UNIVERSITY OF UDINE

Doctorate in Mathematics and Physics

OPTION PRICING AND PERFORMANCE EVALUATION

Supervisor:
Prof MARCELLINO GAUDENZI

Doctoral candidate:
ALICE SPANGARO

2017

Ever tried
Ever failed
No matter
Try again
Fail again
Fail better

Samuel Beckett

Contents

Abstract

This dissertation deals with two aspects of mathematical finance. The first is the pricing of options in a jump-diffusion setting, with lognormal jumps, according to the model proposed by Merton in 1976 [50]. The American option pricing procedure by Hilliard and Schwartz [38], that has been modified by Dai et al. [21] reducing the computational complexity from $O(n^3)$ to $O(n^{2.5})$, is here further improved to a computational complexity of $O(n^2 \log n)$, by trimming of the bivariate tree while keeping the error in check, and then to an $O(n^2)$ unidimensional procedure. These results are discussed in the joint works by Gaudenzi, Spangaro and Stucchi [33] and [34]. The other issue addressed in this dissertation is the performance evaluation of investments under different reward to risk ratios. Different portfolio performance measures have been compared applied to asset class indexes with a distribution far from normality: Omega, Sortino, Reward-to-VaR, STARR, Rachev ratio are correlated. Even though the values the various performance measures attribute to each investment differ, both in the cases analysed by Eling and Schuhmacher [28] and in that by Spangaro and Stucchi [73] many of them express good rank correlation with Sharpe ratio. The results draw heavily on the joint work Spangaro and Stucchi [73].

Acknowledgements

I would like to thank my supervisor professor Gaudenzi and professor Stucchi for their constant advice and precious support through all the years of my Ph. D. studies; this dissertation would never have seen the light of day without their encouragement and their help.

I would also like to thank my family, for being near from afar, and Pietro, for challenging me to be happy.

Introduction

The usefulness of the tree structure in the discretisation and also visualisation of random processes, and the simplicity of its concept, are probably the reasons for its popularity in Mathematical Finance. In 1979, Cox, Ross and Rubinstein [20] introduced this powerful tool in the option pricing scenario, giving a discrete model which allowed at the same time both to derive the Black and Scholes model [10] and to construct a procedure for the pricing of those options, such as the American ones, where the possibility of early exercises influences the evaluation.

Since then, many tree structures and lattices have been developed in order to accomodate different situations, increase the precision of the results and recover other models of the dynamic of a stock or interest rates.

The jump-diffusion model proposed by Merton in 1976 [50], where the not-deterministic behaviour of the logreturn of the underlying is due to the presence of both a Brownian motion and a Poissonian arrival of discontinuities with lognormal amplitude, captures the leptokurtic behaviour one can observe in the market data, which can hardly be explained with the Black and Scholes assumptions. Merton's model has been discretised by Amin [5] with the use of an infinite lattice of possible states for the underlying, while Hilliard and Schwartz [38] interpreted it as a bivariate tree where the binomial branching of the CRR tree is combined with a multiple branching which mimes the jumps. Dai et al. [21] modified this bivariate tree introducing a trinomial structure, in order to limit the growth in the number of the nodes, reducing the computational complexity from $O(n^3)$ to $O(n^{2.5})$. Gaudenzi, Spangaro and Stucchi [33] formulated a procedure that draws on the algorithm by Hilliard and Schwartz, and provides a method for the trimming of the bivariate tree while keeping the error in check, further reducing the computational complexity to $O(n^2 \log n)$. Moreover, we devised a solution for the unification of the effects of the jump and the diffusion branching, which allows for an even faster computation (due to the reduction of the nodes of the tree) while maintaining the accuracy (see [34]).

This aim of simplification and optimisation is also the main point in the performance evaluation issue.

In the development of finance as a science, the concept of a rational choice of investment has evolved, from the initial comparison between the returns of the different alternatives of investment (eventually taking into account the different times at which the profits would be obtained) to the estimation of the value of an investment that also depends on the risk one needs to undertake when making that specific investment instead of another. This means we are not comparing the return as it is anymore, but a risk adjusted return.

In literature, the risk has been incorporated in the evaluation of an investment (be it assets, hedge funds, mutual funds, etc.) in many different ways. The first contributions, in the mean-variance setting proposed by Markowitz [48], considered the standard error (see [59]) or the volatility (see [76]) as the selected measure of risk. The Sharpe ratio, introduced by Sharpe [63] both for portfolio selection (in its *ex ante* formulation) and for performance measure (in its *ex post* version), encountered rapid diffusion but also criticism for its normal-distribution setting and its use of a risk-free asset as a touchstone for the analysed portfolio, which brought to an update of the definition by the same author [66], but also to the exploration of different measures of risk that could better suit non-normal distributions. Amongst those, the VaR and the CVaR, chosen for their focus on the left side of the return distribution, gave rise to new performance ratios ([24], [4], [57], [9]), and the higher and lower partial moments (of the return with respect to a minimum acceptable return) allowed for the definition of a whole new family of performance ratios, e.g. the Sortino and the Omega ratio ([71], [62]). While the normality assumption made Sharpe ratio obsolete in contrast to the more refined measures, it has been observed by Eling and Schuhmacher [28] that many of these different measures appear positively correlated to each other and give approximately the same ranking in the case of hedge funds. Spangaro and Stucchi [73] investigate further this hypothesis in the case of 12 asset class indexes from Europe, US, Russia and China in the period 2003-2015. We come to the conclusion that the evaluation performed by more refined performance measures does not conflict with the traditional Sharpe ratio, which stands adequate and preferable in its simplicity.

The plan of the thesis is as follows.

Chapter 1 gives a brief overview of models for the dynamics of a stock and numerical methods adopted in option pricing, contextualises Johnson's system of transformations and its implementations and then focuses on the pricing of options via lattices when the underlying follows Merton's jump-diffusion model, starting with the Johnson binomial tree by Simonato [68]. Then various discretisations of Merton's process, the works by Amin [5], Hilliard and

Schwartz [38], Dai *et al.* [21] are described and compared, and the method developed by Gaudenzi, Spangaro and Stucchi ([33], [34]) upon the Hilliard and Schwartz's one is detailed.

Chapter 2 analyses the performance evaluation for asset class indices: we start with a description of a various range of ranking criteria, from the historical (much simpler) ones to the more recent and refined. This work draws on the paper by Spangaro and Stucchi [73].

The Appendix includes a collection of well-known formulas which are necessary to understand ours and other authors' work.

Chapter 1

Option pricing in a Merton jump-diffusion model

1.1 Options

An option is a financial derivative instrument, that is, its value depends on the changes in the price of some other underlying asset. It is a contract between two parts, the buyer of the option and the writer of the option, that guarantees the buyer a right, while constitutes an obligation for the writer: this is the reason why the buyer shall correspond to the writer a price, or premium, for the option. The most simple financial options are called vanilla options and can be divided into European and American, call and put option.

A European call option is a contract that allows the owner of the option to purchase a prescribed asset, which is referred to as the underlying, at a prescribed time, the maturity or expiry date of the option, at a prescribed price, the strike.

On the converse, a European put option is a contract that allows the owner of the option to sell the underlying, at a given maturity, at a prescribed strike.

American call and put options differ from their European counterparts in that the right of buying/selling the underlying can be exercised at any given time before maturity, and not only at the expiry date.

Options can be used as an hedging technique: their possess in combination with long or short positions in shares of the underlying can limit losses.

The fundamental characteristic of the option contract is that the owner can choose to exercise their right if convenient (if the strike price is preferable to the current price of the underlying) while simply letting the option expire if not.

More refined "exotic" kinds of options have been formulated, such as binary or compound options, and also path dependent options, whose price depends on the whole history of the underlying from the writing of the contract (with barriers, Asian, lookback), but the focus of this work is on European and American options.

The major problem is how to determine the premium of an option, knowing its style and characteristics (maturity, strike, underlying asset). Since the value of the underlying is random (in the sense that it can't be computed deterministically), so is the gain one can obtain from exercising the option. Therefore the price of the option must take care of the different possible evolutions in the price of the underlying. We know the current price of the underlying, but we cannot predict what the price of the underlying will be in the future. We will need to take into account its volatility, that is a measure of the standard deviation of the returns. The price of the option will also be a function of the time from the writing of the contract to maturity, since we can suppose that given a larger amount of time, a larger variation in the price of the underlying is possible.

Often the absence of transaction costs and taxes is taken as an hypothesis, and arbitrage-free or equilibrium assumptions are made.

In order to find a fair price for the option, we need to devise a stochastic model for the behaviour of the underlying.

1.2 Models for the dynamics of the underlying

In order to study financial phenomena, in which a large part of what happens appears to be random, many different hypothesis, some largely shared among the experts in the field, some not, have been made on the characteristics of the random components. Also, many suggestions have taken place, trying to define the dynamics that should appropriately model the objects we deal with in the market.

First of all, the process that informs the behaviour of the stock price is supposed to be a Markov chain, i.e. its future is only supposed to depend on the present value, regardless of what happened in the past (we would say, "with no memory").

Imposing other characteristics has given birth to different models, some preferable for their tractability, others for a better fit of specific situations.

1.2.1 Black and Scholes

The first attempt to model a stock price dynamics with a random walk has been made by a French mathematician, Louis Bachelier, in his doctoral thesis in 1900. Afterwards the model has been modified by Samuelson attributing that same characteristic to the logreturn.

The classical hypothesis is for the logreturns of the underlying to be normally distributed with constant mean and variance. The dynamic of the underlying asset price is then to be represented by the following stochastic differential equation:

$$dS = S\mu dt + \sigma S dB$$

where μdt gives the deterministic part of the price variation, with μ the (constant) continuously compounded rate, while the Brownian motion B is responsible for the random variation, and σ indicates the (constant) volatility. The stock price S is a stochastic process that is said to follow a Geometric Brownian motion.

Assuming that the underlying follows a GMB and pays no dividends, Black and Scholes [10] obtain via Itô's Lemma that the value V of an option on S must satisfy the following partial differential equation:

$$\frac{\partial V}{\partial t} + \frac{1}{2}\sigma^2 S^2 \frac{\partial^2 V}{\partial S^2} + rS\frac{\partial V}{\partial S} - rV = 0 \tag{1.1}$$

for $0 \leq t \leq \tau$ with τ the maturity, r the riskless rate, σ the standard deviation, and $S(t) \geq 0$ is the price of the underlying at time t. Given the boundary condition that fixes the payoff of the option, that in the call case is $V(\tau) = (S(\tau) - K_0)^+$, where K_0 is the strike, they identify an explicit formula for the European call option price on a stock:

$$C(t, S, \tau, K_0, r, \sigma^2) = S(t)N(d_1) - K_0 e^{-r(\tau - t)} N(d_2) \tag{1.2}$$

where d_1, d_2 are computed according to the following equations

$$d_1 = \frac{\log(S(t)/K_0) + (r + \sigma^2/2)(\tau - t)}{\sigma\sqrt{\tau - t}}$$

$$d_2 = \frac{\log(S(t)/K_0) + (r - \sigma^2/2)(\tau - t)}{\sigma\sqrt{\tau - t}}.$$

The price of the European put option on the same underlying, with the same maturity and strike price, is then obtained via the put-call parity, which states that for the values of the call option $C(t)$ and the put option $P(t)$ written on the same underlying $S(t)$ with the same maturity τ and strike K one has:

$$C(t) - P(t) = S(t) - K_0 e^{-r(\tau - t)}$$

for any $0 \leq t \leq \tau$.

The simplicity of the model and the comfort of a solid theoretical result explain why it is still largely used both among scholars and practitioners (cf. [82] and [85]).

Appreciate that the riskless rate r, which has also been taken as a constant, appears in Equations (1.1)-(1.2), whereas the expected return μ does not impact on the correct price for the option. This is a consequence of the definition of a risk-neutral measure, that grants the independence of the price from the risk preferences of the investors.

Black and Scholes model can be generalised to the cases of non constant but deterministic interest rate $r(t)$ and with a dividend yield $d(t)$.

Nevertheless, there are some drawbacks to assuming the hypotheses from the Black and Scholes model; namely, the following three characteristics observed in the market cannot be theoretically justified:

- the probability distribution of the logreturn of the stock, according to the market data, shows asymmetric leptokurtic features and higher peaks and thicker tails than one would expect from a normal distribution;
- the volatility doesn't appear to be constant, but to oscillate, creating the so called volatility smile;
- new informations can have a strong impact on the market, causing ample variations in the prices.

These real world difficulties have spurred the creation and development of many different models, for example the introduction of stochastic volatility models (see Heston [35]), GARCH models, or the Constant Elasticity of Variance model (see [19]), which are not in the scope of this dissertation. Hereafter, we will focus on the jump-diffusion models, and we will work on the case where the jump dynamics is regulated by a compound Poisson of i.i.d. lognormal variables. Jump-diffusion models are a special case of Lévy processes (for which we refer to Cont, Tankov [16]) and can also be expanded to include stochastic volatility (see Duffie *et al.* [25]) and inserted in regime-switching models (see Costabile *et al.* [18]), that are also not in the scope of this work.

1.2.2 Merton's jump-diffusion model

When the observations of the underlying asset display oscillations, spikes or heavy-tailed distributions, the lognormal diffusion process does not provide an adequate model for the data.

The market data suggest that the variations in the value of stocks shouldn't be considered continuous, since external events, such as new arrivals of information, can cause significant variations in small amounts of time, which we will call "jumps".

Cox and Ross [19] proposed to model the dynamics of the underlying as a series of discontinuous jumps, which could be expressed via a simple or compound Poisson process.

Merging the jump process approach with the traditional diffusion one, in order to describe a dynamic that could incorporate the possibility of discontinuities, but also the continuous small variations, Merton [50] provided a jump-diffusion model, where the dynamics of the price of the underlying are not only subjected to variations due to a Brownian process, but also to possible, if rare, greater variations.

To model the random arrival of rare events, a random variable distributed as a Poisson, independent from the Brownian motion, is used. Each event causes a jump in the value of the price. The random variables that model the amplitude of the single jumps are supposed to be each one independent from the other and identically distributed. Various distributions for the jump have been studied in literature, for reasons of simplicity and relevance. In the following, the major results obtained by Merton [50], Kou [42] with different specifications of the distribution are highlighted; though the distribution we will be especially interested in is the lognormal one, as in [50], [5], [38].

The dynamics of the underlying according to the Merton model, under the assumptions stated above, is given by the equation:

$$\frac{dS}{S} = (r - v - \lambda \bar{j})dt + \sigma dz + J dq \tag{1.3}$$

where we used the same notation as in the previous section for the diffusion part of the dynamics: r is the risk-free interest rate, v is the continuous dividend yield, σ^2 is the Brownian portion of the variance (i.e. the variance of the return under the condition that no jumps occur). The additional parameters characterise the jump part of the dynamics: λ is the intensity of the Poisson process that models the arrival of jumps (i.e. λ equals the average number of arrivals in a time unit), dq assumes values 1 or 0 according to the presence or absence of a jump, J is the random amplitude of the jump, whose expected value is equal to $\bar{j} = E(J)$.

Applying Itô's Lemma in a formulation that also accounts for jumps (see [49]), we can express the solution of Equation (1.3) as:

$$S = S_0 e^{(r - v - \lambda \bar{j} - \frac{\sigma^2}{2})t + \sigma z(t)} \prod_{i=0}^{m(t)} (1 + J_i) \tag{1.4}$$

where $m(t)$ is a Poisson process of parameter λ, $J_0 = 0$ and the J_i's, for $i \geq 1$, are independent identically distributed random variables.

In the following, for clarity purposes we will consider the logarithmic return as divided in two components X_t and Y_t, where

$$X_t = \alpha t + \sigma z(t),$$

with $\alpha = r - v - \lambda \bar{j} - \frac{\sigma^2}{2}$ the so-called "drift", is the diffusion component, while

$$Y_t = \sum_{i=0}^{m(t)} \log(1 + J_i) \tag{1.5}$$

is the jump component. We will focus on understanding the behaviour of Y_t.

Y_t is a compound Poisson process, whose probability density function can be recovered via the characteristic functions:

$$\varphi_{Y_t}(x) = e^{\lambda t(\varphi_{\log(1+J)}(x) - 1)}. \tag{1.6}$$

The presence of the jumps causes the market to be incomplete, thus there is not a unique choice for the risk-neutral probability measure for the pricing of the options. One cannot apply directly the no arbitrage argument that allows for the pricing of the options in the Black and Scholes case.

Nevertheless, under the Black and Scholes [10] assumptions on the Brownian component and considering a compound Poisson process for the jump part, Merton reasons that the jump will be uncorrelated with the market (representing non-systematic risk) and obtains a series solution for the price of a European call option with time to maturity τ and strike K_0:

$$V = \sum_{j=0}^{+\infty} \frac{e^{-\lambda\tau}(\lambda\tau)^j}{j!} E_{Y_{(j)}}(C(0, SY_{(j)}e^{-\lambda\bar{j}\tau}, \tau, K_0, r, \sigma^2)) \tag{1.7}$$

with $C(0, SY_{(j)}e^{-\lambda\bar{j}\tau}, \tau, K_0, r, \sigma)$ as defined in Equation (1.2), $Y_{(0)} \equiv 1$, $Y_{(j)}$ is the sum of j i.i.d. random variables $Y_{(j)} = \sum_{i=0}^{j} \log(1 + J_i)$ and $E_{Y_{(j)}}$ denotes the expectation over the distribution of $Y_{(j)}$.

For example, if the random variables $\log(1 + J_i)$ represent immediate ruin, that is $\log(1 + J_i) \equiv 0$, the Equation (1.7) reduces to a Black and Scholes solution with a larger interest rate:

$$V = C(0, S, \tau, K, r + \lambda, \nu_j^2).$$

Lognormal jumps

From here on, except when explicitly stated otherwise, we will consider the random variable J of Equation (1.3) to follow a lognormal distribution such that $\log(J+1) \sim N(\gamma', \delta^2)$.

Under this hypothesis, in the solution

$$S = S_0 e^{(r-v-\lambda\bar{j}-\frac{\sigma^2}{2})t+\sigma z(t)} \prod_{i=0}^{m(t)} (1+J_i) \tag{1.8}$$

of Equation 1.3 we have $J_0 = 0$ and $\log(1+J_i) \sim N(\gamma', \delta^2)$ for $i \geq 1$, therefore the compound Poisson process Y_t of Equation 1.5 has the following characteristic function:

$$\varphi_{Y_t}(x) = e^{\lambda t\left(e^{ix\gamma' - \frac{\delta^2 x^2}{2}} - 1\right)} \tag{1.9}$$

and the following function of probability density:

$$p(x) = \sum_{k=1}^{+\infty} e^{-\lambda t} \frac{(\lambda t)^k}{k!} \frac{1}{\sqrt{2\pi k\delta^2}} e^{-\frac{(x-k\gamma')^2}{2k\delta^2}} + \delta_0(x) e^{-\lambda t} \tag{1.10}$$

where $\delta_0(x)$ is the Dirac delta function.

If the random variables $\log(1+J_i)$ are normally distributed, the Equation (1.7) articulates into:

$$V = \sum_{j=0}^{+\infty} \frac{e^{-\lambda'\tau}(\lambda'\tau)^j}{j!} C(0, S, \tau, K_0, r_j, \nu_j^2)$$

with $\lambda' = \lambda(1+\bar{j})$, $\nu_j^2 = \sigma^2 + \frac{j\delta^2}{\tau}$ and $r_j = r - \lambda\bar{j} + \frac{j\gamma}{\tau}$, where $\gamma = \gamma' + \frac{\delta^2}{2}$.

Since determining the price of several types of path-dependent options requires the study of the first passage of an horizontal level, the adoption of a jump-diffusion process for the dynamics of the underlying gives rise to the problem of dealing with the correlation between this first passage and the possible overshoot of the barrier, therefore an analytic solution is not available.

There have been a lot of attempts to obtain a discretisation of the Merton model with tree or lattice methods in order to price American or more complex options, which we will see in the dedicated Section (cfr. Section 1.5).

Double exponential jumps

As an example of another popular jump-diffusion model, we present the double exponential model proposed by Kou [42]. Here the stock price is supposed to satisfy the Equation (1.3) but the jump amplitudes are no more regulated by lognormal variables.

The value of the stock price S is given by

$$S = S_0 e^{(r-v-\lambda\bar{j}-\frac{\sigma^2}{2})t+\sigma z(t)} \prod_{i=0}^{m(t)} (1 + J_i)$$

where the processes $(1+J_i)$ are i.i.d. such that $W_i = \log(1+J_i)$ have an asymmetric double exponential distribution, represented by the following probability density function:

$$f_W(w) =:= \begin{cases} p\eta_1 e^{-\eta_1 w} & \text{for } w \geq 0, \\ (1-p)\eta_2 e^{\eta_2 w} & \text{for } w < 0. \end{cases}$$

with $0 \leq p \leq 1$ and $\eta_1 > 1$ (to ensure that the underlying price has finite expectation) and $\eta_2 > 0$.

Within this model, Kou and Wang [43] recover an analytic approximation of the finite-horizon American option, for the put and the call case, and analytical solutions for lookback, barrier and perpetual American options.

1.3 Numerical methods for option pricing

As we have seen in the previous Section, Black and Scholes [10] develop their momentous option pricing formula when the underlying asset is a stock and its price follows a lognormal diffusion process. Nevertheless, frequently, financial derivatives cannot be priced by closed-form formulas.

When a closed-form formula for the option price is not available or computationally costly, the alternative is to resort to numerical methods that allow for an approximation of the solution.

In this section we offer a brief overview of the numerical methods that are usually applied for the evaluation of option prices.

We will be interested specifically in the last of the family of methods presented, the lattice methods. Other numerical methods that do not belong to these families but have been specifically introduced in the evaluation of options with an underlying following Merton's jump-diffusion process exist, but are not in the scope of this work; for a brief panoramic see Chiarella and Ziogas [15].

The concerns in using a numerical method are the following:

- we need to discretise a continuous model in order to construct an appropriate algorithm;

- we would like the algorithm to provide an approximate evaluation that is arbitrarily close to the correct one, that is, we want some convergence results; usually we divide the time to expiry in n steps, and we would like for the algorithm to converge to the exact solution for $n \to +\infty$.

