Already Cox and Rubinstein[85] presented an adjustment for the valuation of Down–and–out calls. Hull and White[93] modified the original CRR–model for the pricing of path dependent exotic options by linear or quadratic interpolation. Recently, Cheuk and Vorst[94] presented a model where the payoff of Lookback options itself is modelled in a lattice, thus resolving the path dependency. Whereas our paper does not focus directly on the pricing of complex payoff themes, we view our contribution as a starting point for the derivation of methods with superior accuracy there.

Further extensions to the field of lattice approaches involve the transfer to the pricing of derivative contracts with multiple underlying securities (see He[90]). Other authors devote research to the construction of "simple" binomial lattices, that is construction principles where pricepaths recombine properly even when more complex models such as those with state varying volatility functions are considered (e.g. see Nelson and Ramaswamy[90], Li[92]).

2. On discrete and continuous models

Black and Scholes assume that trading at financial markets proceeds continuously in time. Asset-price dynamics are described by

$$dS_t = r\, S_t\, dt + \sigma\, S_t\, dW_t \tag{1}$$

where r is the instantaneous expected return of the underlying asset $S = (S_t)_{t\geq 0}$ if immediately the risk-neutrality argument of Harrison and Pliska[81] is used, σ^2 is the instantaneous variance of the return, and $(W_t)_{t\geq 0}$ is a standard Gauss–Wiener process on a suitable probability space $(\Omega, \mathcal{F}, P)$.

Within this model, for any fixed European call option, having strike price K and maturity date T, a hedge portfolio can be constructed containing solely the underlying asset S and a savings account with riskless borrowing and lending at r, which perfectly replicates the value of this option [1] at each instant t of time.

It is this equilibrium connection which results in the Black-Scholes differential equation for the unknown function c:

$$\frac{\partial c}{\partial t} + r\, S\, \frac{\partial c}{\partial S} + \frac{1}{2}\, \sigma^2\, S^2\, \frac{\partial^2 c}{\partial S^2} = r\, c \tag{2}$$

The solution to the differential equation with boundary condition $f : x \mapsto (x - K)^+$ as payoff function is given by the Black-Scholes option pricing formula:

$$\begin{aligned} c(t, S) &= S \cdot \mathcal{N}(d_1) - K \cdot e^{-r(T-t)} \mathcal{N}(d_2) \\ d_{1,2} &= \frac{\ln(S/K) + (r \pm \frac{1}{2}\sigma^2)(T-t)}{\sigma\sqrt{T-t}} \end{aligned} \tag{3}$$

[1] The value depends on time t and asset price S under consideration. We adopt the common notation $c(t, S)$ for this function.

where $\mathcal{N}(\cdot)$ is the cumulative standard normal distribution function. Due to Harrison and Pliska[81] we know that this can be written as

$$c(t,S) := e^{-r(T-t)} E_W\left[f(S_T)\right] \tag{4}$$

Transferring this framework into the simplifying binomial structure induces several adjustments. Suppose there is given a second probability space $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{P})$. Given a prespecified number n of trading dates, we suppose that trading occurs only at equidistant spots of time $t_i^n \in \{0 = t_0^n, \cdots, t_n^n = T\}$ with $t_{i+1}^n - t_i^n = \frac{T}{n}$ $(i = 0, \ldots, n-1)$. The one–period returns $\overline{R}_{n,i} \quad (i = 1, \ldots, n)$ are modelled by two point iid binomial random variables

$$\overline{R}_{n,i} = \begin{cases} u_n & \text{with probability} & p_n \\ d_n & \text{with complementary probability} & 1 - p_n \equiv q_n \end{cases} \tag{5}$$

on $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{P})$. Thus for all $k = 0, \ldots, n$ the discrete asset price dynamics at time t_k are described by

$$\overline{S}_{n,k} = S_0 \cdot \prod_{i=1}^{k} \overline{R}_{n,i} \tag{6}$$

The specification of the one–period returns is a complete description of the discrete dynamics $\overline{S}_n$. We call such a finite sequence $\overline{R}_n = (\overline{R}_{n,i})_{i=1,\ldots,n}$ a **lattice (tree)**.
The parameters u_n, d_n, p_n, n differ from lattice to lattice, but remain constant throughout a specific lattice. We call a method which assigns to each refinement n a lattice a **lattice approach**. In the sequel we will suppose always that there is given a whole sequence of lattices. One should think of it as a triangular array

$$\begin{array}{llll} \overline{R}_{1,1} & & & \\ \overline{R}_{2,1} & \overline{R}_{2,2} & & \\ \overline{R}_{3,1} & \overline{R}_{3,2} & \overline{R}_{3,3} & \\ \vdots & \vdots & \vdots & \ddots \end{array}$$

where each row represents a lattice.
Several different lattice–approaches have been proposed. They all take into account that the risk–neutrality argument of Harrison and Pliska[81] implies that the expected one–period return $\overline{E}[\overline{R}_{n,1}]$ must equal the one period return of the riskless bond $r_n = \exp\{r\Delta t_n\}$. The remaining degrees of freedom are resolved in different ways as shown in the following table:

CRR[79]	JR[83]	TIAN [93]
$u_n = \exp\left\{\sigma\sqrt{\frac{T}{n}}\right\}$ $d_n = \exp\left\{-\sigma\sqrt{\frac{T}{n}}\right\}$	$u_n = \exp\left\{\mu'\frac{T}{n} + \sigma\sqrt{\frac{T}{n}}\right\}$ $d_n = \exp\left\{\mu'\frac{T}{n} - \sigma\sqrt{\frac{T}{n}}\right\}$	$u_n = \frac{r_n v_n}{2}\left(v_n + 1 + \sqrt{v_n^2 + 2v_n - 3}\right)$ $d_n = \frac{r_n v_n}{2}\left(v_n + 1 - \sqrt{v_n^2 + 2v_n - 3}\right)$
	$\mu' = r - \frac{1}{2}\sigma^2$	$r_n = \exp\left\{r\frac{T}{n}\right\}$ $v_n = \exp\left\{\sigma^2\frac{T}{n}\right\}$

TABLE 1. remaining degrees of freedom are resolved in different ways in the considered lattice–approaches of CRR, JR, and Tian; thus there are alternative definitions of tree parameters

Within all these models, a hedge portfolio can be constructed which perfectly replicates the value of a European call option at each discrete point of time t_i^n regardless whether the asset price increases to $\overline{S}_{n,i+1} = u_n\overline{S}_{n,i}$ or decreases to $\overline{S}_{n,i+1} = d_n\overline{S}_{n,i}$. Nodewise payoff replication with respect to up- or down-movements requires

$$\begin{aligned}\Delta_i \cdot u_n \cdot \overline{S}_{n,i} + r_n \cdot B_i &= c_n\left(t_{i+1}^n, u_n \cdot \overline{S}_{n,i}\right)\\ \Delta_i \cdot d_n \cdot \overline{S}_{n,i} + r_n \cdot B_i &= c_n\left(t_{i+1}^n, d_n \cdot \overline{S}_{n,i}\right)\end{aligned}$$

to hold.

From this, the proportion Δ_i of the underlying asset and amount B_i of savings in the replicating portfolio can be derived, the value at t_i^n is interpretable as discounted expected value of prices at t_{i+1}^n with martingale measure equal to $p_n = (r_n - d_n)/(u_n - d_n)$.

Here, this equilibrium connection is subsumed for all discrete steps in a binomial formula, of which the first was presented by CRR:

$$c_n(0 = t_0^n, S_0) = r_n^{-n}\overline{E}[f(\overline{S}_{n,n})] \tag{7}$$

$$= r_n^{-n}\sum_{j=0}^{n}\binom{n}{j}p_n(1-p_n)^{n-j}[\overline{S}_{n,n} - K]^+ \tag{8}$$

and equivalently:

$$c_n(0 = t_0^n, S_0) = S_0\,\Phi[a; n, p_n'] - K\,r_n^{-n}\,\Phi[a; n, p_n] \tag{9}$$
$$\text{where}\quad p_n = \frac{r_n - d_n}{u_n - d_n} \qquad p_n' = \frac{u_n}{r_n}\cdot p_n \qquad a = \left\lfloor \frac{\ln(K/S_0) - n\,\ln d_n}{\ln u_n - \ln d_n}\right\rfloor$$

where $\Phi[\cdot]$ denotes the binomial distribution function.

3. Examining convergence

By an easy application of the central limit theorem one immediately proofs convergence of the binomial distribution terms in (9) to its respective normal distributed terms in (3). Essentially, existing lattice approaches only differ in the way how this limit result is acquired. The proceeding involves differing definition of the tree parameters u_n and d_n. For all three models, the requirements to achieve the same limit distribution are fulfilled. But beyond that, the distinct approaches do not reveal properties suggesting superiority or inferiority in terms of convergence quality. Essentially, equating moments merely assures convergence to a distribution with matching parameters. Yet, computing option prices within a tree constructed this way does not lead to the best achievable estimation results. Figure 1 depicts a typical pattern resulting from option price calculations along iteration of the tree refinement. The straight line indicates the Black–Scholes solution. For the binomial calculations we mark each computation and draw a connecting line to emphasize the changing results. Equally for all three models we find oscillations and wave patterns. There occur intervals with decreasing error but followed by intervals with again increasing error. [2] Crucially, the accuracy of approximation is influenced by a property, inherent in all these tree models, illustrated by an example. In figure 2 the vertical dashed lines mark the terminal nodes of a 10–step binomial tree. The smooth line gives the continuous density function along a continuous set of asset prices. The rectangulars surrounding the terminal nodes form a histogram indicating the connection of continuous and discrete density. But contrary to this illustration these probability masses are concentrated at the terminal nodes, implying that splitting the probability mass with the exogenous fixed strike price occurs in whole rectangulars only. Translating into a binomial distribution function we have a piecewise constant function with jumps at each terminal node. But the set of binomial asset

[2]We didn't depict calculations of an at–the–money option with the CRR–model, because the homogenous convergence pattern there represents a unique exception.

