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. Even with the valuation of European call options distorting irregularities occur. For this case, sources of convergence patterns are explained. Furthermore, 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 remarkably, we even 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. Approximation of American type options with the new models exhibits order of convergence one but smaller initial error than previously existing binomial models.

1. INTRODUCTION

In the virtue of arbitrage pricing theory, the present value of a derivative security is derived by calculating the initial cost of some dynamic perfectly replicating portfolio, consisting of proportions in the underlying security and amounts of cash which change with time. Hence this portfolio is risklessly interchangeable with the option itself regardless of the actually occuring states during the lifetime of the option. However, an infinite number of possible future states translating into supposed price movements of the underlying security define the portfolio's composition and variations. Now assume that these price movements are described mathematically by a stochastic diffusion process. Then technically, any price change can be decomposed into an infinite sequence. Consequently, an artificial dynamic portfolio with a corresponding sequence of proportion adjustments achieves the duplication task. Consistently, idealized financial markets must be assumed, where a continuous flow of supply and demand to assets arrives and clears at a sequence of equilibrium prices instantly quoted. Evidently, it is this very assumption of continuous and frictionless price movements in continuous time markets which allows for the duplication of any uncertain income stream. Of course, trading at financial markets does not occur continuously in the strict sense. Rather, the flow of multiple single deals translating into prices is modelled in this way mathematically. Alternatively, describing the formation of uncertain future prices can be carried out much more simplified. Imagine, that a set of future prices is obtained by a decomposition consisting of a Bernoulli sequence. Naturally, only a limiting lattice structure with an infinite number of Bernoulli steps succeeds to capture infinite states of nature. But if this lattice is constructed correspondingly to the continuous framework, both models may coincide in the limit. Marginally quoted, technically only models where vertices recombine permanently are tracktable. This problem we assume away here.

Date. March 1995. This version: March 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, and Dieter Sondermann a.o. for comments and assistance. Financial support from the Deutsche Forschungsgemeinschaft, Sonderforschungsbereich 303 at Rheinische Friedrich-Wilhelms-Universität Bonn is gratefully acknowledged. Document typeset in LaTeX. Please address correspondence to the second author.

Actually, a binomial tree of fixed length approximates the continuous set of trading occurences and security prices always by covering but some finite range of security prices with a discrete grid structure of constant log–steps at discrete equidistant trading instances. With every application of binomial trees, one inevitably must examine the approximation quality by careful consideration of the approximation result with changing tree refinements. Incidentally stated, to our opinion such simulations do not describe how discontinuities resolve at financial markets. Rather, solely properties of the approximation theme are shown.

The approximation result is expressed by the parameters describing the dynamic replication strategy. This study is restricted to the examination of the option price only.

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.

Our paper goes beyond the findings of the existing literature in several ways. 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 which is measured by the degree of change with iterated refinement in absolute difference of binomial price and continuous solution. Furthermore, with smoothness of convergence the approximation results improve with each increase of the refinement regardless of the given parameter constellation, especially independently of a special choice to the strike price.

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, although there exists a vast mathematical literature on convergence speed in connection with central limit theorems. Furthermore, reasons for unsatisfactory convergence patterns are discussed. Then, the presentation of methods to achieve models with smooth convergence patterns follows. Next, a model with a higher order of convergence and smooth convergence pattern is presented. Moreover, all the presented models have the same computation speed given the same tree refinement, because only the formulas to calculate the tree parameters change.

But 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 characterised clearly, whereas this theoretical construction is somewhat concealed in the continuous setup. Notably, these 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 approximate prices 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. 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. Therefore, we must excuse that not all of them can be mentioned here.

Jarrow and Rudd[83] constructed a binomial model where the first two moments of the discrete and continuous model coincide by construction. Furthermore, the probability measure is equal to one half. Since in the CRR-model, the variance of the asset return converges towards the variance in the Black-Scholes model only in the limit, they claim that their model should have a "better" convergence behavior. The essence of better convergence behavior is not tackled.

Boyle[88] constructed a trinomial lattice, which is fixed up to some arbitrary parameter lambda, 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.

A totally different approach to improve the accuracy of binomial models was inaugurated by Hull and White[88], who transfered the control–variate technique from the Monte Carlo method. An interesting approach, but it differs entirely from the line of thought persued here.

Numerous adjustments have been introduced to apply lattice approaches to various types of options. There is the broad field of exotic options. Cox, Rubinstein[85] presented an adjustment for the valuation of Down–and–out calls. Hull and White[93] modified the original CRR-modell 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 this 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 models with state varying volatility functions are considered (e.g. see Nelson, Ramaswamy[90], Li[92]).

2. Chopping up the continuous framework – construction principles of some previous lattice approaches

Once again, Black and Scholes assume that trading at financial markets proceeds continuously in time, the market rate of interest r is commonly known and fixed over time, payments out of underlying securities, e.g. dividends, do not exist. Stock-price dynamics are described by

$$dS(t) = rS(t)dt + \sigma S(t)dW(t) \tag{1}$$

where r is the instantaneous expected return of the underlying asset S if immediately the risk-neutrality argument of Harrison and Pliska[81] is used, σ^2 is the instanteneous variance of the return, and dW is a standard Gauss–Wiener process.

Within their model, 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 a European call option at each instant of time, having strike price K and time to maturity $T-t_0$.
It is this equilibrium connection which results in the Black-Scholes differential equation where the expected return on the option is expressed in terms of the option price function and its partial derivatives.

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

The solution to the differential equation with boundary condition $f=[S(T)-K,0]^+$ as payoff function is given by the Black-Scholes option pricing formula:

$$\begin{aligned} c(t_0,S) &= S\cdot\mathcal{N}(d_1)-K\cdot e^{[-r(T-t_0)]}\mathcal{N}(d_1-\sigma\sqrt{T-t_0}) \\ d_1 &= \frac{[ln(\frac{S(t_0)}{K})+(r+\frac{1}{2}\sigma^2)(T-t_0)]}{\sigma\sqrt{T-t_0}} \end{aligned} \tag{3}$$

where $\mathcal{N}$ is the cumulative standard normal distribution function.
Transferring this framework into the simplifying binomial structure induces several adjustments. Here, the model evolves step by step at certain spots of time $t_i\in\{t_0,\cdots,t_n\}$. Δt characterises a fixed length of time passing by between sequential model events. Model events consist of price changes in the underlying security and option price. Consistently, the replicating portfolio requires adjustments there only. Continuous parameters, i.e. r and $\sigma^2 t$ are translated to per period variables $\hat{r}$, $\sigma^2(T/n)$. Central significance devolves upon the construction of price movements. Generally, within all binomial models the stochastic price behavior is modelled by

$$S(t_i,j,n)=S(t_0)\,u(n)^j\,d(n)^{i-j} \tag{4}$$

where $S(t_i,j,n)$ denotes the asset price at time t_i after j up-movements in a n-step binomial tree, with $T-t_0$ as length of the time axis. When we define $\hat{R}_i$ as the randomly distributed one-period return for the underlying asset we have

$$\hat{R}_i=\begin{cases} u(n) & \text{with probability} \quad p(n) \\ d(n) & \text{with complementary probability} \quad 1-p(n)\equiv q(n) \quad i=0,\cdots,n \end{cases} \tag{5}$$

where n denotes that these parameters belong to a specific binomial tree and change with refinement n. Apart from that, these parameters remain constant throughout a tree. Evidently, a special ordering of up–and down–movements does not affect terminal asset prices.
Within 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 regardless whether the asset price increases to $S(t_{i+1})=u(n)S(t_i)$ or decreases to $S(t_{i+1})=d(n)S(t_i)$. Payoff replication with respect to up- or down-movements requires

$$\begin{aligned} \Delta\cdot u(n)\cdot S(t_i,j)+rB &= C\left(S(t_{i+1},j+1)\right) \\ \Delta\cdot d(n)\cdot S(t_i,j)+rB &= C\left(S(t_{i+1},j)\right) \end{aligned}$$

to hold.
From this, the proportion Δ of the underlying asset and amount B of cash in the replicating portfolio can be derived, the value at t_i is interpretable as discounted expected value of prices at t_{i+1} with martingale measure equal to $p(n)=(\hat{r}-d(n))/(u(n)-d(n))$.
Here, this equilibrium connection is subsumed for all discrete steps in a binomial formula, of which the first was presented by CRR:

$$C(t_0,S(t_0),n)=\hat{r}^{-n}\sum_{j=0}^{n}\binom{n}{j}p(n)(1-p(n))^{n-j}[S(T,j,n)-K]^+ \tag{6}$$

and equivalently:

$$(7) \qquad C(t_0, S(t_0), n) = S(t_0)\Phi[a; n, p'(n)] - Kr^{-n}\Phi[a; n, p(n)]$$
$$\text{where} \quad p = \frac{\hat{r} - d(n)}{u(n) - d(n)} \qquad p' = \frac{u(n)}{\hat{r}} \cdot p(n) \qquad a = \text{Int}\left[\frac{ln(K/Sd(n)^n}{ln(u(n)/d(n))}\right]$$

and $\Phi[\cdot]$ denotes the complementary binomial distribution.
Transition to the continuous model eventually implies transition from the binomial components to the standard normal components which result from the terminal distribution of asset prices. The terminal asset price distribution results from return sequences randomly changing between $u(n)$ and $d(n)$:

$$S(T) = S(t_0) \cdot \hat{R}_i \Leftrightarrow \frac{S(T)}{S(t_0)} = \prod_{i=1}^{n} \hat{R}_i$$

Consider a random Bernoulli variable with

$$X(i,n) = \begin{cases} 1 & \text{when } \hat{R}_i = u(n) \\ 0 & \text{when } \hat{R}_i = d(n) \end{cases}$$

which counts the number of up-movements. Defining a sum variable we have

$$(8) \qquad j(i,n) = X(1,n) + X(2,n) + \ldots + X(i,n) + \ldots + X(n,n) \cdot$$

Now consider

$$\frac{S(t_n, j, n)}{S(t_0)} = u(n)^j d(n)^{n-j}$$

By transformation to

$$(9) \qquad \ln\frac{S(t_n, j, n)}{S(t_0)} = j \ln\frac{u(n)}{d(n)} + n \ln d(n)$$

we have $j = B(n;p)$, which is a sum of independent identically distributed variables to which a central limit theorem can be applied. By simple linear transformation the limit distribution for $\ln S(t_n, j, n)/S(t_0)$ can be derived.
Consistently, due to the assumed asset price dynamics, we must have that $\ln S(T)/S(t_0)$ is normally distributed with

$$\ln\left(\frac{S(T)}{S(t_0)}\right) = \mathcal{N}\left[\left(r - \frac{\sigma^2}{2}\right)(T - t_0), \sigma\sqrt{T - t_0}\right]$$

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 $d(n)$.

CRR[79]	JR[83]	TIAN[93]
$u = \exp\left\{\sigma\sqrt{\Delta t}\right\}$ $d = \exp\left\{-\sigma\sqrt{\Delta t}\right\}$	$u = \exp\left\{\mu'\Delta t + \sigma\sqrt{\Delta t}\right\}$ $d = \exp\left\{\mu'\Delta t - \sigma\sqrt{\Delta t}\right\}$	$u = \frac{MV}{2}\left(V + 1 + \sqrt{V^2 + 2V - 3}\right)$ $d = \frac{MV}{2}\left(V + 1 - \sqrt{V^2 + 2v - 3}\right)$
	$\mu' = r - \frac{1}{2}\sigma^2$	$M = \exp\{r\Delta t\}$ $V = \exp\left\{\sigma^2\Delta t\right\}$

The way of deduction for the parameters in the CRR-model is unknown, though dependency an volatility and stepsize seems straightforward.

The parameters of Jarrow and Rudd can be derived by solving an equation system equating first and second moments of continuous and discrete model having fixed $p(n) = \bar{p} = 1/2$.
Tian explicitly derives the parameters by equating the first three moments and using $p + q = 1$. Interestingly, rearranging the equation to derive the martingale measure yields

$$E\left[\frac{S(T)}{S(t_0)}\right] = [p(n) \cdot u(n) + q(n) \cdot d(n)]^n = \hat{r}^n \tag{10}$$

which corresponds to

$$E\left[\log\left(\frac{S(T)}{S(t_0)}\right)\right] = n\left\{p(n)\log\frac{u(n)}{d(n)} + \log d(n)\right\} \quad \longrightarrow \quad \left(\log r - \frac{1}{2}\sigma^2\right) t \tag{11}$$

when considering the logarithm of the return. Consequently, the first moment is fixed instantly for every model which fulfills the noarbitrage condition.
Equivalence with the second moment holds either by construction as well (see Jarrow, Rudd and Tian) or is achieved in the limit (see CRR). With that, all requirements to acquire the same limit distribution are given.
But beyond that, the distinct approaches do not reveal properties suggesting superiority or inferiority in terms of convergence quality expect in some cases[1]. On the contrary, simulations indicate that the three models behave similarily with respect to convercence speed, a result which will be stated and proved in the next section. 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.
Crucially, also the accuracy of approximation is influenced by a second fact, inherent in all tree models. The entire probability mass is concentrated exclusively at the terminal nodes of the tree presuming that outcomes deviating from these nodes cannot occur. Translating into a binomial distribution function we have a piecewise constant function with jumps at each terminal node. Illustrating the discrete structure take the set of probabilities connected with a corresponding set of terminal asset prices to construct a histogram outlining the approximation of the continuous density.
Notice that the probability mass of each rectangular is confined to a terminal node.

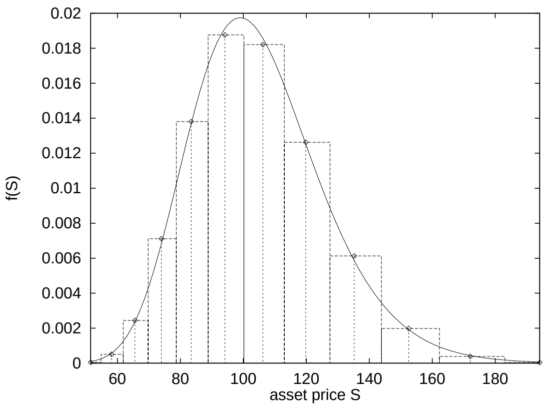

FIGURE 1. histogram vs. density function

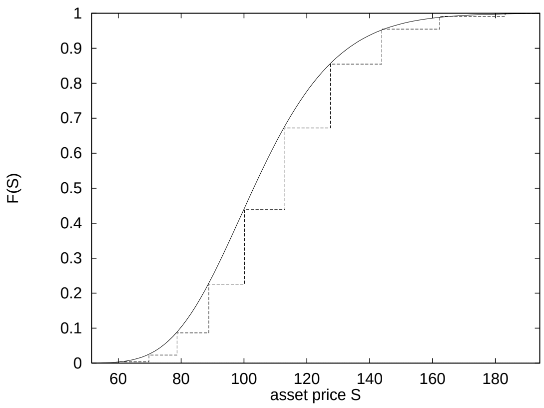

FIGURE 2. step– vs. distribution function

But the set of possible asset prices shifts with each iteration of the tree refinement. Having fixed exogenous strike price and fluctuating locations of jumps in the distribution function, the separation of the probability mass bounces back and forth with changing step patterns, because arbitrary locations can occur.
Thus, the binomial structure itself in combination with a strike price located independently of the tree

[1] Trigeorgis [91] shows that in the CRR-model negative option prices may occur when there is a very long time to maturity combined with low tree refinement.

grid induces irregularities which distort the approximation result. Except of special cases, the computed option prices oscillate[2], and convergence wavy to the Black-Scholes solution[3]. The remaining part of the paper is devoted exactly to these two deficiencies.
At the same time, actually these findings inspired to search for conditions under which fluctuations in option price approximation can be avoided and beyond that high accurancy achieved. In numerous cases the approximations oscillate almost exactly around the true value. Furthermore, with changing amplitude of the waves, the approximations accidentically approach the true value repeatedly.