- we are interested in efficient algorithms, that provide an acceptable result in as short a time as possible, and using the smallest amount of computer memory as possible. The speed of the algorithm depends on the number of operations needed to perform the algorithm, and is usually expressed as a $O(f(n))$ where n is the number of steps.

1.3.1 Monte Carlo simulations

Monte Carlo methods are based on the numerical generation of a large quantity of realisations of random walks that satisfy the equation of the considered model. The series of stock value trajectories can then be used to estimate the option value (with the risk-neutral assumption that allows to calculate the price as the discounted expected value).

Monte Carlo simulation has been first applied to option pricing by Boyle [12], who also suggested its use in the evaluation of options where the underlying follows a mixture of stochastic processes. In particular, Boyle provided a description of how to treat specifically Merton's one, generating realisations for the three different random processes involved in Merton's model: the Brownian motion, the Poisson process regulating the number of the jumps occurring, and the amplitude of the jumps (that need not be restricted to the lognormal case thanks to the flexibility of Monte Carlo simulations).

Monte Carlo simulation has the advantage of being a general and convergent method, but is also computationally expensive and slow (the error is of the order of the inverse square root of the number of simulations) and there is the concern of the implementation of a non-biased random number generator. There are also difficulties in incorporating in the algorithms the early exercise feature that is typical of American options: Longstaff and Schwarz proposed a solution to this problem with the Least Squares Monte Carlo method [46].

1.3.2 Finite difference methods

Finite difference methods have been introduced in the pricing of options by Brennan and Schwartz in 1978.

These methods involve starting from a partial differential equation that the function we are interested in must satisfy and substituting the partial derivatives with finite differences. In other words, where we would have $\frac{\partial f}{\partial x}$, that is defined as a limiting difference

$$\frac{\partial f}{\partial x}(x,t) = \lim_{h\to 0} \frac{f(x+h,t) - f(x,t)}{h},$$

we would consider a difference quotient with a small but not infinitesimal $h > 0$. Several choices are available, there is the forward difference

$$\frac{f(x+h,t) - f(x,t)}{h},$$

the backward difference

$$\frac{f(x,t) - f(x-h,t)}{h},$$

and the central difference

$$\frac{f(x+h/2,t) - f(x-h/2,t)}{h}.$$

Central differences in particular are used in the Crank-Nicholson method. Since there are three possible ways of defining the finite difference for the first derivative, and the second derivative itself can be defined in the same three possible ways with respect to the first derivative, there is a plethora of possibilities for the definition of the finite difference representing the second derivative $\frac{\partial^2 f}{\partial x^2}$; usually the choice falls on the formulas that provide simmetry, as:

$$\frac{f(x+h,t) - 2f(x,t) + f(x-h,t)}{h^2}.$$

Accuracy of the approximation provided by the finite difference can be obtained via Taylor's theorem (for a sufficiently differentiable function f).

The (x,t) space is then divided into a grid, not necessarily equally spaced, and we look for the solutions of the finite differences equations relative to the points of the grid. When the technique allows for a direct computation of the solution, we are dealing with explicit finite differences schemes, when instead the solution of a system of algebraic equations is required we are implementing an implicit finite difference scheme.

Finite difference schemes need to be developed according to the problem at hand. For example, in order to simplify the application of the method to the Black and Scholes model, one can employ an appropriate change of variable in order to reduce the partial differential equation relative to the option value into the diffusion equation:

$$\frac{\partial v}{\partial t} = \frac{\partial^2 v}{\partial x^2}.$$

Finite difference methods have been applied to the pricing of options with jump-diffusion underlying for example by Zhang [84] and Cont and Voltchkova [17].

1.3.3 Lattice methods

The umbrella name of lattice methods includes all the methods where lattices and recombining trees are implemented in order to price a derivative. According to the branching of each node, these methods are categorised into binomial, trinomial or multinomial trees. They spring from the idea of discrete time stochastic models, that can be written as a list of difference equations.

These methods are efficient for simple call and puts, less efficient when we are dealing with more complicated options. Results of convergence are not obvious and need to be investigated in the various models and discretisations.

Binomial trees have been proposed for the first time by Cox, Ross and Rubinstein [20].

Their framework is based on an interpretation of the stock price as a discrete multiplicative binomial process that in a single time step, from t to $t+\Delta t$, can only acquire two possible values. Starting from the value S at the beginning of the period, at the end the stock price will either assume value uS with probability p or value dS with probability $1-p$.

Their main concern is the pricing of a call option on such an underlying S with strike K_0. They adopt the same assumptions of the Black and Scholes model: a constant riskless interest rate of return r, the absence of transaction costs or taxes; they allow for the investor to sell short any security. Based on the no-arbitrage principle, they require $u > e^{r\Delta t} > d$.

Supposing that the option expires after only one period, they obtain a pricing for the option by considering the composition of a replicating portfolio containing only Δ shares of the underlying stock S and a certain amount B of bonds: the quantity of bonds and shares are obtained by imposing that the value of the portfolio after a period coincides with the value of the option in any case, whether S has an up-tick and assumes value uS (and then the option will assume value $C_u = (uS-K)^+$, since the holder of the option would choose to exercise it only when it is profitable, that is when $uS > K$) or has a down-tick and assumes value dS (and then the option will have value $C_d = (dS-K)^+$) .

Solving the resulting system of two equations in two unknowns, one obtains that $\Delta = \frac{C_u - C_d}{(u-d)S} \geq 0$ and $B = \frac{uC_d - dC_u}{(u-d)e^{r\Delta t}} \leq 0$.

The value of the call option, provided that it can only be exercised at maturity (European option), needs to be equal to the value of this hedging portfolio, that is

$$\Delta S + B = \frac{C_u - C_d}{(u-d)} + \frac{uC_d - dC_u}{(u-d)e^{r\Delta t}} = e^{-r\Delta t}\left(\frac{e^{r\Delta t} - d}{u-d}C_u + \frac{u - e^{r\Delta t}}{u-d}C_d\right).$$

By defining $\pi = \frac{e^{r\Delta t} - d}{u-d}$ the equation above can be written as $C = e^{-r\Delta t}(\pi C_u +$

$(1-\pi)C_d)$ and therefore the price of the European call option appears as the discounted expected value of the option at maturity under a measure given by π, which can be interpreted as a probability, for it is $0 \leq \pi \leq 1$. This measure does not depend on the probability p we assigned to the up-tick of the underlying, therefore it does not depend on the suppositions the investors may have on such a probability, neither does it depend on their risk-aversion, instead it is the value p would have in a risk-neutral world (where the expected rate of return of the stock is equal to the riskless interest rate).

The authors extend recursively the same reasoning to a European call option with expiry date after an arbitrary number n of time intervals, supposing the up and down factors and the branching probabilities are constant in time and do not depend on the value of the stock.

Since $udS = duS$, if the stock has an up tick and a down tick in the following period it reaches the same value that it would with a down tick followed by an up tick. This means that the possible developments in the value of the option may be represented with a recombining binomial tree, where every parent node has two children, that will have a children in common. After n steps, the terminal nodes of the tree are not 2^n, as we would expect from a structure where the possibilities are doubled at every step, but $n+1$, due to the recombination effect.

In such a setting, the European call option price turns out to be the discounted expected value in a risk-neutral world of the payoffs at maturity $\tau = n\Delta t$, that is

$$C = e^{-r\tau} \sum_{j=0}^{n} \binom{n}{j} \pi^j (1-\pi)^{n-j} (Su^j d^{n-j} - K_0)^+,$$

since the possible values of the stock at maturity S_n are given by $Su^j d^{n-j}$.

One can show that for $r > 0$ American call options (where exercise of the option is allowed at any time before expiry) do not differ in price from the European ones, since the early exercise strategy is never to be preferred.

If we are interested instead in the price of an American put, then we are not going to use only the payoffs at maturity, but we will start from the values at maturity and proceed backwards, at every node assigning to the put option the maximum value between the payoff at that node and the discounted expected value of the children. This comparison between the value of holding the option and that of immediate exercise is the way of taking into account the possibility of early exercise embedded in American options.

If instead of adding up periods, we consider a fixed time to maturity τ and subdivide it in a number n of time intervals $\Delta t = \frac{\tau}{n}$, Cox, Ross and Rubinstein show that for $n \to +\infty$ the price of the call option obtained via the binomial method for an appropriate choice of the values u, d, p, approximates the price given by the Black and Scholes formula.

Such values are found by asking the logreturn of the discrete variable represented by the binomial process to have an expected value and a variance that approach those of the logreturn of the stock of the Black and Scholes model. The expected value and variance of the discrete logreturn are as follows:

$$E\left(\log\frac{S_n}{S}\right) = E(j\log u + (n-j)\log d) = \log\frac{u}{d}E(j) + n\log d = n\left(\log\frac{u}{d}p + \log d\right)$$
$$Var\left(\log\frac{S_n}{S}\right) = E\left(\left(j\log\frac{u}{d} - \log\frac{u}{d}E(j)\right)^2\right) = \log\frac{u}{d}Var(j) = \log\frac{u}{d}np(1-p)$$

since $\log\frac{S_n}{S} = j\log u + (n-j)\log d$, $E(j) = np$ and $Var(j) = np(1-p)$.

The desired conditions are

$$n\left(\log\frac{u}{d}p + \log d\right) \to \mu\tau$$
$$n\log\frac{u}{d}p(1-p) \to \sigma^2\tau$$

for $n \to +\infty$. A set of values for u, d and p that attains these conditions is

$$u = e^{\sigma\sqrt{\Delta t}} \qquad d = e^{-\sigma\sqrt{\Delta t}} \qquad p = \frac{1}{2} + \frac{1}{2}\frac{\mu}{\sigma}\sqrt{\Delta t}$$

Note that $ud = 1$.

Only approaching the mean and variance would not guarantee that the probability distribution of the discrete process approaches that of the continuously compounded return, but Cox, Ross and Rubinstein show that a version of the central limit theorem is applicable in this case, making higher-order properties negligible as $n \to +\infty$.

Alternative choices for u, d and p have been proposed. As an example, we recall the Jarrow Rudd risk-neutral model, which imposes:

$$u = e^{\left(r-\frac{\sigma^2}{2}\right)\Delta t + \sigma\sqrt{\Delta t}} \qquad d = e^{\left(r-\frac{\sigma^2}{2}\right)\Delta t - \sigma\sqrt{\Delta t}} \qquad p = \frac{e^{r\Delta t} - d}{u - d}.$$

This work focuses on the application of lattice methods to the pricing of options in Merton's jump-diffusion model with jumps of lognormal amplitude, where the simple CRR binomial tree is no longer trustworthy. In the next Section, after an introduction on parametric models necessary to understand it, we will describe Simonato's [68] Johnson trees. Then, in the following

Section, we will describe the "rectangular" lattice provided by Amin [5], the multinomial lattice by Hilliard and Schwartz [38], the trinomial structure by Dai *et al.* [21] and our methods.

1.4 Johnson's system and option evaluation

In this Section we will describe the major parametric models that have been introduced and became popular in mathematical finance for the modelling of the underlying distributions, and then we will focus on the characteristics, usefulness and application of the Johnson's family of distributions in the pricing of options, as studied by Simonato [68].

1.4.1 Parametric models

We start by recalling that parametric modelling pertains to statistical inference: we are interested in identifying from the data what sort of probability distribution may have generated them. One of the possible ways to do this is to refer to a parametric model, which is a family of distributions that can be described using a finite number of parameters. Common examples of such a family are, in the discrete, the Bernoulli distribution (where all elements in the family are described by the probability p of a positive result) and the Binomial distribution (two parameters: the probability p and the number n of tosses); in the continuous the Poisson distribution (where in order to specify a distribution we only need to know the intensity of the arrival of events, λ) and the normal distribution (where every element of the family can be fully described by two parameters, the mean μ of the distribution and its variance σ^2). The parameters are chosen to belong to a given parameter space.

The importance of the parametric models is that, provided the model can satisfactorily fit the data - that is, provided we are correct in the assumption on the family the distribution should belong to - we can draw on the theory of the corresponding model and the predictions we obtain are extremely accurate and precise. Historically, the well-known theory about normal distribution has made it one of the preferred parametric models, probably also due to its usage in the social sciences.

Antithetical to parametric models are non parametric models, which are collections of distributions that cannot be described by a finite set of parameters: this means they are more flexible but also less accurate in their predictions.

1.4.2 Pearson's family of curves

The Pearson's system is a family of continuous distribution curves. They belong to the group of the skew frequency curves, created in order to treat cases where the sampling distribution of the data seems too skewed to be satisfactorily interpreted with a normal distribution. The Pearson system, as the Charlier and Johnson systems we will see later on, can be used in order to approximate theoretical distributions whose first four (and in some special case, even fewer) moments are known, for example in order to create simulations by generating random values from the fitted distribution.

We will give a brief scan of the characteristics of the family, based on Stuard and Ord [74].

Let a, b_0, b_1, b_2 be real numbers. The frequency curves of the Pearson system are defined via the following differential equation regarding the probability density function f:

$$\frac{df(x)}{dx} = \frac{(x-a)f(x)}{b_0 + b_1 x + b_2 x^2}. \tag{1.11}$$

This request may seem less obscure if we think that the discrete version of this condition,

$$\frac{f(x_{j+1}) - f(x_j)}{f(x_j)} \frac{df(x)}{dx} = \frac{x_j - a}{b_0 + b_1 x_j + b_2 x_j^2},$$

is satisfied by binomial, Poisson and hyperbolic distributions (see [74] vol. 1 pp. 171).

All solutions of Equation (1.11) have the following properties:

- their derivative vanishes at $x = a$, therefore a is a stationary point (the mode), and at $f(x) = 0$, which means that the pdf has smooth contact with the x-axis
- $\beta_2 > \beta_1 + 1$
- every moment can be recursively computed from the lower order ones according to the following formula:

  $$\mu'_{n+1} = \frac{(a - b_1(n+1))\mu'_n - nb_0\mu'_{n-1}}{(n+2)b_2 + 1}$$

  which implies that the four parameters a, b_0, b_1 and b_2 can be determined from the first four (central) moments:

  $$b_1 = a = -\frac{\sqrt{\mu_2}\gamma_1(\beta_2 + 3)}{A'} \tag{1.12}$$

$$b_0 = -\frac{\mu_2(4\beta_2 - 3\beta_1)}{A'} \tag{1.13}$$

$$b_2 = -\frac{(2\beta_2 - 3\beta_1 - 6)}{A'} \tag{1.14}$$

where $A' = 10\beta_2 - 12\beta_1 - 18$, $\beta_1 = \gamma_1^2$, where γ_1 is the third standardized moment, or Fisher-Pearson coefficient of skewness, and β_2, the fourth standardized moment, also introduced by Pearson as a measure of kurtosis. From these Equations one can see that $b_1 \neq 0$, while $b_0 = 0$ if $\beta_1 = 0$ (for example when we deal with a symmetric distribution)

- differentiating Equation (1.11) we obtain that possible points of inflection must satisfy $(x-a)^2 = \dfrac{b_0 + a^2(1+b_2)}{(b_2-1)}$, which implies that in any case there are no more than two and, if both solutions to the previous equation fall in to the range of variation of the distribution, they must be at equal distance from the mode a.

The distributions belonging to this family assume many different aspects: unimodal (either with a fall on either side of the mode of the probability density function, or J-shaped, that is with a monotonic density function) or U-shaped (with maxima at both ends of the distribution and a minimum in between).

Rewriting Equation (1.11) with the change of variable $X = x - a$ we obtain

$$\frac{df(x)}{dx} = \frac{Xf(x)}{B_0 + B_1X + B_2X^2} \tag{1.15}$$

with

$$B_0 = b_0 + a^2(1+b_2) \tag{1.16}$$

$$B_1 = a(1+2b_2) \tag{1.17}$$

$$B_2 = b_2. \tag{1.18}$$

The polynomial $B_0 + B_1X + B_2X^2$ allows for a classification of the types the Pearson curves are divided into. Its solutions may be real of the same sign, real of opposite sign, complex, which can be determined from the value of

$$K = \frac{B_1^2}{4B_0B_2}. \tag{1.19}$$

Some of the Types individuated by Pearson have fallen out of use, here we will see the most implemented ones.

We call Type I, or Beta distributions of the first kind, the variables that satisfy (1.15) where the polynomial in the denominator has real roots of opposite sign, that is $K = \frac{B_1^2}{4B_0B_2} \leq 0$.

Their range is limited, and their name is due to the fact that their distribution function is the Incomplete Beta function, while the probability density function is given by:

$$f(x) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} x^{p-1}(1-x)^{q-1} \tag{1.20}$$

for $0 \leq x \leq 1$, $p, q > 0$.

All moments of Type I distributions exist, and can be expressed by:

$$\mu'_n = \frac{\Gamma(p+q)\Gamma(p+n)}{\Gamma(p)\Gamma(p+q+n)}.$$

Type II is the special subcase of Type I with $K = 0$, that is $B_1 = 0$, or $\beta_1 = 0$.

Type III, which we obtain for $B_2 = 0$, is the so called Gamma distribution, since its distribution function is the Incomplete Gamma function.

Its pdf is given by

$$f(x) = \frac{1}{\Gamma(p)} x^{p-1} e^{-x}$$

for $x \geq 0$ and $p > 0$. If $0 < p \leq 1$, the distribution is J-shaped.

We call Type IV the variables that satisfy (1.15) where the polynomial in the denominator has complex roots, that is $0 \leq K < 1$.

All distributions in this group are unimodal and of unlimited range, and their density function is given by:

$$f(x) = k\left(1 + \frac{x^2}{a^2}\right)^{-m} e^{-\nu \arctan \frac{x}{a}}$$

with $m > \frac{1}{2}$. The existence of the first four moments is guaranteed by $m > \frac{5}{2}$.

Type VII is the special subcase of Type IV with $K = 0$, and coincides with (non-standardised) Student's t distribution.

Another special case ($B_1 = B_2 = 0$) is the normal distribution.

Type V random variables have probability density function:

$$f(x) = \frac{1}{\Gamma(p)} \frac{e^{-\frac{1}{x}}}{x^{p-1}}$$

for $x > 0$.

We call Type VI, or Beta distributions of the second kind, the variables that satisfy (1.15) where the polynomial in the denominator has distinct real roots of the same sign, that is $K = \frac{B_1^2}{4B_0B_2} > 0$.

Its name is due to the fact that the probability density function

$$f(x) = \frac{\Gamma(p+q)}{\Gamma(p)\Gamma(q)} \frac{x^{p-1}}{(1+x)^{p+q}}$$

for $x \geq 0$, $p, q > 0$ can be related to (1.20) by a change of variable. Its moment of order n is only available for $n < q$, when it can be obtained by:

$$\mu'_n = \frac{\Gamma(p+q)\Gamma(q-n)}{\Gamma(p)\Gamma(q)}.$$

The fitting of a distribution can be then implemented as such: starting from the data, estimates of the first four moments and indexes of skewness and kurtosis are computed. Then, Equations (1.16)-(1.19) are used to classify the distribution in the appropriate Type. The parameters of the specific Type are then recovered either by matching of the moments or by maximum likelihood estimation.

1.4.3 Gram-Charlier's system of frequency curves

Also Gram-Charlier's system of curves has different subspecies of distributions, but the most used is Type A, which has been defined as follows.

Let $\alpha(x)$ be the standardized normal distribution function $\alpha(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, and let $H_j(x)$ for $j \in \mathbb{N}$ be the Chebyshev-Hermite polynomials, that is: $\alpha(x)H_j(x) = -\frac{\partial^j}{\partial x^j}\alpha(x)$.

Gram-Charlier Type A is the family of distributions whose pdf can be written as a series of derivatives of $\alpha(x)$, that is

$$f(x) = \sum_{j=0}^{+\infty} c_j H_j(x)\alpha(x) \tag{1.21}$$

The evaluation of the constants c_j by means of a procedure similar to that used with Fourier series provides the following expression for the first terms of the series:

$$f(x) = \alpha(x)\left(1 + \frac{1}{2}(\mu_2 - 1)H_2(x) + \frac{1}{6}\mu_3 H_3(x) + \frac{1}{24}(\mu_4 - 6\mu_2 + 3)H_4(x) + \dots\right)$$

The problems with such a representation are the inadequacy in obtaining a useful approximation of a pdf considering only a finite number of terms in the series. In particular, there is no guarantee that the series, stopped after a finite number of terms, gives a non-negative function. Moreover, increasing the number of terms does not necessarily give a better approximation. The expression we have reported above is proved to provide a unimodal density function only for values of skewness and kurtosis close enough to 0 [74]. This also means that when our interest is in a better approximation of the behaviour on the tails, other families of functions are preferable.

Gram-Charlier's Type B is defined as the family of distributions whose pdf can be expressed as a series of derivatives of the Poisson distribution (with respect to the λ parameter) $\alpha(\lambda) = e^{-\lambda}\frac{\lambda^x}{x!}$.

1.4.4 Johnson's family of probability distributions

As we have already highlighted, the usefulness of the normal distribution, the importance of theoretical results like the central limit theorems, and the availability of the tabulations for functions associated to it, made it an interesting tool in statistical modelling. Nevertheless, in many situations the data suggest a distribution far from normality, be it "skewed" or with heavier or thinner tails. In those cases, one of the viable options is to study how big a deviation there is from the normality results. Another is to deal with those distributions referring them back to the Gaussian bell by applying some sort of transformation.

In 1898, Edgeworth [26] introduced the Method of Translation, which consisted in using transformations on random variables such that the probability distribution of the output variable would be a normal distribution. This allowed to draw, from the many results on normal distributions (e.g. the statistical significance test), conclusions on a more wide group of functions, thanks to the relation between the normal random variable and the transformed. The drawback was that Edgeworth's method only considered polynomial types of transformation, and this restriction meant that only a limited number of curves could be included in this analysis. Therefore, the Pearson's method was preferred to the Method of Translation. The latter has been since then improved and generalised, from the first proposal of only using polynomials as transformations, to considering more general transformations. Kapteyn and Van Uven [41] and Baker [7] showed that every continuous distribution can be transformed into a normal distribution.

This means that, given a standard normal variable Z, one could consider the system of all random variables X such that

$$f(X) = Z \tag{1.22}$$

with f being a not better specified function; nevertheless, in the interest of obtaining something treatable, the function f needs to be specified and bound to a finite number of parameters.