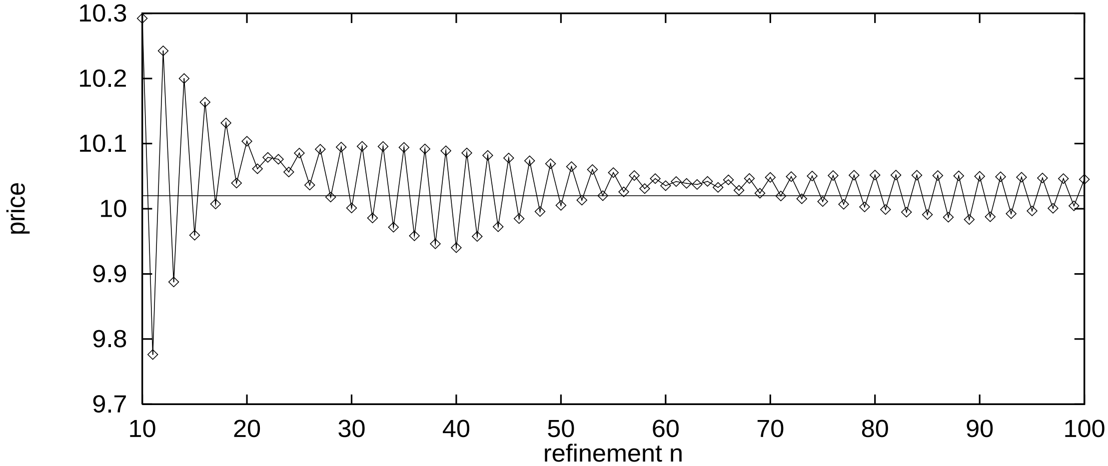

FIGURE 1. typical pattern resulting from option price calculations with binomial models: example with CRR-Model and the following selection of parameters: $S = 100$, $K = 110$, $T = 1$, $r = 0.05$, $\sigma = 0.3$, $n = 10, \ldots, 150$

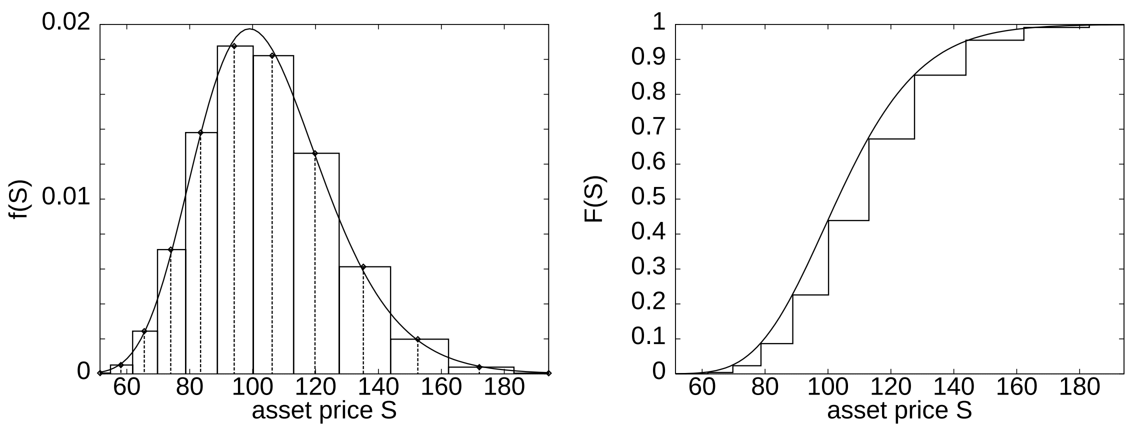

FIGURE 2. a histogram usually depicts the connection to the continuous density function; separation of the probability mass only with whole rectangulars gives a distribution function with steps; example drawn from a 10–step binomial tree

prices shifts with each iteration of the tree refinement. Fluctuating locations of jumps in the distribution function result. The separation of the probability mass bounces back and forth with changing step patterns. Thus, the binomial structure itself in combination with a strike price located independently of the tree grid induces irregularities which distort the approximation result. Except of special cases, the computed option prices oscillate and converges wavy to the Black-Scholes solution[3].
Alternatively, in order to visualize that solely the fluctuating separation causes existing convergence patterns take the Black-Scholes formula and adapt the strike price in the normal components to the implicite separation rule of the binomial models, which is a continuity correction of one half. Interestingly, any existing convergence patterns can be reproduced, displaced with respect to the distribution error only. Consequently, merely the separation problem signs responsible for all existing convergence patterns. The distribution error can be quantified indeed by adapting the Black-Scholes formula once more, now with a series expansion. Thus, binomial option prices can be duplicated by this twofold adaptation. Oscillation is produced because the strike jumps over rectangulars with even and odd refinements, waves describe the relative movement between surrounding terminal nodes, which itself change in value though.

[3]this property is sometimes called "even-odd-problem"

Now suppose we take one sequence of lattices in particular. We see already in figure 1 that discrete and continuous prices do not match in general; we find an error $e_n = |c(0, S_0) - c_n(0, S_0)|$.
Due to the central limit theorem we have $\lim_{n\to\infty} e_n = 0$, which means prices calculated by a sequence of lattices converges to its respective Black–Scholes price.
From (4) and (7) our error takes the form:

$$e_n := e^{-rT} \; \left| E\left[f(S_T)\right] - \overline{E}\left[f(\overline{S}_{n,n})\right] \right| \tag{10}$$

We remember when observing calculations with refinement n, one typically observes wavy patterns. The adequate formal description of the approximation quality is to characterize convergence by convergence of the worst case. That means we are looking for a description of a reasonable upper bound for error e_n. For this we use the mathematical concept of "order of convergence". Restated in our specific case here, we make the following

Definition. Let $f : x \mapsto (x - K)^+$ be a European call option. A sequence of lattices **converges with order** $\rho > 0$ if there exists a constant $\kappa > 0$ such that

$$\forall n \in \mathbb{N} : e_n \leq \frac{\kappa}{n^\rho} \tag{11}$$

A lattice–approach **converges with order** $\rho > 0$ if for all S_0, K, r, σ, T the specified sequence of lattices converges with order $\rho > 0$.
In the sequel we will often write shortly $e_n = \mathcal{O}(\frac{1}{n^\rho})$ for this, too.
Please note that convergence of prices is implied by any order greater than 0. Moreover we remark that higher order means "quicker" convergence. Thus the theoretical concept of order of convergence is not unique: a lattice approach with order ρ has also order $\tilde{\rho} \leq \rho$.
At first glance the concept of order of convergence may seem very theoretical. In fact it is very easy to observe in actual simulations: in figures we plot the error e_n against the refinement n on a log-log-scale. Because of $\log \frac{\kappa}{n^\rho} = \log \kappa - \rho \log n$ the bounding function $\frac{\kappa}{n^\rho}$ becomes a straight line with slope equal $(-\rho)$ and shift κ.
For example in figure 3 obviously order of convergence 1 is suggested. We could present lots of these

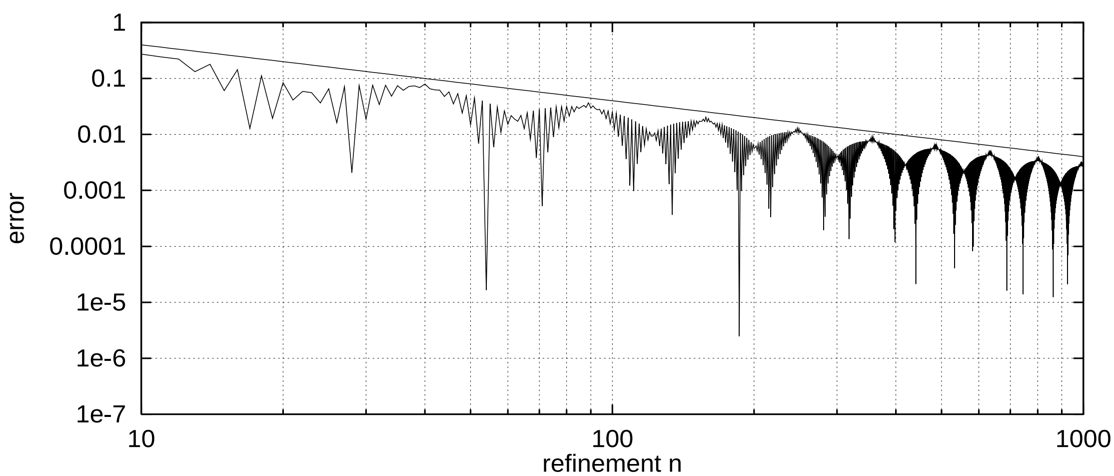

FIGURE 3. graphical representation and examination of the error bound; x–axis and y–axis with log–scale; example with CRR-Model and the following selection of parameters: $S = 100$, $K = 110$, $T = 1$, $r = 0.05$, $\sigma = 0.3, n = 10, \ldots, 1000$

figures. All suggest order of convergence 1. Our aim is to identify unique criteria which determine the order of convergence of a specific lattice–approach under consideration in mathematical terms. As a starting point we make the following

Definition. For a sequence of lattices $(\overline{R}_n)_{n\in\mathbb{N}}$ we call for all $n \in \mathbb{N}$:

$$\mathfrak{m}_n^1 := \overline{E}\left[\overline{R}_{n,1} - 1\right] - E\left[R_{n,1} - 1\right] \quad (12)$$

$$\mathfrak{m}_n^2 := \overline{E}\left[\left(\overline{R}_{n,1} - 1\right)^2\right] - E\left[\left(R_{n,1} - 1\right)^2\right] \quad (13)$$

$$\mathfrak{m}_n^3 := \overline{E}\left[\left(\overline{R}_{n,1} - 1\right)^3\right] - E\left[\left(R_{n,1} - 1\right)^3\right] \quad (14)$$

moments and

$$\mathfrak{p}_n := \overline{E}\left[\left(\ln \overline{R}_{n,1}\right)\left(\overline{R}_{n,1} - 1\right)^3\right] \quad (15)$$

pseudo–moment

Here and in the sequel for any $n \in \mathbb{N}$ we denote by $R_n = (R_{n,i})_{i=1,\dots,n}$ the continuous return between times t_i^n and t_{i+1}^n. They are iid random variables on $(\Omega, \mathcal{F}, P)$ such that $S_{t_k^n} = S_0 \prod_{i=1}^k R_{n,i} \quad \forall k = 0, \dots, n$.

Our moments are a generalisation of the ordinary moments as they are mainly the difference between the ordinary moments of discrete and continuous approach. The form of the pseudo–moment is of technical nature as it results from our proof of the following theorem. We defined $\mathfrak{m}_n^1$, too. In the case here, however, we always get $\mathfrak{m}_n^1 = 0$ because of the risk neutrality argument of Harrison and Pliska[81].