3. Examining the order of convergence

In this chapter, we will first define the order of convergence. Then we will prove a general result on the order of convergence. The known lattice approaches fit into the framework of the main theorem; so it will be applied to them. At the end, we will furthermore give some simulations and explain, how the results could be used.
We will adopt the following notation:

1. We remember that the security follows the price process $X_t^{s,x}$ which is the solution of the stochastic differential equation

$$X_t^{s,x} = x + \int_s^t r X_{t'}^{s,x} dt' + \int_s^t \sigma X_{t'}^{s,x} dW_{t'} \tag{12}$$

   The probability measure is denoted by P_W.
2. $f \ : \ x \longmapsto (x-K)^+$ shall be the payoff–function, K be the strike.
3. The discrete process is described by $Y_0, \ldots, Y_n$ where Y_k is the value of the discrete security price process at time t_k. They are random variables for $k > 0$. The probability measure is P_B.

If examining a certain lattice approach for a specific security, the only changing parameter is the refinement n or equivalently the step size Δt. In fact, the option price in a discrete lattice approach is a function of n. The Black-Scholes-value $c(t,S)$ is the expected payoff at time T discounted to time t that is

$$c(t,S) := e^{-r(T-t)} E_W \left[f(X_T^{t,S}) \right] \tag{13}$$

The lattice value is the expected payoff at time T discounted to time t that is

$$e^{-r(T-t)} E_B \left[f(Y_n) \right] \tag{14}$$

Since we are only interested in the option price, we will examine e_n as error at refinement n defined by the absolute value of the difference between discrete and continuous price.

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

[3]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, put on top 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 change in value continuously though.

Definition. Let

$$e_n := e^{-rT} \; \left| E_W \left[f(X_T^{0,S}) \right] - E_B \left[f(Y_n) \right] \right| \tag{15}$$

be the **error in price**.
A lattice approach **converges** if and only if for all parameters K, r, σ, T, Y_0 we have

$$\lim_{n \to \infty} e_n = 0$$

This concept of convergence is exactly that used in mathematical literature. In observing convergence in the refinement n, one typically observes wavy patterns. This was already discussed in the previous section. An approximation which is rather close to the Black-Scholes value may follow another good or even a worse approximation.
However we remember that convergence exactly says that in giving a certain error bound, we can find a refinement n_0 such that each finer one has an error which does not exceed the bound. But how does the refinement depend exactly ? To make things precise, we shall say:

Definition. European call options, computed with a lattice approach **converge with order** $r > 0$ if there exists a constant $C > 0$ such that

$$\forall n : e_n \leq \frac{C}{n^r} \tag{16}$$

Remark.

1. In our Theorem below we will see that the estimation of the error can be decomposed into a constant $C > 0$ dependant of the specific option and the order r dependant of the chosen lattice approach.
2. Please note that convergence is implied by any order greater than 0. Moreover we remark that a lattice approach with order r has also order $\tilde{r} \leq r$. A higher order means "quicker" convergence. To achieve a certain precision level, the constant C and the order is of importance.
3. The most important fact is, that in plotting e_n against the refinement n on a log-log-scale, the bounding function $\frac{C}{n^r}$ becomes a straight line with slope equal $(-r)$ and shift C. In notifying this, it becomes easy to observe the order-of-convergence in simulations.

Example.

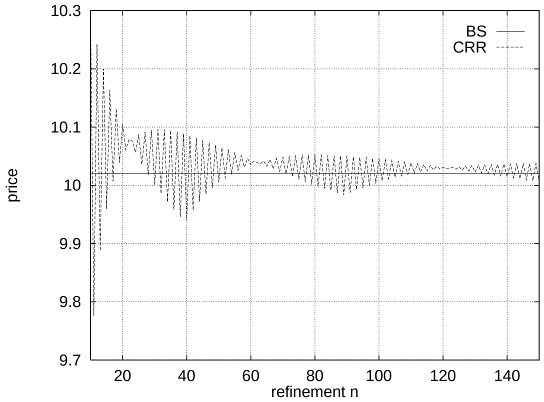

FIGURE 3. CRR–price

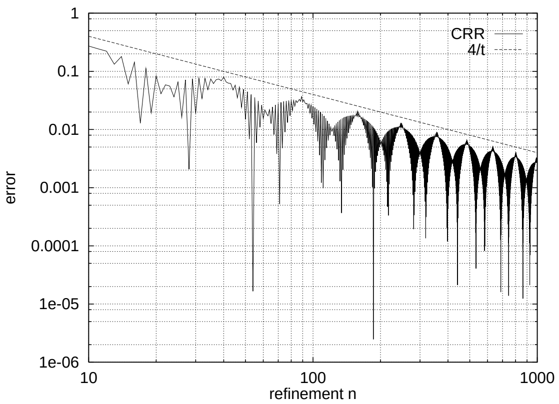

FIGURE 4. CRR–error

$$S = 100, K = 110, T = 1, r = 0.05, \sigma = 0.3$$

In this case, the order–of–convergence is equal 1. The proof is given later in this chapter.
How could one determine the order of a lattice-approach mathematically?

Using the representation of the discrete price with the cumulative binomial distribution mentioned in the previous chapter, one could examine the order of approximation of the respective distribution function. Berry [41] & Esséen [45] examined this: thus getting order $\frac{1}{2}$.
However we have already suggested order 1 in the above example of the CRR approach. Since we prove the better result by our theorem we are not going to present this idea in more detail.
Other approaches such as Ibragimov[66] are examining the characteristic function, thus using Fourier–Analysis of the distribution function. This yields conditions which are difficult to verify.
Since the option price is the discounted expectation of the final payoff, and the logarithms of the security are normally (N) and binomially (B) distributed random variables respectively, one is led to examine the order of the "weak convergence" of the DeMoivre-Laplace Theorem, that is formulas of the kind

$$E[g(N)] - E[g(B)] \quad \text{with } g : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$$

By a more detailed examination, as was done by Butzer & Hahn [75], using the Operator-Method of Trotter [59] to prove the central limit theorem, one gets the order of convergence of above terms. In essence this requires the function g to be sufficiently smooth, in order to make a Taylor-expansion. Since our payoff function is not differentiable at $S = K$ at all, this idea is not applicable directly.
However these approaches do not make use of the fact, that we are in the special situation of a stochastic process. In the case of a stochastic process and a function g with polynomial bounded derivatives (of some order), Kloeden & Platen [92] proceed differently : They discretize the time axis. In this way they are able to represent the above as a sum of the differences in each step. In each time step they evaluate it by making a Taylor-Expansion as above. This is possible because in their case they can make use of a theorem of Miculevicius ensuring sufficient differentiability for their purposes.
Unfortunately this is not the case here: however, observing that the Black–Scholes price is smooth, this is the approach that allows to circumvent the problem of nondifferentiability of our payoff function.
Remarkably, although we are valuing European call-options in this paper only, the approach here allows to extend it to path–dependent options easily, because we make use of the whole price–process.
Distributions are completely characterized by their moments. For example the normal distribution by its first moment (the mean) and second moment (the variance). Therefore in the central limit theorem the moments of the discrete random variable at least need to approximate those of the normal distribution. Convergence is ensured by the Ljapunoff condition, which is a sufficient condition on the convergence of one higher moment.
In a lattice approach, the order of convergence is completely determined by the following factors:

Definition. We call

$$\bar{m}_n^1 := E_B\left[\frac{Y_{k+1}}{Y_k} \;\Big|\; \mathcal{A}_k\right] - E_W\left[\frac{X_{k+1}}{Y_k} \;\Big|\; \mathcal{A}_k\right] \tag{17}$$

$$\bar{m}_n^2 := E_B\left[\left(\frac{Y_{k+1}}{Y_k}\right)^2 \;\Big|\; \mathcal{A}_k\right] - E_W\left[\left(\frac{X_{k+1}}{Y_k}\right)^2 \;\Big|\; \mathcal{A}_k\right] \tag{18}$$

$$\bar{m}_n^3 := E_B\left[\left(\frac{Y_{k+1}}{Y_k}\right)^3 \;\Big|\; \mathcal{A}_k\right] - E_W\left[\left(\frac{X_{k+1}}{Y_k}\right)^3 \;\Big|\; \mathcal{A}_k\right] \tag{19}$$

our **moments** and

$$p_n := E_B\left[\left(\ln\frac{Y_{k+1}}{Y_k}\right)\left(\frac{Y_{k+1}}{Y_k} - 1\right)^3 \;\Big|\; \mathcal{A}_k\right] \tag{20}$$

$$\bar{p}_n := E_W\left[\left(\ln\frac{X_{k+1}}{Y_k}\right)\left(\frac{X_{k+1}}{Y_k} - 1\right)^3 \;\Big|\; \mathcal{A}_k\right] \tag{21}$$

our **pseudo–moments**

Remark. Notice that moments and pseudomoments donot depend on specific k. $\bar{m}_n^1 = 0$ because of the risk neutrality argument of Harrison & Pliska[81]

Theorem. Let $\{Y_0^n, \dots, Y_n^n\}$ with $Y_0^n = Y_0 = S$ denote the discrete price process of a lattice approach. The order of convergence is the smallest order contained in $\bar{m}_n^2, \bar{m}_n^3$ or p_n reduced by 1, but not smaller than 1, that is :
There exists a constant C, only depending on S, K, r, σ, T such that : $e_n \leq C\left(\frac{\bar{m}_n^2 + \bar{m}_n^3 + p_n + (\Delta t)^2}{\Delta t}\right)$

Proof. is given in Appendix B We just note that the above mentioned Taylor–Expansion yields :

$$\begin{aligned} e_n \quad \leq \quad & \left|\sum_{k=0}^{n-2} e^{-rt_{k+1}} E_B\left[\tilde{c}_1(t_{k+1}, Y_k) \;\Big|\; \mathcal{A}_0\right]\right| \cdot \left|m_n^1\right| + \left|\sum_{k=0}^{n-2} e^{-rt_{k+1}} E_B\left[\tilde{c}_2(t_{k+1}, Y_k) \;\Big|\; \mathcal{A}_0\right]\right| \cdot \left|m_n^2\right| \\ & + \left|\sum_{k=0}^{n-2} e^{-rt_{k+1}} E_B\left[\tilde{c}_3(t_{k+1}, Y_k) \;\Big|\; \mathcal{A}_0\right]\right| \cdot \left|m_n^2\right| + \left|E_B\left[\sum_{k=0}^{n-2} e^{-rt_{k+1}} \Big\{ E_B\left[R_3(t_{k+1}, Y_{k+1}, Y_k) \;\Big|\; \mathcal{A}_k\right]\right.\right. \\ & \left.\left. + E_W\left[R_3(t_{k+1}, X_{k+1}, Y_k) \;\Big|\; \mathcal{A}_k\right] \Big\} \;\Big|\; \mathcal{A}_0\right]\right| + \mathcal{O}\left(\frac{1}{n}\right) \end{aligned}$$

where R_3 are remainder–terms.

Remark.

1. The Theorem separates convergence into two parts:
   - the constant is dependent on the type of option: here a Call-Option
   - the moments and pseudomoments contain the different lattice approach
2. Mainly the Theorem states that order of convergence one is inherently contained in all binomial lattice approaches. Interestingly the next section shows how the order of convergence can be tuned by a very specific construction principle.
3. In the later simulations we will see that order 1 cannot be improved in the approaches known in literature. Moreover we will explain later in the simulations how to conduct efficiently convergence speed measurements in applications. Notably we proved only that order of convergence equals at least one, possibly higher order could be contained, though simulations indicate the opposite, entirely.
4. To achieve order of convergence one, the theorem states that the approximating moments of the discrete process must converge with order two toward the moments of the continuous security process, because one degree is lost with summation over time. Furthermore, one needs the same order of convergence in the pseudomoments. This is not only a technical matter but explains why the model proposed by Tian does not perform better.

Proposition. The CRR[79] model converges with order 1.

Proposition. The JR[83] model converges with order 1.

Proposition. The Tian[93] model converges with order 1.

The proofs all use the above Theorem and will be given in Appendix C.

Simulation. Notifying that we are getting always similar pictures for convergence in price, we will present only two pictures with this topic.

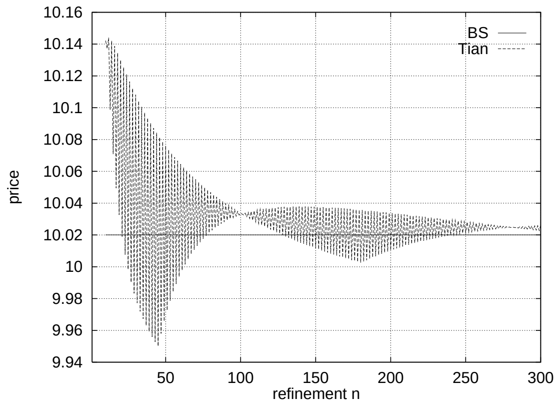

FIGURE 5. Tian–price
$S = 100, K = 110, T = 1, r = 0.05, \sigma = 0.3$

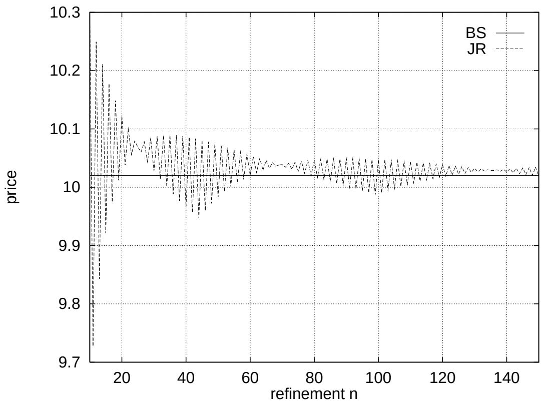

FIGURE 6. JR–price
$S = 100, K = 110, T = 1, r = 0.05, \sigma = 0.3$

In what follows we will demonstrate the three lattice approaches with three different strikes. One can recognize in the simulations that the error can always be dominated by an upper bound of order $\frac{1}{n}$.