Johnson [39] proposed an application of Edgeworth's Method to transform the observed distributions into normal distributions, when the data do not allow to consider the analysed variables as normally distributed (for example in the case of non zero skewness, i.e. when the data distribution is asymmetric with respect to its average), focusing in particular on the logarithmic transformations, which had previously been investigated, among others, by Wicksell [81] and Rietz [58].

The reason for the interest in logarithmic transformation is the properties of the distributions that can be obtained in this way from the normal distribution.

Johnson restricts Eq. 1.22 to the case

$$a + b \cdot f\left(\frac{X - c}{d}\right) = Z \tag{1.23}$$

with f (for practical convenience) a non-decreasing function of x, not depending on any parameters; and b, $d > 0$.

The parameters c and d permit a shift and a dilation / contraction of the random variable X, providing a function $Y = \frac{X-c}{d}$ with a distribution of the same shape of X, a mean equal to $E(Y) = \frac{E(X)-c}{d}$ and a variance $V(Y) = \frac{V(X)}{d^2}$. The parameters a and b are instead related to the skewness and kurtosis of Y.

The estimation of the parameters a ,b, c, d is obtained by fitting the moments of the observed distribution up to the fourth: the realised skewness and kurtosis bind the values of a and b, and - once those are fixed - the appropriate values c and d are recovered from the average and the actual variance.[1]

Obviously, a bad estimation of the parameters (provided that the model is correct) brings to the wrong distribution.

In order for the system to be worthwhile, Johnson asks of f to be a function easy to calculate, with image equal to $\mathbb{R}$ and domain appropriate to interact with the distributions usually encountered in the data.

[1] Johnson [39] notes that the method of percentile points and the method of maximum likelihood can also be used.

S_L family and selection criteria

When we choose $f(y) = \log y$ in Equation (1.59), we obtain the lognormal transformation:

$$a + b \log\left(\frac{X - c}{d}\right) = Z. \tag{1.24}$$

In addition to the results in mathematical finance, the application of the lognormal transformation are several, in many fields, since it can be considered whenever there is a growth process that can be described as the cumulative result of percentage modifications of the previous value (see for example Limpert et al. [45]).

Applying a change of variable to the normal probability density function, we can obtain the probability density function for $Y = e^{\frac{Z-a}{b}}$ using the formula $f_Y(y) = f_Z(v(y))|v'(y)|$, where $v(y) = a + b \log y$:

$$f_Y(y) = \frac{b}{\sqrt{2\pi} y} e^{-\frac{(b \log y + a)^2}{2}}. \tag{1.25}$$

Wicksell [81] provided the formulas for the moments of Y. The n^{th}-moment of Y is given by:

$$\begin{aligned} \mu'_n &= \int_{-\infty}^{+\infty} y^n(z) f_Z(z)\, dz = \int_{-\infty}^{+\infty} e^{\frac{n(z-a)}{b}} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}\, dz = e^{\frac{n^2}{2b^2} - \frac{na}{b}} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-\frac{n}{b})^2}{2}}\, dz \\ &= e^{\frac{n^2}{2b^2} - \frac{na}{b}} \end{aligned} \tag{1.26}$$

Taking $\omega = e^{\frac{1}{b^2}}$, with the same notation used by Johnson [39], and considering the standard relationship between the central moments μ_j and the ordinary ones μ'_j, which is given by:

$$\mu_n = \sum_{j=0}^{n} (-1)^j \binom{n}{j} \mu'_{n-j} \mu'^j_1$$

where we consider $\mu'_0 = 1$, we can write:

$$\begin{aligned}
\mu_n &= \sum_{j=0}^{n} (-1)^j \binom{n}{j} e^{\frac{(n-j)^2}{2b^2} - \frac{(n-j)a}{b}} e^{\frac{j}{2b^2} - \frac{ja}{b}} = \\
&= \sum_{j=0}^{n} (-1)^j \binom{n}{j} e^{\frac{(n-j)^2+j}{2b^2} - \frac{na}{b}} = e^{-\frac{na}{b}} \sum_{j=0}^{n} (-1)^j \binom{n}{j} e^{\frac{(n-j)^2+j}{2b^2}} = \\
&= e^{-\frac{na}{b}} \sum_{j=0}^{n} (-1)^j \binom{n}{j} \omega^{\frac{(n-j)^2+j}{2}}
\end{aligned} \tag{1.27}$$

from which we have

$$\begin{aligned}
\mu_2 &= \mu'_2 - \mu'^2_1 = e^{-\frac{2a}{b}} \sum_{j=0}^{2} (-1)^j \binom{2}{j} \omega^{\frac{(2-j)^2+j}{2}} \\
&= e^{-\frac{2a}{b}} \left(\omega^2 - \omega\right) = e^{-\frac{2a}{b}} \omega \left(\omega - 1\right) \\
\mu_3 &= \mu'_3 - 3\mu'_2\mu'_1 + 2\mu'^3_1 = e^{-\frac{3a}{b}} \sum_{j=0}^{3} (-1)^j \binom{3}{j} \omega^{\frac{(3-j)^2+j}{2}} \\
&= e^{-\frac{3a}{b}} \left(\omega^{\frac{9}{2}} - 3\omega^{\frac{5}{2}} + 2\omega^{\frac{3}{2}}\right) = e^{-\frac{3a}{b}} \omega^{\frac{3}{2}} (\omega - 1)^2 (\omega + 2) \\
\mu_4 &= \mu'_4 - 4\mu'_3\mu'_1 + 6\mu'_2\mu'^2_1 - 3\mu'^4_1 = e^{-\frac{4a}{b}} \sum_{j=0}^{4} (-1)^j \binom{4}{j} \omega^{\frac{(4-j)^2+j}{2}} \\
&= e^{-\frac{4a}{b}} \left(\omega^8 - 4\omega^5 + 6\omega^3 - 3\omega^2\right) = e^{-\frac{4a}{b}} \omega^2 (\omega - 1)^2 (\omega^4 + 2\omega^3 + 3\omega^2 - 3).
\end{aligned}$$

We can express the skewness and the kurtosis of the resulting distribution Y via the indicators $\beta_1 = \gamma_1^2$, where γ_1 is the third standardized moment, for the skewness, and β_2, that is the fourth standardized moment, for the kurtosis, in the following way:

$$\begin{aligned}
\beta_1 &= \left(\frac{\mu_3}{\sigma^3}\right)^2 = \left(\frac{\mu_3}{\mu_2^{\frac{3}{2}}}\right)^2 = \left(\frac{e^{-\frac{3a}{b}} \omega^{\frac{3}{2}} (\omega - 1)^2 (\omega + 2)}{\left(e^{-\frac{2a}{b}} \omega (\omega - 1)\right)^{\frac{3}{2}}}\right)^2 = \\
&= \frac{(\omega - 1)^4 (\omega + 2)^2}{(\omega - 1)^3} = (\omega - 1)(\omega + 2)^2 \\
\beta_2 &= \frac{\mu_4}{\sigma^4} = \frac{e^{-\frac{4a}{b}} \omega^2 (\omega - 1)^2 (\omega^4 + 2\omega^3 + 3\omega^2 - 3)}{e^{-\frac{4a}{b}} \omega^2 \left(\omega - 1\right)^2}
\end{aligned} \tag{1.28}$$

$$= \omega^4 + 2\omega^3 + 3\omega^2 - 3. \tag{1.29}$$

By the properties of the logarithm, the functions in the lognormal system only need three parameters to be described, since:

$$a + b\log\left(\frac{X-c}{d}\right) = a - b\log d + b\log(X-c) = \hat{a} + b\log(X-c) \tag{1.30}$$

and we have made no request on a.

In giving formulas for the estimation of the $\hat{a}$, b, c parameters, in his 1949 paper Johnson [39] draws on a procedure by Wicksell [81] that uses the connection between the moments μ'_j of $Y = X - c$ and the central moments $\mu_{i,X}$ of X. For simplicity we will keep using the indexes β_1, β_2 and the moments $\mu'_{i,X}$, $\mu_{i,X}$ of X even when their estimates based on the data shall be used instead; it will be clear from the context when we are considering the latter or the former ones.

$$\begin{aligned}
\mu'_1 &= \mu'_{1,X} - c \\
\mu'_2 &= \mu_{2,X} + (\mu'_{1,X} - c)^2 \\
\mu'_3 &= \mu_{3,X} + 3\mu_{2,X}(\mu'_{1,X} - c) + (\mu'_{1,X} - c)^3
\end{aligned}$$

We can write the following system of equations:

$$\begin{cases}
k\omega^{\frac{1}{2}} = \mu'_{1,X} - c \\
k^2\omega^2 = \mu_{2,X} + (\mu'_{1,X} - c)^2 \\
k^3\omega^{\frac{9}{2}} = \mu_{3,X} + 3\mu_{2,X}(\mu'_{1,X} - c) + (\mu'_{1,X} - c)^3
\end{cases}$$

where μ'_1, μ'_2 and μ'_3 have been substituted from the previous equations with the respective values obtained with Equation (1.26), $k = e^{-\frac{\hat{a}}{b}}$ and we recall that $\hat{a} = a - b\log c$.

By isolating ω and k in the first two equations and substituting them in the following yields:

$$\frac{(\mu_{2,X} + (\mu'_{1,X} - c)^2)^3}{(\mu'_{1,X} - c)^3} = \mu_{3,X} + 3\mu_{2,X}(\mu'_{1,X} - c) + (\mu'_{1,X} - c)^3$$

which gives:

$$\mu^3_{2,X} + 3\mu^2_{2,X}(\mu'_{1,X} - c)^2 = \mu_{3,X}(\mu'_{1,X} - c)^3.$$

Dividing by $\mu_{2,X}^{\frac{3}{2}}$ we obtain that we can retrieve

$$t = \frac{\sqrt{\mu_{2,X}}}{\mu'_{1,x} - c}$$

as the only real root in the third degree equation: $t^3 + 3t = \gamma_1$, where we know there is only one real root since the discriminant $\frac{\beta_1}{4} + 1$ is positive. Obtaining t for the γ_1 of the data distribution allows to compute the estimates for the $\hat{a}$, b, c parameters.

A more immediate strategy of estimating the parameters is expressed in [27] and then implemented in the algorithm by Hill, Hill and Holder [37]. Considering Equation (1.30), for the moments of X we can write:

$$\mu'_{1,X} = e^{-\frac{\hat{a}}{b}}\sqrt{\omega} + c \tag{1.31}$$

$$\mu_{2,X} = e^{-\frac{2\hat{a}}{b}}\omega\,(\omega - 1) \tag{1.32}$$

$$\mu_{3,X} = e^{-\frac{3\hat{a}}{b}}\omega^{\frac{3}{2}}(\omega - 1)^2(\omega + 2) \tag{1.33}$$

$$\mu_{4,X} = e^{-\frac{2\hat{a}}{b}}\omega^2(\omega - 1)^2(\omega^4 + 2\omega^3 + 3\omega^2 - 3) \tag{1.34}$$

and we can see that β_1 and β_2 of X are exactly the same as those of Y.

This means that we can retrieve the value of ω for the X distribution by solving the equation $x^3 + 3x^2 - 4 = \beta_1$ or equivalently $x^3 - 3x - 2 = \beta_1$. Applying Cardano's formula to the latter equation we can observe that the discriminant $\Delta = \frac{\beta_1^2}{4} + 4\beta_1$ is positive, therefore there is only one real solution: $\omega = \sqrt{1 + \frac{1}{\beta_1} + \sqrt{\beta_1}(1 + \frac{\beta_1}{4})} + \sqrt{1 + \frac{1}{\beta_1} - \sqrt{\beta_1}(1 + \frac{\beta_1}{4})} - 1$.

Once we obtain ω, the $b = \frac{1}{\sqrt{(\log \omega)}}$ parameter is determined, and c and $\hat{a}$ can be recovered by Equations (1.31)-(1.32).

As we can see from the sign of μ_3, Y is always right-skewed. In order to retrieve also left-skewed functions, Hill, Hill and Holder [37] consider the d parameter to assume values ± 1 in accordance to the sign of the third moment of the observed distribution. With this modification, the c and a parameters can be retrieved by:

$$a = \frac{b}{2}\log\frac{\omega(\omega - 1)}{\mu_{2,X}}$$
$$c = \mu'_{1,X} - d\sqrt{\omega}e^{-\frac{a}{b}}.$$

Note that Equation (1.29) has for now not been used in determining the parameters, since the value of β_2 shall not give additional information on the distribution if this is lognormal.

Graphically, this means that the S_L family is individuated by a curve in the (β_1, β_2) plane: Johnson takes this as a starting point for the definition of his other two families, one for each of the regions the curve divides the plane into. Excluding the area where $\beta_2 < \beta_1 + 1$, in analogy with Pearson's family of curves, the (β_1, β_2) plane is divided into an S_B region, that contains the distributions with range of variation bound at both extremities, and an S_U region, that includes the distributions that have a range unbounded at either extremity.

In practice, this condition can serve as a criterium for the selection of the appropriate family starting from the data, and is implemented as such in the algorithm by Hill, Hill and Holder [37]. If the ω we obtain from Equation (1.28), substituted in the right-hand side of 1.29, returns a value acceptably close to the β_2 we compute from the data, then the observed distribution is classified as an element of the S_L family, the distributions that can be transformed into normal ones with the logarithmic function.

When the relation between the β_1 and β_2 estimated from the data does not fit into the scheme highlighted by Equations (1.28), (1.29), one of the other systems of curves is deemed more effective for representing the data distribution.

In addition to the S_L, S_U and S_B systems introduced by Johnson, Hill, Hill and Holder [37] also consider the special cases S_N of a normal distribution (which is selected if β_1 and β_2 are suitably close to 0 and 3 respectively, and corresponds to the limit of the S_L case when $b \to +\infty$) and S_T, the distributions for which $\beta_2 = \beta_1 + 1$, which correspond to the frontier of the S_B region but can profitably be seen as a stand-alone situation for they allow a straightforward identification of the four parameters.

When the β_2 estimated from the data exceeds the value of the right-hand side of (1.29), computed using the ω obtained from Equation (1.28), than the S_U transformation is selected, while if the β_2 estimated from the data is inferior to the computed value via the estimated ω, we must choose S_B or S_T transformations.

Slifker and Shapiro [70] point out some drawbacks in such a method, namely the fact that the β_1 and β_2 estimates have high variance and are biased for small samples and greatly affected by outliers. Hence, they propose instead a condition on percentiles in order to establish the family the sampling distribution belongs to and to provide an estimate for the parameters. Starting from an arbitrary value $z > 0$, which shall be chosen empirically in order to catch the area of most interest of the data distribution (for samples of moderate size, the authors suggest a value near $\frac{1}{2}$), the percentages $P_\zeta = \int_{-\infty}^{\zeta} f_Z(x)\, dx$ for $\zeta = -3z$, $-z$, z, $3z$ are computed and the corresponding percentiles from the data are selected: if we set the data in ascending order, x_ζ is going to be the observation in place $nP_\zeta + \frac{1}{2}$ (eventually interpolating).

Once we have obtained x_z, x_{-z}, x_{3z}, x_{-3z}, the following differences are defined:

$$m = x_{3z} - x_z \tag{1.35}$$

$$n = x_{-z} - x_{-3z} \tag{1.36}$$

$$p = x_z - x_{-z} \tag{1.37}$$

Slifker and Shapiro are guided by the intuition that, while the intervals $[-3z\,,\,-z]$, $[-z\,,\,z]$ and $[z\,,\,3z]$ are equally spaced, not only the intervals $[x_{-3z}\,,\,x_{-z}]$, $[x_{-z}\,,\,x_z]$ and $[x_z\,,\,x_{3z}]$ are not, but the difference in the spacing allows to distinguish between the bounded family, where the $[x_{-z}\,,\,x_z]$ interval is larger than the outer ones, and the unbounded family, where the $[x_{-z}\,,\,x_z]$ interval is smaller than the outer ones.

In terms of m, n and p, the above conditions translate into the following selection criterium:

- if $\dfrac{mn}{p^2} = 1$ then the S_L family is selected;
- if $\dfrac{mn}{p^2} > 1$ then the S_U is selected;
- if $\dfrac{mn}{p^2} < 1$ then the S_B is selected.

Since the condition $\dfrac{mn}{p^2} = 1$ has probability 0 to happen, in practice we will set a tolerance level η around 1.

After the family selection, for each situation Slifker and Shapiro express the parameters estimate in terms of m, n and p, generalising the results of Aitchinson and Brown [3] on the S_L system.

S_B family: bounded system

The bounded system is obtained by choosing $f(y) = \log \frac{y}{1-y} = 2\tanh^{-1}(2y-1)$ in Equation (1.59):

$$a + b \log\left(\frac{X-c}{d-X+c}\right) = Z \tag{1.38}$$

The function f satisfies the previous requirements of being increasing and having range $\mathbb{R}$. Moreover, the graph of $f(y)$ is symmetrical with respect to

$y = \frac{1}{2}$, which implies that the transformed distribution Y is symmetrical too if and only if $a = 0$, while Y is positively skewed when $a > 0$.

In order for this transformation to be well-defined, the range of values of the distribution $Y = \frac{1}{1+e^{-\frac{Z-a}{b}}}$ must satisfy $0 < y < 1$, which translates into $c < x < c + d$.

With a further change of variable from Equation (1.25), the density probability function for the distribution Y can be expressed as:

$$f_Y(y) = \frac{b}{\sqrt{2\pi}y(1-y)} e^{-\frac{(b \log \frac{y}{1-y} + a)^2}{2}}. \tag{1.39}$$

The n^{th} moment of the Y distribution is given by the following integral:

$$\mu'_n(a, b) = \int_{-\infty}^{+\infty} y^n(z) f_Z(z)\, dz = \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} \frac{e^{-\frac{z^2}{2}}}{1 + e^{-\frac{z-a}{b}}}\, dz \tag{1.40}$$

which does not yield a simple formula for their calculation. The mean of the Y process is expressed by Johnson with the following formula involving a series:

$$\begin{aligned}
\mu'_1(a, b) =& \frac{1}{\sqrt{2\pi}} e^{-\frac{a^2}{2}} \frac{1}{1 + 2\sum_{j=1}^{+\infty} e^{-2j^2\pi^2 b^2} \cos(2j\pi ab)} \cdot \\
& \cdot \left(\frac{1}{2b} + \frac{1}{b} \sum_{j=1}^{+\infty} e^{-\frac{j^2 b^2}{2}} \cosh \frac{j(1-2ab)}{2b^2} \operatorname{sech} \frac{j}{2b^2} + \right. \\
& \left. -2\pi b \sum_{j=1}^{+\infty} e^{-\frac{(2j-1)^2 \pi^2 b^2}{2}} \sin((2j-1)\pi ab) \operatorname{csch}((2j-1)\pi^2 b^2) \right)
\end{aligned} \tag{1.41}$$

and the successive moments are computed applying the following recursive formula involving partial derivatives:

$$\mu'_{n+1}(a, b) = \mu'_n(a, b) + \frac{b}{n} \frac{\partial \mu'_n(a, b)}{\partial a} \tag{1.42}$$

or the following recursive formula involving forward differences and steps of amplitude $\frac{1}{b}$ in the value of a:

$$\mu'_0(0, b) = 1 \tag{1.43}$$

$$\mu'_1(0, b) = \frac{1}{2} \qquad \text{due to the symmetry of } Y \text{ when } a = 0 \tag{1.44}$$

$$\mu'_{n+1}\left(\frac{m}{b},b\right) = -e^{-\frac{2k-1}{2b^2}}\left(\mu'_n\left(\frac{m-1}{b},b\right) - \mu'_{n+1}\left(\frac{m-1}{b},b\right)\right) \tag{1.45}$$

When $m > n+1$ the previous formula invokes negative moments, which are easier to compute from Equation (1.40). With further calculations (see the Appendix in [39]) Johnson concludes that the asymptotic behaviour of $\mu'_n(a,b)$ for $a \to +\infty$ is given by:

$$\mu'_n(a,b) \simeq e^{-\frac{na}{b}+\frac{n^2}{2b^2}} = e^{-n\Omega}\omega^{\frac{n^2}{2}} \tag{1.46}$$

where we use the notation introduced above of $\omega = e^{\frac{1}{b^2}}$ and we write $\Omega = \frac{a}{b}$.

The previous equation can be used to clarify that the bounded system occupies in the (β_1, β_2) plane the whole region between the S_L line and the "impossibility"line $\beta_2 = \beta_1+1$. In fact, Equation (1.46) and the standard relationship between central and non-central moments give the following asymptotic behaviours for the central moments and the skewness and kurtosis indices:

$$\mu_2(a,b) \simeq e^{-2\Omega}\omega(\omega-1) \tag{1.47}$$

$$\mu_3(a,b) \simeq e^{-3\Omega}\omega^{\frac{3}{2}}(\omega^3-3\omega+2) = e^{-3\Omega}\omega\frac{3}{2}(\omega-1)^2(\omega+2) \tag{1.48}$$

$$\mu_4(a,b) \simeq e^{-4\Omega}\omega^2(\omega^6-4\omega^3+6\omega-3) = e^{-4\Omega}\omega^2(\omega^6-4\omega^3+6\omega-3) \tag{1.49}$$

$$\beta_1(a,b) = \frac{\mu_3^2(a,b)}{\mu_2^3(a,b)} \simeq \frac{(\omega-1)^4(\omega+2)^2}{(\omega-1)^3} = (\omega+2)^2(\omega-1) \tag{1.50}$$

$$\beta_2(a,b) = \frac{\mu_4(a,b)}{\mu_2^2(a,b)} \simeq \omega^4+2\omega^3+3\omega^2-3. \tag{1.51}$$

We can differentiate Equation (1.39) and set the result to zero in order to find the modal values of the Y distribution:

$$f'_Y(y) = \frac{b}{\sqrt{2\pi}y^2(1-y)^2}e^{-\frac{(b\ln\frac{y}{1-y}+a)^2}{2}}\left(-1+2y-ab-b^2\log\frac{y}{1-y}\right) = 0 \tag{1.52}$$

which yields $g(y) = -1+2y-ab-b^2\log\frac{y}{1-y} = 0$.