The ordinary moments are of great importance in the probabilistic literature. For example they play a central role in the central limit theorem, as by means of the Ljapunoff condition weak convergence can be checked easily by checking convergence of the first two and one higher moment.

In the case here, where we have a discrete approximation of a continuous framework, it is not surprising that order of convergence is determined by the difference of the ordinary moments, that is by that of our moments. This is exactly what the following theorem says.

Theorem 1. Let $(\overline{R}_n)_{n\in\mathbb{N}}$ be a sequence of lattices and $\mathfrak{m}_n^2, \mathfrak{m}_n^3, \mathfrak{p}_n$ its respective (pseudo–) moments. The order of convergence is the smallest order contained in $\mathfrak{m}_n^2, \mathfrak{m}_n^3$ and $\mathfrak{p}_n$ reduced by 1, but not smaller than 1, that is :

$$\exists\ \kappa(S_0, K, r, \sigma, T) : e_n \le \kappa \cdot \left\{ n \cdot \left(m_n^2 + m_n^3 + \mathfrak{p}_n\right) + \frac{1}{n} \right\} \quad (16)$$

Proof. [4] Denote by $(\hat{\Omega}, \hat{\mathcal{F}}, \hat{P})$ the product space of $(\Omega, \mathcal{F}, P)$ and $(\overline{\Omega}, \overline{\mathcal{F}}, \overline{P})$. For all $n \in \mathbb{N}$ and $k = 0, \dots, n$ let $\mathcal{A}_{n,k} = \sigma(\overline{S}_{n,i} | i \le k)$.

Since $f(\overline{S}_{n,n}) = c(T, \overline{S}_{n,n})$ we can write

$$\begin{aligned} \hat{E}\left[f(\overline{S}_{n,n})\right] &= \hat{E}\left[c(T, \overline{S}_{n,n})\right] \quad \text{and} \\ e^{-rT}\hat{E}\left[f(S_T)\right] &= c(0, S_0) \; = \; \hat{E}\left[c(0, S_0)\right] \end{aligned}$$

Therefore we have:

$$e_n = \left| \hat{E}\left[e^{-rT} c(T, \overline{S}_{n,n}) - c(0, S_0) \mid \mathcal{A}_{n,k}\right] \right|$$

Discretization of time–axis :

$$e_n = \left| \hat{E}\left[\sum_{k=0}^{n-1} e^{-rt_k^n} \left\{ e^{-r\Delta t} c(t_{k+1}^n, \overline{S}_{n,k+1}) - c(t_k^n, \overline{S}_{n,k}) \right\} \;\Big|\; \mathcal{A}_{n,k} \right] \right|$$

[4] At first glance an easy description of the order of convergence seems possible as follows: With the theorems of Berry[41] and Esséen[45] the order of convergence of the approximation of a normal by a binomial distribution–function can be calculated. Then, immediately by formulas (3) and (9) we get the order of convergence of e_n. However from this we get the order $\frac{1}{2}$ for CRR, whereas we observe higher order of convergence in simulations. Other approaches as that by Butzer and Hahn[75] apply a Taylor–expansion. But this approach is not applicable since the payoff–function of the European call option is well known to be not differentiable at $S = K$. In our proof we proceed in a very specific way to circumvent this problem. It relies on the observation that for times $t < T$ $c(t, \cdot) \in C^\infty(\mathbb{R})$. Therefore we apply a trick used by Kloeden and Platen[92] to a different setting: a discretization of time–axis and the representation of the error e_n as a sum of relative errors at each time t_i^n.

The Black–Scholes price is locally riskless, that means :

$$\hat{E}\left[e^{-r\Delta t}c(t_{k+1}^n, R_{n,k+1}\overline{S}_{n,k}) - c(t_k^n, \overline{S}_{n,k}) \mid \mathcal{A}_{n,k}\right] = 0$$

Now it follows:

$$\begin{aligned} e_n &= \left|\hat{E}\left[\sum_{k=0}^{n-1} e^{-rt_{k+1}^n} \left\{\hat{E}\left[c(t_{k+1}^n, \overline{S}_{n,k+1}) \,\Big|\, \mathcal{A}_{n,k}\right] - \hat{E}\left[c(t_{k+1}^n, \overline{S}_{n,k}\cdot R_{n,k+1}) \,\Big|\, \mathcal{A}_{n,k}\right]\right\}\right]\right| \\ &= \left|\hat{E}\left[\sum_{k=0}^{n-1} e^{-rt_{k+1}^n} \left\{\hat{E}\left[c(t_{k+1}^n, \overline{S}_{n,k+1}) - c(t_{k+1}^n, \overline{S}_{n,k}) \,\Big|\, \mathcal{A}_{n,k}\right]\right.\right.\right. \\ &\qquad \left.\left.\left. -\hat{E}\left[c(t_{k+1}^n, \overline{S}_{n,k}\cdot R_{n,k+1}) - c(t_{k+1}^n, \overline{S}_{n,k}) \,\Big|\, \mathcal{A}_{n,k}\right]\right\}\right]\right| \end{aligned}$$

The last time point $k = n-1 \Leftrightarrow t_{k+1}^n = T$ must be evaluated separately as $\mathcal{O}\left(\frac{1}{n}\right)$. This is an easy, however lengthy and technical matter.
All other time points $0 \le k < n-1$ can be evaluated by a one–dimensional Taylor series expansion of $c(t_k^n, \cdot)$ around $\overline{S}_{n,k}$. If we denote its remainder term by R_3, the error e_n is bounded by $\mathcal{O}(1/n)$ plus the following term:

$$\begin{aligned} \Bigg|\hat{E}\Bigg[\sum_{k=0}^{n-2} e^{-rt_{k+1}^n}\Bigg\{ \quad & \hat{E}\Big[\ \frac{\partial c}{\partial S}(t_{k+1}^n, \overline{S}_{n,k})(\overline{S}_{n,k+1} - \overline{S}_{n,k}) + \frac{\partial^2 c}{\partial S^2}(t_{k+1}^n, \overline{S}_{n,k})(\overline{S}_{n,k+1} - \overline{S}_{n,k})^2 \\ & + \frac{\partial^3 c}{\partial S^3}(t_{k+1}^n, \overline{S}_{n,k})(\overline{S}_{n,k+1} - \overline{S}_{n,k})^3 + R_3(t_{k+1}^n, \overline{S}_{n,k+1}, \overline{S}_{n,k}) \,\Big|\, \mathcal{A}_{n,k}\Big] \\ - \quad & \hat{E}\Big[\ \frac{\partial c}{\partial S}(t_{k+1}^n, \overline{S}_{n,k})(\overline{S}_{n,k}\, R_{n,k+1} - \overline{S}_{n,k}) + \frac{\partial^2 c}{\partial S^2}(t_{k+1}^n, \overline{S}_{n,k})(\overline{S}_{n,k}\, R_{n,k+1} - \overline{S}_{n,k})^2 \\ & + \frac{\partial^3 c}{\partial S^3}(t_{k+1}^n, \overline{S}_{n,k})(\overline{S}_{n,k}\, R_{n,k+1} - \overline{S}_{n,k})^3 + R_3(t_{k+1}^n, \overline{S}_{n,k}\, R_{n,k+1}, \overline{S}_{n,k}) \,\Big|\, \mathcal{A}_{n,k}\Big]\Bigg\}\Bigg]\Bigg| \end{aligned}$$

One has:

$$\begin{aligned} \hat{E}\left[\frac{\partial c}{\partial S}(t_{k+1}^n, \overline{S}_{n,k})(\overline{S}_{n,k+1} - \overline{S}_{n,k}) \,\Big|\, \mathcal{A}_{n,k}\right] &= \overline{S}_{n,k}\frac{\partial c}{\partial S}(t_{k+1}^n, \overline{S}_{n,k})\ \hat{E}\left[\overline{R}_{n,k+1} - 1\right] \\ \hat{E}\left[\frac{\partial c}{\partial S}(t_{k+1}^n, \overline{S}_{n,k})(\overline{S}_{n,k}\cdot R_{n,k+1} - \overline{S}_{n,k}) \,\Big|\, \mathcal{A}_{n,k}\right] &= \overline{S}_{n,k}\frac{\partial c}{\partial S}(t_{k+1}^n, \overline{S}_{n,k})\ \hat{E}\left[R_{n,k+1} - 1\right] \end{aligned}$$

Analogous results can be obtained for terms with functions $S^2\dfrac{\partial^2 c}{\partial S^2}(\cdot, S)$ and $S^3\dfrac{\partial^3 c}{\partial S^3}(\cdot, S)$.
Therefore we can bound e_n by $\mathcal{O}(1/n)$ plus the following:

$$\begin{aligned} \Bigg|\hat{E}\Bigg[\sum_{k=0}^{n-2} e^{-rt_{k+1}^n}\Bigg\{ \qquad & \overline{S}_{n,k}\frac{\partial c}{\partial S}(t_{k+1}^n, \overline{S}_{n,k})\cdot \mathfrak{m}_n^1 + \overline{S}_{n,k}^2\frac{\partial^2 c}{\partial S^2}(t_{k+1}^n, \overline{S}_{n,k})\cdot \mathfrak{m}_n^2 \\ + \quad & \overline{S}_{n,k}^3\frac{\partial^3 c}{\partial S^3}(t_{k+1}^n, \overline{S}_{n,k})\cdot \mathfrak{m}_n^3 + \hat{E}\Big[R_3(t_{k+1}^n, \overline{S}_{n,k+1}, \overline{S}_{n,k}) \,\Big|\, \mathcal{A}_{n,k}\Big] \\ - \quad & \hat{E}\Big[R_3(t_{k+1}^n, \overline{S}_{n,k}\cdot R_{n,k+1}, \overline{S}_{n,k}) \,\Big|\, \mathcal{A}_{n,k}\Big]\Bigg\}\Bigg]\Bigg| \end{aligned}$$

By a detailed (very lengthy and technical) observation of the derivatives one proves:

$$\left|\sum_{k=0}^{n-2} e^{-rt_{k+1}^n}\hat{E}\left[\overline{S}_{n,k}\frac{\partial c}{\partial S}(t_{k+1}^n,\overline{S}_{n,k})\right]\right| = \mathcal{O}(n)$$

$$\left|\sum_{k=0}^{n-2} e^{-rt_{k+1}^n}\hat{E}\left[\overline{S}_{n,k}^2\frac{\partial^2 c}{\partial S^2}(t_{k+1}^n,\overline{S}_{n,k})\right]\right| = \mathcal{O}(n)$$

$$\left|\sum_{k=0}^{n-2} e^{-rt_{k+1}^n}\hat{E}\left[\overline{S}_{n,k}^3\frac{\partial^3 c}{\partial S^3}(t_{k+1}^n,\overline{S}_{n,k})\right]\right| = \mathcal{O}(n)$$

$$\left|\hat{E}\left[\sum_{k=0}^{n-2} e^{-rt_{k+1}^n}\hat{E}\left[R_3(t_{k+1}^n,\overline{S}_{n,k+1},\overline{S}_{n,k})\mid \mathcal{A}_{n,k}\right]\right]\right| = \mathcal{O}(\mathfrak{p}_n\cdot n)$$

$$\left|\hat{E}\left[\sum_{k=0}^{n-2} e^{-rt_{k+1}^n}\hat{E}\left[R_3(t_{k+1}^n,\overline{S}_{n,k}\cdot R_{n,k+1},\overline{S}_{n,k})\mid \mathcal{A}_{n,k}\right]\right]\right| = \mathcal{O}(\frac{1}{n})$$

These proofs and the evaluation of the last time point — mentionend above — are available upon request from the authors.