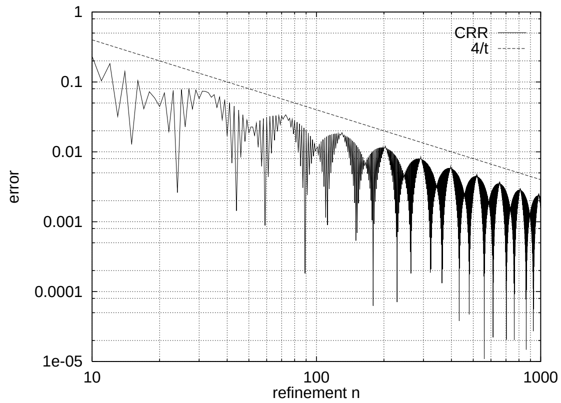

FIGURE 7. CRR–error

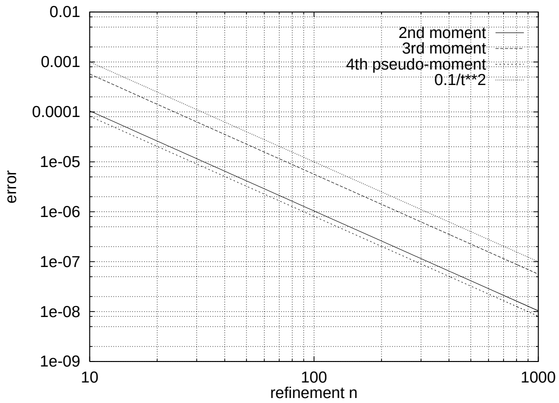

FIGURE 8. CRR–moments

$S = 100, K = 90, T = 1, r = 0.05, \sigma = 0.3$

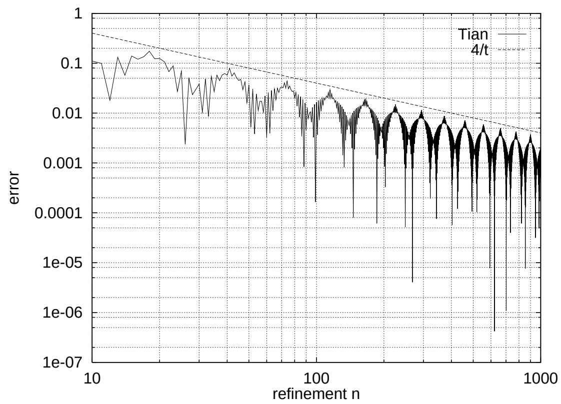

FIGURE 9. Tian–error

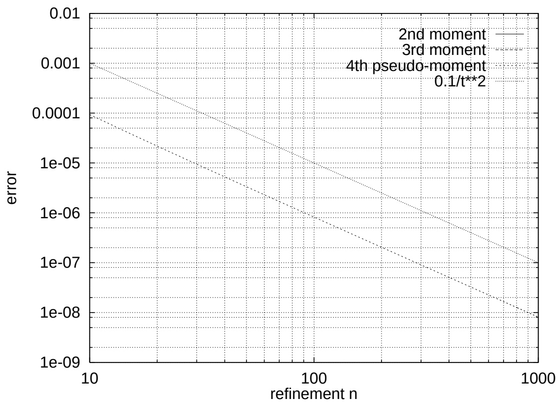

FIGURE 10. Tian–moments

$S = 100, K = 100, T = 1, r = 0.05, \sigma = 0.3$

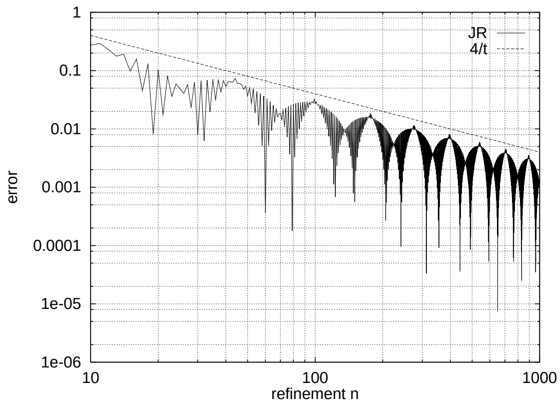

FIGURE 11. JR–error

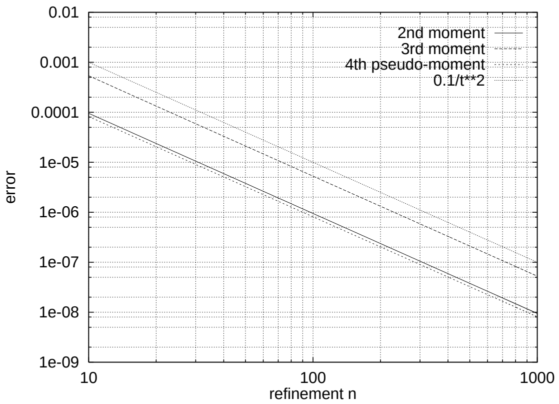

FIGURE 12. JR–moments

$S = 100, K = 110, T = 1, r = 0.05, \sigma = 0.3$

Please note, that the lattice approach by Tian has explicitely fixed the first two 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.

4. Construction of binomial models with improved convergence properties

So far, we described the sources of convergence patterns in existing lattice approaches. Moreover, we derived the order of convergence in the previous section. Here alternative methods to construct binomial trees will be discussed. Our aim: definition of binomial trees, which beforehand obey the sources of irregularities simply by different but only slightly modified definition of tree parameters $u(n)$ and $d(n)$. In section 2 we saw that irregularities evolve because the relative position of the strike within the tree varies. Now, we suggest that this relative position should be fixed in some sense.

The construction of trees to achieve convergence to Black-Scholes does not depend on any particular grid or grid sequence whatsoever. Vice versa we postulate that the construction of sequential trees may be linked to achieve homogeneity with the location of the strike.

This concept brings up the question to the correct relative position of the strike. Without extending this question, some reflections speak in favor of a location precisely in the middle of two surrounding nodes. This commitment results in fixing the random variable j to integer numbers with continuity correction of one half. Moreover, in line with the histogram concept, rectangulars remain undivided.

But where should we fix the strike overall? Why not in the center of the tree? This proceeding implies that the strike is always contained in the binomial tree grid. Consequently, the paths surrounding quantify explicitly the probability for those asset prices which finish in the near of the strike. Notably, here actually the most sensible situations arise. Consequently, the model itself speaks in favor of this construction. Besides, since most of the time trading in options occurs in at-the-money and near-the-money options only, actually this requirement changes only minorly the structure. These reflections leed to the definition of the following model.

$$u(n)p(n) + d(n)q(n) = \exp\{rT/n\} \equiv M \tag{22.1}$$

$$u(n)^2 p(n) + d(n)^2 q(n) = \exp\{\sigma^2 T/n\} \equiv M^2 V \tag{22.2}$$

$$p(n) + q(n) = 1 \tag{22.3}$$

$$u(n) \cdot d(n) = \exp\{2/n \ln(K/S)\} \tag{22.4}$$

In accordance to the approach of Tian we define an equation system where (22.1) and (22.2) fix the first two moments as sufficient but not necessary conditions to achieve convergence to the given continuous distribution. Equation (22.3) expresses that the point probabilities sum up to one. Importantly, in difference to Tian, who wasted the remaining degree of freedom to fix the third moment, we implement a condition guaranteeing that the strike is positioned at the center of the tree for every tree refinement at maturity. With even number of steps this position contains a terminal node and with odd number of steps this position precisely separates two terminal nodes. Thus, we may restrict ourselfes to even or odd refinements only depending an the desired separation rule. This proceeding does not devaluate our approach since tree calculations do not depend on any specific choice of refinement. Besides, even and odd refinements converge monotonically, respectively. Below, we present the explicit expressions for the tree parameters as unique solution of the equation system above.

$$u(n) = [g + (K + M^2\,V)\,\sqrt{g}]/[2\,M\,\sqrt{g}] \tag{23.1}$$

$$d(n) = K/u \tag{23.2}$$

$$p(n) = \frac{\hat{r} - d(n)}{u(n) - d(n)} \tag{23.3}$$

$$q(n) = 1 - p(n) \tag{23.4}$$

$$\text{where} \quad g = K^2 - 4\,K\,M^2 + 2\,K\,M^2\,V + M^4\,V^2$$

Consider, the convergence pattern of the CRR-model for at-the-money options. There is merely oscillation of the option price without any waves. Along even and odd refinements alone, we have a monotonical convergence pattern. Notably, this convergence pattern is conserved for any choice of the strike here. Remarkably, such smooth convergence patterns can serve for the application of extrapolation methods.

Remember our reflections earlier. Above all, we desired to improve the accuracy of approximation. Unfortunately, the model above does not succeed in accelerating convergence speed. Here, we propose an entirely new approach using fairly old findings of mathematical approximation theory. Because of its simplicity, the binomial ditribution always has served as a very popular distribution. Notably, actually the first central limit theorem was proved for this distribution type. Despite of the simplicity, the application of the formula is cumbersome, because the computation might involve factorials of large integers or the sumation of a large numer of individual terms. Therefor, normal approximations to the binomial distribution were derived. Especially, the Camp-Paulson[51] method and the approximations of Peizer and Pratt[68] reveal a remarkable quality of accuracy[4]. Summarising, eventually these normal approximations determine the input of the standard normal function which supposedly approximates the binomial formula with small and decreasing error. But here, our problem represents the opposite direction. Computation of binomial option prices eventually involves that normal components are approximated by binomial components. 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[5]. Now, we will demonstrate how these findings can be used to construct CRR-like binomial models.
For a given refinement the inversion formulas above specify the distribution parameter p to approximate $N(z)$ with $B(n;p)$ when the separating variable j is fixed[6]. Consistently, fixing j implies positioning the strike somewhere within a binomial tree. Once again, we locate the strike at the center of the tree as we justified in the reflections earlier. Moreover, this principle allows the usage of the Peizer-Pratt method with explicit inversion rule, when we restrict the set of refinements to odd integers[7].

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

[5] in the usual understanding of the binomial distribution

[6] In the usual setup j gives the number of successes in n trials; here, j is identified with the number of up-movements.

[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.

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

$$p = \left(\frac{b}{a}\right)^2 \left(\frac{[9a-1][9b-1] + 3z[a(9b-1)^2 + b(9a-1)^2 - 9abz^2]^{\frac{1}{2}}}{[9b-1]^2 - 9bz^2}\right)^{\frac{1}{3}}$$

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

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

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

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

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

Notably, the Camp-Paulson formula can be applied for arbitrary choice of refinement.
Using approximation rule A, B, or C we obtain p and p' as distribution parameters of the two binomial components in the binomial option pricing formula. Then, we derive tree parameters $u(n)$ and $d(n)$ by a simple trick. The noarbitrage condition implies that $p(n) = (\hat{r} - d(n))/(u(n) - d(n))$ holds. Furthermore, p' is defined to $p' = u/\hat{r} \cdot p$. 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. Notice, that $f(z, j(n))$ denotes the chosen inversion function.

$$p' = f(d_1, j(n)) \tag{24.1}$$

$$p = f(d_2, j(n)) \tag{24.2}$$

$$u = r \cdot \frac{p'(n)}{p(n)} \tag{24.3}$$

$$d = \frac{\hat{r} - p(n) \cdot u(n)}{1 - p(n)} \tag{24.4}$$

Seemingly, 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 changes dramatically. Nevertheless, within this class of models the particular theoretical building blocks for which the CRR-model became famous are entirely transfered by construction. 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.

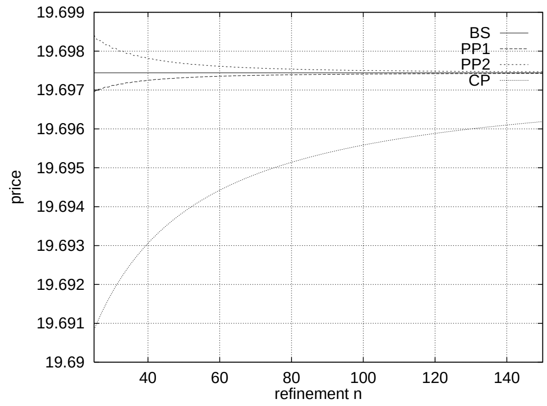

FIGURE 13. price

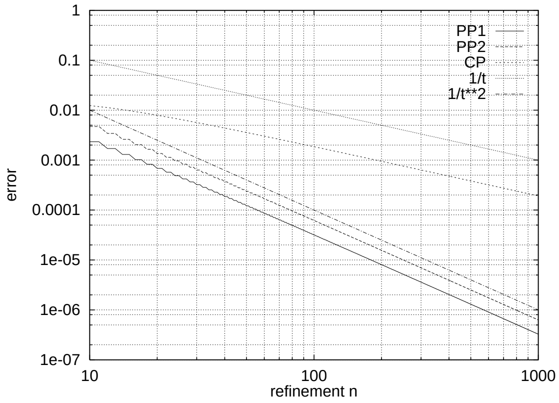

FIGURE 14. error

$S = 100, K = 90, T = 1, r = 0.05, \sigma = 0.3$

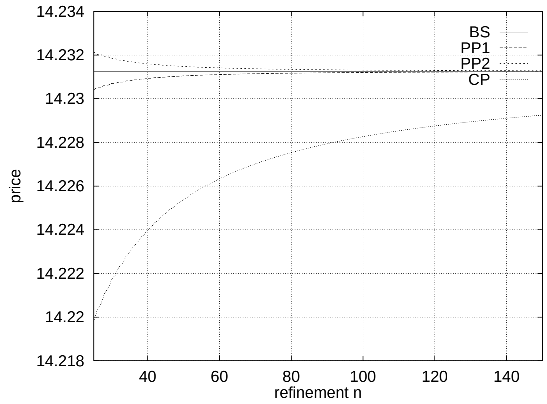

FIGURE 15. price

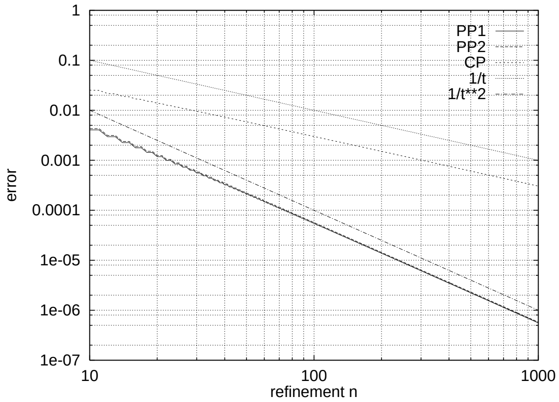

FIGURE 16. error

$S = 100, K = 100, T = 1, r = 0.05, \sigma = 0.3$

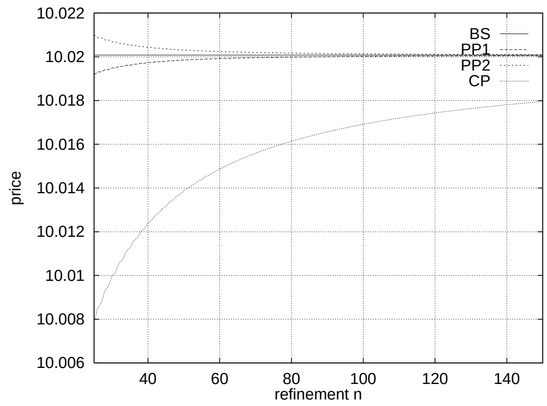

FIGURE 17. price

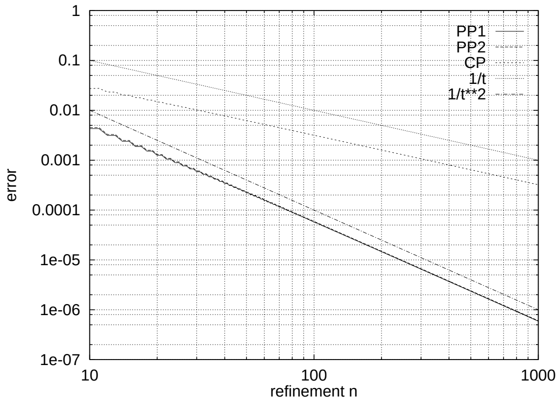

FIGURE 18. error

$S = 100, K = 110, T = 1, r = 0.05, \sigma = 0.3$

At the moment we are not able to give a strict proof of the greater order of convergence. However, we believe that it has become clear from the above simulations — especially if one compares them with the previous simulations of the models in literature — , that our models behave much better :