The function $g(y)$ turns out to have either one or three zeros: in the first case the Y distribution is unimodal, otherwise it is bimodal.

Johnson [39] shows that the Y distribution is bimodal if and only if the following requirements are satisfied by the parameters:

$$|a| < \frac{\sqrt{1-2b^2}}{b} - b\log\frac{\sqrt{1-2b^2}}{1-\sqrt{1-2b^2}} \qquad 0 < b < \frac{1}{\sqrt{2}}.$$

Considering the above conditions we can divide the bounded region of the (β_1, β_2) plane in two sections, the unimodal and the bimodal one.

Johnson [39] proposes different ways to fit the data to the model, depending on the kind of information at our disposal:

- matching of mean and variance, if both end-points of the range of variation of the variable are known (which means we know parameters c and d) and the data are given singularly: we only need to apply the transformation $g(x) = \log\left(\frac{x-c}{c+d-x}\right)$ to the data and determine a and b by asking that the estimated mean and variance of the transformed observation match those of $\frac{Z-a}{b}$

- percentiles, if we are dealing with grouped data or we do not know both end-points of the range. The suggested method is similar to that described above. If we know one of the end-points then we need three equations: we fix z and compute $P = \frac{1}{2\pi}\int_z^{+\infty} e^{-\frac{1}{2}t^2}\,dt$, we estimate from the data the points x_M, x_U and x_L corresponding to the median, the upper $100\%P$ and lower $100\%P$ percentiles of the observed distribution and then estimate the remaining three parameters by imposing x_M, x_U and x_L to be translated into 0, z and $-z$ via the transformation $z = a + b\log\left(\frac{x-c}{c+d-x}\right)$. If we know neither one of the end-points then we need four equations: we consider four different values for $\zeta = z_1$, z_2, z_3 z_4 and compute their P_ζ, then we estimate the corresponding x_ζ from the data and impose $\zeta = a + b\log\left(\frac{x_\zeta - c}{c+d-x_\zeta}\right)$.

Hill, Hill and Holder [37] instead propose fitting by moments in any situation, for two main reasons: the practicality when theoretically computed moments are known; and the possibility of reaching, by the matching of moments, if not the best estimate of the parameters at least a good starting point for a maximum likelihood iteration (e.g. Newton-Raphson method) or Marquardt non-linear regression.

In Hill, Hill and Holder [37] algorithm, as we have already said above, unless β_2 is near $\beta_1 + 1$, S_B is the selected transformation family when the the right-hand side of (1.29), computed using the ω obtained from Equation (1.28), exceeds the β_2 estimated from the data.

In this case, the calculation of the first six moments is necessary for the determination of the fourth parameters. A first approximation of the b parameter is retrieved by numerical interpolation in the following way:

$$b_0 = \begin{cases} 0.8 \cdot (\beta_2 - 1) & \text{if } \beta_2 < 1.8 \\ \frac{0.626\beta_2 - 0.408}{(3-\beta_2)^{0.479}} & \text{otherwise.} \end{cases}$$

The constants intervening in the previous expressions are obtained by the shape of the curve for b as a function of β_2 when $\beta_1 = 0$. Similarly, a first approximation for a is found depending on the initial value b_0 and the estimated β_1 (using a formula by Draper). Let us call η the tolerance we allow our estimate of β_1 (in Hill, Hill and Holder [37] $\eta = 10^{-4}$).

$$a_0 = \begin{cases} 0 & \text{if } \beta_1 < \eta \\ (0.7466 b_0^{1.7973} + 0.5955)\beta_1^{0.485} & \text{if } b_0 \leq 1 \\ (1.0614 b_0^2 - 0.7077 b_0 + 0.9281)\beta_1^{0.0623 b_0 + 0.4043} & \text{if } 1 < b_0 \leq 2.5 \\ (1.0614 b_0^2 - 0.7077 b_0 + 0.9281)\beta_1^{0.0124 b_0 + 0.5291} & \text{if } b_0 > 2.5 \end{cases}$$

Then we use a_0 and b_0 as starting values for the Newton-Raphson method, asking for the matching of the moments of Y μ'_i up to the sixth (using the derivatives for the moments obtained through the recursive formula by Mordell we have seen above (1.42)).

After a and b have been successfully estimated, c and d are obtained by matching of the first two moments:

$$d = \frac{\mu_{2,X}}{\sqrt{\mu_2}} \tag{1.53}$$

$$c = \mu'_{1,X} - d\mu'_1. \tag{1.54}$$

If β_2 is larger than $\beta_1 + 1$ but close to it, Hill, Hill and Holder [37] label it as an S_B "two-ordinate" case, and set the parameters as

$$a = 0 \tag{1.55}$$

$$b = \frac{1}{2}\left(1 + \text{sgn}(\mu_3)\sqrt{\frac{\beta_1}{\beta_1 + 4}}\right) \tag{1.56}$$

$$d = \mu'_{1,X} - \mu_{2,X}\sqrt{\frac{1-b}{b}} \tag{1.57}$$

$$c = \mu'_{1,X} - \mu_{2,X}\sqrt{\frac{b}{1-b}}. \tag{1.58}$$

S_U family: unbounded system

The unbounded system is obtained by choosing $f(y) = \log(y + \sqrt{y^2+1}) = \sinh^{-1} y$ in Equation (1.59):

$$a + b\log\left(\frac{X-c}{d} + \sqrt{\left(\frac{X-c}{d}\right)^2 + 1}\right) = Z \tag{1.59}$$

The function f satisfies the previous requirements of being increasing and having range $\mathbb{R}$ and is defined for all $y \in \mathbb{R}$.

Moreover, f is odd, since

$$f(-y) = \log(-y+\sqrt{y^2+1}) = -\log \frac{1}{-y+\sqrt{y^2+1}} = -\log \frac{y+\sqrt{y^2+1}}{-y^2+y^2+1} = -f(y)$$

for all $y \in \mathbb{R}$, and it is convex for $y > 0$.

Also in this case we focus on the moments of the distributions involved. The n^{th}-moment of Y is given by:

$$\begin{aligned}
\mu'_n &= \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \left(\frac{e^{\frac{z-a}{b}} - e^{-\frac{z-a}{b}}}{2} \right)^n dz = \\
&= \frac{1}{2^n} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \sum_{i=0}^{n} (-1)^i \binom{n}{i} \left(e^{\frac{z-a}{b}} \right)^{n-i} \left(e^{-\frac{z-a}{b}} \right)^i dz = && (1.60) \\
&= \frac{1}{2^n} \sum_{i=0}^{n} (-1)^i \binom{n}{i} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} e^{z\frac{n-2i}{b} - a\frac{n-2i}{b}} dz = && (1.61) \\
&= \frac{1}{2^n} \sum_{i=0}^{n} (-1)^i \binom{n}{i} e^{\frac{(n-2i)^2}{b^2} - a\frac{n-2i}{b}} \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{(z-\frac{n-2i}{b})^2}{2}} dz = && (1.62) \\
&= \frac{1}{2^n} \sum_{i=0}^{n} (-1)^i \binom{n}{i} e^{\frac{n-2i}{b}\left(\frac{n-2i}{2b} - a\right)} && (1.63) \\
&= \frac{1}{2^n} \sum_{i=0}^{n} (-1)^i \binom{n}{i} \omega^{\frac{(n-2i)^2}{2}} e^{-\Omega(n-2i)} && (1.64)
\end{aligned}$$

where we used $\omega = e^{\frac{1}{b^2}}$ and $\Omega = \frac{a}{b}$ as before, which gives:

$$\mu'_1 = \frac{1}{2}\sqrt{\omega}\left(e^{-\Omega} - e^{\Omega}\right) = \sqrt{\omega}\sinh(-\Omega) \tag{1.65}$$

$$\mu'_2 = \frac{1}{2}\left(\omega^2\cosh(2\Omega) - 1\right) \tag{1.66}$$

$$\mu'_3 = -\frac{1}{4}\sqrt{\omega}\left(\omega^4\sinh(3\Omega) - 3\sinh\Omega\right) \tag{1.67}$$

$$\mu'_4 = +\frac{1}{16}\left(\omega^8\cosh(4\Omega) - 4\cosh(2\Omega) + 6\right). \tag{1.68}$$

From the previous Equations we can obtain the values for the central moments and therefore the indices for skewness and kurtosis.

$$\mu_2 = \frac{1}{2}(\omega - 1)\left(\omega \cosh(2\Omega) + 1\right) \tag{1.69}$$

$$\mu_3 = -\frac{1}{4}\sqrt{\omega}(\omega - 1)^2 \left(\omega(\omega + 2)\sinh(3\Omega) + 3\sinh\Omega\right) \tag{1.70}$$

$$\mu_4 = \frac{1}{8}(\omega - 1)^2 \left(\omega^2(\omega^4 + 2\omega^3 + 3\omega^2 - 3)\cosh(4\Omega) + 4\omega^2(\omega + 2)\cosh(2\Omega) + 3(2\omega + 1)\right) \tag{1.71}$$

$$\beta_1 = \frac{1}{2}\frac{\omega(\omega - 1)\left(\omega(\omega + 2)\sinh(3\Omega) + 3\sinh\Omega\right)^2}{\left(\omega\cosh(2\Omega) + 1\right)^3} \tag{1.72}$$

$$\beta_2 = \frac{1}{2}\frac{(\omega^4 + 2\omega^3 + 3\omega^2 - 3)\cosh(4\Omega) + 4\omega^2(\omega + 2)\cosh(2\Omega) + 3(2\omega + 1)}{\left(\omega\cosh(2\Omega) + 1\right)^2}. \tag{1.73}$$

From Equation (1.70), we have that the sign of a determines the sign of μ_3, hence the skewness of the distribution: for a positive a we will have a negative μ_3, and the inequality mode>median>mean will be true, therefore the function Y will be left-skewed; on the other hand a negative a will give a positive μ_3, and the converse inequality for the central tendencies will hold: mode<median<mean, therefore the function Y will be right-skewed.

As usual, β_1 and β_2 are the same for X and Y. The formulas for β_1 and β_2 need to be studied in order to show that the S_U family covers the region of the (β_1, β_2) plane delimited by the β_2 axis and the S_L line.

As Tuenter [78] proves, for β_1 and β_2 the following conditions must hold:

$$0 \leq \beta_1 < (\omega - 1)(\omega + 2)^2 \tag{1.74}$$

$$\frac{1}{2}(\omega^4 + 2\omega^2 + 3) \leq \beta_2 < \omega^4 + 2\omega^3 + 3\omega^2 - 3 \tag{1.75}$$

From the condition (1.75), considering that $\omega > 1$, we have that $\beta_2 > 3$, which means that X is a leptokurtic distribution.

Following Johnson's [39] justification, we can focus on the behaviour of the β_1 and β_2 indices for extreme values of the a and b parameters.

For $a = 0$, we have $\Omega = 0$, therefore $\sinh(k\Omega) = 0$ and $\cosh(k\Omega) = 1$ for all $k \in \mathbb{R}$, which means:

$$\beta_1 = 0$$

$$\beta_2 = \frac{1}{2}\frac{\omega^2(\omega^4 + 2\omega^3 + 3\omega^2 - 3) + 4\omega^2(\omega + 2) + 3(2\omega + 1)}{(\omega + 1)^2} = \frac{1}{2}(\omega^4 + 2\omega^2 + 3).$$

We can therefore identify the point $(0, \frac{1}{2}(\omega^4 + 2\omega^2 + 3))$ as the starting point of the unbounded region. If $b \to +\infty$ then $\omega \to 1$ and $\beta_2 \to 3$ which agrees with the corresponding lognormal couple $(\beta_1, \beta_2) = (0, 3)$. If $b \to 0$ then $\omega \to +\infty$, and $\beta_2 \to +\infty$.

Fixing b instead and analysing β_1 and β_2 behaviours for $a \to +\infty$, provides for the other limitation of the region. Using the equalities and asymptotic behaviours concerning the hyperbolic functions reported in the Appendix (Section 3.2):

$$\lim_{\Omega\to+\infty} \beta_1 = \lim_{\Omega\to+\infty} \frac{\omega(\omega-1)}{2} \frac{\omega^2(\omega+2)^2 e^{6\Omega}/4}{\omega^3 e^{6\Omega}/8} = (\omega-1)(\omega+2)^2$$

$$\lim_{\Omega\to+\infty} \beta_2 = \lim_{\Omega\to+\infty} \frac{1}{2} \frac{\omega^2(\omega^4 + 2\omega^3 + 3\omega^2 - 3)e^{4\Omega}/2}{\omega^2 e^{4\Omega}/4} = \omega^4 + 2\omega^3 + 3\omega^2 - 3$$

which means that for $a \to +\infty$ we reach the S_L line.

Moreover, as Tuenter [78] points out, inequalities (1.74) and (1.75) translate into boundaries for ω, additional to the obvious one $\omega > 1$:

$$\max\{w_0, w_1\} < \omega \leq w_2 \tag{1.76}$$

where w_0, w_1 and w_2 are the unique positive roots of equations $(\omega-1)(\omega+2)^2 = \beta_1$ (which can be solved via Cardano's method), $\omega^4 + 2\omega^3 + 3\omega^2 - 3 = \beta_2$ (which Tuenter solves by Ferrari's method) and $\frac{1}{2}(\omega^4 + 2\omega^2 + 3) = \beta_2$ (which is straightforward). $w_1 < w_2$ since the function $f_1(\omega) = \omega^4 + 2\omega^3 + 3\omega^2 - 3 - \beta_2$ is increasing for $\omega > 0$ and $f_1(w_1) = 0$ by definition, while $f_1(w_2) = \beta_2 + 2w_2^3 + w_2^2 - 6 > \beta_2 - 3$ since we are only considering $\omega > 1$. Since, as we have pointed out above, $\beta_2 > 3$, we have $w_1 < w_2$.

In Tuenter's algorithm, the condition (1.74) is checked for w_1 in order to confirm that the S_U transformation is the appropriate one.

In order to estimate the value of the parameters a, b, c, d from the data, Johnson [39] proposes the use of the matching of the mean, standard deviation, β_1, β_2. Specifically, using Equations (1.72) and (1.72), he provides an abac linking the couples (β_1, β_2) to the pairs (ω, Ω) that give rise to the skewness and kurtosis estimated from the data.

Hill, Hill and Holder [37] treat the case as an S_U one when the β_2 estimated from the data exceeds the corresponding value in the right-hand side of Equation (1.29), computed using the ω obtained from Equation (1.28). Their algorithm provides the parameters for the $\beta_1 = 0$ case:

$$a = 0 \tag{1.77}$$

$$\omega = \sqrt{\sqrt{2\beta_2 - 2} - 1} \quad \Rightarrow \quad b = (\log \omega)^{-\frac{1}{2}} \tag{1.78}$$

$$c = \mu_{1,X} \tag{1.79}$$

$$d = \sqrt{\frac{2\mu_{2,X}}{(\omega - 1)(\omega + 1)}} \tag{1.80}$$

while for the $\beta_1 \neq 0$ case they implement the iteration presented in [40]. The iteration adopts the formulas for the indices of skewness and kurtosis as modified by Leslie [44]:

$$\beta_1 = \frac{(\omega - 1)t[4(\omega + 2)t + 3(\omega + 1)^2]^2}{2(2t + \omega + 1)^3} \tag{1.81}$$

$$\beta_2 = \frac{(\omega - 1)(A_2(\omega)t^2 + A_1(\omega)t + A_0(\omega)}{2(2t + \omega + 1)^2} + 3 \tag{1.82}$$

that come from Equations (1.72) and (1.73), using the hyperbolic equalities (3.7)-(3.9) in the Appendix, with the substitution $t = \mu_1'^2 = \omega \sinh^2 \Omega$ and

$$\begin{aligned} A_2(\omega) &= 8(\omega^3 + 3\omega^2 + 6\omega + 6) \\ A_1(\omega) &= 8(\omega^4 + 3\omega^3 + 6\omega^2 + 7\omega + 3) \\ A_0(\omega) &= \omega^5 + 3\omega^4 + 6\omega^3 + 10\omega^2 + 9\omega + 3) \end{aligned}$$

and develops as follows:

- the starting point is $\omega_0 = \sqrt{\sqrt{2\beta_2 - 2.8\beta_1 - 2} - 1}$;
- given some ω_n, the solution t_n of Equation

  $$\beta_2 - 3 = \frac{(\omega_n - 1)(A_{2,n}t^2 + A_{1,n}t + A_{0,n}}{2(2t + \omega_n + 1)^2},$$

  where $A_{i,n} = A_i(\omega_n)$ for $i = 0, 1, 2$, is computed;
- an approximation b_n of β_1 based on ω_n and t_n is found:

  $$b_n = \frac{t_n(\omega_n - 1)[4(\omega_n + 2)t_n + 3(\omega_n + 1)^2]^2}{2(2t_n + \omega_n + 1)^3}$$

- the subsequent term ω_{n+1} is obtained from Equation

  $$\frac{\beta_2 - 0.5(\omega_{n+1}^4 + 2\omega_{n+1}^2 + 3)}{\beta_1} = \frac{\beta_2 - 0.5(\omega_n^4 + 2\omega_n^2 + 3)}{b_n}$$

- the procedure stops when b_n is satisfactorily close to β_1.

The procedure allows to estimate b and $|a|$. The sign we need to attribute to a is the opposite of μ_3. Again, once the first two parameters a and b have been estimated, the matching of the first two moments provides with the estimates for c and d.

Johnson's algorithm [40] has been generalised by Shenton and Bowman [67], providing a Lagrange expansion for $(\omega^2+1)^2$ which allows still another iterative procedure for the approximation of ω:

- the starting point is $\omega_0 = 1$ or $\omega_0 = w_2$;
- given some ω_n, the functions

$$H(\beta_1, \omega_n) = \sum_{i=1}^{+\infty} \frac{2^{2i}(3i-3)!}{3^{3i}(2i-1)!} \left(\frac{\beta_1}{(\omega_n+1)^3} \right)^i$$

  and

$$g(\omega_n) = (\omega_n+1)^2(\omega_n^2 + 2\omega_n + 3)$$

  are computed;

- ω_{n+1} is obtained by

$$\omega_{n+1} = \sqrt{-1 + \sqrt{2\beta_2 - 2 - 6g(\omega_n)H(\beta_1, \omega_n)}}$$

Bowman and Shenton [11] also proposed a different procedure to estimate ω. They followed the linearity argument used by Johnson with regard to the contours of b and Ω in the (β_1, β_2) plane and showed that the function

$$\psi(\beta_1, \beta_2, \omega) = \frac{\beta_2 - 0.5(\omega^4 + 2\omega^2 + 3)}{\beta_1} \tag{1.83}$$

used in the iteration described above also has nearly linear contours in the (β_1, β_2) plane for $3 < \beta_2 \leq 75$, covering a wider area than that spanned by Johnson's abac ($\beta_2 \leq 15$). They proposed to approximate the function ψ with the ratio of two third degree polynomials in β_1 and β_2, and they estimated the polynomial coefficients via linear least squares using 350 points in the region of the (β_1, β_2) plane delimited by the S_L curve, the axis $\beta_1 = 0$ and the line $\beta_2 = 50$.

Hence, Bowman and Shenton [11] procedure for the estimation of the Johnson's S_U parameters involves plugging the values β_1 and β_2 obtained from the data in the polynomials, obtaining an approximation of $\psi(\beta_1, \beta_2, \omega)$ from which an approximation of ω can be computed solving (1.83); afterwards Ω is obtained by (1.82) and c and d are recovered as usual. The error in the ω approximation results in at most 0.006% over the spanned domain.

The obvious drawbacks of both the previous methods are the empirical foundation on the observed linearity properties, which may only hold locally, and also that the coefficients of the polynomials involved depend on the regions of the (β_1, β_2) plane that have been investigated, and their accuracy influences the accuracy of the parameters.

Trying to overcome these difficulties, Tuenter [78] proposes an algorithm that also starts from Equations (1.72) and (1.73) but proposing a different substitution:

$$t = \frac{\omega^2 - 1}{\omega + 1 + 2\omega \sinh^2 \Omega}. \tag{1.84}$$

With the substitution (1.84) the formulas for β_1 and β_2 become polynomial:

$$\beta_1 = (\omega - 1 - t)\left(\omega + 2 + \frac{t}{2}\right)^2 \tag{1.85}$$

$$\beta_2 = (\omega^2 + 2\omega + 3)\left(\omega^2 - \frac{t^2}{2} - 2t\right) - 3 \tag{1.86}$$

and an expression for t can be isolated from (1.86) and plugged into (1.85).

$$t = -2 + \sqrt{4 + 2\left(\omega^2 - \frac{\beta_2 + 3}{\omega^2 + 2\omega + 3}\right)} \tag{1.87}$$

$$\beta_1 = \left(\omega + 1 - \sqrt{4 + 2\left(\omega^2 - \frac{\beta_2 + 3}{\omega^2 + 2\omega + 3}\right)}\right)\left(\omega + 1 + \frac{\sqrt{4 + 2\left(\omega^2 - \frac{\beta_2+3}{\omega^2+2\omega+3}\right)}}{2}\right)^2 \tag{1.88}$$

From Equation (1.84) we can see that t is necessarily positive (since $\omega > 1$) and not greater than $\omega - 1$. Manipulating Equation (1.87) one can see that these conditions on t are equivalent to the conditions (1.75). Moreover, one could observe that the proposed substitution is not injective, but if we express ω in terms of t we obtain:

$$\omega = \frac{t(1 + 2\sinh^2 \Omega) \pm \sqrt{t^2(1 + 2\sinh^2 \Omega)^2 + 4 + 4t}}{2}.$$

which gives only one acceptable $\omega > 0$ for $t > 0$.

Therefore the proposed substitution is bijective for ω satisfying the conditions (1.76).

Tuenter proves (in the Appendix, pages 342-345) that the function $f(\omega)$ given by the right-hand side of (1.88) is a function strictly decreasing in ω, therefore in order to estimate the parameters one can implement the bisection method or the Newton-Raphson method to find ω which satisfies (1.88). Such an ω must be searched in the interval $(w_1, w_2]$ (where w_1 and w_2 are the unique positive roots of $\omega^4+2\omega^3+3\omega^2-3 = \beta_2$ and $\frac{1}{2}(\omega^4+2\omega^2+3) = \beta_2$ respectively). The Newton-Raphson iteration function is given by

$$g(\omega) = \omega - \frac{f(\omega) - \beta_1}{f'(\omega)} \tag{1.89}$$

and w_2 is an appropriate starting point. Once ω has been estimated, t is obtained by Equation (1.84), and from t also Ω, taking care as usual to attribute to Ω a sign opposite to that of the third moment. c and d are obtained from the matching of the first two moments.