If we set $v_n := \exp\{\sigma^2\frac{T}{n}\}$ we know that $E\left[R_{n,1}^k\right] = r_n^k\, v_n^{k(k-1)/2}$. Applying the binomial–formulas $(x+y)^k = \sum_{j=0}^{k}\binom{k}{j}x^j y^{n-j}$ one immediatly gets the following equivalent form of theorem 1 which is the useful form in concrete applications:

Theorem 2. Conditions as in Theorem 1. Let $\bar{\mathfrak{m}}_n^2 := \overline{E}[(\overline{R}_{n,1})^2] - r_n^2\, v_n$, $\bar{\mathfrak{m}}_n^3 := \overline{E}[(\overline{R}_{n,1})^3] - r_n^3\, v_n^3$ Then the conclusions of Theorem 1 hold with $\bar{\mathfrak{m}}_n^2, \bar{\mathfrak{m}}_n^3$ instead of $\mathfrak{m}_n^2, \mathfrak{m}_n^3$.

Notably we proved only that order of convergence equals at least one, possibly higher order could be contained, though simulations indicate the opposite, entirely. To achieve order of convergence one, the theorem states that the approximating moments of the discrete asset dynamics must converge with order two towards the moments of the continuous asset dynamics, because one degree is lost with summation over time. Furthermore, one needs the same order of convergence in the pseudomoments[5].

With the theorem we obtain a characterisation of the error boundary which only relies on for all binomial models equally well measurable properties. Thus, there is a unique criterium to compare these models with respect to convergence speed. Virtually, the theorem reduces the answer to the order of convergence of the moments and pseudo–moments.

Now, the results of the theorem can be applied readily to the binomial models under consideration.

Proposition 1. The lattice–approach proposed by CRR[79] converges with order 1.

Proposition 2. The lattice–approach proposed by JR[83] converges with order 1.

Proposition 3. The lattice–approach proposed by Tian[93] converges with order 1.

Using theorem 2 all these proofs are an easy application of the series expansion of the exponential function. They are given in Appendix B. In the following the three lattice approaches are depicted with three different strikes. One can recognize in the simulations that the error can always be dominated by an upper bound of order $\frac{1}{n}$. Please note, that the lattice approach by Tian has explicitely fixed the first three moments, thus being exactly 0. Moreover we remark that the method of determining the order of convergence from that of the moments and pseudomoments works very well. Remarkably, with the moments there occur no distorting irregularities whatsoever.

[5] This is not only a technical matter but explains why the model proposed by Tian actually does not perform better.

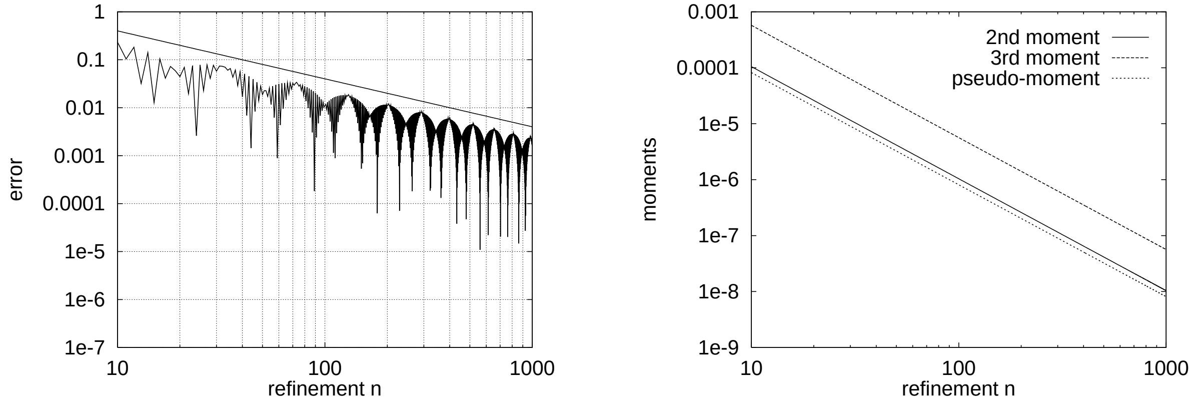

Figure 4. illustration to the result of proposition 1: order of convergence one because second and third moment and pseudo–moment have order of convergence two: example with CRR-Model and the following selection of parameters: $S = 100$, $K = 90$, $T = 1$, $r = 0.05$, $\sigma = 0.3$, $n = 10, \ldots, 1000$

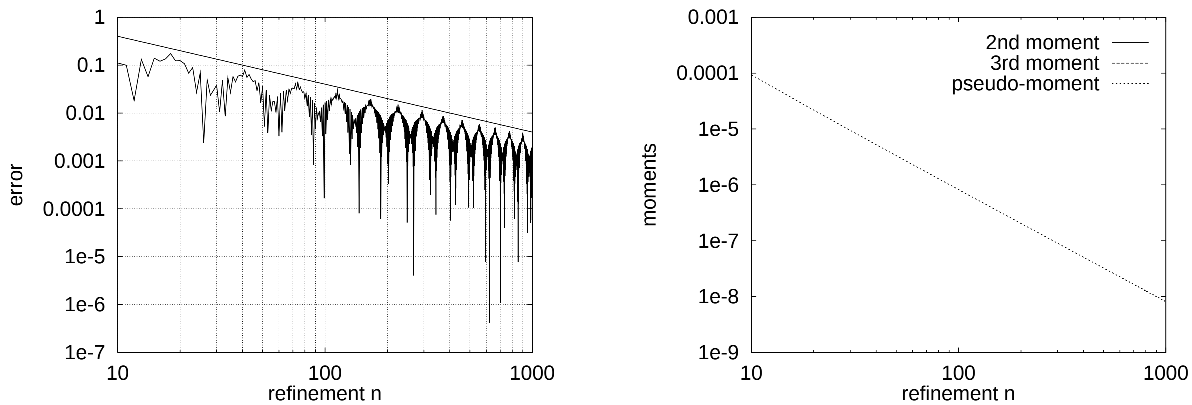

Figure 5. illustration to the result of proposition 2: order of convergence one because pseudo–moment has order of convergence two: example with Tian-Model and the following selection of parameters: $S = 100$, $K = 100$, $T = 1$, $r = 0.05$, $\sigma = 0.3$, $n = 10, \ldots, 1000$

4. Construction of binomial models with improved convergence properties

So far, we examined the problem of convergence in existing lattice approaches. We derived a general theorem to measure the order of convergence. Applying the theorem to the considered binomial models we see no significant difference in terms of convergence to the Black–Scholes solution. In this second main part of the paper we present a method to achieve improved approximation trees for the calculation of option prices. Here our aim is twofold: On the one hand, we desire to obtain binomial trees without irregularities, on the other hand we want faster convergence speed. Unfortunately, the theorem indicates that improved order of convergence as resulting from the properties of the binomial process would require higher order of convergence for the moments and pseudomoment. In principle, this cannot be achieved with a two–point random variable. Now, the only possible recipy consists of utilizing structural properties of the considered option contract.

We call the resulting method an **extended lattice approach** to emphasize the difference to the standard lattice approach.

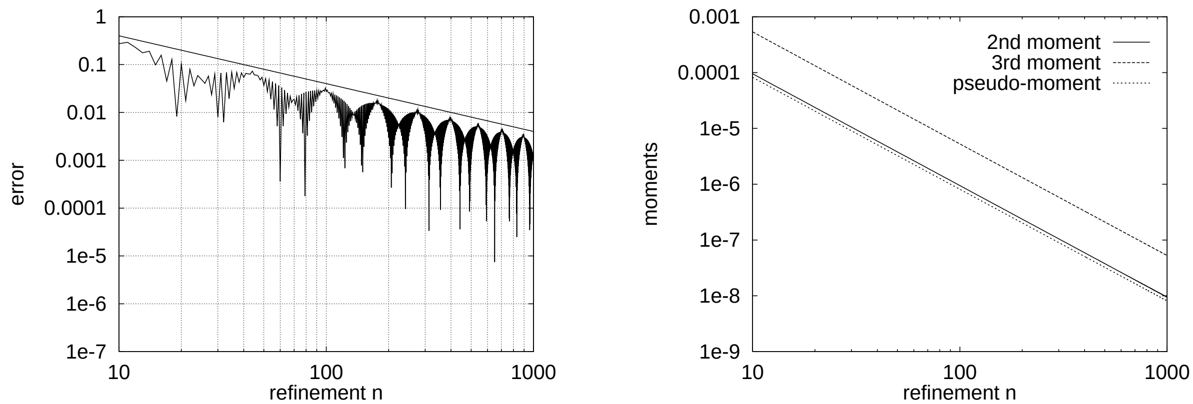

FIGURE 6. illustration to the result of proposition 3: order of convergence one because second and third moment and pseudo–moment have order of convergence two: example with JR-Model and the following selection of parameters: $S = 100$, $K = 110$, $T = 1$, $r = 0.05$, $\sigma = 0.3$, $n = 10, \ldots, 1000$