- the order of convergence is increased by one
- the constant C is about $\frac{1}{10}$ of the usual constant
- the convergence shows very little oscillating patterns and is in fact monotonically converging to the Black-Scholes price

Our theorem in the last section does not yield a proof for better convergence for technical reasons. However, we can use it to explain the better convergence. The proof is mainly just using a Taylor-Expansion. Since this one is exact, up to the unknown remainder terms, one expects that all convergence patterns, such as oscillation and order are reflected in the derivatives. We remember from the proof that the error e_n is dominated by

$$\begin{aligned} e_n \quad \leq \quad & \left| \sum_{k=0}^{n-2} e^{-rt_{k+1}} E_B\left[\tilde{c}_1(t_{k+1}, Y_k) \;\middle|\; \mathcal{A}_0\right] \right| \cdot \left| m_n^1 \right| + \left| \sum_{k=0}^{n-2} e^{-rt_{k+1}} E_B\left[\tilde{c}_2(t_{k+1}, Y_k) \;\middle|\; \mathcal{A}_0\right] \right| \cdot \left| m_n^2 \right| \\ & + \left| \sum_{k=0}^{n-2} e^{-rt_{k+1}} E_B\left[\tilde{c}_3(t_{k+1}, Y_k) \;\middle|\; \mathcal{A}_0\right] \right| \cdot \left| m_n^2 \right| + \Bigg| E_B\Bigg[\sum_{k=0}^{n-2} e^{-rt_{k+1}} \Big\{ E_B\left[R_3(t_{k+1}, Y_{k+1}, Y_k) \;\middle|\; \mathcal{A}_k \right] \\ & + E_W\left[R_3(t_{k+1}, X_{k+1}, Y_k) \;\middle|\; \mathcal{A}_k \right] \Big\} \;\Big|\; \mathcal{A}_0 \Bigg] \Bigg| + \mathcal{O}\left(\frac{1}{n}\right) \end{aligned}$$

where

$$\begin{aligned} \tilde{c}_2(t,S) &= \frac{Ke^{-r(T-t)}}{\sqrt{2\pi}\sigma\sqrt{T-t}} \exp\{-\frac{(d_1 - \sigma\sqrt{T-t})^2}{2}\} \\ \tilde{c}_3(t,S) &= -\frac{Ke^{-r(T-t)}}{\sqrt{2\pi}\sigma^2(T-t)} \left((d_1 - \sigma\sqrt{T-t}) + 2\sigma\sqrt{T-t} \right) \exp\{-\frac{(d_1 - \sigma\sqrt{T-t})^2}{2}\} \end{aligned}$$

If we set $\tilde{K} := Ke^{-(r+\frac{\sigma^2}{2})(T-t)}$ for $t \in [0,T]$ (see Lemma 2 in Appendix B) we note that they are

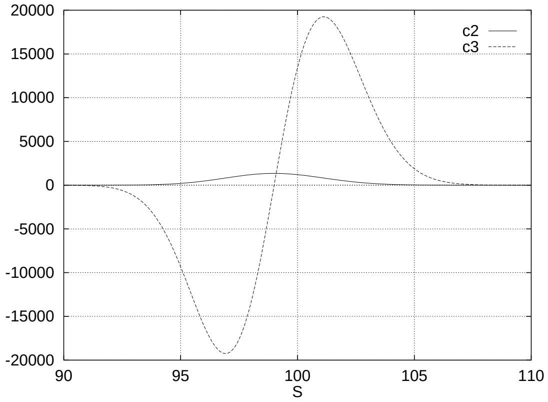

FIGURE 19. $\tilde{c}_2(t,S)$ and $\tilde{c}_3(t,S)$ for 1/100

symmetrical around $\tilde{K}$.

Moreover we remark, that $\tilde{c}_2(t,S)$ and $\tilde{c}_3(t,S)$ are critical in t as $t \longrightarrow T$, that is as the remaining time–to–maturity becomes 0.

The time–point t rules the maximum of the two functions; moreover, since the exponential function becomes dominating very quickly, it gives the "width" of the functions.

We also recognize, that $\tilde{c}_3(t,S)$ is positive for $S > \tilde{K}$ and negative otherwise. The maxima are lying in

a range less than $\tilde{K}e^{\sigma\sqrt{t}}$.
We believe that this changing sign is responsible for oscilliating convergence patterns. It is extremely critical for little values of t, since then they are even amplified by $\frac{1}{\sqrt{t}}$ respexctively $\frac{1}{t}$

To summarize : the behaviour is extremely critical in a range of order $\tilde{K}e^{\sigma\sqrt{t}}$.
At our last but one time point $t_{n-1} = (n-1)\frac{T}{n}$ we have that this critical range is of order of u and d.

However, this does not need to present a problem, moreover correctly adjusted it is a chance to get monotonically and quick convergence. Because of the symmetry of the functions we may even hope to get $E_B[\tilde{c}_3(t, Y_k)] = 0$. Actually this is not possible, but there is need of the knowledge how to adjust the parameters u, d properly, such that they fit best. This is done by the adjusting function within the normal approximations (see Appendix A).

We are not going to extend this examination. But we want to present the effect in some simulations at the most critical time point t_{n-1} :

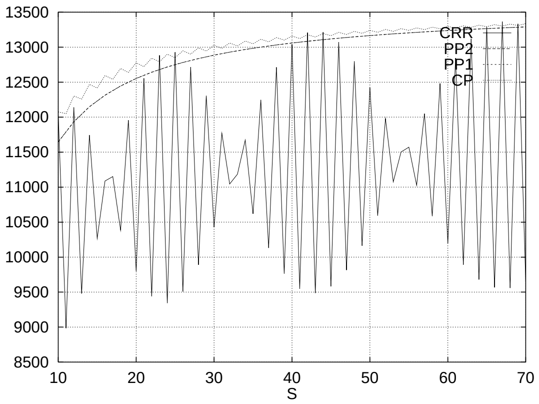

FIGURE 20. $E_B[\tilde{c}_2(t_{n-1}, Y_{n-1})]$

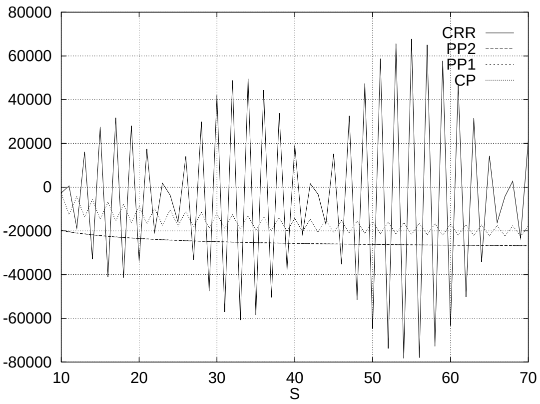

FIGURE 21. $E_B[\tilde{c}_3(t_{n-1}, Y_{n-1})]$

One sees very well that our approaches exhibit only neglecting oscillations in comparision to CRR.

5. Numerical results

In this section we present computational results. We compare 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 prosecuting 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 . Remarkably, 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 maximal error 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	smo.	CP	PP1	PP2	True value
European Call Options								
80	23.74082	23.76300	23.70657	23.86642	23.76050	23.75822	23.75875	23.75799
90	16.13376	16.08486	16.12494	16.21076	16.09619	16.09941	16.10037	16.09963
100	10.21317	10.20142	10.20418	10.22651	10.12545	10.13316	10.13440	10.13377
110	6.01218	6.02481	6.01304	6.03261	5.94162	5.94889	5.95015	5.94946
120	3.31890	3.33429	3.33318	3.37003	3.27993	3.28258	3.28366	3.28280
European Put Options								
80	0.98926	1.01143	0.95500	1.11485	1.00893	1.00665	1.00719	1.00642
90	3.03825	2.98934	3.02943	3.11524	3.00068	3.00390	3.00486	3.00412
100	6.77371	6.76196	6.76472	6.78705	6.68599	6.69370	6.69494	6.69431
110	12.22878	12.24141	12.22963	12.24920	12.15821	12.16548	12.16675	12.16606
120	19.19155	19.20694	19.20583	19.24268	19.15258	19.15523	19.15631	19.15545
American Put Options								
80	1.01842	1.03864	0.98396	1.15261	1.04231	1.04264	1.04317	1.037
90	3.16580	3.12447	3.14640	3.24107	3.11786	3.12832	3.12928	3.123
100	7.10823	7.10415	7.08701	7.12158	7.00982	7.02858	7.02981	7.035
110	13.00108	13.01511	12.98978	13.00907	12.90304	12.93136	12.93253	12.955
120	20.73344	20.74479	20.73566	20.73510	20.65254	20.67576	20.67649	20.717

Table 1. 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 exactly matches the choice of Broadie, Detemple.
Figure 22 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

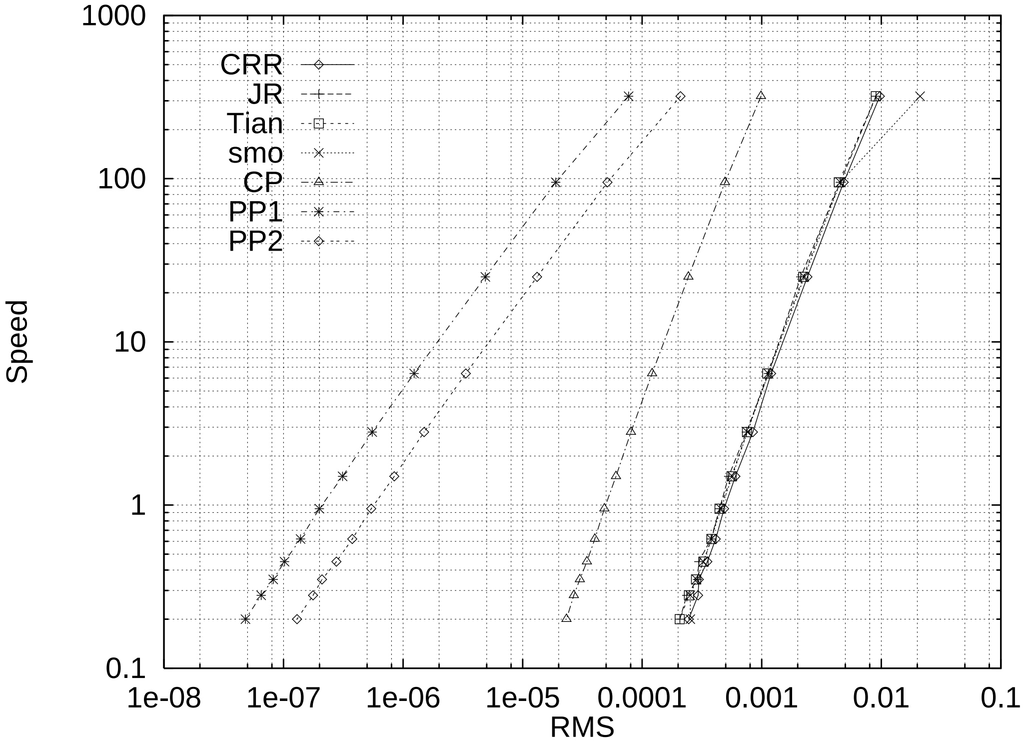

FIGURE 22. *testing efficiency of binomial models for European call options with* $n_i = \{25, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000\}$

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 approach for smooth convergence. 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.