The boundaries w_1 and w_2 also allow Tuenter to analyse the accuracy of the convergence of his method to the parameter value, giving for the ω computed with n iterations of the bisection method a relative error of $0.6818 \cdot 2^{-n}$. [78]

Relation between Pearson's and Johnson's curves

The S_U portion of the (β_1, β_2)-plane covers entirely Pearson's Type IV and part of Pearson's Type VI curves [27]. Johnson's curves give a good approximation of Pearson's Type IV curves [39].

The S_B system provides transformation of Pearson's Types I and II that approximate normality better than the transformation obtained via the S_L system, while the same cannot be said for Type III transformation. The S_L system was also suggested by Pearson himself as a feasible approximation of his Type IV curves.

Since Johnson's systems all provide distributions with all finite moments, these system do not capture the situations where higher moments than the fourth are infinite, while Pearson's family can accomodate such cases.

Johnson's system for the pricing of options

Simonato [68] started from the Edgeworth binomial tree to propose the evaluation of options with an underlying characterised by Merton's jump diffusion process with lognormal jump amplitude.

Simonato starts from the solution for Merton's process (see page 7)

$$S = S_0 e^{(r-v-\lambda\bar{j}-\frac{\sigma^2}{2})t+\sigma z(t)} \prod_{i=0}^{m(t)} (1+J_i)$$

with $m(t)$ Poisson of parameter λ, $J_0 = 0$ and $\log(1+J_i) \sim N(\gamma', \delta^2)$ for $i \geq 1$, and uses the formulas provided by Das and Sundaram [22] for the annualised higher moments of the cumulative asset return $\log(S_\tau/S_0)$.

$$\begin{aligned}
\mu_2 &= \sigma^2 + \lambda(\gamma'^2 + \delta^2) \\
\gamma_1 &= \frac{1}{\sqrt{\tau}} \left(\frac{\lambda\gamma'(\gamma'^2 + 3\delta^2)}{\mu_2^{3/2}} \right) \\
\beta_2 &= 3 + \frac{1}{\tau} \left(\frac{\lambda(\gamma'^4 + 6\gamma'^2\delta^2 + 3\delta^4)}{\mu_2^2} \right)
\end{aligned}$$

Values γ_1 and β_2 are given as an input for the Hill et al. algorithm [37], together with $\mu_{1,X} = 0$ and $\mu_{2,X} = 1$ (the process is standardised). The Hill algorithm selects the function $f(y)$ of the appropriate family in accordance to the (β_1, β_2) values, and gives as an output the four parameter estimates a, b, c, d for the Johnson transformation. Once these values are at our disposal, we consider all possible values assumed by a binomial random variable after n steps, that is:

$$z_{n,i} = \frac{2i-n}{\sqrt{n}} \qquad \text{with probability } p_{n,i} = \binom{n}{i} \left(\frac{1}{2}\right)^n$$

and we translate them (leaving the probabilities untouched) into $x_{n,i} = c + df^{-1}\left(\frac{z_{n,i}-a}{b}\right)$

The discrete distribution X_n given by the $x_{n,i}$ and $p_{n,i}$ is then standardised, and the so found new values $x'_{n,i}$ are used to compute the possible values of the stock at maturity:

$$S(n,i) = S_0 e^{\mu\tau + \nu\sqrt{\tau}x'_{n,i}}$$

where $\mu = r - d - \frac{1}{\tau}\sum_{i=0} n p_{n,i} e^{\nu\sqrt{\tau}x'_{n,i}}$ and $\nu = \sqrt{\mu_2}$ is the annualised volatility of $\log(S_\tau/S_0)$.

Backward recursion allows the computation of all $S(i,j)$ for $i \leq n$, $0 \leq j \leq i$:

$S(i,j) = e^{-(r-d)\Delta t \frac{S(i+1,j+1)+S(i+1,j)}{2}}$

and at the same time permits to recover an approximation of the price of an option written on S. The procedure's main advantage is that the computation is fast, alas the accuracy in the determination of the price cannot be improved by increasing the number of steps (that is, we do not have convergence results) since we are using a one-dimensional lattice, that cannot adequately represent information from both sources of variation of the price, the Brownian motion and the jump, that are featured in the model.

This springs the research for a lattice model that can better represent the underlying process, while allowing for a fast pricing of the derivatives.

1.5 On the discretisation of the Merton model

The aim of this Section is to analyse already known lattice procedures for the discretisation of Merton's jump-diffusion model. In the following Sections we will highlight the improvements brought to the procedure by the modifications introduced by Gaudenzi, Spangaro and Stucchi ([33] and [34]).

1.5.1 Amin

Amin proposes the following procedure for derivative pricing in a discretisation of Merton's jump-diffusion setting.

The jump process is considered lognormal, which means that a jump provides a variation of the stock price from S to $S(1+J)$, where $\log(J+1) \sim N(\gamma', \delta^2)$. In Amin's calculations the γ' parameter, which Merton [50] defined as $\gamma - \frac{\delta^2}{2}$, is taken equal to $-\frac{\delta^2}{2}$, i.e. $\gamma = 0$.

Consider the parameters of the logreturn as usual: r the risk-free rate, v the dividend yield, λ the Poisson parameter that gives the frequency of the arrival of jumps, $\bar{j}$ the expectation of the jump process J, σ the volatility of the diffusion component, $\alpha = r - v - \lambda\bar{j} - \frac{\sigma^2}{2}$ the drift. All these values are supposed to be constant throughout the lattice.

The time to maturity τ of the option we want to evaluate is divided in n intervals, each of amplitude $\Delta t = \frac{\tau}{n}$. In the following, we will often make an abuse of notation and indicate by i the time $i\Delta t$. We will also take $R = e^{r\Delta t}$ and $D = e^{v\Delta t}$, while $u = e^{\alpha\Delta t + \sigma\sqrt{\Delta t}}$ and $d = e^{\alpha\Delta t - \sigma\sqrt{\Delta t}}$ will stand for the amplitude of an up and a down Brownian move respectively.

In order to recover the backward recursion formula for the price of the option at time i, known the possible prices at time $i+1$, Amin follows the hedging argument by CRR [20] for the value $V_j(i+1)$ at time $i+1$ of a portfolio made of an option, risk-free bonds and shares of the underlying stock. At first he considers the case of absence of jumps ($j = \pm 1$), obtaining

$$\pi = \frac{(R-D)-d}{u-d}$$

for the transition probability of an up Brownian move; then imposes the diversifiability of the jump risk, which gives that the expectation of the portfolio in the next period must be equal to zero when taken with respect to the distribution of the rare event.[2]

Therefore,

$$0 = \lambda E_{J+1}[V_j(i+1)] + (1-\lambda)V_{\pm 1}(i+1).$$

This gives that the option value $C(i)$ at time i, provided that we don't exercise it, depends on the option values at time $i+1$ according to the following formula:

$$C(i) = \frac{1}{R}\left\{\lambda E_{J+1}[C_j(i+1)] + (1-\lambda)p' C_{+1}(i+1) + (1-\lambda)(1-p')C_{-1}(i+1)\right\} \tag{1.90}$$

where

$$p' = \frac{\frac{(R-D)-\lambda E_{J+1}[J+1]}{1-\lambda} - d}{u-d}.$$

Let us call $C(i)$ the continuation value of the option at time i.

The probabilities p' and λ are then those which define a risk-neutral measure Q under which the value of the option $C(i)$ is the discounted expected return of the values at the next time step. The distribution of J under Q is taken to be the same as under the original probability measure.

Equation (1.90), coupled with the underlying values at maturity, allows for a backward recursive evaluation of the option, with the following standard caveat: if the option supports early exercise, at every time step the current payoff of the option may exceed the expected value of detaining the option still in the next period, which means that the rational investor would choose to exercise it.

Therefore, while for a European option it suffices (1.90), for an American call option with strike K_0 we will need

$$A_C(i) = \max\{S(i) - K_0, C(i)\} \tag{1.91}$$

and likewise for an American put option

$$A_P(i) = \max\{K_0 - S(i), C(i)\}. \tag{1.92}$$

The definition of the probability p' imposes some conditions on the value of R, D, u, d, to ensure that $0 < p' < 1$.

[2] For details, see [5].

It is necessary to define a discretisation J_n of the J process, and more conditions are needed in order for the discrete process to weakly converge to Merton's jump-diffusion process.

At each time step i, for $i = 0, \dots, n$, the possible states for the value of the stock with initial value S_0 are of the kind $S_0 e^{\alpha i \Delta t + j\sigma\sqrt{\Delta t}}$, which means that jumps of the logreturn have a random amplitude which must be a multiple of the Brownian move $\sigma\sqrt{\Delta t}$. In each time interval, the stock price can move accordingly to the Brownian motion, i.e. the variation in the logreturn will be $\alpha\Delta t \pm \sigma\sqrt{\Delta t}$, or accordingly to a jump, which allows for any variation of the kind $\alpha\Delta t + j\sigma\sqrt{\Delta t}$ except $j = \pm 1$.

As a simplified approach, Amin supposes the two possible moves of the stock price, the Brownian one and the jump one, cannot both occur in the same interval, but this is not a restrictive request when we consider the limit for $n \to \infty$.

The probability of the arrival of a jump in a given period Δt is taken equal to $\lambda\Delta t$. Conditional to the happening of a jump, the probability q_k of going from state $S_0 e^{\alpha i\Delta t + j\sigma\sqrt{\Delta t}}$ to state $Se^{\alpha(i+1)\Delta t + (j+k)\sigma\sqrt{\Delta t}}$ is defined as follows:

- if $k \neq \pm 1$, we assign to the discrete jump of $k\sigma\sqrt{\Delta t}$ the probability $J+1$ has to fall into the interval $[\alpha\Delta t + (k - \frac{1}{2})\sigma\sqrt{\Delta t}, \alpha\Delta t + (k + \frac{1}{2})\sigma\sqrt{\Delta t}]$;
- if $k = \pm 1$, the probability is taken equal to 0 (this implies the ability for the observer to distinguish between the Brownian and the rare moves: if $k = \pm 1$, the variation is due to the diffusion process);
- if $k = 0$, we assign to the discrete jump of $k\sigma\sqrt{\Delta t}$ the probability $J+1$ has to fall into the interval $[\alpha\Delta t - \frac{3}{2}\sigma\sqrt{\Delta t}, \alpha\Delta t + \frac{3}{2}\sigma\sqrt{\Delta t}]$ (in this way, the "no jump" situation absorbs the probability that has been deflected from the neighbourhoods of the local moves).

With this specification for J_n, the value of p' is modified into:

$$p = \frac{\frac{(R-D) - \lambda\Delta t E_{J_n+1}[J_n+1]}{1-\lambda} - d}{u - d} = \frac{1}{2} + o(\Delta t),$$

which means the error will be negligible if we take $p = \frac{1}{2}$, as soon as n is big enough (Amin estimates an error on the value of p less than 10^{-3} for $n \geq 100$).

With this notation, the transition probabilities from the state $S_0 e^{\alpha i\Delta t + j\sigma\sqrt{\Delta t}}$ to whichever of the states $S_0 e^{\alpha(i+1)\Delta t + (j+k)\sigma\sqrt{\Delta t}}$ of the next period are:

- $(1 - \lambda\Delta t)p$ if $k = +1$,
- $(1 - \lambda\Delta t)(1 - p)$ if $k = -1$,

- $\lambda\Delta t q_k$ if $k \neq \pm 1$.

With these values, Amin [5] obtains the weak convergence of the (interpolated) discrete time process to the continuous time process, under a risk-neutral probability measure, which in turn guarantees the convergence of the European option prices computed via the discrete model to the continuous ones.

The model as described so far is not immediately applicable, for we have at every time step an infinite number of possible states: Amin restricts the number of possibilities to a fixed $M = 2n+1$ (recall that n is the number of steps) around the value $S_0 e^{\alpha i \Delta t}$, which results in a "rectangular" lattice, where at time i there are M possible states spanning from $S_0 e^{\alpha i\Delta t - n\sigma\sqrt{\Delta t}}$ to $S_0 e^{\alpha i\Delta t + n\sigma\sqrt{\Delta t}}$, and then proceeds to truncate the jump distribution in the following ways:

- for each state at time i, the jump will move the price of the underlying at most n places upwards and n places downwards, in any case respecting the borders of the "rectangular" lattice;
- given $k = \left\lceil \max\{\frac{\gamma'+3\delta-\alpha\Delta t}{\sigma\sqrt{\Delta t}}, \frac{3\delta-\gamma'+\alpha\Delta t}{\sigma\sqrt{\Delta t}}\}\right\rceil$, the distribution of the process J_n+1 is truncated outside the interval $I = [\alpha\Delta t - k\sigma\sqrt{\Delta t}, \alpha\Delta t - k\sigma\sqrt{\Delta t}]$. This is sensible since, for the definition of k, the interval I contains $[\gamma' - 3\delta, \gamma' + 3\delta]$, which means that we are considering the 99.7% of the normal distribution. The probability of the tails we are cutting out is assigned to the extremal nodes.

At every date, the price of the option when the underlying has the lowest possible or the highest possible value are either computed with the closed form formula for the European option price or set at their intrinsic value.

The previous procedure for the pricing of an American option is $O(n^3)$, since for every node we need to consider the value of the nodes it is parent to (which can amount to a minimum of $n+1$ for the border nodes up to a maximum of $2n+1$ for the central nodes), and this needs to be repeated for every node in the lattice, which means $n(2n+1)$ times.

Amin suggests to reduce the computational time via Newton-Cotes integration: for every node, instead of considering the whole set of n to $2n+1$ nodes we can reach in the following period, we divide the interval in three equally spaced subintervals (possibly enlarged with respect to the space state in order to fit specifications), and we consider the approximation of the expected value of the option in each of these intervals given by the 10-points Newton Cotes integral formula. The result would be an $O(n^2)$ procedure, which nevertheless looses in accuracy when compared to the Merton closed formula benchmark.

1.5.2 Hilliard and Schwarz

Hilliard and Schwartz [38] point out that Amin's discretisation is negatively affected by the limitation imposed on the jump, to only take values which are multiple of the Brownian move.

Considering the independency of the two processes involved in the Merton model, the authors develop a multinomial lattice: one variable mimicking the diffusion process X_t and the second one the log-normal jumps in the compound Poisson process Y_t.

Given τ as the time to maturity, n the number of time steps, $\Delta t = \frac{\tau}{n}$ as the time interval, Hilliard and Schwartz consider $\sigma\sqrt{\Delta t}$ as the amplitude of the Brownian step, as per usual in binomial trees, whereas they establish a value h, for the minimum possible amplitude of the jump, that is not dependent on the amplitude of the Brownian step, but depends on Merton's parameters for the distribution of the continuous lognormal jump process.

The minimal amplitude of a jump, h, is set to:

$$h = b\sqrt{\gamma'^2 + \delta^2}, \qquad \text{with } 0 < b \leq 1 \tag{1.93}$$

which, for $b = 1$, ensures weak convergence in the special case of fixed jump amplitude (i.e. $\delta = 0$). This encouraged the authors to use $h = \sqrt{\gamma'^2 + \delta^2}$ also in the general case, a choice which is supported by accurate numerical results.

A non-negative integer N, constant throughout the tree, determines the maximum amplitude Nh of a single jump.

In order to recover a structure as faithful as possible to the jump dynamics, Hilliard and Schwartz introduce a node for the "no jumps" situation, and then $2N$ additional nodes to take into consideration the possibility of a jump of amplitude $\pm h$, $\pm 2h$, …, $\pm Nh$.

The discrete counterparts X_n and Y_n of X_τ and Y_τ are the algebraic sum of n i.i.d. processes X_Δ and Y_Δ respectively defined as follows.

The discretisation of the Brownian motion follows the classical CRR one: an up move of X_Δ gives a $+\sigma\sqrt{\Delta t}$ variation in the logreturn, a down move a $-\sigma\sqrt{\Delta t}$ variation; the probability p of an up move is taken as in the work by Nelson and Ramaswamy [52]:

$$p = \frac{1}{2}\left(1 + \frac{\alpha\sqrt{\Delta t}}{\sigma}\right).$$

For p to be well defined, we need to impose $-1 \leq \frac{\alpha\sqrt{\Delta t}}{\sigma} \leq 1$.

As for the jump part, the discretisation J_n of the process J which gives the amplitude of a single jump is then obtained in the following way: at every time step, the jump can assume the value kh for $k = -N, \ldots, 0, \ldots, N$ with a probability q_k.

The values for the q_k are also constant in time, and are found by imposing a moment-matching condition: the variation of the logreturn due to the jump in the Δt time interval shall have the same moments as its continuous counterpart $Y_{\Delta t}$ (as defined in 1.6):

$$\begin{cases} \sum_{k=-N}^{N} q_k = 1 \\ \sum_{k=-N}^{N} (kh)^{i-1} q_k = E\{Y_{\Delta t}^{i-1}\} \quad \text{for } i = 2, \ldots, 2N+1 \end{cases} \tag{1.94}$$

Considering the relations intervening between moments and cumulants (cf. Stuart and Ord [74], pp. 85 ff.), and the dependence on the amplitude of the time interval Δt of the cumulants of $Y_{\Delta t}$, which can be derived from the cumulant generating function (see Section 3.3 in the Appendix, with $t = \Delta t$), the moments $E(Y_{\Delta t}^{i-1})$ are proved to be equal to the relative cumulants k_{i-1}, up to a negligible $O(\Delta t^2)$ term. This allows Hilliard and Schwartz to use, in substitution of the moments, the first $2N$ cumulants. Hilliard and Schwartz test their procedure for $N = 1, 2, 3, 4$, therefore the cumulants they need are:

$$\begin{aligned} k_1 &= \lambda \Delta t \gamma' \\ k_2 &= \lambda \Delta t (\gamma'^2 + \delta^2) \\ k_3 &= \lambda \Delta t (\gamma'^3 + 3\gamma'\delta^2) \\ k_4 &= \lambda \Delta t (\gamma'^4 + 6\gamma'^2\delta^2 + 3\delta^4) \\ k_5 &= \lambda \Delta t (\gamma'^5 + 10\gamma'^3\delta^2 + 15\gamma'\delta^4) \\ k_6 &= \lambda \Delta t (\gamma'^6 + 15\gamma'^4\delta^2 + 45\gamma'^2\delta^4 + 15\delta^6) \\ k_7 &= \lambda \Delta t (\gamma'^7 + 21\gamma'^5\delta^2 + 105\gamma'^3\delta^4 + 105\gamma'\delta^6) \\ k_8 &= \lambda \Delta t (\gamma'^8 + 28\gamma'^6\delta^2 + 210\gamma'^4\delta^4 + 420\gamma'^2\delta^6 + 105\delta^8) \end{aligned} \tag{1.95}$$

Note that the q_k obtained as a solution of the system above, which requires the inversion of a Vandermonde matrix, already contain the information about the intensity of the Poisson process, λ.

The higher the choice for N, the more precise J_n will be as an approximation of J, therefore the more refined the results. A choice of $N = 1$ gives a rough approximation of the underlying and therefore of the price of the derivatives; by choosing $N = 2, 3, 4$ (that is, respectively, a five-, seven-, nine-node tree) the results are more refined. In the usual trade-off between precision and computational costs, the choice $N = 3$ seems to be the best option (see [38]).

The discretisation Y_Δ is defined as:

$$Y_\Delta := \begin{cases} kh & \text{with probability } q_k \text{ for } -N \le k \le N, \\ 0 & \text{with probability } 1-\lambda\Delta t. \end{cases}$$

The weak convergence of X_n to X_τ, and its consequence on the convergence of European put and call option prices, has already been proven in literature (see [52]); Hilliard and Schwartz prove the weak convergence of Y_n to Y_τ in the special case with a fixed jump amplitude, by showing that the characteristic function of Y_n converges to the characteristic function of Y_τ.

For the general case, there is no such result; we must be satisfied with the convergence of the first $2N$ cumulants of Y_n to those of Y_τ, which the authors show via the cumulant functions of the two processes. While comforting, this does not guarantee convergence of the discrete option prices to the continuous ones; we need to rely on a numerical justification only.

Hilliard and Schwartz construct a bivariate tree for the dynamic of the logreturn.

At every time step, the nodes of the tree represent the values of the return of the underlying considering that this is influenced both by the Brownian motion and by the lognormal jumps.

Fixed N, at every interval we can have two possibilities for the Brownian move ($+\sigma\sqrt{\Delta t}$ and $-\sigma\sqrt{\Delta t}$) and $2N+1$ possibilities for the jump. Every node (but the terminal ones) in the tree is parent to the $2\cdot(2N+1)$ nodes given by the combination of the up or down Brownian move with the $2N+1$ possible moves for the jump.

We can label every node of the tree with a triplet (i,j,k), where the first index i keeps track of the time ($0 \le i \le n$), j describes the effect of the Brownian moves up to that time ($0 \le j \le i$) and k the result of the jump moves ($-Ni \le k \le Ni$).

Let us denote by $S(i,j,k)$ the value of the underlying on the node (i,j,k). This means $S(i,j,k) = S_0 e^{(-i+2j)\sigma\sqrt{\Delta t}+kh}$.

The chosen dynamics entails that, from the state $S_0 e^{(-i+2j)\sigma\sqrt{\Delta t}+kh}$ we can move to:

$$\begin{aligned} &S_0 e^{[-(i+1)+2(j+1)]\sigma\sqrt{\Delta t}+(k+l)h} && \text{with probability } pq_l, \\ &S_0 e^{[-(i+1)+2j]\sigma\sqrt{\Delta t}+(k+l)h} && \text{with probability } (1-p)q_l \end{aligned}$$

for $l = -N, \ldots, N$.