CRR showed convergence of their model to the Black–Scholes formula by examining the two binomial terms in (9) separately. Equivalently, one can view the valuation of a European call option as two approximations of type $\Phi[a; n, p] \longrightarrow \mathcal{N}(z)$, which we desire to improve here.
Despite of the simplicity, the calculation of binomial probabilities is cumbersome, because it might involve factorials of large integers or the sumation of a large number of individual terms. Therefore, normal approximations to the binomial distribution were derived. Especially, the method by Camp and Paulson[51] and the approximations of Peizer and Pratt[68] reveal a remarkable quality of accuracy[6]. With all these approaches, basically a binomially calculated true probability P is approximated with the standard normal function $\mathcal{N}(z)$, where the input is determined by some adjustment function $z = h(a; n, p)$, where in our setting a is the number of up–movements of the asset price in a n–step binomial tree with martingale measure equal to p. In the simplest case we have the de Moivre–Laplace theorem where $P = 1 - \Phi[a; n, p]$ is approximated by $P \approx \mathcal{N}\left(h(a; n, p) = \dfrac{a - n\,p}{\sqrt{n\,p\,(1-p)}}\right)$, where $\Phi[a; n, p]$ is the complementary binomial distribution function of (9). But our option pricing problem here represents the opposite direction. Computation of binomial option prices eventually involves that normal components are approximated by binomial components. Thus, for a given binomial tree refinement the adjustment function h(a;n,p) from above inverted specifies the distribution parameter $h^{-1}(z) = p$ to approximate $P = \mathcal{N}(z)$ with $P \approx 1 - \Phi[a; n, p]$. Peizer and Pratt derived the inversion formula to the Camp-Paulson method and specified the inversion formula of their method in the case with identical number of successes and fails[7].

(A) Camp–Paulson–Inversion: (universally valid)

$$h^{-1}(z) = \left(\frac{y}{x}\right)^2 \left(\frac{[9x-1][9y-1] + 3z[x(9y-1)^2 + y(9x-1)^2 - 9xyz^2]^{\frac{1}{2}}}{[9y-1]^2 - 9yz^2}\right)^{\frac{1}{3}}$$

with $x = n - a$, $y = a + 1$, z as input of the standard normal function.

[6] The reader will find some remarks to the derivation of these approximations and the citation of literature in the appendix

[7] Otherwise the inversion could be achieved numerically. Since this inversion is generally valid to any parameter selection, it could be tabulated or approximated polynomially for fixed n similar to the proceeding with the standard normal function.

(B) Peizer–Pratt–Method–1–Inversion $\left[\text{case: } a+\frac{1}{2}=n-\left(a+\frac{1}{2}\right),\ n=2a+1\right]$

$$h^{-1}(z) = 0.5 \mp \left[0.25 - 0.25 \cdot \exp\left\{-\left(\frac{z}{n+\frac{1}{3}}\right)^2 \cdot \left(n+\frac{1}{6}\right)\right\}\right]^{\frac{1}{2}}$$

(C) Peizer–Pratt–Method–2–Inversion $\left[\text{case: } a+\frac{1}{2}=n-\left(a+\frac{1}{2}\right),\ n=2a+1\right]$

$$h^{-1}(z) = 0.5 \mp \left[0.25 - 0.25 \cdot \exp\left\{-\left(\frac{z}{n+\frac{1}{3}+\frac{0,1}{(n+1)}}\right)^2 \cdot \left(n+\frac{1}{6}\right)\right\}\right]^{\frac{1}{2}}$$

So far, we talked about approximation only. Now, we will demonstrate how these findings can be used to construct CRR-like binomial models. For the construction of a binomial tree we continue similarily to the Tian approach. We setup an equation system to determine uniquely tree parameters which assure convergence. First, with d_1 and d_2 as inputs from equation (3) in $h^{-1}(z)$, we obtain p and p' as distribution parameters of the two binomial components in the binomial option pricing formula using approximation rule A, B, or C.

Then, we derive tree parameters u_n and d_n by a simple trick. The noarbitrage condition implies that $p_n = (r_n - d_n)/(u_n - d_n)$ holds. Furthermore, p'_n is defined to $p'_n = u_n/r_n \cdot p_n$. Taking these two relations as equation system which can be solved uniquely with respect to u_n and d_n, we succeed to acquire a new binomial model. The formulas below sum up the model parameters.

$$\begin{aligned} p'_n &= h^{-1}(d_1) \\ p_n &= h^{-1}(d_2) \\ u_n &= r_n \cdot \frac{p'_n}{p_n} \\ d_n &= \frac{r_n - p_n \cdot u_n}{1-p_n} \end{aligned}$$

In the paper here, we consider only three inverted normal approximations. The method is general with respect to the chosen method. In particular one could invert the simple continuity correction and apply it as approximation rule.

Taking adjustment function $h(a;n,p)$ the parameter z for the standard normal function is uniquely determined with given a,n, and p. But vice versa having parameter z, for each pair of (a,n) we get a distribution parameter p. Thus we can make a choice of the parameter $a(n)$ ex ante. The parameter $a(n)$ is chosen such that the strike price K is located in the center of the binomial tree. Regarding that usual interest focuses almost entirely on at–the–money and near–the–money options, this selection doesn't really distort the tree significantly.

In example calculations, the resulting binomial tree parameters diverge only very little from those of previous models, but astonishingly, the convergence properties with the computation of option prices change dramatically. Nevertheless, within this class of models the particular theoretical building blocks for which the CRR-model became famous are entirely preserved by construction. We have completeness and a unique replicating strategy can be derived. But moreover, this model construction profits from the attributes of the chosen normal approximation. Below, the figures demonstrate the strength of the method in approximating option prices in comparison to previously existing models.

Figure 7 shows an example of the results for the binomial model using the Camp–Paulson method. Notably, the y–axis of the lefthand figure has a range of two pennies. The approximation error declines monotonically with order of convergence one and the constant is approximately equal to one instead of four (compare with figures 4, 5, and 6). Beforehand, the parameter $a(n)$ can be chosen freely since it is determined implicitely from the , we decided to locate the strike in the center of the binomial tree. Regarding that usual interest focuses almost entirely on at–the–money and near–the–money options, this

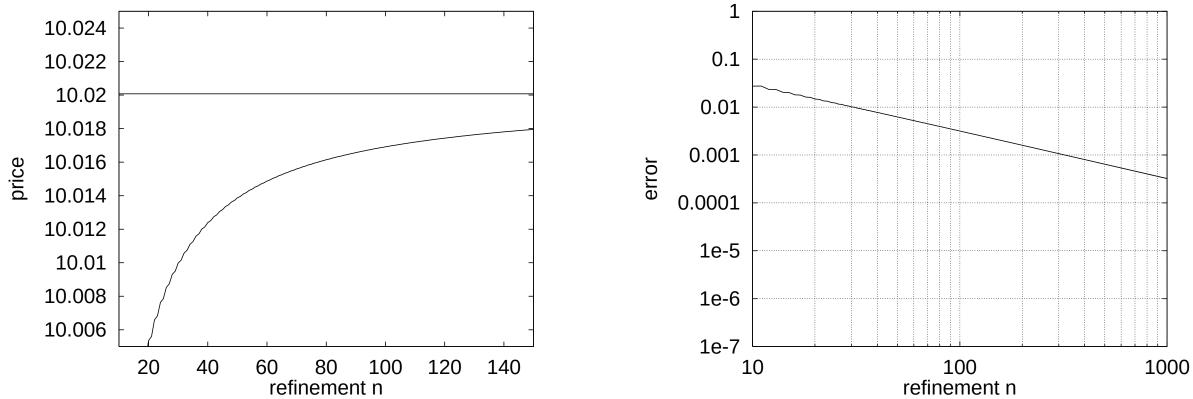

FIGURE 7. illustration of the results with the improved binomial construction theme using the inverted normal approximation of Camp–Paulson; following selection of parameters: $S = 100, K = 110, T = 1, r = 0.05, \sigma = 0.3, n = 10, \ldots, 150 \quad (n = 10, \ldots, 1000)$

selection doesn't really distort the tree significantly. Checking the results with strike location relatively to the deviation to the initial asset price no significant differences occur. Figure 8 depicts an example

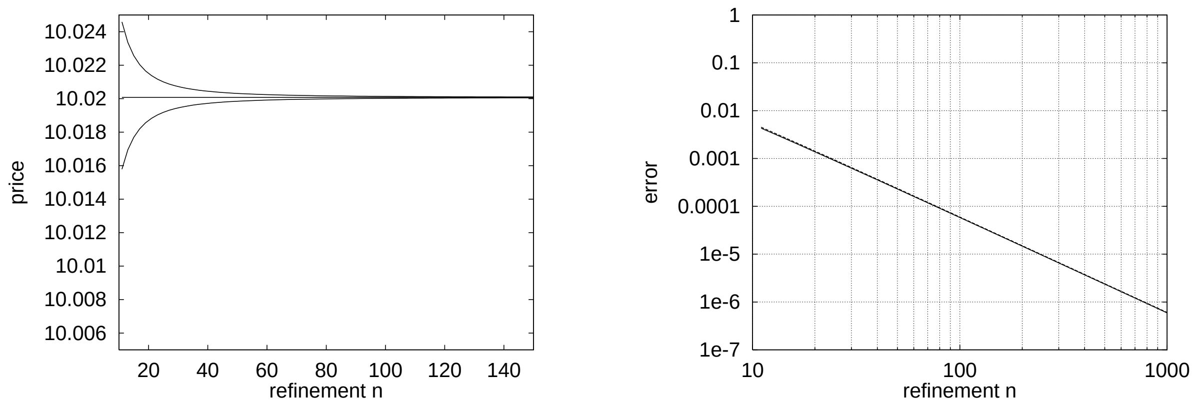

FIGURE 8. illustration of the results with the improved binomial construction theme using both inverted normal approximations of Peizer–Pratt; following selection of parameters: $S = 100, K = 110, T = 1, r = 0.05, \sigma = 0.3, n = 10, \ldots, 150 \quad (n = 10, \ldots, 1000)$

of the results with the Peizer–Pratt methods. There are two methods, which don't show significant differences in the application. Here, we take odd refinements only due to the restricted domain of the explicit inversion formula. This doesn't diminish the results because there is no meaning to any specific refinement except serving for approximation purposes. Here, we need also the center location of the strike. Again the approximation error declines monotonically but now with order of convergence two and constant approximately one. Indeed, the use of the superior inverted normal approximation improves the calculation of option prices. The remarkable performance here allows for a considerable reduction of the tree refinement to save computation time. Looking at the moments and pseudomoments, the error only declines with order of convergence one. This confirms that the improvement is achieved by utilisation of structural properties of the chosen contract type. Namely, the location of the strike within the tree grid is controlled. Consequently, we are not able to give a strict proof of the greater order of convergence in line with our theorem.