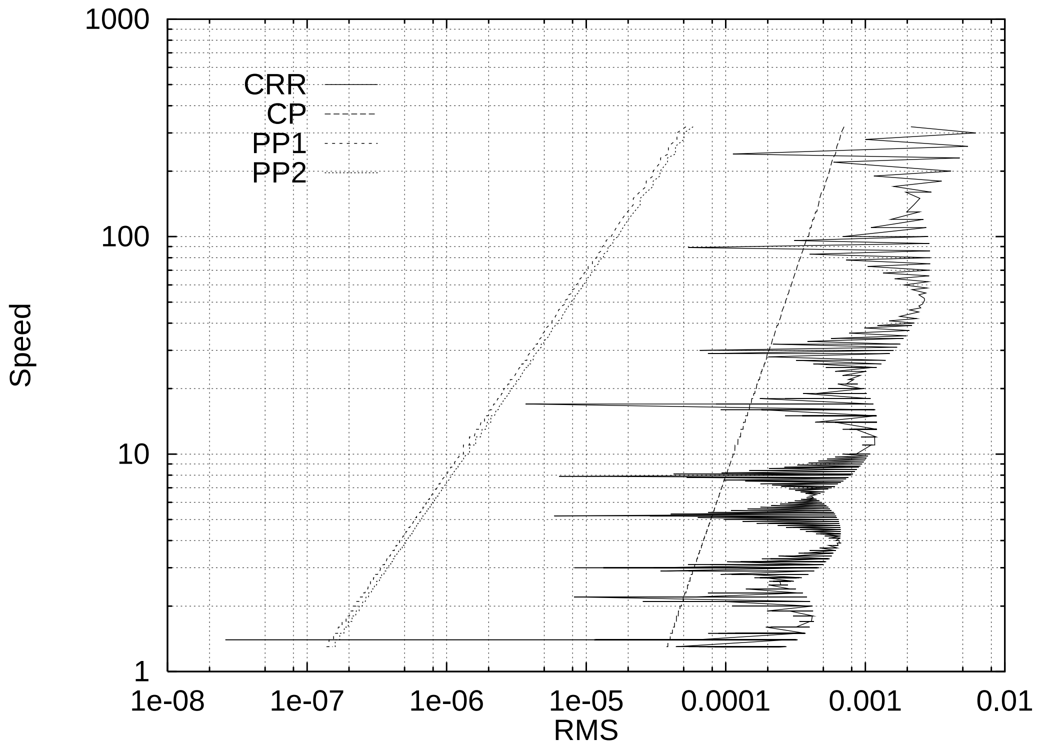

FIGURE 23. *single selection of parameters with* $n_i = [25, 500]$

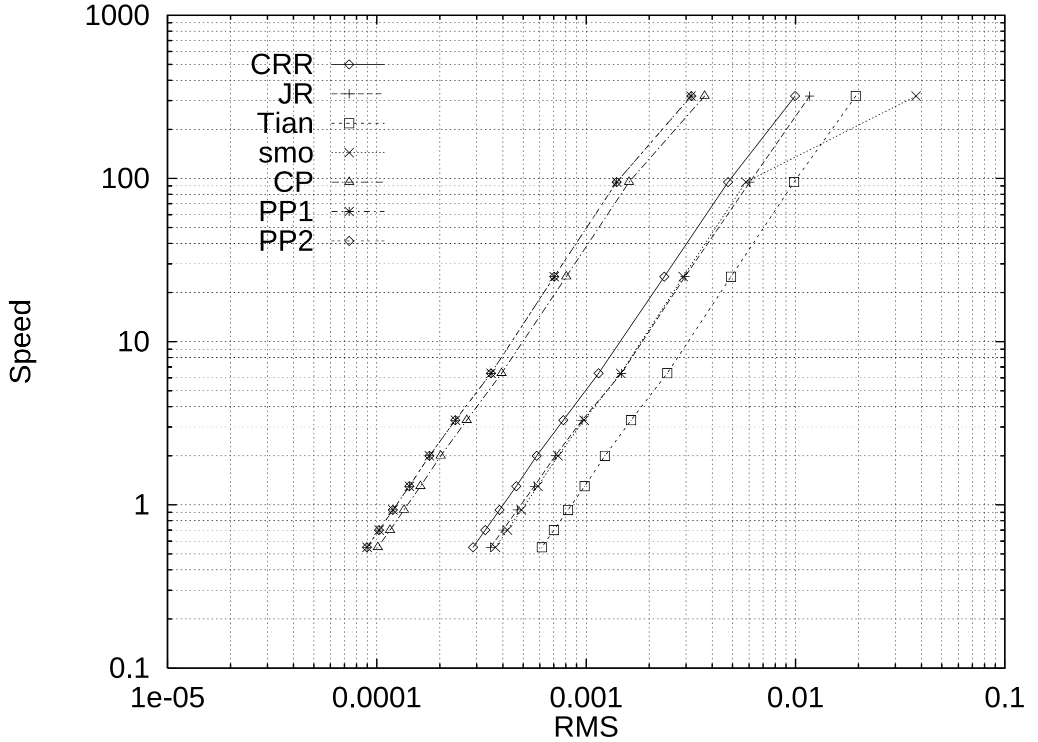

FIGURE 24. *testing efficiency of binomial models for American put options with* $n_i = \{25, 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000\}$

Finally, figure 24 reports speed–accuracy properties with the calculation of American type options. Whereas, similar results are obtained for the previous models, the new models contain order of convergence one here only but with small initial error. Naturally, the design solely assured high accuracy with respect to the terminal payoff distribution. Approximation of early exercise premiums involve original sources of irregularities. Nevertheless, the unexpectable stability shows that the approximation error chiefly arises from deficiencies in connection with the terminal payoff distribution.

6. Conclusion

Convergence speed and convergence patterns of three previously existing lattice approaches were examined. Generally, we find order of convergence one. Unfortunately, convergence is distorted by over tree refinements fluctuating relative positions of the strike price. We succeeded to construct a binomial model which exhibits smooth convergence. Moreover, we presented a smoothly converging model with convergence order one but improved coefficient. Chiefly, we presented a smoothly converging model with order of convergence two. Finally, we listed simulation results. Especially we conducted an examination of computation speed and accuracy for a large sample of randomly selected parameter constellations. Remarkably, price computation of American type option is improved. Transfering these findings to the valuation of complex option types remains for future research.

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:

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

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 $z^* = [(j+\frac{1}{2})-np]/\sqrt{rpq}$, where $j+\frac{1}{2}$ denotes the number of successes in n Bernoulli trials with continuity correction, they correct for misplacement of the median by

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

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:

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

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

Further modification to the first part delivers a second approximation

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

In the case where $j+\frac{1}{2} = n-(j+\frac{1}{2})$ 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:

In this Appendix we will prove the theorem from section 3 about order–of–convergence. The conduction of the proof is rather cumbersome and sophisticated mathematics are involved. Unfortunately, further introductory notation is needed. The proof is separated into several parts. The actual proof is ending after making the Taylor–Expansion. Then convergence is already separated into an option dependent and a lattice dependent part.
The following lemmata are rather technical. They give estimates for the "constant" in the definition of order–of–convergence (see equation 16). We will adopt the following

Notation.

1. $X_t^{s,x}$ shall be the solution of the stochastic differential equation

$$X_t^{s,x} = x + \int_s^t r X_{t'}^{s,x} dt' + \int_s^t \sigma X_{t'}^{s,x} dW_{t'} \ .$$

   The probability measure is denoted by P_W.
2. $c(t,S) := e^{-r(T-t)} E_W \left[f(X_T^{t,S}) \right]$ is the Black–Scholes–Price of a call–option with stock value S at time t and maturity at time T.
3. The time axis will be discretized in steps of length $\Delta t = \frac{T}{n}$. The discrete time points will be denoted by $t_i := i \cdot \Delta t$.
4. We use the abbrevation $X_{k+1} = X_{t_{k+1}}^{t_k, Y_k}$ for notational simplicity.
5. The information structure is modelled by $\mathcal{A}_k = \sigma(Y_i \mid j \leq k)$.
6. Let

$$\begin{array}{lclllcl} c_1(t,S) & := & \dfrac{\partial c}{\partial S}(t,S) & & \tilde{c}_1(t,S) & := & S \cdot c_1(t,S) \\ c_2(t,S) & := & \dfrac{\partial^2 c}{\partial S^2}(t,S) & & \tilde{c}_2(t,S) & := & S^2 \cdot c_2(t,S) \\ c_3(t,S) & := & \dfrac{\partial^3 c}{\partial S^3}(t,S) & & \tilde{c}_3(t,S) & := & S^3 \cdot c_3(t,S) \end{array}$$

   Please note : $d_1(S) = \dfrac{\ln \frac{S}{K} + \left(r + \frac{\sigma^2}{2} \right) t}{\sigma \sqrt{t}}$
   Subsequently we will make no difference between $c_i(t,S), \tilde{c}_i(t,S)$ and $c_i(t, d_1^{-1}(d_1))$ resp. $\tilde{c}_i(t, d_1^{-1}(d_1))$ as functions in (t, d_1). However, when taking partial derivatives one needs to pay some attention: In our notation we have by the chain rule : $\dfrac{\partial c_i}{\partial S} = \dfrac{\partial c_i}{\partial d_1} \dfrac{\partial d_1}{\partial S}$ Since $c_4(t,S) := \dfrac{\partial c_3}{\partial d_1} \left(t, d_1^{-1}(S) \right)$ becomes more significant later, we will denote it by a special letter.
7. Let $R_3(t, z_1, z_0) := \displaystyle\int_{z_0}^{z_1} (z_1 - S)^3 \frac{\partial c_3}{\partial S}(t,S) dS \quad \forall \ z_0, z_1 \in \mathbb{R}^+ \quad \forall \ t \in [0,T]$.
8. Let

$$\begin{array}{lcl} M_2(t) & := & \dfrac{K e^{-rt}}{\sqrt{2\pi}\sigma}, \quad M_3(t) \ := \dfrac{K e^{-rt}}{\sqrt{2\pi}\sigma^2}(1+\sigma) \\ M_4(t) & := & 4 \dfrac{e^{(2r+3\sigma^2)t}}{\sqrt{2\pi} \ K^2 \sigma^2} \end{array}$$

9. Let

$$\begin{aligned}
m_n^1 &:= E_B\left[\frac{Y_{k+1}}{Y_k}-1 \;\Big|\; \mathcal{A}_k\right] - E_W\left[\frac{X_{k+1}}{Y_k}-1 \;\Big|\; \mathcal{A}_k\right] \\
m_n^2 &:= E_B\left[\left(\frac{Y_{k+1}}{Y_k}-1\right)^2 \;\Big|\; \mathcal{A}_k\right] - E_W\left[\left(\frac{X_{k+1}}{Y_k}-1\right)^2 \;\Big|\; \mathcal{A}_k\right] \\
m_n^3 &:= E_B\left[\left(\frac{Y_{k+1}}{Y_k}-1\right)^3 \;\Big|\; \mathcal{A}_k\right] - E_W\left[\left(\frac{X_{k+1}}{Y_k}-1\right)^3 \;\Big|\; \mathcal{A}_k\right] \\
p_n &:= E_B\left[\left(\ln\frac{Y_{k+1}}{Y_k}\right)\left(\frac{Y_{k+1}}{Y_k}-1\right)^3 \;\Big|\; \mathcal{A}_k\right] \\
\bar{p}_n &:= E_W\left[\left(\ln\frac{X_{k+1}}{Y_k}\right)\left(\frac{X_{k+1}}{Y_k}-1\right)^3 \;\Big|\; \mathcal{A}_k\right]
\end{aligned}$$

Notice that they donot depend on specific k.

10. We often need to evaluate intervals $[Y_k, Y_{k+1}]$ with Y_{k+1} at random. So maybe there is even $Y_{k+1} < Y_k$; thus we use for simplicity the definition

$$[Y_k, Y_{k+1}] := [Y_{k+1}, Y_k] \text{ if } Y_{k+1} < Y_k \ .$$

We will now state the theorem; please note that we state it here in a sligtly more **general** form using m_n^2, m_n^3 instead of $\bar{m}_n^2, \bar{m}_n^3$

Theorem. Let $\{Y_0^n, \dots, Y_n^n\}$ be a lattice approach with $Y_0^n = Y_0 \quad \forall \quad n$.
Let e_n be the error in the price of a European call option that is

$$e_n := e^{-rT}\left|E_B\left[f(Y_n) \mid \mathcal{A}_0\right] - E_W\left[f(X_T^{0,Y_0}) \mid \mathcal{A}_0\right]\right|$$

Then there exists a constant C, only depending on S, K, r, σ, T such that :

$$e_n \le C\left(\frac{m_n^2 + m_n^3 + p_n + (\Delta t)^2}{\Delta t}\right)$$

Proof. Since $f(Y_n) = c(T, Y_n)$ we get

$$\begin{aligned}
E_B\left[f(Y_n) \mid \mathcal{A}_0\right] &= E_B\left[c(T, Y_n) \mid \mathcal{A}_0\right] \quad \text{and} \\
e^{-rT} E_W\left[f(X_T^{0,Y_0}) \mid \mathcal{A}_0\right] &= c(0, Y_0) \;=\; E_B\left[c(0, Y_0) \mid \mathcal{A}_0\right]
\end{aligned}$$

Therefore :

$$\begin{aligned}
e_n &= \left|E_B\left[e^{-rT} c(T, Y_n) - c(0, Y_0) \mid \mathcal{A}_k\right]\right| \\
&= \left|E_B\left[\sum_{k=0}^{n-1} e^{-rt_k}\left\{e^{-r\Delta t} c(t_{k+1}, Y_{k+1}) - c(t_k, Y_k)\right\} \;\Big|\; \mathcal{A}_k\right]\right|
\end{aligned}$$

We next observe that the Black–Scholes price is riskless, which means

$$E_W\left[e^{-r\Delta t} c(t_{k+1}, X_{k+1}) - c(t_k, X_{t_k}^{t_k, Y_k}) \mid \mathcal{A}_k\right] = 0$$

It follows :

$$
\begin{aligned}
e_n &= \Bigg| E_B\Bigg[\sum_{k=0}^{n-1} e^{-rt_k}\Big\{E_B\Big[e^{-r\Delta t}c(t_{k+1},Y_{k+1}) - c(t_k,Y_k)\;\Big|\;\mathcal{A}_k\Big] \\
&\qquad -E_W\Big[e^{-r\Delta t}c(t_{k+1},X_{k+1}) - \underbrace{c(t_k,X_{t_k}^{t_k,Y_k})}_{=c(t_k,Y_k)}\;\Big|\;\mathcal{A}_k\Big]\Big\}\;\Bigg|\;\mathcal{A}_0\Bigg]\Bigg| \\
&= \Bigg| E_B\Bigg[\sum_{k=0}^{n-1} e^{-rt_{k+1}}\Big\{E_B\Big[c(t_{k+1},Y_{k+1})\;\Big|\;\mathcal{A}_k\Big] - E_W\Big[c(t_{k+1},X_{k+1})\;\Big|\;\mathcal{A}_k\Big]\Big\}\;\Bigg|\;\mathcal{A}_0\Bigg]\Bigg| \\
&= \Bigg| E_B\Bigg[\sum_{k=0}^{n-1} e^{-rt_{k+1}}\Big\{E_B\Big[c(t_{k+1},Y_{k+1}) - c(t_{k+1},Y_k)\;\Big|\;\mathcal{A}_k\Big] \\
&\qquad -E_W\Big[c(t_{k+1},X_{k+1}) - c(t_{k+1},Y_k)\;\Big|\;\mathcal{A}_k\Big]\Big\}\;\Bigg|\;\mathcal{A}_0\Bigg]\Bigg|
\end{aligned}
$$

The last time point $k = n-1$ is evaluated separatedly in Lemma 8 and 9 as $\mathcal{O}\left(\frac{1}{n}\right)$. The other time points are evaluated by a Taylor series expansion around Y_k. This yields:

$$
\begin{aligned}
e_n \;\le\; \mathcal{O}\left(\frac{1}{n}\right) &+ \Bigg| E_B\Bigg[\sum_{k=0}^{n-2} e^{-rt_{k+1}}\Big\{E_B\Big[c_1(t_{k+1},Y_k)(Y_{k+1}-Y_k) + c_2(t_{k+1},Y_k)(Y_{k+1}-Y_k)^2 \\
&+ c_3(t_{k+1},Y_k)(Y_{k+1}-Y_k)^3 + R_3(t_{k+1},Y_{k+1},Y_k)\;\Big|\;\mathcal{A}_k\Big] \\
&- E_W\Big[c_1(t_{k+1},Y_k)(X_{k+1}-Y_k) + c_2(t_{k+1},Y_k)(X_{k+1}-Y_k)^2 \\
&+ c_3(t_{k+1},Y_k)(X_{k+1}-Y_k)^3 + R_3(t_{k+1},X_{k+1},Y_k)\;\Big|\;\mathcal{A}_k\Big]\Big\}\;\Bigg|\;\mathcal{A}_0\Bigg]\Bigg|
\end{aligned}
$$