Note that the Hilliard and Schwartz's bivariate tree incorporates CRR's binomial tree: indeed, from a node $S_0 e^{(-i+2j)\sigma\sqrt{\Delta t}}$, with probability q_0 (associated with the "no jump" situation) we go to $S_0 e^{[-(i+1)+2(j+1)]\sigma\sqrt{\Delta t}}$ or $S_0 e^{[-(i+1)+2j]\sigma\sqrt{\Delta t}}$.

In order to calculate the price of an option on the stock we are considering, we can use the following formula of backward recursion for the continuation value of the derivative at time i, in the node (i, j, k):

$$V(i,j,k) = e^{-r\Delta t} \sum_{l=-N}^{N} (V(i+1, j+1, k+l)p + V(i+1, j, k+l)(1-p))q_l. \quad (1.96)$$

Provided the values at maturity, which are equal to $V(n,j,k) = (S(n,j,k) - K_0)^+$ in the case of the call option and $V(n,j,k) = (K_0 - S(n,j,k))^+$ in the case of the put option, for any j integer between 0 and n and k integer between $-Nn$ and Nn, Equation (1.96) is sufficient for the evaluation of the European options. Alternatively, the price of the option in the European case can be evaluated as the discounted expected value on all the possible payoffs at maturity.

Given $Q_N(k)$ the probability of reaching jump value k at maturity, and

$$P(j) = \binom{n}{j} p^j (1-p)^{n-j},$$

the formula for the European call option is:

$$V_E = e^{-r\tau} \sum_{j=0}^{n} \sum_{k=-Nn}^{Nn} (S_0 e^{(-n+2j)\sigma\sqrt{\tau}+hk} - K_0)^+ P(j) Q_N(k) \quad (1.97)$$

and similarly for the European put option we have:

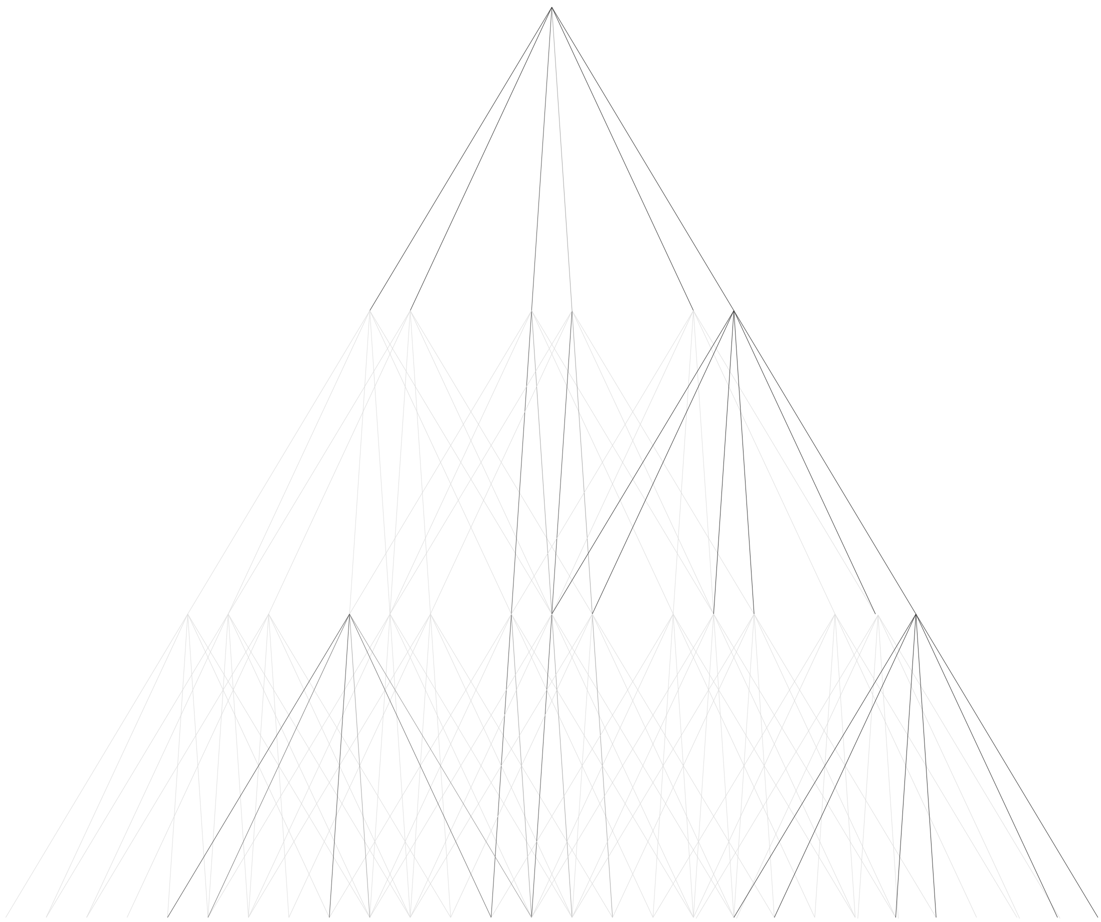
$$P_E = e^{-r\tau} \sum_{j=0}^{n} \sum_{k=-Nn}^{Nn} (K_0 - S_0 e^{(-n+2j)\sigma\sqrt{\tau}+hk})^+ P(j) Q_N(k). \quad (1.98)$$

These formulas have a computational complexity of $O(n^2)$.

For the American call options we will use the backward recursion with the analogous formula:

$$\begin{aligned} C_V(i,j,k) &= e^{-r\Delta t} \sum_{l=-N}^{N} (V_A(i+1, j+1, k+l)p + V_A(i+1, j, k+l)(1-p))q_l \\ V_A(i,j,k) &= \max\{C_V(i,j,k), S(i,j,k) - K_0\} \end{aligned} \quad (1.99)$$

Figure 1.1: Hilliard and Schwartz tree for the $N = 1$ case. In green, the CRR tree starting from the current value of the underlying. In black, the children of a single node. In different colours, the $+h$, 0, $-h$ jumps, every pair individuating a ± 1 Brownian move.



with initial data $V_A(n, j, k) = (S(n, j, k) - K_0)^+$, for j integer between 0 and n and k integer between $-Nn$ and Nn, while for the American put option:

$$\begin{aligned} C_P(i,j,k) &= e^{-r\Delta t} \sum_{l=-N}^{N} (P_A(i+1,j+1,k+l)p + P_A(i+1,j,k+l)(1-p))q_l \\ P_A(i,j,k) &= \max\{C_P(i,j,k), K_0 - S(i,j,k)\} \end{aligned} \tag{1.100}$$

with initial data $P_A(n, j, k) = (K_0 - S(n, j, k))^+$, for j integer between 0 and n and k integer between $-Nn$ and Nn.

When $N = 3$, i.e. the second variable is allowed a seven-node branching at every time step, the HS procedure provides more accurate results than the one by Amin (the benchmark in the comparison is the European Merton value). In the application of their bivariate tree to the evaluation of American options, Hilliard and Schwartz provide a backward procedure of time complexity $O(n^3)$.

1.5.3 Dai *et al.*

Dai *et al.* [21] build on the HS procedure reducing the complexity to $O(n^{2.5})$ by dissolving the intermediate nodes introduced by the jumps on the tree in the nearest diffusion nodes, therefore providing a one-dimensional tree.

The theoretical setting is not different from the one described above. As in Hilliard and Schwartz [38], Dai *et al.* [21] start by considering the jump-diffusion process that represents the dynamics of the logreturn as divided into a diffusion process X_t and a jump process Y_t. Indicating as before with τ the time to maturity, n the number of steps and Δt the ratio $\frac{\tau}{n}$, X_τ is discretised in the standard CRR way, while Y_τ as the sum of n i.i.d. random variables Y_Δ, the jump amplitude h and probabilities of which are fixed, as in Hilliard and Schwartz, accordingly to the parameters of the jump distribution, with the matching of the first $2N$ moments (Dai *et al.* [21] use $N = 3$). Contrary to what happens in the Hilliard and Schwartz procedure, the up probability for the Brownian process is fixed exactly as π in the CRR process,

$$p = \frac{e^{r\Delta t} - d}{u - d}$$

where r is the risk-free rate of return, $u = e^{\sigma\sqrt{\Delta t}}$ and $d = e^{-\sigma\sqrt{\Delta t}}$ are the usual up and down Brownian moves, but this variation has negligible effects on the price for the data sets we have considered.

The authors then modified the procedure in the following way. Let us call S the value of the underlying the grid is centred around; this isn't necessarily

S_0, since one can anchor the grid so that one of the maturity nodes is precisely aligned with the value of the strike, in order to avoid nonlinearity error. At every step, only from each of the Brownian nodes (that is, the nodes corresponding to values for the underlying of the kind $Se^{j\sigma\sqrt{\Delta t}}$ for some j) depart $2(2N+1)$ nodes, since the jump allows for N up jumps, N down jumps, and a no jump node, and for each one of these there are both the Brownian up tick and the Brownian down tick. The pair of no jump nodes are the usual Brownian nodes we would obtain with the CRR grid. Each one of the other nodes, which represent the situations where a jump has occurred, are reabsorbed, using a trinomial branching, into three appropriate consecutive Brownian jumps in the following time step.

The individuation of the appropriate triple for a given jump node is determined in accordance with the parameters of the process: the mean of the logreturn $\mu = (r - \lambda\bar{j} - \frac{\sigma^2}{2})\Delta t$ and its variance $Var = \sigma^2\Delta t$. A given jump node X in the tree represents a value for the underlying of the type:

$$S(X) = Se^{j\sigma\sqrt{\Delta t}+kh},$$

where $k = \pm 1, \ldots, \pm N$, since the reabsorption of jump nodes happens before they can further branch. The node X is linked via a trinomial structure to three consecutive nodes (A, B, C) representing the underlying values $S(C) = Se^{(j'-2)\sigma\sqrt{\Delta t}}$, $S(B) = Se^{j'\sigma\sqrt{\Delta t}}$, $S(A) = Se^{(j'+2)\sigma\sqrt{\Delta t}}$ chosen such that the central node B has $S(X)$-log-price

$$\hat{\mu} = \log\left(\frac{S(B)}{S(X)}\right) = (j' - j)\sigma\sqrt{\Delta t} - kh$$

as close as possible to the mean μ.

Once we have identified, for a given node X, the nodes A, B, C it should be branching to, the probabilities p_A, p_B, p_C are found by imposing the matching of the mean and variance of the $S(X)$-log-price:

$$\begin{cases} p_A + p_B + p_C = 1 \\ p_A\alpha + p_B\beta + p_C\gamma = 0 \\ p_A\alpha^2 + p_B\beta^2 + p_C\gamma^2 = Var \end{cases}$$

where $\beta = \hat{\mu} - \mu$, $\alpha = \beta + 2\sigma\sqrt{\Delta t}$, $\gamma = \beta - 2\sigma\sqrt{\Delta t}$. The system above can be solved via Cramer's rule, giving the following formulas for p_A, p_B and p_C:

$$\begin{cases} p_A = \dfrac{(\beta\gamma + Var)(\gamma - \beta)}{(\beta - \alpha)(\gamma - \alpha)(\gamma - \beta)} \\ \\ p_B = \dfrac{(\alpha\gamma + Var)(\alpha - \gamma)}{(\beta - \alpha)(\gamma - \alpha)(\gamma - \beta)} \\ \\ p_C = \dfrac{(\beta\alpha + Var)(\beta - \alpha)}{(\beta - \alpha)(\gamma - \alpha)(\gamma - \beta)} \end{cases} \tag{1.101}$$

which are non negative (with sum 1), therefore valid as probabilities.

Since finding the appropriate triple may not be feasible considering only the Brownian nodes already provided by the CRR grid, the tree will be expanded in order to include, in every step, from the C node of the triple associated with the lowest jump of the lowest brownian node for the time step, up to the A node for the highest jump of the highest brownian node. The time interval Δt, the minimum jump amplitude h and the volatility σ are constant throughout the tree, and the $S(X)$-log-price $(j' - j)\sigma\sqrt{\Delta t} - kh$ only depends on the difference between the Brownian indexes ($j' - j$, which will necessarily be an odd integer) and the value k relative to the amplitude of the jump, which means that the number of nodes we will need to add to the tree can be computed forward at once. Counting the total number of nodes allows to quantify the computational cost of the procedure, which can be shown to be $O(n^{2.5})$.

The same trinomial structure can also be used in order to connect the original node at time 0, corresponding to the initial value of the underlying S_0, with three appropriate Brownian nodes at time Δt of the grid that is aligned with the strike.

These three nodes are obtained as follows: considering that they belong to timestep 1, while the strike K_0 needs to be aligned with the nodes at timestep n, we know that they must have an underlying of the kind $K_0 e^{j\sigma\sqrt{\Delta t}}$ where j is an odd integer when n is even, and vice versa.

Once fixed a value for the strike K_0, we consider, between all such nodes, the one whose S_0-log-price is the closest to μ: that node will serve as the central node B, and those immediately above and below as A and C. The transition probabilities between the initial value of the underlying and the three values represented by these A, B, and C are the solution of the system in (1.101), where $\beta = \log(\frac{S(B)}{S_0}) - \mu$ and α and γ are defined as we did above.

1.6 Establishing an appropriate cut

In order to improve computational efficiency of the Hilliard and Schwartz method, we propose two strategies: the first - illustrated in this Section - is establishing an appropriate cutting of the bivariate tree that prevents spending too much computational time on too thin tails, the second (drawing on Amin's interpretation of the jump amplitude as a multiple of the Brownian move) - in the following Section - is to translate the bivariate tree into an univariate tree.

In this Section, drawn from [33], we propose a $O(n \log n)$ and a $O(n^2 \log n)$ procedure respectively for the evaluation of European and American option prices.

The main idea behind this lies on the way a European option is evaluated: by taking the discounted expected value of the price at maturity.

More in detail, in order to find the proper truncation we start from the evaluation of the European call option as the discounted expected payoff at maturity, derived from the HS backward procedure, we focus on the jump probabilities relative to every ending node and analyse the error obtained by considering only a limited range of accumulated jumps. We consider the backward procedure truncated throughout the whole tree with the same limitations, and we obtain an upper estimation of the error. This also provides an upper estimation of the error we get if we apply the truncation to the forward computation of the jump probabilities of the terminal nodes. We show that an error lower than $\frac{1}{n}$ can be achieved with a number of steps proportional to $\log n$. In this way, we are able to construct a procedure of order $O(n \log n)$ for the European case. Finally, we move on to the American case. The latter procedure is of order $O(n^2 \log n)$.

As a second result we develop a univariate procedure with very different features from [21] and this new procedure improves slightly the complexity ($O(n^2)$ instead of $O(n^2 \log n)$).

The setting is that of Hilliard and Schwartz [38], where the possible prices at maturity are computed via a bivariate tree, and every node of the tree is labelled with both the price and the associated probability. We use the notation already introduced in 1.5.2.

Since the jump process and the diffusion one are supposed to be independent, it is more profitable to us to focus on the effects of the Brownian motion and the compound Poisson jumps separately. The probability of reaching node (n, j, k) is given by the product of the probability $P(j)$ of having Brownian total balance at maturity $-n + 2j$ (regardless of what happens to the jump process) and the probability $Q_N(k)$ of having jump total balance at maturity k (regardless of what happens to the Brownian motion). In other words, $Q_N(k)$ is the probability of reaching whichever of the nodes (n, j', k) for $j' \in \{0, \dots n\}$, or - that is the same - the probability for the discrete process Y_n to take value k.

The contributions of the nodes to the price of the option depend on both the value of the option that is associated to them, and the probability we have to reach them. We ask ourselves when it is possible to neglect part of the nodes of the tree.

In this section we see that we can appropriately find "jump levels" on the tree such that cutting the tree at those frontiers provides us with an $O(n \log n)$ number of nodes at maturity, which allows for a $O(n \log n)$ procedure for the evaluation of the European option prices and a $O(n^2 \log n)$ procedure for the evaluation of the American option prices.

First we prove theoretically the validity of the procedure in the case $N = 1$ with a precise estimation of the error in the European case, then we extend the same reasoning to $N = 2$ and at last to the arbitrary N case. This choice of presentation is due to the fact that the possibility of different amplitudes of jumps allows for multiple different combinations of reaching the same "jump level"at maturity, and analysing the $N = 1, 2$ steps allow for a better understanding of the mechanics.

We first focus on the call option, and afterwards treat the put option.

At last, we discuss the American case.

Obviously, depending on the derivative we choose, there are nodes which do not contribute to the price.

For a European call, Equation (1.97) implies that leaves (n, j, k) such that $S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk} \leq K_0$ have no effect on the price. This establishes a lower bound at maturity for the jumps for call options $\overline{l_C}$: in order to have

$$S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk} > K_0,$$

which can be written as $(-n+2j)\sigma\sqrt{\Delta t} + hk > \log \frac{K_0}{S_0}$, we ask for $-n\sigma\sqrt{\Delta t} + hk > \log \frac{K_0}{S_0}$, which gives

$$k > \frac{\log \frac{K_0}{S_0} + n\sigma\sqrt{\Delta t}}{h} = \overline{l_C}.$$

Similarly for a European put, Equation (1.98) implies that leaves (n, j, k) such that $S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk} \geq K_0$ have no effect on the price. This establishes an upper bound at maturity for the jumps for put options $\overline{u_P}$:

$$k < \frac{\log \frac{K_0}{S_0} - n\sigma\sqrt{\Delta t}}{h} = \overline{u_P}.$$

First approach to the problem

The first idea is a very simple one: since the relevance of the contribution of the single leaf (n, j, k) of the tree to the European option price is subjected to the probability $Q_N(k)$, we compute recursively the probability $Q_N(k)$ using the transition probabilities q_i. When $Q_N(k)$ is smaller than a fixed value, say $\varepsilon = 10^{-8}$, we deem the related price at maturity to be negligible, and ignore it in the computation of our expectation.

The implementation of this technique is straightforward, but it shall not escape our notice that the computational cost of the procedure cannot be less than $O(n^2)$, since this is already the cost of the recursive procedure we need for the evaluation of the $Q_N(k)$ cumulated probabilities. Unfortunately, we do not have at our disposal a formula for Q_N.

Moreover, there are several shortcomings of this strategy that make it - although numerically trustworthy - theoretically unfounded: first of all the price of a call option al level k also involves the calculation of a e^{hk}, which we should consider if we are to say that the contribution of a node can be neglected; secondly, ensuring that for a given $k > 0$ we have $Q_N(k) < \varepsilon$ does not guarantee that $Q_N(l) < \varepsilon$ for all $k \leq l \leq Nn$, so the cut may exclude not negligible (according to our allowed error ε) nodes; lastly, if we are to exclude the tails of the distribution from two given boundaries, we want to be sure that the whole tail, on both sides, is negligible, and not only the nodes taken singularly.

We are interested in the contribution to the derivative price of the values that lie on the nodes (n, j, k) for $|k| \geq \overline{k}$, for some $\overline{k} \in \mathbb{N}$.

Let us consider an European call option. Recall that its value V_E computed with the Hilliard and Schwartz method is:

$$V_E = e^{-r\tau} \sum_{k=-n\cdot N}^{n\cdot N} \sum_{j=0}^{n} (S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk} - K_0)^+ P(j) Q_N(k).$$

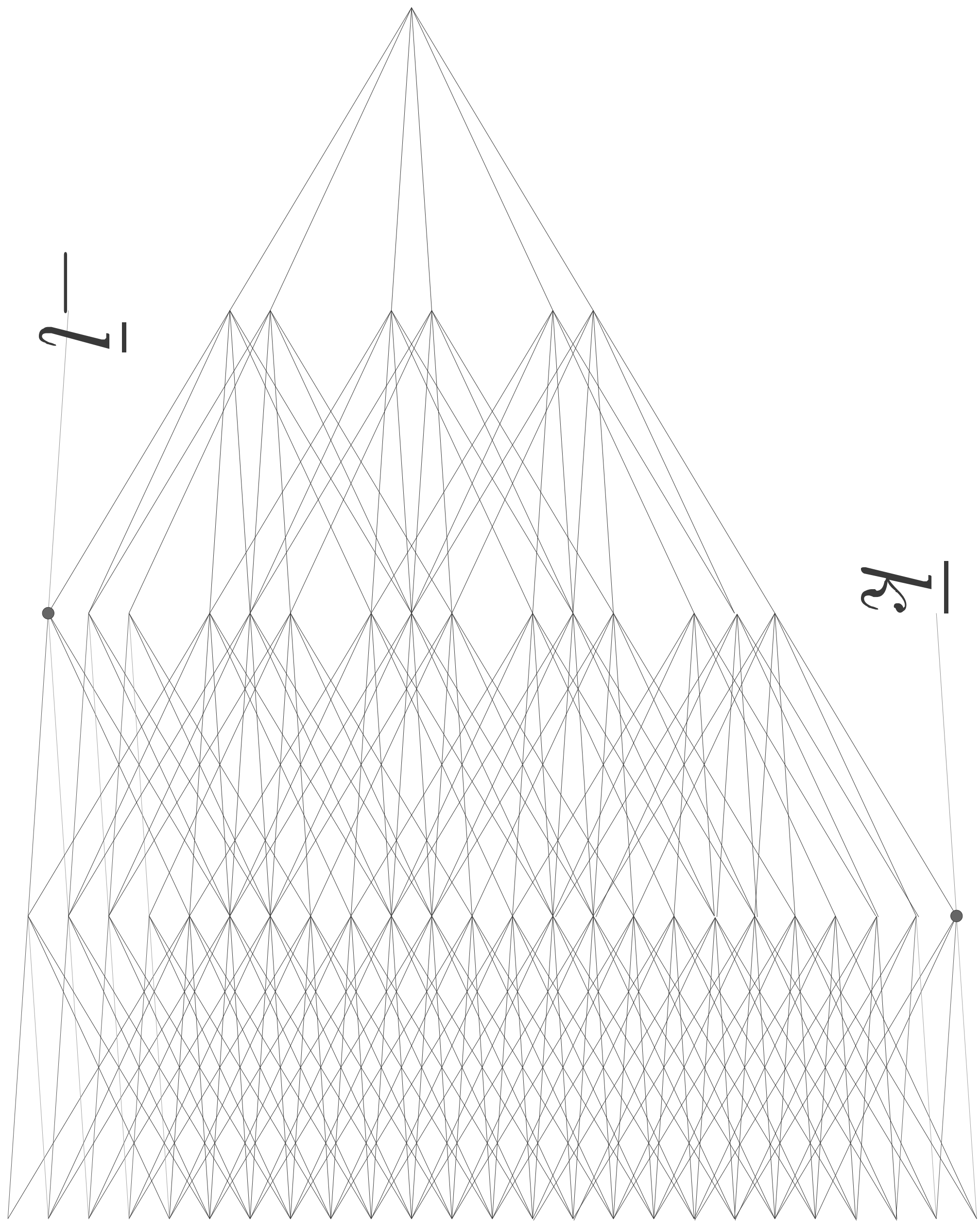
In order to reduce the computational complexity of the previous formula, we want to truncate the computation of V_E to some $k = -\overline{l}$ and $k = \overline{k}$, committing an error smaller than ε.