5. Numerical results

Now we compare directly the three binomial methods considered in section two with those newly developed. Below, there is a table containing example computations for European call and put options with a specific selection of parameters. Computing binomial prices for a fixed tree refinement represents only a small window of the whole approximation theme with accidental degrees of accuracy. Nevertheless, we give a table to the convenience of those readers intending the implementation of methods. Even with the very low tree refinement of $n = 25$, the outstanding performance of models using normal approximations can be recognized . More digits must be displayed to catch the degree of accuracy. Notably, care must be taken of the method to calculate the standard normal function in order to avoid distortion by the supposedly true solution. Thus, the chosen method guarantees accuracy of 7 digits. Although the tree adjustment primarily served for the improved approximation of European standard options, we show that valuable improvements for the pricing of American type options are contained. True American option values were derived using the CRR–method using 15000 tree–steps.

Strike	CRR	JR	Tian	CP	PP1	PP2	True value
European Call Options							
80	23.74082	23.76300	23.70657	23.76050	23.75822	23.75875	23.75799
90	16.13376	16.08486	16.12494	16.09619	16.09941	16.10037	16.09963
100	10.21317	10.20142	10.20418	10.12545	10.13316	10.13440	10.13377
110	6.01218	6.02481	6.01304	5.94162	5.94889	5.95015	5.94946
120	3.31890	3.33429	3.33318	3.27993	3.28258	3.28366	3.28280
European Put Options							
80	0.98926	1.01143	0.95500	1.00893	1.00665	1.00719	1.00642
90	3.03825	2.98934	3.02943	3.00068	3.00390	3.00486	3.00412
100	6.77371	6.76196	6.76472	6.68599	6.69370	6.69494	6.69431
110	12.22878	12.24141	12.22963	12.15821	12.16548	12.16675	12.16606
120	19.19155	19.20694	19.20583	19.15258	19.15523	19.15631	19.15545
American Put Options							
80	1.01842	1.03864	0.98396	1.04231	1.04264	1.04317	1.037
90	3.16580	3.12447	3.14640	3.11786	3.12832	3.12928	3.123
100	7.10823	7.10415	7.08701	7.00982	7.02858	7.02981	7.035
110	13.00108	13.01511	12.98978	12.90304	12.93136	12.93253	12.955
120	20.73344	20.74479	20.73566	20.65254	20.67576	20.67649	20.717

Table 2. a table with numerical results, a small window of the whole approximation theme with accidental degrees of accuracy, but useful for those readers intending the implementation of methods; parameters $S = 100$, $r = 0.07$, $\sigma = 0.3$, $T = 0.5$ years, $n = 25$ for all

Each considered simulation result may depend significantly on an accidentically chosen parameter set. Thus we looked for a procedure to test simultaneously across a whole set of parameters. We stick to an analysis recently conducted by Broadie and Detemple [1994] who tested several methods for the pricing of American options. There, within one analysis several methods using a large sample of randomly selected parameters are compared simultaneously over refinements with measurement of computation speed and approximation error. Computation speed is expressed by the number of option prices calculated per second. Since we stick to tree models with identical structure except for the tree parameters, for all models here, we use the speed results of Broadie, Detemple for CRR. Thus, we need not care on tuning our computer implementation of methods. The approximation error is measured by the relative root–mean–squared (RMS) error. RMS–error is defined by

$$\mathrm{RMS} = \sqrt{\frac{1}{m}\sum_{i=1}^{m} e_i^2}$$

where $e_i = (\hat{c}_i - c_i)/c_i$ is the relative error, c_i ist the true option value. $\hat{c}_i$ ist the estimated option value. To make relative error meaningful, that is to avoid senseless distortions because of very small option prices, the summation is taken over options in the dataset satisfying $c_i \geq 0.50$.

We chose the following distribution of parameters. Volatility is distributed uniformly between 0.1 and 0.6. Time to maturity is, with probability 0.75, uniform between 0.1 and 1.0 years and, with probability 0.25, uniform between 1.0 and 5.0 years. We fix the strike price at $K = 100$ and take the initial asset price $S \equiv S_0$ to be uniform between 70 and 130. Relative errors do not change if S and K are scaled by the same factor, i.e., only the ratio S/K is of interest. The riskless rate r is, with probability 0.8, uniform between 0.0 and 0.10 and, with probability 0.2, equal to 0.0. Each parameter is selected independently of the others. This selection of parameters matches the choice of Broadie, Detemple except for dividends which we donot regard here..

Figure 9 reports the results for European call options to which the analysis was devoted especially so far. Of course, similar results could be presented for European put options. Amazingly, the newly developed

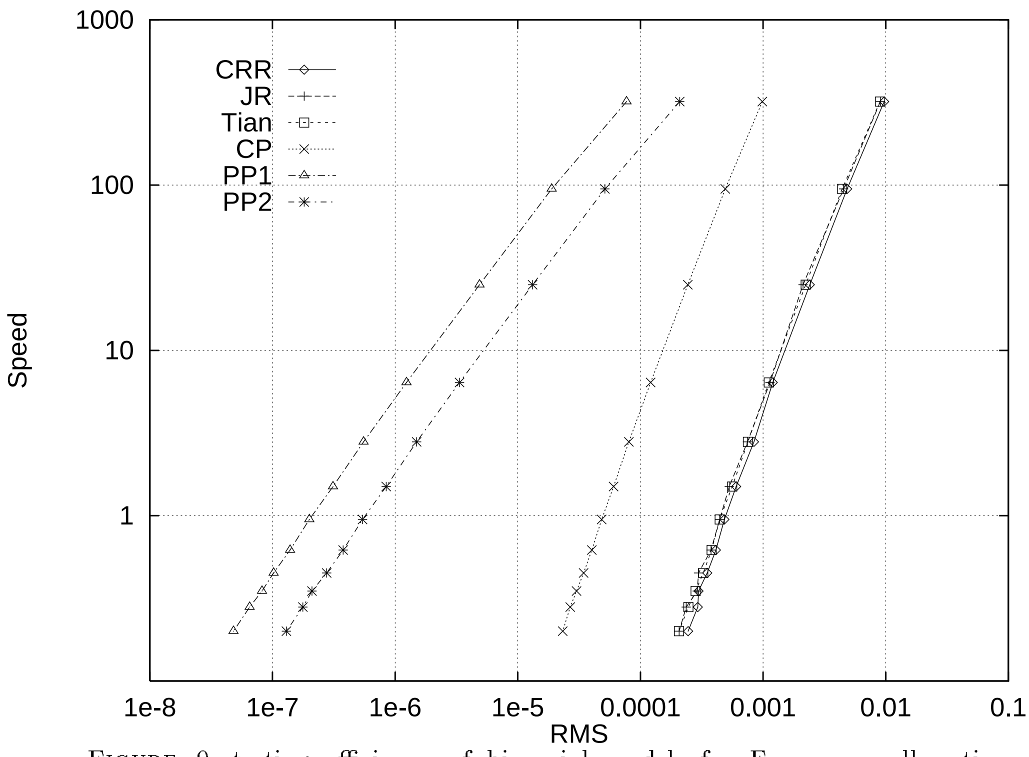

FIGURE 9. testing efficiency of binomial models for European call options: relative mean squared error in calculations with a large parameter sample, sample size $m = 2500$, versus computation speed as expressed by the number of option prices calculated per second, calculations at marked points with tree refinements: $n_i = \{25, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000\}$; x–axis and y–axis with log–scale

methods outperform all tested approximations in terms of accuracy. We reproduced the finding that the speed–accuracy line for the CRR–model is linear in appearance. These findings take over to the JR–model, Tian–model, and our approaches. Objecting, the smooth line develops from the averaging over the results of the whole sample. Taking only a single parameter constellation yields a picture, where the convergence patterns described earlier emerge again, whereas the lines for the new methods remain stable. Finally, figure 10 reports speed–accuracy results with the calculation of American type options.

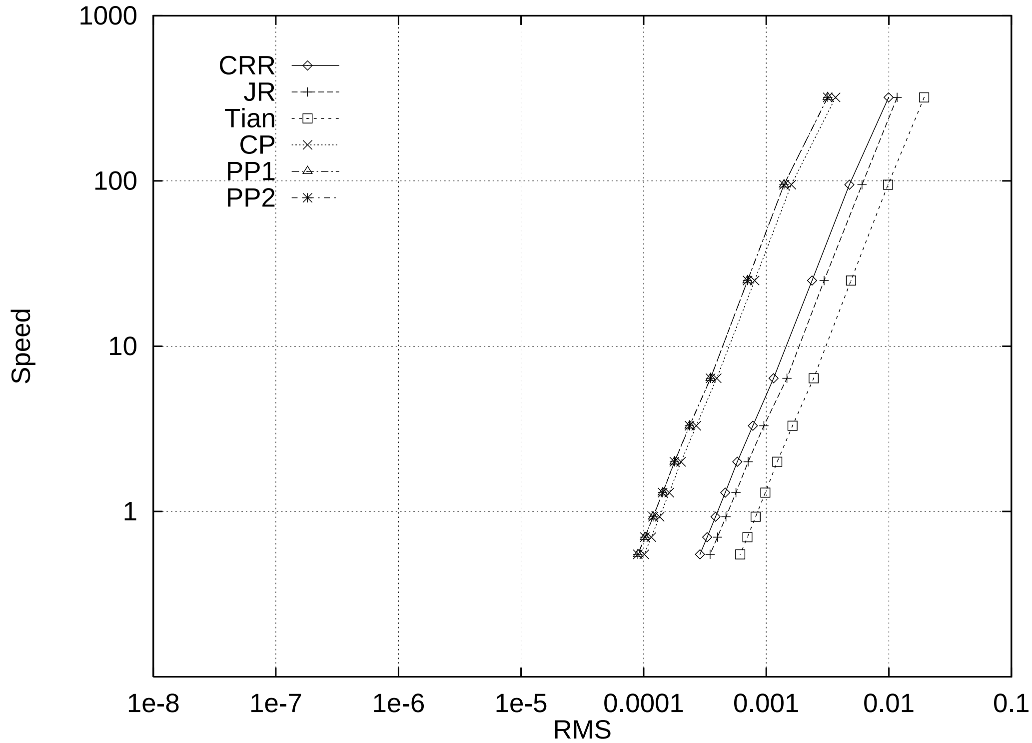

FIGURE 10. testing efficiency of binomial models for American put options: relative mean squared error in calculations with a large parameter sample, sample size $m = 2500$, versus computation speed as expressed by the number of option prices calculated per second, calculations at marked points with tree refinements: $n_i = \{25, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000\}$; x–axis and y–axis with log–scale

Similar results are obtained for the previous models. Interestingly, the CRR dominates JR and Tian. With CRR we have the property $u\,d = 1$. Generally the early exercise boundary approaches a constant line with movement towards t_0 in the tree. Consequently, we expect less distorting crossing of node layers and early exercise boundary. The new models contain order of convergence one here only but with smaller initial error. Naturally, the design solely assured high accuracy with respect to the terminal payoff distribution. Approximation of early exercise premiums produces partially again original sources of irregularities. Nevertheless, the unexpectable stability shows that the approximation error chiefly arises from deficiencies in connection with the terminal payoff distribution.