One has:

$$
\begin{aligned}
E_B\Big[c_1(t_{k+1},Y_k)(Y_{k+1}-Y_k)\;\Big|\;\mathcal{A}_k\Big] &= \tilde{c}_1(t_{k+1},Y_k)\; E_B\left[\frac{Y_{k+1}}{Y_k}-1\;\Big|\;\mathcal{A}_k\right] \\
E_W\Big[c_1(t_{k+1},Y_k)(X_{k+1}-Y_k)\;\Big|\;\mathcal{A}_k\Big] &= \tilde{c}_1(t_{k+1},Y_k)\; E_W\left[\frac{X_{k+1}}{Y_k}-1\;\Big|\;\mathcal{A}_k\right]
\end{aligned}
$$

One gets analogous results for terms with $\tilde{c}_2$ and $\tilde{c}_3$. Therefore :

$$
\begin{aligned}
e_n \;\le\; \mathcal{O}\left(\frac{1}{n}\right) &+ \Bigg| E_B\Bigg[\sum_{k=0}^{n-2} e^{-rt_{k+1}}\Big\{\tilde{c}_1(t_{k+1},Y_k)\cdot m_n^1 + \tilde{c}_2(t_{k+1},Y_k)\cdot m_n^2 + \tilde{c}_3(t_{k+1},Y_k)\cdot m_n^3 \\
&\qquad +E_B\Big[R_3(t_{k+1},Y_{k+1},Y_k)\;\Big|\;\mathcal{A}_k\Big] - E_W\Big[R_3(t_{k+1},X_{k+1},Y_k)\;\Big|\;\mathcal{A}_k\Big]\Big\}\;\Bigg|\;\mathcal{A}_0\Bigg]\Bigg| \\
\le\; \mathcal{O}\left(\frac{1}{n}\right) &+ \left|\sum_{k=0}^{n-2} e^{-rt_{k+1}} E_B\Big[\tilde{c}_1(t_{k+1},Y_k)\;\Big|\;\mathcal{A}_0\Big]\right|\cdot\left|m_n^1\right| \\
&+ \left|\sum_{k=0}^{n-2} e^{-rt_{k+1}} E_B\Big[\tilde{c}_2(t_{k+1},Y_k)\;\Big|\;\mathcal{A}_0\Big]\right|\cdot\left|m_n^2\right| \\
&+ \left|\sum_{k=0}^{n-2} e^{-rt_{k+1}} E_B\Big[\tilde{c}_3(t_{k+1},Y_k)\;\Big|\;\mathcal{A}_0\Big]\right|\cdot\left|m_n^2\right| \\
&+ \Bigg| E_B\Bigg[\sum_{k=0}^{n-2} e^{-rt_{k+1}}\Big\{E_B\Big[R_3(t_{k+1},Y_{k+1},Y_k)\;\Big|\;\mathcal{A}_k\Big] \\
&\qquad +E_W\Big[R_3(t_{k+1},X_{k+1},Y_k)\;\Big|\;\mathcal{A}_k\Big]\Big\}\;\Bigg|\;\mathcal{A}_0\Bigg]\Bigg|
\end{aligned}
$$

The proof now follows immediately from the following lemmata.

Lemma 1. Then

$$\begin{aligned}\max_{b\in\mathbb{R}}\left|e^{-\frac{b^2}{2}}\right| &= 1\\ \max_{b\in\mathbb{R}}\left|be^{-\frac{b^2}{2}}\right| &= e^{-\frac{1}{2}}\\ \max_{b\in\mathbb{R}}\left|(b^2-1)e^{-\frac{b^2}{2}}\right| &= 1\end{aligned}$$

Further for $t<1$:

$$\begin{aligned}\forall\ b\notin\left[-\sqrt{\ln\frac{1}{t}},\sqrt{\ln\frac{1}{t}}\right] &: \quad \frac{1}{\sqrt{t}}e^{-\frac{b^2}{2}} < 1\\ \forall\ b\notin\left[-2\sqrt{\ln\frac{1}{t}},2\sqrt{\ln\frac{1}{t}}\right] &: \quad \left|\frac{b}{\sqrt{t}}e^{-\frac{b^2}{2}}\right| < 1 \text{ and } \left|\frac{b}{t}e^{-\frac{b^2}{2}}\right| < 1\\ \forall\ b\notin\left[-2\sqrt{\ln\frac{1}{t}},2\sqrt{\ln\frac{1}{t}}\right] &: \quad \left|\frac{b^2-1}{t}e^{-\frac{b^2}{2}}\right| < 2\end{aligned}$$

Proof. $H_0\equiv 1$, $H_1(b)=b$, $H_2(b)=b^2-1$ are the first three Hermite–Polynomials. The first part now follows from the properties given in Abramowitz and Stegun [68]. Moreover one gets from there, that the functions in the second part are strictly decreasing in their respective range. Therefore they can be limited by their value at the boundary of their respective range. Since $\forall\ 1>t>0\ :\ \left|2\sqrt{\ln\frac{1}{t}}\sqrt{t}\right|<1$ this proves the lemma.

Lemma 2.

$$\begin{aligned}\tilde{c}_2(t,S) &= \frac{Ke^{-r(T-t)}}{\sqrt{2\pi}\sigma\sqrt{T-t}}e^{-\frac{(d_1-\sigma\sqrt{T-t})^2}{2}}\\ \tilde{c}_3(t,S) &= -\frac{Ke^{-r(T-t)}}{\sqrt{2\pi}\sigma^2(T-t)}\left((d_1-\sigma\sqrt{T-t})+2\sigma\sqrt{T-t}\right)e^{-\frac{(d_1-\sigma\sqrt{T-t})^2}{2}}\\ c_4(t,S) &= \frac{e^{(2r+3\sigma^2)(T-t)}}{\sqrt{2\pi}\ K^2\sigma^2(T-t)}\left((d_1+2\sigma\sqrt{T-t}))^2-\sigma\sqrt{T-t}(d_1+2\sigma\sqrt{T-t})+1-\sigma^2(T-t)\right)e^{-\frac{(d_1+2\sigma\sqrt{t})^2}{2}}\end{aligned}$$

Moreover :

$$\begin{aligned}|\tilde{c}_2(t,S)| &\leq \frac{M_2(T-t)}{\sqrt{T-t}}, \quad |\tilde{c}_3(t,S)|\leq\frac{M_3(T-t)}{T-t}\\ |c_4(t,S)| &\leq \frac{M_4(T-t)}{T-t}\end{aligned}$$

Proof. We remark that

$$d_1(S)=\frac{\ln\frac{S}{K}+(r+\frac{\sigma^2}{2})(T-t)}{\sigma\sqrt{T-t}}\Leftrightarrow S=Ke^{\sigma\sqrt{T-t}\ d_1-(r+\frac{\sigma^2}{2})(T-t)}$$

which implies:

$$\frac{\partial d_1}{\partial S}=\frac{1}{S\sigma\sqrt{T-t}}$$

We secondly remark, that

$$\frac{\partial c_i}{\partial S}=\frac{\partial c_i}{\partial d_1}\frac{\partial d_1}{\partial S}$$

by the chain–rule. Now one gets the derivatives by simple calculations. The estimates follow immediatly from those.

Lemma 3. $|R_3(t,z_1,z_0)|\ \leq\ \left(\ln\frac{z_1}{z_0}\right)(z_1-z_0)^3\max_{S\in[z_0,z_1]}|c_4(t,S)|$

Proof. Writing $\bar{c}_3(t, d_1^{-1}(S)) := c_3(t, S)$ we can understand c_3 as a Function in (t, d_1). In this form we get by the chain rule :

$$\frac{\partial c_3}{\partial S}(t,S) = \frac{\partial(\bar{c}_3 \circ d_1)}{\partial S}(t,S) = \frac{\partial(\bar{c}_3 \circ d_1)}{\partial d_1}(t,S) \cdot \frac{\partial d_1}{\partial S}(S) = c_4(t, d_1) \cdot \frac{\partial d_1}{\partial S}(S)$$

Thus we get :

$$\begin{aligned} |R_3(t, z_1, z_0)| &\overset{\text{Def.}}{=} \left| \int_{z_0}^{z_1} (z_1 - S)^3 \frac{\partial c_3}{\partial S}(t,S) dS \right| \\ &= \left| \int_{z_0}^{z_1} (z_1 - S)^3 \, c_4\left(t, d_1(S)\right) \frac{\partial d_1}{\partial S}(S) dS \right| \\ &= \left| \int_{d_1(z_0)}^{d_1(z_1)} \left(z_1 - d_1^{-1}(S)\right)^3 c_4\left(t, S\right) dS \right| \end{aligned}$$

by the transformation $S \mapsto d_1(S)$.
Since d_1^{-1} is strictly decreasing, up to c_4 the integrand is negative exactly when $d_1(z_0) > d_1(z_1) \Leftrightarrow z_0 > z_1$. However in this case the integration boundaries of the Riemannian integral need to be changed, which compensates the minus sign. So one has:

$$\begin{aligned} |R_3(t, z_1, z_0)| &\le \int_{d_1(z_0)}^{d_1(z_1)} \left(z_1 - d_1^{-1}(S)\right)^3 \left|c_4\left(t, d_1^{-1}(S)\right)\right| dS \\ &\le (z_1 - z_0)^3 \max_{S \in [z_0, z_1]} |c_4(t,S)| \int_{d_1(z_0)}^{d_1(z_1)} dy \end{aligned}$$

This immediately implies the lemma.

Lemma 4. Let $m := \left(r - \frac{\sigma^2}{2}\right) \cdot \frac{T}{n}$, $\mathrm{CR} := \left\{ e^b \;\middle|\; b \in \left[-5\sigma\sqrt{(\ln\frac{1}{\Delta t})\Delta t}, 5\sigma\sqrt{(\ln\frac{1}{\Delta t})\Delta t}\;\right]\right\}$.
Suppose that $m < \sigma\sqrt{\frac{T}{n}}$ and $\sqrt{(\ln\frac{1}{\Delta t})\Delta t} > 4\sigma$.
Let Z be a random variable which is lognormally distributed, specifically $\ln Z \sim N\left(m, \sigma\sqrt{\frac{T}{n}}\,\right)$.
Then : $E_W\left[I_{Z \notin \mathrm{CR}}(\ln Z)(Z-1)^3\right] \le \frac{4e^{4|m|}}{\sqrt{2\pi}} \Delta t^4 \sqrt{\ln\frac{1}{\Delta t}}$

Proof. Let $\tilde{R}_u := \left\{ e^b \;|b > 5\sigma\sqrt{(\ln\frac{1}{\Delta t})\Delta t}\right\}$ $\tilde{R}_d := \left\{ e^b \;|b < -5\sigma\sqrt{(\ln\frac{1}{\Delta t})\Delta t}\right\}$
Let $R_u := \left\{ b \;|b > 4\sigma\sqrt{(\ln\frac{1}{\Delta t})\Delta t}\right\} \in \tilde{R}_u$, $R_d := \left\{ b \;|b < -4\sigma\sqrt{(\ln\frac{1}{\Delta t})\Delta t}\right\} \in \tilde{R}_d$
Then :

$$\begin{aligned} E_W\left[I_{Z \notin \mathrm{CR}}(\ln Z)(Z-1)^3\right] &= \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} \int_{\tilde{R}_u \cup \tilde{R}_d} \underbrace{(\ln e^y)(e^y - 1)^3}_{< e^{4|y|} < e^{4|y-m| + 4|m|}} e^{-\frac{(y-m)^2}{2\sigma^2} \cdot \Delta t} dy \\ &< \frac{e^{4|m|}}{\sqrt{2\pi\sigma^2\Delta t}} \int_{R_u} e^{4y} e^{-\frac{y^2}{2\sigma^2}\Delta t} dy + \frac{e^{4|m|}}{\sqrt{2\pi\sigma^2\Delta t}} \int_{R_d} e^{-4y} e^{-\frac{y^2}{2\sigma^2}\Delta t} dy \end{aligned}$$

by a linear transformation.
Because of the symmetry, the second part is exactly identical to the first part; so we only need to estimate the first part.
Since $\sqrt{(\ln\frac{1}{\Delta t})\Delta t} > 4\sigma$ one has $4\sigma\sqrt{(\ln\frac{1}{\Delta t})\Delta t} > 16\sigma^2\Delta t$. For $y > 16\sigma^2\Delta t$: $1 < \frac{y}{16\sigma^2\Delta t} \Rightarrow y < \frac{y^2}{16\sigma^2\Delta t}$

$\Rightarrow y - \frac{y^2}{8\sigma^2\Delta t} < -\frac{y^2}{16\sigma^2\Delta t} \Rightarrow 4y - \frac{y^2}{2\sigma^2\Delta t} = 4\left(y - \frac{y^2}{8\sigma^2\Delta t}\right) < 4\left(-\frac{y^2}{16\sigma^2\Delta t}\right) = -\frac{y^2}{4\sigma^2\Delta t}$

Therefore:

$$\begin{aligned}
\frac{1}{\sqrt{2\pi\sigma^2\Delta t}} \int_{R_u} e^{4y} e^{-\frac{y^2}{2\sigma^2\Delta t}} &< \frac{1}{\sqrt{2\pi\sigma^2\Delta t}} \int_{R_u} e^{-\frac{y^2}{4\sigma^2\Delta t}} dy = \sqrt{2}\frac{1}{\sqrt{2\pi(\sigma\sqrt{2\Delta t})^2}} \int_{R_u} e^{-\frac{y^2}{2(\sigma\sqrt{2\Delta t})^2}} dy \\
&< \sqrt{2}\,\frac{1}{\sqrt{2\pi}}\, \underbrace{e^{-\frac{z^2}{2(\sigma\sqrt{2\Delta t})^2}}}_{=e^{-\frac{z^2}{4\sigma^2\Delta t}}=(\Delta t)^4} \cdot \frac{\sigma\sqrt{2\Delta t}}{z} \qquad \text{with } z = 4\sigma\sqrt{(\ln\frac{1}{\Delta t})\Delta t} \\
&= \frac{2}{\sqrt{2\pi}}\Delta t^4 \sqrt{\ln\frac{1}{\Delta t}}
\end{aligned}$$

The last inequality is an estimate for the "tail–probability" of the normal distribution. It is proven in Feller [57] for variance = 1. The case needed here follows easily by a linear transformation.