Fixed $\overline{k} > 0$ and $\overline{l} > 0$, the sum

$$V_E^T = e^{-r\tau} \sum_{k=-\overline{l}}^{\overline{k}} \sum_{j=0}^{n} (S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk} - K_0)^+ P(j) Q_N(k),$$

then, would appear as the perfect candidate for an approximation of the value V_E, for appropriately chosen $\overline{l}$ and $\overline{k}$. The problem with this is that,

Figure 1.2: The cutting of the tree



even if the sum over k has at most a number of terms proportional to $\log n$, the computational complexity of Q_N is $O(n^2)$. In order to reduce it, we need to substitute Q_N with a computationally cheaper approximation.

This springs a new idea: we need to determine an *a priori* upper estimation for $Q_N(k)$ at level $-Nn \leq k \leq Nn$, which can assure us that certain nodes and branches are negligible. We are going to prove that the probability $Q_N(k)$ is negligible for $|k| > a \log n + b$ (for constants a and b) .

We need to define the discrete function $\widetilde{Q}_N(k)$, which is constructed from $Q_N(k)$ by substitution of all $q_{\pm i}$ for $i = 1, \ldots, N$ with the maximum between q_i and q_{-i}.

Obviously, $Q_N(\pm k) \leq \widetilde{Q}_N(k) = \widetilde{Q}_N(-k)$ for all $-Nn \leq k \leq Nn$.

For brevity purposes, we will refer to $\widetilde{Q}_N(k)$ as to an "enlarged probability", even though it is not a probability measure, since the sum of all the values for k from $-Nn$ to Nn exceeds one.

We are going to prove that $\widetilde{Q}_N(k)$ is negligible for $|k| > a \log n + b$ (for some constants a and b depending on the derivative and on the dynamics of the underlying), and that this translates into the negligibility of some of the possible payoffs in the evaluation of the option, therefore providing a theoretical justification for the elimination of such values in the pricing procedure.

Let us call $\widehat{Q}_N(k)$ the probability of reaching level k computed recursively forwards without taking into account the nodes (i, j, l) with $l < -\bar{l}$ or $l > \bar{k}$. This means $\widehat{Q}_N(k)$ is the sum of the probabilities associated with the paths (regardless of the Brownian moves, that is, only with respect to the jump moves) that don't trespass the $\bar{k}$ and $-\bar{l}$ levels. For brevity, we will refer to the allowed region as $[-\bar{l}, \bar{k}]$, i. e. all the nodes (i, j, l) on the tree with $-\bar{l} \leq l \leq \bar{k}$.

Let us define

$$\widehat{V_E^T} = e^{-r\tau} \sum_{k=-\bar{l}}^{\bar{k}} \sum_{j=0}^{n} (S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk} - K_0)^+ P(j) \widehat{Q}_N(k). \tag{1.102}$$

The following is true:

$$\widehat{V_E^T} \leq V_E^T \leq V_E.$$

We will show that we can choose $\bar{k}$ and $\bar{l}$ such that computing $\widehat{V_E^T}$ is less computationally expensive than V_E and the error $V_E - \widehat{V_E^T}$ is less than an arbitrary ε.

If we focus, instead, on the European value obtained via backward procedure, we remark that substituting the HS value in a node of the tree with 0 is equivalent to nullify the value of all the trajectories that include that node.

Let us call $V^T(i,j,l)$ the value obtained with the following recursion formula:

$$V^T(i,j,k) = e^{-r\Delta t}\sum_{l=-N}^{N}(V^T(i+1,j+1,k+l)p + V^T(i+1,j,k+l)(1-p))q_l \tag{1.103}$$

with initial data $V^T(n,j,k) = (S(n,j,k) - K_0)^+$, for j integer between 0 and n and k integer such that $-\bar{l} \le k \le \bar{k}$, $V^T(n,j,k) = 0$ for j integer between 0 and n and k integer such that $-nN \le k \le -\bar{l}-1$ or $\bar{k}+1 \le k \le nN$, and imposing $V^T(i,j,k) = 0$ for $-iN \le k \le -\bar{l}-1$ and $\bar{k}+1 \le k \le iN$.

For the put options, in the same we define the value $P^T(i,j,k)$ obtained via backward procedure according to the following formula: $P^T(i,j,k) = e^{-r\Delta t}\sum_{l=-N}^{N}(P^T(i+1,j+1,k+l)p + P^T(i+1,j,k+l)(1-p))q_l$ if $k \in [-\bar{l},\bar{k}]$, 0 otherwise; with initial data $P^T(n,j,k) = (K_0 - S(n,j,k))^+$, for j integer between 0 and n and k integer such that $-\bar{l} \le k \le \bar{k}$, $P^T(n,j,k) = 0$ for j integer between 0 and n and k integer such that $-nN \le k \le -\bar{l}-1$ or $\bar{k}+1 \le k \le nN$, and the value $\widehat{P_E^T}$, defined as

$$\widehat{P_E^T} = e^{-r\tau}\sum_{k=-\bar{l}}^{\bar{k}}\sum_{j=0}^{n}(K_0 - S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk})^+ P(j)\widehat{Q}_N(k). \tag{1.104}$$

We want to show that the value $V^T(0,0,0)$ ($P^T(0,0,0)$ for the put case) obtained via the backward truncated procedure coincides with $\widehat{V_E^T}$ ($\widehat{P_E^T}$, respectively).

Lemma 1.6.1. *$V^T(0,0,0) = \widehat{V_E^T}$ and $P^T(0,0,0) = \widehat{P_E^T}$*

Proof. We will write the proof for the call case; for the put options the proof is analogous. It will help to write $\widehat{V_E^T}$ in the following way:

$$\widehat{V_E^T} = e^{-r\tau}\sum_{\substack{\text{paths that reach } \tau\\ \text{and do not trespass}}}\text{prob(path)}\cdot\text{value(path)}. \tag{1.105}$$

where prob(path) identifies the probability of a single path and value(path) the value of the option in the node at the end of the path.

The two expressions identify the same sum: every path that does not go out of the borders needs to end at a level $-\bar{l} \le k \le \bar{k}$, all the path that end in a node (n, j, k) share the same value for the option so if we collect in expression (1.105) all the addenda that end in the same node we obtain $(S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk} - K_0)^+ P(j)\widehat{Q}_N(k)$.

We will show that the $\widehat{V_E^T}$ as in (1.105) coincides with $V^T(0,0,0)$ for induction on the number of steps n.

Let us start with $n = 1$. Our tree has only one step, which means that the values at maturity of the option are given by the $2(2N+1)$ children of $(0,0,0)$. $\Delta t = \tau$. Let $0 \le \bar{l}, \bar{k} \le N$, that means that $(0,0,0)$ is surely in the allowed zone, while some of its children may be not. Since the value of the option on the nodes $(1, j, k)$ with $k \notin [-\bar{l}, \bar{k}]$ is 0, we can write:

$$
\begin{aligned}
V^T(0,0,0) &= e^{-r\tau} \sum_{l=-N}^{N} (V^T(1, j+1, l)p + V^T(1, j, l)(1-p))q_l = \\
&= e^{-r\tau} \sum_{l=-\bar{l}}^{\bar{k}} V^T(1, j+1, l)pq_l + V^T(1, j, l)(1-p)q_l = \\
&= e^{-r\tau} \sum_{\substack{\text{paths that reach } \tau \\ \text{and do not trespass}}} \text{prob(path)} \cdot \text{value(path)} = \widehat{V_E^T}
\end{aligned}
$$

where the last equality is due to the fact that in a single step the paths that trespass are those that end outside the boundary.

Let us now suppose the thesis is true for all trees in $n-1$ steps. Let us consider a tree of n steps. $\Delta t = \tau/n$. We focus on the first step and compute the value of the option in $(0,0,0)$, with the backward procedure: $V^T(0,0,0) = e^{-r\Delta t} \sum_{l=-N}^{N} (V^T(1,1,l)p + V^T(1,0,l)(1-p))q_l$.

If $l \notin [-\bar{l}, \bar{k}]$ $V^T(1,1,l) = V^T(1,0,l) = 0$. Otherwise, we consider the $n-1$ trees that start at $(1, j, l)$ with $j = 0, 1$ and $l \in [-\bar{l}, \bar{k}]$ and end at τ. On these smaller trees we apply induction and write that the values $V^T(1, j, l)$ as

$$
\begin{aligned}
&V^T(1, j, l) = \\
&= e^{-r\tau'} \sum_{\substack{\text{paths that go from } (1,j,l) \text{ to } \tau \\ \text{and do not trespass}}} \text{prob(path')} \cdot \text{value(path')}
\end{aligned}
$$

where we indicated with τ' the time interval $\tau' = \Delta t(n-1)$ and with path' the generic path from $(1, j, l)$ to maturity τ.

Therefore we can write

$$
\begin{aligned}
&V^T(0,0,0) = \\
&= e^{-r\Delta t} \sum_{\substack{l=-N \\ l\in[-\bar{l},\bar{k}]}}^{N} (V^T(1,1,l)p + V^T(1,0,l)(1-p))q_l \\
&= e^{-r\tau} \sum_{\substack{l=-N \\ l\in[-\bar{l},\bar{k}]}}^{N} \left(\sum_{\substack{\text{paths that go from } (1,1,l) \text{ to } \tau \\ \text{and do not trespass}}} \text{prob(path')} \cdot \text{value(path')} p q_l + \right. \\
&\left. \sum_{\substack{\text{paths that go from } (1,0,l) \text{ to } \tau \\ \text{and do not trespass}}} \text{prob(path')} \cdot \text{value(path')}(1-p) q_l \right) = \\
&= e^{-r\tau} \sum_{\substack{\text{paths that go from } (0,0,0) \text{ to } \tau \\ \text{and do not trespass}}} \text{prob(path)} \cdot \text{value(path)}
\end{aligned}
$$

where we used the fact that $\Delta t + \tau' = \tau$, and we considered that if a path that connects the node $(0,0,0)$ to a node at maturity τ (without trespassing) visits node $(1,0,l)$ and is afterwards identical to path', we will have value(path)=value(path') and prob(path) $= (1-p)q_l\cdot$ prob(path'), while if a path that connects the node $(0,0,0)$ to a node at maturity τ (without trespassing) visits node $(1,1,l)$ and is afterwards identical to path', we will have value(path) = value(path') and prob(path) $= pq_l\cdot$ prob(path').

$\square$

If we take $\widehat{V_E^T}$ as the option price, that is if we are truncating the tree at levels $\bar{k}$ and $-\bar{l}$, we are losing probability contributions in two different ways:

(a) neglecting the paths that would reach - at maturity - a node outside the allowed region, i.e. a node (n,j,k) with $k > \bar{k}$ or $k < -\bar{l}$, for any $0 \le j \le n$;

(b) neglecting the paths that, even though ending at maturity in a node inside the allowed region, have at some point before maturity trespassed at least one of the boundaries.

In order to show that the difference between $\widehat{V_E^T}$ and V_E is arbitrarily small, we need not only to establish $\overline{k}$ and $\overline{l}$ according to the probabilities of the paths that end above level $\overline{k}$ or below level $\overline{-l}$: we also need to understand the difference between $\widehat{Q}_N(k)$ and $Q_N(k)$, which is made of the probabilities of the paths that reach level k, with $-\overline{l} \leq k \leq \overline{k}$, at maturity having previously gone outside the $[-\overline{l}, \overline{k}]$ region.

Let us consider separately the two differences: $V_E - V_E^T$ and $V_E^T - \widehat{V_E^T}$.

In both of them appears the Brownian motion; first of all we establish an upper estimate for its contribution to these differences.

Proposition 1.6.2. *With the notation introduced above,*

$$e^{-r\tau} \sum_{j=0}^{n} e^{\sigma\sqrt{\Delta t}(-n+2j)} P(j) \leq e^{(\alpha-r)\tau}.$$

Proof. We can write

$$\begin{aligned}\sum_{j=0}^{n} \binom{n}{j} e^{\sigma\sqrt{\Delta t}(-n+2j)} p^j (1-p)^{n-j} &= \sum_{j=0}^{n} \binom{n}{j} (e^{\sigma\sqrt{\Delta t}} p)^j [e^{-\sigma\sqrt{\Delta t}}(1-p)]^{n-j} \\ &= [e^{\sigma\sqrt{\Delta t}} p + e^{-\sigma\sqrt{\Delta t}}(1-p)]^n.\end{aligned}$$

Since the greater the probability p, the higher is the value $e^{\sigma\sqrt{\Delta t}}p+e^{-\sigma\sqrt{\Delta t}}(1-p)$, the worst case scenario (since we would like to find an upper bound) is $p = 1$.

As we defined p as $\frac{1}{2}\left(1 + \frac{\alpha\sqrt{\Delta t}}{\sigma}\right)$, this translates into $\alpha\Delta t = \sigma\sqrt{\Delta t}$, therefore:

$$\begin{aligned}\sum_{j=0}^{n} \binom{n}{j} e^{\sigma\sqrt{\Delta t}(-n+2j)} p^j (1-p)^{n-j} &\leq (e^{\sigma\sqrt{\Delta t}})^n \\ &\leq (e^{\alpha\Delta t})^n = e^{\alpha\tau},\end{aligned}$$

hence the thesis. □

This means we have the following limitation for the difference $V_E - V_E^T$:

$$V_E - V_E^T = e^{-r\tau} \left[\sum_{k=-Nn}^{-\overline{l}-1} \sum_{j=0}^{n} (S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk} - K_0)^+ P(j) Q_N(k) \right.$$

$$
\begin{aligned}
&\qquad + \sum_{k=\overline{k}+1}^{Nn} \sum_{j=0}^{n} (S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk} - K_0)^+ P(j) Q_N(k) \Bigg] \\
&\le e^{-r\tau} \Bigg[\sum_{k=\overline{k}+1}^{Nn} S_0 e^{hk} \sum_{j=0}^{n} e^{\sigma\sqrt{\Delta t}(-n+2j)} P(j) Q_N(k) \\
&\qquad + \sum_{k=\overline{l}+1}^{Nn} S_0 e^{-hk} \sum_{j=0}^{n} e^{\sigma\sqrt{\Delta t}(-n+2j)} P(j) Q_N(-k) \Bigg] \\
&\le e^{-r\tau} S_0 \sum_{j=0}^{n} e^{\sigma\sqrt{\Delta t}(-n+2j)} P(j) \left(\sum_{k=\overline{k}+1}^{Nn} e^{hk} \widetilde{Q}_N(k) + \sum_{k=\overline{l}+1}^{Nn} e^{-hk} \widetilde{Q}_N(k) \right) \\
&\le e^{(\alpha-r)\tau} S_0 \left(\sum_{k=\overline{k}+1}^{Nn} e^{hk} \widetilde{Q}_N(k) + \sum_{k=\overline{l}+1}^{Nn} e^{-hk} \widetilde{Q}_N(k) \right) \qquad (1.106)
\end{aligned}
$$

with $\overline{k}, \overline{l} > 0$.

As for the difference $V_E^T - \widehat{V_E^T}$, for simplicity's sake, we consider separately the probability of reaching k having gone over the $\overline{k}$ level and the probability of reaching k having gone under the $-\overline{l}$ level; the sum of the two is obviously greater than the probability of reaching k having surpassed at least one of the two boundaries.

Let us define, for any $-\overline{l} \le k \le \overline{k}$, the value $Q_N^{\overline{k}}(k)$ of the probability of a net balance of k jumps at maturity while reaching at some point a net balance higher than $\overline{k}$ and the value $Q_{N\overline{l}}(k)$ of the probability of a net balance of k jumps at maturity while reaching at some point a net balance lower than $-\overline{l}$.

According to these definitions and by Proposition 1.6.2, the difference between V_E^T and $\widehat{V_E^T}$ becomes:

$$
\begin{aligned}
V_E^T - \widehat{V_E^T} &= e^{-r\tau} \sum_{k=-\overline{l}}^{\overline{k}} \sum_{j=0}^{n} (S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk} - K_0)^+ P(j)(Q_N(k) - \widehat{Q}_N(k)) \\
&\le e^{-r\tau} S_0 \sum_{j=0}^{n} e^{\sigma\sqrt{\Delta t}(-n+2j)} P(j) \sum_{k=-\overline{l}}^{\overline{k}} e^{hk} (Q_N(k) - \widehat{Q}_N(k)) \\
&\le e^{(\alpha-r)\tau} S_0 \sum_{k=-\overline{l}}^{\overline{k}} e^{hk} (Q_N^{\overline{k}}(k) + Q_{N\overline{l}}(k)). \qquad (1.107)
\end{aligned}
$$

Putting Equations (1.106) and (1.107) together, we obtain that

$$V_E - \widehat{V_E^T} \leq e^{(\alpha-r)\tau} S_0 \left(\sum_{k=\bar{k}+1}^{Nn} e^{hk} \widetilde{Q}_N(k) + \sum_{k=\bar{l}+1}^{Nn} e^{-hk} \widetilde{Q}_N(k) + \sum_{k=-\bar{l}}^{\bar{k}} e^{hk} (Q_N^{\bar{k}}(k) + Q_{N\bar{l}}(k)) \right). \tag{1.108}$$

To proceed any further, we will relate $Q_N^{\bar{k}}(k)$ and $Q_{N\bar{l}}(k)$ to our "enlarged probability" $\widetilde{Q}_N$. This relation is best treated separately for the $N = 1, 2, 3, \ldots$ cases, but before moving to the separate cases, we state a pair of results that are applicable to all of them.

Lemma 1.6.3. *If* $0 \leq x \leq \dfrac{n+1}{j}$ *for some* $j, n \in \mathbb{N}$, $j > 1$,

then $\displaystyle\sum_{i \geq n} \frac{x^i}{i!} \leq \frac{j}{j-1} \frac{x^n}{n!}$.

Proof. Let all the terms of a summation $\displaystyle\sum_{n=0}^{\infty} a_n$ be such that $a_{i+1} \leq \dfrac{1}{j} a_i$ $\forall i$ and $a_i \geq 0$. Then

$$\sum_{i=0}^{\infty} a_i \leq a_0 \sum_{i=0}^{\infty} \frac{1}{j^i} = \frac{j}{j-1} a_0.$$

In order to apply this to the sequence $a_i = \frac{x^i}{i!}$, we need to show that $\frac{x^{i+1}}{(i+1)!} \leq \frac{x^i}{ji!}$.
If $0 \leq x \leq \dfrac{n+1}{j}$ then we have

$$a_{i+1} = \frac{x^{i+1}}{(i+1)!} = \frac{x^i}{i!} \frac{x}{i+1} = a_i \frac{x}{i+1} \leq a_i \frac{n+1}{(i+1)j},$$

which gives

$$a_{i+1} \leq \frac{1}{j} a_i$$

for $i \geq n$.
Therefore,

$$\sum_{i \geq n} \frac{x^i}{i!} = \sum_{i=0}^{+\infty} \frac{x^{n+i}}{(n+i)!} \leq \frac{j}{j-1} \frac{x^n}{n!}.$$

$\square$

Lemma 1.6.4. *Given* $c > 0$, $n \in \mathbb{N} - \{0\}$

$$\log \frac{c^n}{n!} < n(\log c + 1) - n \log n < -n + ce$$

Proof. We can write $\log \frac{c^n}{n!}$ as $n \log c - \log n!$.

By the Stirling series of $\log n!$ we can express it as

$$\log n! = n \log n - n + \frac{\log(2\pi n)}{2} + \frac{1}{12n} - \frac{1}{360n^3} + \dots .$$

Recall that the error committed by truncating the series is of the same sign of the first term omitted, hence we have:

$$\log n! > n \log n - n + \frac{\log(2\pi n)}{2} > n \log n - n$$

Therefore

$$n \log c - \log n! < n \log c - n \log n + n = n(\log c + 1) - n \log n$$

which is the first inequality we wanted to prove. In order to prove the second inequality, we set $a = \log c + 1$ and consider the function $f(x) = xa - x \log x$. This is a concave function, therefore its graph lies below the tangent line in $x = e^a$. Since the derivative of f is $f'(x) = a - 1 - \log(x)$, $f(e^a) = ae^a - ae^a = 0$ and $f'(e^a) = a - 1 - a = -1$, the equation of the tangent line is $y = -x + e^a$ and we get the following inequality:

$$na - n \log n \leq -n + e^a = -n + e^{\log c + 1} = -n + ce,$$

hence the thesis. □

1.6.1 European call options

European call option, $N = 1$

When $N = 1$, the only possible values for the jump in a Δt period are $-h$, 0, h, with probabilities q_{-1}, q_0, q_1, which are the solutions of the following linear system:

$$\begin{pmatrix} 1 & 1 & 1 \\ -1 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} q_{-1} \\ q_0 \\ q_1 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{k_1}{h} \\ \frac{k_2}{h^2} \end{pmatrix}$$

where k_1 and k_2 are the first two cumulants of the compound Poisson distribution (cf. Equation (1.95)). Therefore:

$$
\begin{aligned}
q_{-1} &= \frac{\lambda\tau}{2n}\left(1-\frac{\gamma'}{h}\right) = \frac{c_{-1}}{n} \\
q_0 &= 1-\frac{\lambda\tau}{n} \\
q_1 &= \frac{\lambda\tau}{2n}\left(1+\frac{\gamma'}{h}\right) = \frac{c_1}{n}
\end{aligned}
$$

Defining $w = \max\{c_1, c_{-1}\}$, we have $q_1, q_{-1} \leq \frac{w}{n}$.