Similarily, the so–called modified binomial model (e.g. see Broadie, Detemple) corrects for the discontinuity with the strike by taking the Black–Scholes continuous prices at the second–final node layer instead the first backward induction step with the terminal payoffs. But for certain, this procedure is slower than our method. Furthermore, this mixture of discrete and continuous models does not represent a consistent model from the theoretical point of view.

6. Conclusion

For the application of numerical methods to option pricing it is of prime importance to classify the approximation quality. We point to unsatisfactory convergence patterns and speed. Our theorem characterises the order of convergence in terms of the by one reduced highest order contained in first, second, and third moments and a pseudomoment. This allows for a comparison of methods with well measurable and quantifiable criteria. For three previously existing lattice approaches we find order of convergence one generally. Then we present a general approach for the construction of binomial trees where the cumulated probabilities approximate better the standard normal functions of the Black–Scholes formula. We acquire a smoothly converging model with order of convergence two. The neat theoretical properties, namely the arbitrage relationships which imply the replicating portfolio, are retained entirely in the newly established models with improved approximation properties. Finally, we list simulation results. Especially we conduct an examination of computation speed and accuracy for a large sample of randomly selected parameter constellations. The true usage of binomial models consists in the flexibility to cover the valuation of numerous option types. But transfer of our findings to the valuation of other possibly complex option types remains to be done.

Appendix A: Normal Approximations

Camp-Paulson Approximation. [8] This approximation proceeds from the equivalence of a cumulative binomial probability to an incomplete beta-function ratio (Kendell, Stuart[77], p. 131) and thence to a probability integral of the variance ratio , F (Kendall, Stuart[77], p. 407). Using an approximation to the integral of F developed by Paulson[42] (Kendall, Stuart[77], p. 410) who in turn used Wilson and Hilferty's[31] approximation for the distribution of chi-square(Kendall, Stuart[77], p. 399) and the result obtained by Fieller[32] and Geary[30] concerning the ratio of two normally distributed variates (Kendall, Stuart[77], p. 288), Camp[51] developed an explicit expression which may be written as follows:

$$\begin{aligned} B(a,n;p) &= \mathcal{N}\left(\frac{-x_1}{3\cdot\sqrt{x_2}}\right) \\ x_1 &= \left[(n-a)\frac{p}{a+1}\cdot(1-p)\right]^{\frac{1}{3}}\cdot\left[9-\frac{1}{n-a}\right]+\frac{1}{a+1} \\ x_2 &= \left[(n-a)\frac{p}{a+1}(1-p)\right]^{\frac{2}{3}}\frac{1}{n-a}+\frac{1}{a+1} \end{aligned}$$

see also Gebhardt[69] and Peizer, Pratt[68]. Peizer and Pratt[68] derived the inversion formula presented in section 4.

Peizer - Pratt Approximations. [9] Let z be the true but functionally unknown input of the standard normal function to approximate a value of the cumulative binomial distribution function. Starting with approximation $z^* = [(a+\frac{1}{2})-np]/\sqrt{npq}$, where $a+\frac{1}{2}$ denotes the number of successes in n Bernoulli trials with continuity correction, they correct for misplacement of the median by

$$z^* = \frac{\left[a+\frac{1}{2}-np+\frac{q-p}{6}\right]}{\sqrt{npq}}$$

and further investigations suggest replacing n by $n+\frac{1}{6}$ in the denominator. Thorough investigation of approximation patterns with z/z^* reveal functionally expressable simple patterns, which eventually lead to the following adjustment:

$$z_1 = \frac{\left[(a+\frac{1}{2})-np+\frac{q-p}{6}\right]}{\sqrt{(n+\frac{1}{6})pq}}\cdot\underbrace{\left[1+q\cdot g\left(\frac{a+\frac{1}{2}}{n\cdot p}\right)+p\cdot g\left(\frac{n-(a+\frac{1}{2})}{n\cdot q}\right)\right]^{\frac{1}{2}}}_{G}$$

$$\text{where}\quad g(x) = (1-x)^{-2}(1-x^2+2x\cdot\ln x)$$

Further modification to the first part delivers a second approximation

$$z_2 = \frac{\left\{\left[(a+\frac{1}{2})-np+\frac{q-p}{6}\right]+0,02\left(\frac{q}{a+1}-\frac{p}{n-a}\frac{q-0.5}{n+1}\right)\right\}}{\sqrt{(n+\frac{1}{6})pq}}\cdot G$$

In the case where $a+\frac{1}{2} = n-(a+\frac{1}{2})$ these formulas reduce to

$$\begin{aligned} z_1 &= \pm(n+\frac{1}{3})\left\{\frac{-\ln(4pq)}{n+\frac{1}{6}}\right\}^{\frac{1}{2}} \\ z_2 &= \pm(n+\frac{1}{3})\left\{\frac{-\ln(4pq)}{n+\frac{1}{6}}\right\}^{\frac{1}{2}} \end{aligned}$$

where the sign is to be chosen to agree with the sign of $q-0.5$.
Only then, inversion formulas as presented in section 4 can be derived.

---

[8] this description is obtained mainly from Raff[56], especially references to Kendall, Stuart[77] were supplemented
[9] see Peizer, Pratt[68] and Pratt[68] for these approximations

Appendix B: Proofs of Proposition 1, 2, and 3

We remember $r_n := \exp\{r\Delta t_n\}$, $v_n := \exp\{\sigma^2 \Delta t_n\}$ and for all $k \in \mathbb{N} : E\left[R_{n,1}^k\right] = r_n^k \, v_n^{k(k-1)/2}$.

Proof of Proposition 1. From theorem 2 we know that we only have to verify the following three terms:

(1) $\bar{\mathfrak{m}}_n^2 = \mathcal{O}\left(\frac{1}{n^2}\right)$

Since $p_n u_n + q_n d_n = r_n$ one gets:

$$\begin{aligned} p_n u_n^2 + q_n(u_n d_n) &= u_n r_n \qquad \text{and} \qquad p_n(u_n d_n) + q_n d_n^2 = r_n d_n \\ \Rightarrow \quad p_n u_n^2 + q_n d_n^2 &= r_n(u_n + d_n) - (u_n d_n)(p_n + q_n) = r_n(u_n + d_n) - 1 \\ \Rightarrow \quad \bar{\mathfrak{m}}_n^2 &= \left|p_n u_n^2 + q_n d_n^2 - r_n^2 v_n\right| = r_n \left|u_n + d_n - r_n^{-1} - r_n v_n\right| \\ &= r_n \cdot \mathcal{O}\left(\frac{1}{n^2}\right) \\ &= \mathcal{O}\left(\frac{1}{n^2}\right) \end{aligned}$$

from a series expansion of the Exponential–Function.

(2) $\bar{\mathfrak{m}}_n^3 = \mathcal{O}\left(\frac{1}{n^2}\right)$

is proven the same way.

(3) $\mathfrak{p}_n = \mathcal{O}\left(\frac{1}{n^2}\right)$

From a series expansion of the Exponential–Function one immediately gets:

$$u_n - 1 = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right), \qquad d_n - 1 = \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$$

Since $\ln u_n = -\ln d_n = \sigma\sqrt{\Delta t_n}$ one gets $\mathfrak{p}_n = \mathcal{O}\left(\frac{1}{n^2}\right)$

Proof of Proposition 2. From theorem 2 we know that we only have to verify the following three terms:

(1) $\bar{\mathfrak{m}}_n^2 = \mathcal{O}\left(\frac{1}{n^2}\right)$

As in the previous theorem one gets:

$$\begin{aligned} p_n u^2 + q_n d_n^2 &= r_n(u_n + d_n) - (u_n d_n) = r_n(u_n + d_n) - e^{2\mu' \Delta t_n} \\ \Rightarrow \quad \bar{\mathfrak{m}}_n^2 &= \left|p_n u_n^2 + q_n d_n^2 - r_n^2 v_n\right| \\ &= r_n e^{\mu' \Delta t_n} \left|u_n e^{-\mu' \Delta t_n} + d_n e^{-\mu' \Delta t_n} - e^{\mu' \Delta t_n} e^{-r\Delta t_n} e^{-(2r+\sigma^2)\Delta t_n}\right| \\ &= r_n e^{\mu' \Delta t_n} \cdot \mathcal{O}\left(\frac{1}{n^2}\right) \\ &= \mathcal{O}\left(\frac{1}{n^2}\right) \end{aligned}$$

by a series expansion of the Exponential–Function.

(2) $\bar{\mathfrak{m}}_n^3 = \mathcal{O}\left(\frac{1}{n^2}\right)$

is proven the same way.

(3) $\mathfrak{p}_n = \mathcal{O}\left(\frac{1}{n^2}\right)$

From a series expansion of the Exponential–Funktion one immediately gets:

$$u_n - 1 = \mathcal{O}\left(\sqrt{\Delta t_n}\right), \qquad d_n - 1 = \mathcal{O}\left(\sqrt{\Delta t_n}\right)$$

Since $\ln u_n = \mathcal{O}\left(\sqrt{\Delta t_n}\right)$, $\ln d_n = \mathcal{O}\left(\sqrt{\Delta t_n}\right)$ this proves: $\mathfrak{p}_n = \mathcal{O}\left(\Delta t_n^2\right)$.