Lemma 5.

$$\sum_{k=0}^{n-2} e^{-rt_{k+1}} E_B\left[|\tilde{c}_3(t_{k+1}, Y_k)| \;\Big|\; \mathcal{A}_0\right] \le \left(\frac{8\sigma M_3}{\sqrt{pq}}2 + M_3\right)\cdot n \quad \text{where} \quad M_3 = M_3(0)$$

Proof. For the moment, suppose $0 < t < \min\{T, 1\}$.
As shown in Lemma 2 : $|\tilde{c}_3(t,S)| = \frac{Ke^{-r(T-t)}}{\sqrt{2\pi}\ \sigma^2}\left|\frac{(d_1-\sigma\sqrt{T-t})+2\sigma\sqrt{T-t}}{T-t}e^{-\frac{(d_1-\sigma\sqrt{T-t})^2}{2}}\right|$.

Let $\tilde{K} := Ke^{-(r+\frac{\sigma^2}{2})(T-t)}$.

One can easily verify: $d_1(S) - \sigma\sqrt{T-t} = b \Leftrightarrow S = \tilde{K}e^{b\sigma\sqrt{T-t}}$.

Let CR $:= \left\{\tilde{K}e^{b\sigma\sqrt{T-t}} \Big| b \in \left[-2\sqrt{\ln\frac{1}{T-t}}, 2\sqrt{\ln\frac{1}{T-t}}\right]\right\}$.

Let $\alpha := -2\sigma\sqrt{(\ln\frac{1}{T-t})(T-t)}, \beta := 2\sigma\sqrt{(\ln\frac{1}{T-t})(T-t)}, x_\alpha := \frac{\alpha-np}{\sqrt{npq}}, x_\beta := \frac{\beta-np}{\sqrt{npq}}$, Φ be the normal distribution function. Since $\ln Y_k$ is binomially bistributed, we know from the DeMoivre-Laplace Theorem (Feller [57], p. 172), that $P[\alpha \le lnY_k \le \beta]$ can be approximated by $\Phi(x_\beta) - \Phi(x_\alpha)$.
Surely we have

$$\begin{aligned}
P[\alpha \le \ln Y_k \le \beta] &\le 2\cdot(\Phi(x_\beta) - \Phi(x_\alpha)) = 2\int_{x_\alpha}^{x_\beta} \underbrace{e^{\frac{-x^2}{2}}}_{\le 1} dx \le 2(x_\beta - x_\alpha) = 2\frac{\beta-\alpha}{\sqrt{npq}} = 8\frac{\sigma\sqrt{(T-t)\ln\frac{1}{T-t}}}{\sqrt{npq}} \\
&\le 8\frac{\sigma}{\sqrt{pq}}\sqrt{T-t}
\end{aligned}$$

For $S \notin$ CR one gets from Lemma 1 and 2:

$$|\tilde{c}_3(t,S)| < M_3$$

Then: $E_B\left[|\tilde{c}_3(t,Y_k)| \;\Big|\; \mathcal{A}_0\right] = E_B\left[I_{Y_k\in \mathrm{CR}}\cdot \underbrace{|\tilde{c}_3(t,Y_k)|}_{\le M_3\cdot\frac{1}{T-t}} \;\Big|\; \mathcal{A}_0\right] + E_B\left[I_{Y_k\notin \mathrm{CR}}\cdot \underbrace{|\tilde{c}_3(t,Y_k)|}_{\le M_3} \;\Big|\; \mathcal{A}_0\right]$

$\le \frac{M_3}{T-t}\cdot P[Y_k \in \mathrm{CR}] + M_3 \le \frac{8\sigma M_3}{\sqrt{pq(T-t)}} + M_3$

Since $\sum_{k=1}^{n} \frac{1}{\sqrt{k}} < 2\sqrt{n}$ we have proven the lemma.

Lemma 6.

$$\sum_{k=0}^{n-2} e^{-rt_{k+1}} E_B\left[E_B\left[|R_3(t, Y_{k+1}, Y_k)| \;\Big|\; \mathcal{A}_k\right] \;\Big|\; \mathcal{A}_0\right] \le \left(\frac{12\sigma M_4}{\sqrt{pq}}2 + M_4 E_B[Y_k^3]\right)\cdot n \quad \text{where} \quad M_4 = M_4(0)$$

Proof. As shown in Lemma 2 :

$$|c_4(t,S)| = \frac{e^{(2r+3\sigma^2)t}}{\sqrt{2\pi}\ K^2\sigma^2}\left|\frac{(d_1+2\sigma\sqrt{T-t})^2-\sigma\sqrt{T-t}(d_1+2\sigma\sqrt{T-t})+(1-\sigma^2(T-t))}{T-t}e^{-\frac{(d_1+2\sigma\sqrt{T-t})^2}{2}}\right|$$

.

Let $\tilde{K} := Ke^{-(r-3\frac{\sigma^2}{2})(T-t)}$.

One easily verifies: $d_1(S)+2\sigma\sqrt{T-t}=b \Leftrightarrow S=\tilde{K}e^{b\sigma\sqrt{T-t}}$.

Let $\mathrm{CR} := \left\{\tilde{K}e^{b\sigma\sqrt{T-t}} \;\middle|\; b\in\left[-3\sqrt{\ln\frac{1}{T-t}}-\frac{\ln u}{\sigma\sqrt{T-t}}, \frac{\ln d}{\sigma\sqrt{T-t}}+3\sqrt{\ln\frac{1}{T-t}}\ \right]\right\}$

For $Y_k\notin\mathrm{CR}$ one has: $Y_{k+1}\notin\left\{\tilde{K}e^{b\sigma\sqrt{T-t}} \;\middle|\; b\in\left[-3\sqrt{\ln\frac{1}{T-t}},3\sqrt{\ln\frac{1}{T-t}}\right]\right\}$

so for $y\in[Y_k,Y_{k+1}]$: $\max\limits_{y\in[Y_k,Y_{k+1}]}|c_4(t,y)|\le M_4$ by Lemma 1 and 2.

Therefore, with Lemma 3:

$$\begin{aligned}
&E_B\left[E_B\left[|R_3(t,Y_{k+1},Y_k)| \;\middle|\; \mathcal{A}_k\right] \;\middle|\; \mathcal{A}_0\right]\\
&= E_B\left[I_{Y_k\in\mathrm{CR}}\cdot Y_k^3\cdot\underbrace{E_B\left[|R_3| \;\middle|\; \mathcal{A}_k\right]}_{\le\frac{M_4}{T-t}p_n} \;\middle|\; \mathcal{A}_0\right]+E_B\left[I_{Y_k\notin\mathrm{CR}}\cdot Y_k^3\cdot\underbrace{E_B\left[|R_3| \;\middle|\; \mathcal{A}_k\right]}_{\le M_4\cdot p_n} \;\middle|\; \mathcal{A}_0\right]
\end{aligned}$$

Since $\max\limits_{Y_k\in\mathrm{CR}}Y_k^3 < (2Y_0)^3$ one immediately gets the Lemma as in the proof of Lemma 5.

Lemma 7.

$$\begin{aligned}
&\sum_{k=0}^{n-2}e^{-rt_{k+1}}E_B\left[E_W\left[|R_3(t,X_{k+1},Y_k)| \;\middle|\; \mathcal{A}_k\right] \;\middle|\; \mathcal{A}_0\right]\\
&\le\frac{12\sigma M_4}{\sqrt{pq}}\cdot 2\cdot(2Y_0)^3\cdot\bar{p}_n+E_B\left[Y_k^3\right]\cdot\left(\frac{4M_4e^{4|m|}}{\sqrt{2\pi}(T-t)}(\Delta t)^4\sqrt{\ln\frac{1}{\Delta t}}+M_4\bar{p}_n\right)
\end{aligned}$$

where $M_4=M_4(0)$ as in Lemma 6, $\quad m:=\left(r-\frac{\sigma^2}{2}\right)\frac{T}{n}$.

Proof. Let $\widetilde{\mathrm{CR}}(y) := \left\{ye^b \;\middle|\; b\in\left[-5\sigma\sqrt{(\ln\frac{1}{\Delta t})\Delta t}, 5\sigma\sqrt{(\ln\frac{1}{\Delta t})\Delta t}\right]\right\}$

$$\mathrm{CR} := \bigcup_y \widetilde{\mathrm{CR}}(y), \quad \text{with } y\in\left\{\tilde{K}e^{b\sigma\sqrt{T-t}} \;\middle|\; b\in\left[-3\sqrt{\ln\frac{1}{T-t}},3\sqrt{\ln\frac{1}{T-t}}\right]\right\}$$

where $\tilde{K}$ as in Lemma 6.
As in Lemma 6 one easily proofs: For $Y_k\notin\mathrm{CR}$ and $X_{k+1}\in\widetilde{\mathrm{CR}}(Y_k)$:

$$\max_{y\in[Y_k,Y_{k+1}]}|c_4(t,y)| < M_4$$

Therefore and with Lemma 3 and 4 for $Y_k\notin\mathrm{CR}$:

$$E_W\left[|R_3(t,X_{k+1},Y_k)|\;\middle|\;\mathcal{A}_k\right]$$

$$\begin{aligned}
&\leq\; Y_k^3\left(E_W\left[\max_{y\in[Y_k,Y_{k+1}]}|c_4(t,y)|\cdot\left(\ln\frac{X_{k+1}}{Y_k}\right)\left(\frac{X_{k+1}}{Y_k}-1\right)^3\;\middle|\;\mathcal{A}_k\right]\right) \qquad \text{(by Lemma 3)}\\
&=\; Y_k^3\Bigg(\underbrace{E_W\left[I_{X_{k+1}\in\widetilde{\mathrm{CR}}(Y_k)}\underbrace{\max_{y\in[Y_k,Y_{k+1}]}|c_4|}_{\leq M_4\cdot\frac{1}{T-t}}\cdot\left(\ln\frac{X_{k+1}}{Y_k}\right)\left(\frac{X_{k+1}}{Y_k}-1\right)^3\;\middle|\;\mathcal{A}_k\right]}_{\leq\frac{4M_4e^{4|m|}}{\sqrt{2\pi}(T-t)}\Delta t^4\sqrt{\ln\frac{1}{\Delta t}}\quad\text{by Lemma 4}}\\
&+\;\underbrace{E_W\left[I_{X_{k+1}\notin\widetilde{\mathrm{CR}}(Y_k)}\underbrace{\max_{y\in[Y_k,Y_{k+1}]}|c_4|}_{<M_4}\cdot\left(\ln\frac{X_{k+1}}{Y_k}\right)\left(\frac{X_{k+1}}{Y_k}-1\right)^3\;\middle|\;\mathcal{A}_k\right]}_{<M_4\cdot\bar{p}_n}\Bigg)\\
&<\; Y_k^3\left(\frac{4M_4e^{4|m|}}{\sqrt{2\pi}(T-t)}\Delta t^4\sqrt{\ln\frac{1}{\Delta t}}+M_4\bar{p}_n\right)
\end{aligned}$$

For $Y_k\in\mathrm{CR}$ one gets from Lemma 3 and 4 as above

$$\begin{aligned}
E_W\left[|R_3(t,X_{k+1},Y_k)|\;\middle|\;\mathcal{A}_k\right] &< \frac{M_4}{T-t}\cdot\bar{p}_n\cdot\max_{Y_k\in\mathrm{CR}}Y_k^3\\
&< \frac{M_4}{T-t}\bar{p}_n(2Y_0)^3
\end{aligned}$$

As in Lemma 6 one now immediately gets the Lemma.

Lemma 8.

$$\left|E_B\left[E_W\left[f(X_n)-f(Y_{n-1})\mid\mathcal{A}_{n-1}\right]\;\middle|\;\mathcal{A}_0\right]\right| = \mathcal{O}(\frac{1}{n})$$

Proof. Let $m:=(r-\frac{\sigma^2}{2})\frac{T}{n}$. Let $\tilde{K}_u:=Ke^{3\sigma\sqrt{(\ln\frac{1}{\Delta t})\Delta t}}$, $\tilde{K}_d:=Ke^{-3\sigma\sqrt{(\ln\frac{1}{\Delta t})\Delta t}}$.
Let Z be a log–normally distributed random variable, $\ln Z\sim N\left(m,\sigma^2\frac{T}{n}\right)$.
We have

$$\begin{aligned}
E_W\left[\max\{Z^2,Z^{-2}\}\right] &\leq e^{2|m|}E_W\left[\max\{(Ze^{-m})^2,(Ze^{-m})^{-2}\}\right]\\
&\leq 2e^{2|m|}E_W\left[(Ze^{-m})^2\right]\\
&\leq 2e^{2|m|}e^{\sigma^2\Delta t}
\end{aligned}$$

where the second inequality follows because $\ln(Ze^{-m})$ has mean 0 and is therefore distributed symmetrically.
Since $\forall\ x\in\mathbb{R}\ :\ |e^x-1|\ \leq\ |x|e^{|x|}$ one has:

$$\begin{aligned}
E_W\left[|X_n-Y_{n-1}|\;\middle|\;\mathcal{A}_{n-1}\right] &= Y_{n-1}E\left[|Z-1|\right]\;\leq\;Y_{n-1}E\left[|\ln Z|\max\{Z,\frac{1}{Z}\}\right]\\
&\overset{\substack{\text{Hölder–}\\ \text{Inequality}}}{\leq} Y_{n-1}\cdot\underbrace{E\left[(\ln Z)^2\right]^{\frac{1}{2}}}_{=(\sigma^2\frac{T}{n}+m^2)^{\frac{1}{2}}}\underbrace{E\left[\max\{Z^2,Z^{-2}\}\right]^{\frac{1}{2}}}_{\leq 2e^{2|m|}e^{\sigma^2\Delta t}}\\
&\leq Y_{n-1}\sqrt{\sigma^2+(r-\frac{\sigma^2}{2})^2}\;2e^{2|m|}e^{\sigma^2\Delta t}\sqrt{\frac{T}{n}}
\end{aligned}$$

Exactly as in Lemma 6 one proofs :

$$I_{Y_{n-1}<\tilde{K}_d} \cdot E_W\left[I_{X_n>K} \cdot X_n \;\middle|\; \mathcal{A}_{n-1}\right] \;<\; \frac{e^{|m|}}{\sqrt{2\pi}} \cdot \frac{T}{n} \cdot Y_{n-1}$$

and

$$I_{Y_{n-1}>\tilde{K}_u} P\left[X_n < K\right] \;<\; \frac{e^{|m|}}{\sqrt{2\pi}} \cdot \frac{T}{n}$$