Proof of Proposition 3. Since Tian has explicitely fixed the first three moments: $\bar{\mathfrak{m}}_n^1 = \bar{\mathfrak{m}}_n^2 = \bar{\mathfrak{m}}_n^3 = 0$, we only need to verify $\mathfrak{p}_n = \mathcal{O}\left(\frac{1}{n^2}\right)$.
From Abramowitz and Stegun [68] we know the following:

$$\forall\ 0 < z \leq 0.5828 : |\ln(1-z)| < \frac{3}{2}z \quad \Rightarrow \quad \forall\ 0.4172 \leq z < 1 : |\ln z| \leq \frac{3}{2}|z-1|$$

$$\forall z > 0 : \ln(1+z) < z \quad \Rightarrow \quad \forall\ z > 1\ :\ |\ln z| < \frac{3}{2}(z-1)$$

This is fulfilled for sufficiently high n with $0.4172 < d_n < 1 < u_n$ thus yielding $(\ln u_n)^2 \leq \frac{9}{4}(u_n-1)^2$, $(\ln d_n)^2 \leq \frac{9}{4}(d_n-1)^2$.
Therefore

$$\begin{aligned}
\mathfrak{p}_n &= \overline{E}\left[\left(\ln \overline{R}_{k+1}\right)\left(\overline{R}_{k+1}-1\right)^3\right] \\
&\leq \overline{E}\left[\left(\ln \overline{R}_{k+1}\right)^2\right]^{\frac{1}{2}} \cdot \overline{E}\left[\left(\overline{R}_{k+1}-1\right)^3\right]^{\frac{1}{2}} \quad \text{by Hölders Inequality} \\
&\leq \frac{9}{4}\overline{E}\left[\left(\overline{R}_{k+1}-1\right)^2\right]^{\frac{1}{2}} \cdot \overline{E}\left[\left(\overline{R}_{k+1}-1\right)^3\right]^{\frac{1}{2}}
\end{aligned}$$

Thus it is sufficient to prove: $u_n - 1 = \mathcal{O}\left(\sqrt{\Delta t_n}\right)$, $d_n - 1 = \mathcal{O}\left(\sqrt{\Delta t_n}\right)$
Forming a series expansion:

$$v_n^2 + 2v_n - 3 \ = \ 1 + \mathcal{O}(\Delta t_n) + 2 + \mathcal{O}(\Delta t_n) - 3 \ = \ \mathcal{O}(\Delta t_n)$$

and

$$\begin{aligned}
u_n - 1 &= \frac{r_n\, v_n}{2}\left(v_n + 1 + \sqrt{v_n^2 + 2v_n - 3}\right) - 1 \\
&= \frac{r_n\, v_n}{2}(v_n + 1 + \mathcal{O}(\sqrt{\Delta t_n})) - 1 \ = \ \frac{r_n\, v_n^2}{2} + \frac{r_n\, v_n}{2} + \mathcal{O}(\sqrt{\Delta t_n}) - 1 \\
&= \frac{1 + \mathcal{O}(\sqrt{\Delta t_n})}{2} + \frac{1 + \mathcal{O}(\sqrt{\Delta t_n})}{2} + \mathcal{O}(\sqrt{\Delta t_n}) - 1 \\
&= \mathcal{O}(\sqrt{\Delta t_n})
\end{aligned}$$

In the same way one proves: $d_n - 1 = \mathcal{O}(\sqrt{\Delta t_n})$.

References

Abramowitz, M.; Stegun, I.(eds.) [1968]: , *Handbook of Mathematical Functions*, 5th Dover Printing, 1968.

Berry, A.C. [1941]: "The Accuracy of the Gaussian Approximation to the Sum of Independent Variables", *Trans. Amer. Math. Soc.* 49, 1941, pp.122–136.

Boyle, P.P. [1988]: "Option Valuation Using a Three–Jump Process", *International Options Journal* 3, 1988, pp. 7–12.

Broadie, M., Detemple, J. [1994]: "American Option Evaluation: New Bounds, Approximations, and a Comparison of Existing Methods", *Working Paper Columbia University, New York, NY, 10027.*

Butzer, P.L.; L. Hahn; Westphal, U. [1975]: "On the Rate of Approximation in the Central Limit Theorem", *Journal of Appr. Theory* 13, 1975, pp.327–340.

Camp, Burton H. [1951]: "Approximation to the point binomial", *Annals of Mathematical Statistics*, 22, 1951, pp. 130–31.

Cheuk, T.H.F.; Vorst, T.C.F. [1994]: "Binomial Models for Some Path–Dependent Options", *Working Paper Tinbergen Institute, Erasmus University Rotterdam, Report 9422/A.*

Cox, J.; Ross, S.A. [1975]: "The Valuation of Options for Alternative Stochastic Processes", *Journal of Financial Economics* 3 1976, pp. 145–166.

Cox, J.; Ross, S.A., Rubinstein M. [1979]: "Option Pricing: A Simplified Approach", *Journal of Financial Economics* 7, 1979, pp. 229–263.

Cox, J.; Rubinstein M. [1985]: "Options Markets", Prentice–Hall, New Jersey.

Esséen, C.G. [1945]: "Fourier Analysis of Distribution Functions", *Acta Math.* 77, 1945, pp. 1–125.

Fieller, E. C. [1932]: "The Distribution of the Index in a Normal Bivariate Population", *Biometrika*, 24, 1932, pp. 428–40.

Feller, W. [1957]: "An Introduction to Probability Theory and Its Applications", *Wiley Series in Probability and Mathematical Statistics*, Vol I+II, Wiley & Son, Inc. New York, London and Sidney.

Geary, R. C. [1930]: "The Frequency Distribution of the Quotient of Two Normal Variates", *Journal of the Royal Statistical Society*, 93, 1930, pp. 442–6

Gebhardt, F. [1969]: "Some Numerical Comparisons of Several Approximations to the Binomial Distribution", *Journal of the American Statistical Association*, Bd. 64, 1969, pp. 1638–1646.

Harrison, J.; Pliska S. [1981]: "Martingales and Stochastic Integrals in the Theory of Continuous Trading", *Stochastic Processes and their Applications*, 1981.

He, Hua [1990]: "Convergence from Discrete- to Continuous-Time Contingent Claims Prices", *Review of Financial Studies* 3, 1990, No. 4, pp. 523–546.

Hull, J. [1993]: "Options, Futures, and Other Derivative Securities", Prentice–Hall, New Jersey.

Hull, J.; White, A. [1992]: "Efficient Procedures for Valuing Path–Dependent Options", *Working Paper, Faculty of Management University of Toronto.*

Hull, J.; White, A.: "The Accelerated Binomial Option Pricing Model", *Journal of Financial and Quantitative Analysis* 26, 1991, pp. 153–164.

Ingersoll, J. [1987]: *Theory of Financial Decision Making*, Rowman & Littlefield, Totowa, New Jersey.

Jarrow, R.; Rudd A. [1983]: *Option Pricing*, Homewood, Illinois, 1983, pp.183–188.

Johnson, N.L.; Kotz S. [1969]: "Discrete Distributions", *Wiley Series in Probability and Mathematical Statistics*, Wiley & Son, Inc. New York, London and Sidney.

Karlin, S. [1969]: "A First Course in Stochastic Processes", *Academic Press*, New York and London.

Kendall, Sir M.; A. Stuart [1977]: *The Advanced Theory of Statistics*, Vol.1, Charles Griffin & Company, London.

Kloeden, P.E.; Platen, E. [1992]: *Numerical Solution of Stochastic Differential Equations*, Springer Verlag, 1992.

Li, A. [1992]: "Binomial Approximation: Computational Simplicity and Convergence", *Working Paper #9201, Federal Reserve Bank of Cleveland*, 1992.

Nelson, D. B.; Ramaswamy, K. [1990]: "Simple Binomial Processes as Diffusion Approximations in Financial Models", *Review of Financial Studies* 3,1990, pp.393–430.

Omberg, E. [1988]: "Efficient Discrete Time Jump Process Models in Option Pricing", *Journal of Financial and Quantitative Analysis*, Vol. 23, No.2, June 88, pp. 161–174.

Paulson, Edward [1942]: "An Approximate Normalization of the Analysis of Variance Distribution", *Annals of Mathematical Statistics*, 13, 1942, pp. 233–235.

Peizer, D.B.; Pratt, J.W. [1968]: "A Normal Approximation for Binomial, F, Beta, and Other Common Related Tail Probabilities, I", *The Journal of the American Statistical Association*, Bd. 63, 1968, pp. 1416–1456.

Pratt, J.W. [1968]: "A Normal Approximation for Binomial, F, Beta, and Other Common, Related Tail Probabilities, II", *The Journal of the American Statistical Association*, Bd. 63, 1968, pp. 1457–1483.

Raff, M.S. [1956]: "On Approximating the Point Binomial", *American Statistical Association Journal*, June 1956.

Raff, M.S. [1955]: "The Comparative Accuracy of Several Approximations to the Cumulative Binomial Distribution", Unpublished Master's thesis. Washington: *The American University*, 1955.

Rendleman, R.; B. Bartter [1979]: "Two–State Option Pricing", *Journal of Finance* 34 December 1979, pp. 1093–1110.

Rubinstein, M. [1990]: "Exotic Options", *Working Paper University of California at Berkley.*

Tian, Y. [1993]: "A Modified Lattice Approach to Option Pricing", *Journal of Futures Markets*, Vol. 13, No.5, pp. 563–577.

Trigeorgis, Lenos [1991]: "A Log-transformed Binomial Numerical Analysis Method for Valuing Complex Multi-Option Investments", *Journal of Financial and Quantitative Analysis* 26, No. 3, September 1991, pp. 309–326.

Trotter, H.F. [1959]: "An Elementary Proof of the Central Limit Theorem" *Arch. Math.* 10, 1959, pp. 226–234.

Wilson, E. B.; Hilferty, M. M.: "The Distribution of Chi-Square", *Proceedings of the National Academy of Sciences*, 17, 1931, pp. 684–688.

Department of Statistics, Rheinische Friedrich-Wilhelms-Universität Bonn, Adenauerallee 24–42, D-53113 Bonn, Germany

E-mail address: leisen@addi.or.uni-bonn.de

Department of Statistics, Rheinische Friedrich-Wilhelms-Universität Bonn, Adenauerallee 24–42, D-53113 Bonn, Germany

E-mail address: reimer@addi.or.uni-bonn.de