Therefore, since $E\left[X_n - Y_{n-1}\middle| \mathcal{A}_{n-1}\right] = \left(e^{r\frac{T}{n}} - 1\right) Y_{n-1}$:

$$\begin{aligned}
&\left|E_B\left[E_W\left[f(X_n) - f(Y_{n-1}) \;\middle|\; \mathcal{A}_{n-1}\right] \;\middle|\; \mathcal{A}_0\right]\right| \\
\leq\;& E_B\Bigg[I_{Y_{n-1}>\tilde{K}_u} \cdot \underbrace{\left|E_W\left[f(X_n) - f(Y_{n-1}) \;\middle|\; \mathcal{A}_{n-1}\right]\right|}_{\leq E_W[X_n - Y_{n-1} \;|\; \mathcal{A}_{n-1}] + Y_{n-1}\cdot P[X_n<K]} + I_{\tilde{K}_d \leq Y_{n-1} \leq \tilde{K}_u} \cdot \underbrace{E_W\left[\left|X_n - Y_{n-1}\right| \;\middle|\; \mathcal{A}_{n-1}\right]}_{\leq Y_{n-1}\sqrt{\sigma^2 + \left(r - \frac{\sigma^2}{2}\right)^2} 2e^{2|m|}e^{\sigma^2\Delta t}\sqrt{\frac{T}{n}}} \\
&+ I_{\tilde{K}_d > Y_{n-1}} \cdot \underbrace{\left|E_W\left[f(X_n) - \underbrace{f(Y_{n-1})}_{=0} \;\middle|\; \mathcal{A}_{n-1}\right]\right|}_{\leq E_W[I_{X_n>K}\cdot X_n \;|\; \mathcal{A}_{n-1}]} \;\Bigg|\; \mathcal{A}_0\Bigg] \\
\leq\;& \underbrace{E_B\left[Y_{n-1}\right]}_{=e^{r\frac{n-1}{n}\cdot T}}\Bigg(\underbrace{e^{r\frac{T}{n}} - 1}_{\leq r\frac{T}{n}e^{r\frac{T}{n}}} + \frac{e^{|m|}}{\sqrt{2\pi}} \cdot \frac{T}{n} + \underbrace{E_B\left[I_{\tilde{K}_d \leq K \leq \tilde{K}_u} \cdot Y_{n-1}\right]}_{\leq \tilde{K}_u \cdot P[\tilde{K}_d \leq Y_{n-1} \leq \tilde{K}_u]} \\
&\cdot\sqrt{\sigma^2 + \left(r - \frac{\sigma^2}{2}\right)^2} \cdot 2e^{2|m|}e^{\sigma^2\Delta t}\sqrt{\frac{T}{n}} + \frac{e^{|m|}}{\sqrt{2\pi}}\frac{T}{n}\Bigg) \\
\leq\;& \frac{\text{const.}}{n}
\end{aligned}$$

Since we can estimate as shown in the Proof of Lemma 5 :

$$P[\tilde{K}_d \leq Y_{n-1} \leq \tilde{K}_u] \leq \frac{12\sigma\sqrt{(\ln\frac{1}{\Delta t})\Delta t}}{\sqrt{npq}}$$

and $\sqrt{\ln\frac{1}{\Delta t}} \leq \sqrt{\Delta t}$ the Lemma is proven.

Lemma 9.

$$\left|E_B\left[E_B\left[f(Y_n) - f(Y_{n-1}) \;\middle|\; \mathcal{A}_{n-1}\right] \;\middle|\; \mathcal{A}_0\right]\right| \;= \mathcal{O}(\frac{1}{n})$$

Proof. $\left|E_B\left[E_B\left[f(Y_n) - f(Y_{n-1}) \;\middle|\; \mathcal{A}_{n-1}\right] \;\middle|\; \mathcal{A}_0\right]\right|$

$$\begin{aligned}
\leq\;& E_B\Bigg[I_{Y_{n-1}>K} \cdot I_{dY_{n-1}>K} \cdot \left|E_B\left[Y_n - Y_{n-1} \;\middle|\; \mathcal{A}_{n-1}\right]\right| \\
&+ I_{Y_{n-1}>K} \cdot I_{dY_{n-1}<K} \cdot \left|dY_{n-1} - Y_{n-1}\right| + I_{Y_{n-1}<K} \cdot I_{uY_{n-1}>K} \cdot \left|uY_{n-1} - Y_{n-1}\right| \;\Bigg|\; \mathcal{A}_0\Bigg] \\
\leq\;& \underbrace{E_B\left[Y_{n-1}\right]}_{=e^{r\frac{n-1}{n}T}}\left(e^{r\frac{T}{n}} - 1 + \frac{1}{\sqrt{npq}}\; E_B\left[\left|\frac{Y_n}{Y_{n-1}} - 1\right| \;\middle|\; \mathcal{A}_{n-1}\right]\right) \\
\leq\;& e^{r\frac{n-1}{n}\cdot T}\left(r\frac{T}{n}e^{r\frac{T}{n}} + \frac{1}{\sqrt{npq}}\; E_B\left[\left(\frac{Y_n}{Y_{n-1}} - 1\right)^2 \;\middle|\; \mathcal{A}_{n-1}\right]^{\frac{1}{2}}\right)
\end{aligned}$$

The second inequality follows since the probability of a single outcome can be estimated as $\frac{1}{\sqrt{npq}}$ (Feller [57]). The last inequality is Hölder's.

Since $E_B\left[(\frac{Y_n}{Y_{n-1}}-1)^2|\mathcal{A}_{n-1}\right]^{\frac{1}{2}} \leq m_n^2 + \underbrace{E_W\left[(\frac{X_n}{Y_{n-1}}-1)^2|\mathcal{A}_{n-1}\right]^{\frac{1}{2}}}_{=\mathcal{O}(\frac{1}{\sqrt{n}})\text{follws as statement in Lemma 11}}$ the Lemma is proven.

Lemma 10.

$$E_B\left[Y_k^3 \;\middle|\; \mathcal{A}_0\right] \leq const.$$

Proof.

$$\begin{aligned}
E_B\left[Y_k^3 \;\middle|\; \mathcal{A}_0\right] &= Y_0^3 \prod_{j=0}^{k-1} E_B\left[\left(\frac{Y_j}{Y_{j-1}}-1\right)^3 \;\middle|\; \mathcal{A}_{j-1}\right] \\
&\leq Y_0^3 \cdot E_B\left[\left(\frac{Y_1}{Y_0}-1\right)^3 \;\middle|\; \mathcal{A}_0\right]^n \\
&\leq Y_0^3(1+\frac{A}{n})^{3n} \\
&\leq Y_0^3 e^{3A}
\end{aligned}$$

since $\exists A>0: E_B\left[\left(\frac{Y_1}{Y_0}\right)^3 \;\middle|\; \mathcal{A}_{n-1}\right] \leq \bar{m}_n^3 + \underbrace{E_B\left[\left(\frac{X_1}{Y_0}-1\right)^3 \;\middle|\; \mathcal{A}_0\right]}_{=1+\mathcal{O}(\frac{1}{n})} \leq \frac{A}{n}$

Lemma 11.

$$\bar{p}_n = \mathcal{O}\left(\Delta t^2\right)$$

Proof. Let $m := (r-\frac{\sigma^2}{2})\Delta t$

Let Z be a lognormally distributed RV, $\ln z = N\left(m, \sigma^2\Delta t\right)$.

For $x \in \mathbb{R}\ :\ |x-1| \leq |\ln z|\max\left\{|x|, |\frac{1}{x}|\right\} \Rightarrow \quad |(\ln x)(x-1)^3| \leq (\ln x)^4 \max\left\{x^4, \frac{1}{x^4}\right\}$

Since $\ln(Ze^{-m}) \sim N(0,\sigma^2\Delta t)$ we have

$$\begin{aligned}
\bar{p}_n = E\left[\max\left\{Z^8, \frac{1}{Z^8}\right\}\right] &\leq e^{8|m|}E\left[\max\left\{(Ze^{-m})^8, (Ze^{-m})^{-8}\right\}\right] \\
&\leq e^{8|m|}\cdot 2\cdot E\left[(Ze^{-m})^8\right] \;=\; 2e^{8|m|}e^{28\sigma^2\Delta t}
\end{aligned}$$

where the second inequality follows because $\ln(Ze^{-m})$ has mean 0 and is therefore distributed symmetrically.

This implies:

$$\begin{aligned}
E\left[(\ln Z)(Z-1)^3\right] &\leq E\left[(\ln Z)^4\max\left\{Z^4, Z^{-4}\right\}\right] \\
&\leq E\left[\ln Z)^8\right]^{\frac{1}{2}} \cdot E\left[\max\left\{Z^8, Z^{-8}\right\}\right]^{\frac{1}{2}} \qquad \text{by Hölders Inequality} \\
&= \left((\sigma\sqrt{\Delta t})^8 \cdot 7\cdot 5\cdot 3\cdot 1\right)^{\frac{1}{2}} \cdot \left(2e^{8|m|}e^{28\sigma^2\Delta t}\right)^{\frac{1}{2}} \\
&= \mathcal{O}(\Delta t^2)
\end{aligned}$$

Appendix C:

Proposition. CRR converges with order 1.

Proof. We only have to verify:

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

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

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

For simplicity, set $a := e^{r\Delta t}$, $b := e^{\sigma^2 \Delta t}$

(1) Since $pu + qd = a$ one gets:

$$\begin{aligned} pu^2 + q(ud) &= ua \qquad \text{and} \qquad p(ud) + qd^2 \;=\; ad \\ \Rightarrow \quad pu^2 + qd^2 &= a(u+d) - (ud)(p+q) \;=\; a(u+d) - 1 \\ \Rightarrow \quad \bar{m}_n^2 &= \left|pu^2 + qd^2 - a^2 b\right| \;=\; a\left|u + d - a^{-1} - ab\right| \\ &= a \cdot \mathcal{O}\left(\frac{1}{n^2}\right) \\ &= \mathcal{O}\left(\frac{1}{n^2}\right) \end{aligned}$$

from the series expansion of the Exponential–Function.

(2) $\bar{m}_n^3 = \mathcal{O}\left(\frac{1}{n^2}\right)$ is proven in the same way.

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

$$u - 1 \;=\; \mathcal{O}\left(\frac{1}{\sqrt{n}}\right), \qquad d - 1 \;=\; \mathcal{O}\left(\frac{1}{\sqrt{n}}\right)$$

Since $\ln u \;=\; -\ln d \;=\; \sigma\sqrt{\Delta t}$ one gets $p_n \;=\; \mathcal{O}\left(\frac{1}{n^2}\right)$

Proposition. JR converges with order 1.

Proof. We only have to verify:

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

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

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

For simplicity, set $a := e^{r\Delta t}$, $b := e^{\sigma^2 \Delta t}$

(1) As in the privious theorem one gets:

$$\begin{aligned} pu^2 + qd^2 &= a(u+d) - (ud) \;=\; a(u+d) - e^{2\mu' \Delta t} \\ \Rightarrow \quad \bar{m}_n^2 &= \left|pu^2 + qd^2 - a^2 b\right| \\ &= a e^{\mu' \Delta t}\left|u e^{-\mu' \Delta t} + d e^{-\mu' \Delta t} - e^{\mu' \Delta t} e^{-r\Delta t} e^{-(2r+\sigma^2)\Delta t}\right| \\ &= a e^{\mu' \Delta t} \cdot \mathcal{O}\left(\frac{1}{n^2}\right) \\ &= \mathcal{O}\left(\frac{1}{n^2}\right) \end{aligned}$$

by a series expansion of the Exponential–Function.

(2) $\bar{m}_n^3 = \mathcal{O}\left(\frac{1}{n^2}\right)$ is proven in the same way.

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

$$u - 1 \;=\; \mathcal{O}\left(\sqrt{\Delta t}\right), \qquad d - 1 \;=\; \mathcal{O}\left(\sqrt{\Delta t}\right)$$

Since $\ln u \;=\; \mathcal{O}\left(\sqrt{\Delta t}\right)$, $\ln d \;=\; \mathcal{O}\left(\sqrt{\Delta t}\right)$ this proves: $p_n \;=\; \mathcal{O}\left(\Delta t^2\right)$.

Proposition. Tian converges with order 1.

Proof. Since Tian has explicitely fixed the first three moments: $\bar{m}_n^1 = \bar{m}_n^2 = \bar{m}_n^3 = 0$, we only need to verify $p_n = \mathcal{O}\left(\frac{1}{n^2}\right)$.
For simplicity, set $a := e^{r\Delta t}$, $b := e^{\sigma^2 \Delta t}$
For $0 < z \le 0.5828$ one has: $|\ln(1-z)| < \frac{3}{2}z$ (Abromowitz and Stegun [68]). That means for $0.4172 \le z < 1 : |\ln z| \le \frac{3}{2}|z-1|$. Moreover for $z > 0$ one has $\ln(1+z) < z$ (Abramowitz and Stegun [68])
$\Rightarrow\ \forall\ z > 1\ :\ \ln z < z - 1 \Rightarrow\ |\ln z| < \frac{3}{2}|z-1| \quad \forall\ 0.4172 < z$.
This is the case for $n >$. We have $0.8 < d < u < 1.2$ and therefore $(\ln u)^2 \le 2(u-1)^2$, $(\ln d)^2 \le 2(d-1)^2$.
Since

$$\begin{aligned}
p_n &= E_B\left[\left(\ln\frac{Y_{k+1}}{Y_k}\right)\left(\frac{Y_{k+1}}{Y_k} - 1\right)^3 \;\Big|\; \mathcal{A}_k\right] \\
&\le E_B\left[\left(\ln\frac{Y_{k+1}}{Y_k}\right)^2\right]^{\frac{1}{2}} \cdot E_B\left[\left(\frac{Y_{k+1}}{Y_k} - 1\right) \;\Big|\; \mathcal{A}_k\right]^{\frac{1}{2}} \text{ by Hölders Inequality}
\end{aligned}$$

it is sufficient to prove: $u - 1 = \mathcal{O}\left(\sqrt{\Delta t}\right)$, $d - 1 = \mathcal{O}\left(\sqrt{\Delta t}\right)$
Form a series expansion one gets:

$$V^2 + 2V - 3 \;=\; 1 + \mathcal{O}(\Delta t) + 2 + \mathcal{O}(\Delta t) - 3 \;=\; \mathcal{O}(\Delta t)$$

and

$$\begin{aligned}
u - 1 &= \frac{MV}{2}\left(V + 1 + \sqrt{V^2 + 2V - 3}\right) - 1 \\
&= \frac{MV}{2}(V+1)^{-1} + \mathcal{O}(\Delta t) \;=\; \frac{MV^2}{2} + \frac{MV}{2} - 1 + \mathcal{O}(\Delta t) \\
&= \frac{1 + \mathcal{O}(\Delta t)}{2} + \frac{1 + \mathcal{O}(\Delta t)}{2} - 1 + \mathcal{O}(\Delta t) \\
&= \mathcal{O}(\Delta t)
\end{aligned}$$

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

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.

Ibragimov [1966]: "On the Accuracy of Gaussian Approximation to the Distribution Functions of Sums of Independent Variables", *Teor. Prob. Appl.* 11, 1966, pp. 559–597.

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.

Reimer, M.; Sandmann, K. [1993]: "Down–And–Out Call", SFB 303, Discussion Paper No. B–239, University of Bonn.

Reimer, M.; Sandmann, K. [1994]: "European and American Barrier Options", SFB 303, Discussion Paper No. B–272, University of Bonn.

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 Applied Mathematics, Rheinische Friedrich-Wilhelms-Universität Bonn, Wegelerstrasse 6, D-53115 Bonn, Germany

E-mail address: leisen@indi0.iam.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