The probability $Q_1(k)$ of reaching a node (n, j, k) for some j is given by

$$
Q_1(k) = \sum_{l=0}^{\lfloor \frac{n-k}{2} \rfloor} C_{n,k+2l} C_{k+2l,l} q_1^{k+l} q_{-1}^{l} q_0^{n-k-2l}
$$

which is the sum of all the probabilities of the "jump paths" with $k+l$ up jumps and l down jumps, for l such that $k+2l \leq n$. In the previous Equation we have used the standard notation $C_{n,k}$ for the number of $k-$combinations for a set of n elements.

We also need to define the discrete function $\widetilde{Q}_1(k)$, which is constructed from $Q_1(k)$ by substitution of both q_{-1} and q_1 with the maximum of the two: $\frac{w}{n}$.

$$
\widetilde{Q}_1(k) = \sum_{l=0}^{\lfloor \frac{n-k}{2} \rfloor} C_{n,k+2l} C_{k+2l,l} \left(\frac{w}{n}\right)^{k+2l} q_0^{n-k-2l}.
$$

As we have already said for any N, $Q_1(\pm k) \leq \widetilde{Q}_1(k)$ for all $-n \leq k \leq n$.

From the preliminary discussion, basing on Equation (1.108) articulated in the $N=1$ case, for the difference $V_E - \widehat{V_E^T}$ the following inequality is true:

$$
\begin{aligned}
V_E - \widehat{V_E^T} \leq e^{(\alpha-r)\tau} S_0 \Bigg(& \sum_{k=\bar{k}+1}^{n} e^{hk} \widetilde{Q}_1(k) + \sum_{k=\bar{l}+1}^{n} e^{-hk} \widetilde{Q}_1(k) + \\
& + \sum_{k=-\bar{l}}^{\bar{k}} e^{hk} (Q_1^{\bar{k}}(k) + Q_{1\bar{l}}(k)) \Bigg) .
\end{aligned}
\tag{1.109}
$$

We can relate $Q_1^{\bar{k}}(k)$ and $Q_{1\bar{l}}(k)$ to our "enlarged probability" $\widetilde{Q}_1(t)$.

Lemma 1.6.5.

$$Q_1^{\overline{k}}(k) \leq \widetilde{Q}_1(2\overline{k} - k + 2)$$

$$Q_{1\bar{l}}(k) \leq \widetilde{Q}_1(2\bar{l} + k + 2)$$

for all

$$-\bar{l} \leq k \leq \overline{k}$$

Proof. Let us focus on the first inequality, that concerns the trespassing of the upper boundary $\overline{k}$.

By reflection principle (see [30]), for every path that reaches the $\overline{k}+1$ level at some point before maturity and ends at a level k there is a path that ends at level $2\overline{k} - k + 2$.

Our intent is to recover an upper estimate of $Q_1^{\overline{k}}(k)$ using the probability of the "reflected" paths.

We consider a single path that reaches the $\overline{k}+1$ level at some point before maturity and ends at a level k with $-\bar{l} \leq k \leq \overline{k}$, and we define its reflection as the path that behaves like the original path up until the first time the original path touches the $\overline{k}+1$ level, and afterwards has an $+1$ jump when the other has a -1 jump and viceversa. Time intervals with no jump for the original path are intervals where the reflection has no jump too. The reflection path will end up at $2\overline{k} - k + 2$.

The probabilities of both the original path and the reflection differ in that a number l of q_{-1} factors in the probability of the original path need to be substituted with l q_1 factors to obtain the probability of the reflection. Both probabilities are not greater then the value obtained by substituting all occurrences of q_1 and q_{-1} with $\frac{w}{n}$, and the sum over all paths reaching level $2\overline{k}+2-k$ of these modified probabilities is $\widetilde{Q}_1(2\overline{k}+2-k)$, therefore we can write

$$Q_1^{\overline{k}}(k) \leq \widetilde{Q}_1(2\overline{k} + 2 - k).$$

Similarly, we treat the second inequality: the number of paths that reach the $-\bar{l}-1$ level at some point before maturity and end at a level k with $-\bar{l} \leq k \leq \overline{k}$ is the same as the number of paths that end at level $-2\bar{l}-2-k$, and both the probability of the original path and that of its reflection with respect to the level $-\bar{l}-1$ are not greater than the modified ones, therefore it holds:

$$Q_{1\bar{l}}(k) \leq \widetilde{Q}_1(-2\bar{l} - 2 - k) = \widetilde{Q}_1(2\bar{l} + k + 2).$$

□

The previous Lemma, applied to Equation (1.107), allows us to write:

$$
\begin{aligned}
V_E^T - \widehat{V_E^T} &\leq e^{(\alpha-r)\tau} S_0 \sum_{k=-\bar{l}}^{\bar{k}} e^{hk} (\widetilde{Q}_1(2\bar{k}+2-k) + \widetilde{Q}_1(2\bar{l}+2+k)) \\
&\leq e^{(\alpha-r)\tau} S_0 \left(\sum_{s=\bar{k}+2}^{\min\{2\bar{k}+\bar{l},n\}} e^{h(2\bar{k}+2-s)} \widetilde{Q}_1(s) + \sum_{s=\bar{l}+2}^{\min\{2\bar{l}+\bar{k},n\}} e^{h(s-2\bar{l}-2)} \widetilde{Q}_1(s) \right) \qquad (1.110) \\
&\leq e^{(\alpha-r)\tau} S_0 e^{h\bar{k}} \left(\sum_{s=\bar{k}+2}^{\min\{2\bar{k}+\bar{l},n\}} \widetilde{Q}_1(s) + \sum_{s=\bar{l}+2}^{\min\{2\bar{l}+\bar{k},n\}} \widetilde{Q}_1(s) \right)
\end{aligned}
$$

therefore, if we take $\bar{l} = \bar{k}$ we can write Equation (1.109) as

$$
\begin{aligned}
V_E - \widehat{V_E^T} \leq e^{(\alpha-r)\tau} S_0 &\left(\sum_{k=\bar{k}+1}^{n} e^{hk} \widetilde{Q}_1(k) + \sum_{k=\bar{l}+1}^{n} e^{-hk} \widetilde{Q}_1(k) + \right. \\
&\left. + e^{h\bar{k}} \sum_{k=\bar{k}+2}^{n} \widetilde{Q}_1(k) + e^{h\bar{k}} \sum_{k=\bar{l}+2}^{n} \widetilde{Q}_1(k) \right) \\
\leq 2e^{(\alpha-r)\tau} S_0 &\left(\sum_{k=\bar{k}+1}^{n} (e^{hk} + e^{h\bar{k}}) \widetilde{Q}_1(k) \right). \qquad (1.111)
\end{aligned}
$$

We could use the previous result to determine numerically the largest integer $\bar{l} = \bar{k}$ such that the loss is inferior to an arbitrary ε with an $O(n^2)$ procedure.

The summation (1.111) can be computed with an $O(n^2)$ procedure: it takes the usual recursive $O(n^2)$ procedure to compute the function $\widetilde{Q}_1$; once we have $\widetilde{Q}_1(k)$, we only need to start computing the addends $(e^{hk} + e^{h\bar{k}})\widetilde{Q}_1(k)$ from the highest reachable level (which is n for $N = 1$) and update this level, decreasing, until we surpass $\frac{\varepsilon}{2e^{(\alpha-r)\tau} S_0}$, thus providing the level $\bar{k}$.

In order to get a theoretical bound for $\bar{k}$ and $\bar{l}$ such that $V_E - \widehat{V_E^T}$ is inferior to an arbitrary ε, we are going to prove that $\widetilde{Q}_1(k)$ is negligible for $|k| > a \log n + b$. We will use Lemmas 1.6.3 and 1.6.4.

Proposition 1.6.6. *For k integer, $2w - 1 \leq k \leq n$,*

$$
\widetilde{Q}_1(k) \leq 2e^w \frac{w^k}{k!}. \qquad (1.112)
$$

For $\overline{k}$ integer, $2w-1 \leq \overline{k} \leq n$:

$$\sum_{k=\overline{k}}^{n} \widetilde{Q}_1(k) \leq 4e^w \frac{w^{\overline{k}}}{\overline{k}!}. \tag{1.113}$$

For $\overline{k}$ integer, $2e^h w - 1 \leq \overline{k} \leq n$:

$$\sum_{k=\overline{k}}^{n} e^{hk}\widetilde{Q}_1(k) \leq 4e^w \frac{(e^h w)^{\overline{k}}}{\overline{k}!}. \tag{1.114}$$

and for $\overline{k}$ integer, $2w-1 \leq \overline{k} \leq n$:

$$\sum_{k=\overline{k}}^{n} e^{-hk}\widetilde{Q}_1(-k) \leq 4e^w \frac{(e^{-h} w)^{\overline{k}}}{\overline{k}!}. \tag{1.115}$$

Proof. Let us focus on the definition of $\widetilde{Q}_1$:

$$\widetilde{Q}_1(k) = \sum_{l=0}^{\left\lfloor \frac{n-k}{2} \right\rfloor} C_{n,k+2l} C_{k+2l,l} \left(\frac{w}{n}\right)^{k+2l} q_0^{n-k-2l}.$$

Since $q_0 < 1$ and

$$\begin{aligned}
\frac{C_{n,k+2l}C_{k+2l,l}}{n^{k+2l}} &= \binom{n}{k+2l}\binom{k+2l}{l}\frac{1}{n^{k+2l}} = \frac{n!}{(n-k-2l)!(k+2l)!}\frac{(k+2l)!}{(k+l)!l!}\frac{1}{n^{k+2l}} = \\
&= \frac{n!}{n^{k+2l}(n-k-2l)!(k+l)!l!} \leq \frac{1}{(k+l)!l!}
\end{aligned}$$

for $l = 0, \ldots, \left\lfloor \frac{n-k}{2} \right\rfloor$, we can write:

$$\begin{aligned}
\widetilde{Q}_1(k) &\leq \sum_{l=0}^{\left\lfloor \frac{n-k}{2} \right\rfloor} \frac{C_{n,k+2l}C_{k+2l,l}}{n^{k+2l}} w^{k+2l} \\
&\leq \sum_{l=0}^{\left\lfloor \frac{n-k}{2} \right\rfloor} \frac{w^{k+2l}}{(k+l)!l!} \\
&\leq \frac{w^k}{k!} + \frac{w^{k+1}w}{(k+1)!} + \frac{w^{k+2}w^2}{(k+2)!2!} + \ldots + \frac{w^{\left\lfloor \frac{n+k}{2} \right\rfloor} w^{\left\lfloor \frac{n-k}{2} \right\rfloor}}{\left\lfloor \frac{n+k}{2} \right\rfloor! \left\lfloor \frac{n-k}{2} \right\rfloor} \\
&\leq \left(\frac{w^k}{k!} + \frac{w^{k+1}}{(k+1)!} + \frac{w^{k+2}}{(k+2)!} + \ldots + \frac{w^{\left\lfloor \frac{n+k}{2} \right\rfloor}}{\left\lfloor \frac{n+k}{2} \right\rfloor!}\right)\left(1 + w + \frac{w^2}{2} + \ldots + \frac{w^{\left\lfloor \frac{n-k}{2} \right\rfloor}}{\left\lfloor \frac{n-k}{2} \right\rfloor!}\right)
\end{aligned}$$

$$\leq \sum_{i=k}^{\left\lfloor \frac{n+k}{2} \right\rfloor} \frac{w^i}{i!} e^w \leq 2e^w \frac{w^k}{k!}$$

by Lemma 1.6.3 with $j = 2$, for $0 \leq w \leq \frac{k+1}{2}$.

Applying again Lemma 1.6.3 to $\sum_{k=\overline{k}}^{n} 2e^w \frac{w^k}{k!}$ we obtain Equation (1.113).

A further application of Lemma 1.6.3 with $x = e^h w \leq \frac{\overline{k}+1}{2}$ gives us:

$$\sum_{k=\overline{k}}^{n} e^{hk} Q_1(k) \leq 2e^w \sum_{k=\overline{k}}^{n} \frac{(e^h w)^k}{k!} \leq 4e^w \frac{(e^h w)^{\overline{k}}}{\overline{k}!}.$$

for $\overline{k} \geq 2e^h w - 1$.

Similarly for $\overline{k} \geq 2w - 1$ we obtain inequality (1.115). □

The above results allows us to have a closed-form formula for an upper bound for the difference $V_E - \widehat{V_E^T}$.

Theorem 1.6.7. *Given $\varepsilon > 0$, taking*

$$\begin{aligned} g^- &= w(e^{-h+1} + 1) + (\alpha - r)\tau - 1 + \log(4S_0) + \log(2 + e^h w), \\ g^+ &= w(e^{h+1} + 1) + (\alpha - r)\tau - 1 + \log(4S_0) + \log(2 + e^{-h} w), \end{aligned} \tag{1.116}$$

for

$$\begin{aligned} \bar{l} &\geq \max\{-\log \varepsilon + g^-, 2e^h w - 3, 2w - 2\} \\ \overline{k} &\geq \max\{-\log \varepsilon + g^+, 2e^h w - 2\} \end{aligned} \tag{1.117}$$

we have

$$V_E - \widehat{V_E^T} < \varepsilon$$

Proof. Combining Equations (1.106) and (1.110) we can write:

$$V_E - \widehat{V_E^T} \leq e^{(\alpha-r)\tau} S_0 \left(\sum_{k=\overline{k}+1}^{n} e^{hk} \widetilde{Q}_1(k) + \sum_{k=\bar{l}+1}^{n} e^{-hk} \widetilde{Q}_1(k) + \right.$$
$$\left. + e^{h(2\overline{k}+2)} \sum_{k=\overline{k}+2}^{n} e^{-hk} \widetilde{Q}_1(k) + e^{h(-2\bar{l}-2)} \sum_{k=\bar{l}+2}^{n} e^{hk} \widetilde{Q}_1(k) \right).$$

By Proposition 1.6.6 we obtain:

$$V_E - \widehat{V_E^T} \leq e^{(\alpha-r)\tau} S_0 \left(4e^w \frac{(e^h w)^{\overline{k}+1}}{(\overline{k}+1)!} + 4e^w \frac{(e^{-h} w)^{\bar{l}+1}}{(\bar{l}+1)!} + \right.$$

$$+e^{h(2\bar{k}+2)} \cdot 4e^{w}\frac{(e^{-h}w)^{\bar{k}+2}}{(\bar{k}+2)!} + e^{h(-2\bar{l}-2)} \cdot 4e^{w}\frac{(e^{h}w)^{\bar{l}+2}}{(\bar{l}+2)!}\Bigg)$$

$$\leq 4e^{(\alpha-r)\tau+w}S_0\left(\frac{(e^{h}w)^{\bar{k}+1}}{(\bar{k}+1)!} + \frac{(e^{-h}w)^{\bar{l}+1}}{(\bar{l}+1)!} + e^{-2h}\frac{(e^{h}w)^{\bar{k}+2}}{(\bar{k}+2)!} + e^{2h}\frac{(e^{-h}w)^{\bar{l}+2}}{(\bar{l}+2)!}\right)$$

$$\leq 4e^{(\alpha-r)\tau+w}S_0\left(\left(1+\frac{e^{-h}w}{\bar{k}+2}\right)\frac{(e^{h}w)^{\bar{k}+1}}{(\bar{k}+1)!} + \left(1+\frac{e^{h}w}{\bar{l}+2}\right)\frac{(e^{-h}w)^{\bar{l}+1}}{(\bar{l}+1)!}\right)$$

$$\leq 4e^{(\alpha-r)\tau+w}S_0\left(\left(1+\frac{e^{-h}w}{2}\right)\frac{(e^{h}w)^{\bar{k}+1}}{(\bar{k}+1)!} + \left(1+\frac{e^{h}w}{2}\right)\frac{(e^{-h}w)^{\bar{l}+1}}{(\bar{l}+1)!}\right)$$

for $\bar{k} \geq 2e^{h}w-2$ and $\bar{l} \geq \max\{2e^{h}w-3, 2w-2\}$

Splitting the error ε between the upper and the lower tail, we ask:

$$4S_0e^{(\alpha-r)\tau+w}\left(1+\frac{e^{h}w}{2}\right)\frac{(e^{-h}w)^{\bar{l}+1}}{(\bar{l}+1)!} < \frac{\varepsilon}{2} \tag{1.118}$$

$$4S_0e^{(\alpha-r)\tau+w}\left(1+\frac{e^{-h}w}{2}\right)\frac{(e^{h}w)^{\bar{k}+1}}{(\bar{k}+1)!} < \frac{\varepsilon}{2} \tag{1.119}$$

Taking $g = 4S_0e^{(\alpha-r)\tau+w}$, we obtain that Equation (1.118) is equivalent to

$$\frac{(e^{-h}w)^{\bar{l}+1}}{(\bar{l}+1)!} < \frac{\varepsilon}{(2+e^{h}w)g},$$

which is guaranteed by Lemma 1.6.4 for

$$\bar{l} > -\log\varepsilon + \log[(2+e^{h}w)g] + e^{-h+1}w - 1,$$

while Equation (1.119) is equivalent to

$$\frac{(e^{h}w)^{\bar{k}+1}}{(\bar{k}+1)!} < \frac{\varepsilon}{(2+e^{-h}w)g}$$

which is guaranteed by Lemma 1.6.4 for

$$\bar{k} \geq -\log\varepsilon + \log[(2+e^{-h}w)g] + e^{h+1}w - 1.$$

Taking into account all the previous conditions, for

$$\bar{l} \geq \max\{-\log\varepsilon + w(e^{-h+1}+1) + (\alpha-r)\tau - 1 + \log(4S_0) + \log(2+e^{h}w), 2w-2, 2e^{h}w-3\}$$

and

$$\overline{k} \geq \max\{-\log\varepsilon + w(e^{h+1}+1) + (\alpha - r)\tau - 1 + \log(4S_0) + \log(2 + e^{-h}w), 2e^h w - 2\}$$

we have the thesis. □

If for some reasons we wish for a symmetric cut, since h is positive, hence $e^{-h} < 1 < e^h$, we can take the more stringent condition:

$$\overline{l} = \overline{k} \geq \max\{-\log\varepsilon + w(e^{h+1}+1) + (\alpha - r)\tau - 1 + \log(4S_0) + \log(2 + e^h w), 2e^h w - 2\}.$$

From the previous theorem comes the main result, which guarantees that for appropriately chosen $\overline{k}$ and $\overline{l}$, the value V_E of the European call option can be approximated by

$$\widehat{V_E^T} = e^{-r\tau} \sum_{k=-\overline{l}}^{\overline{k}} \sum_{j=0}^{n} (S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk} - K_0)^+ P(j) \widehat{Q}_1(k)$$

with an efficiency gain.

Theorem 1.6.8. *Given* $\varepsilon = \frac{1}{n} > 0$*, and* $\overline{k}$ *and* $\overline{l}$ *the smallest integers as in Theorem 1.6.7, the proposed procedure for* $\widehat{V_E^T}$ *converges to the HS price and its computational complexity is* $O(n \log n)$*.*

Proof. By taking $\overline{k}$ and $\overline{l}$ the smallest integers as in Theorem 1.6.7, the error is given by

$$\left| V_E - e^{-r\tau} \sum_{k=-\overline{l}}^{\overline{k}} \sum_{j=0}^{n} (S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk} - K_0)^+ P(j) \widehat{Q}_1(k) \right| < \varepsilon$$

for our choice of $\overline{l}$ and $\overline{k}$.

The sum over k has at most a number of terms proportional to $\log n$, the approximate probability distribution $\widehat{Q}_1$ is also computed with an $O(n \log n)$ procedure, therefore the computational complexity of the whole procedure is $O(n \log n)$. □

The results we have seen above can also be used in dealing with the truncated backward procedure, which will prove to be convenient for the extension of our reasoning to the American case.

The value $V^T(0,0,0)$ obtained via the backward truncated procedure coincides with $\widehat{V_E^T}$, hence the Theorem 1.6.7 also provides an upper bound for the difference between $V^T(0,0,0)$ and $V(0,0,0)$.

Theorem 1.6.9. *Given* $\varepsilon = \frac{1}{n} > 0$*, and* $\overline{k}$ *and* $\overline{l}$ *the smallest integers as in Theorem 1.6.7, the backward procedure described above for* $V^T(0,0,0)$ *converges to the HS price and its computational complexity is* $O(n^2 \log n)$*.*

Proof. By taking $\overline{k}$ and $\overline{l}$ the smallest integers as in Theorem 1.6.7, the error is

$$\left|V(0,0,0) - V^T(0,0,0)\right| < \varepsilon.$$

The number of nodes at maturity, depending on $\overline{l}$ and $\overline{k}$, is at most proportional to $n \log n$, therefore the computational complexity of the procedure is $O(n^2 \log n)$. □

European call option, $N = 2$

This part is dedicated to the results for the $N = 2$ case, analogous to those seen above for $N = 1$. Studying this case allows to understand the problems that arise when dealing with different possibilities for the amplitude of the jump in a single step. Afterwards, we will be able to transfer the same mechanics to the arbitrary N situation.

The probabilities q_k associated with jumps of amplitude kh for $k = 0, \pm 1, \pm 2$ are obtained by Hilliard and Schwartz [38] from the solution of a linear system (the formulas are reported in the Appendix). As in case $N = 1$, we will use the fact that $q_0 < 1$ and that the q_k's are inversely proportional to n. We will denote:

c_k the constant $q_k \cdot n$ for $k = \pm 1, \pm 2$, such that $q_k = \frac{c_k}{n}$;

w_k the maximum between c_{+k} and c_{-k} for $k = 1, 2$;

$W = w_1^2 + w_2$;

$Q_2(k)$ the probability of reaching at maturity level $-2n \leq k \leq 2n$ with the jumps, i.e. of reaching a node (n, j, k) for some $j = 0, \dots, n$

$\widetilde{Q}_2(k)$ the function obtained from $Q_2(k)$ by substituting both q_1 and q_{-1} with their maximum $\frac{w_1}{n}$, and q_2 and q_{-2} with their maximum $\frac{w_2}{n}$;

$\widehat{Q}_2(k)$ the probability of reaching jump level k but excluding the branches outside the allowed zone $[-\overline{l}, \overline{k}]$;

$Q_2^{\overline{k}}(k)$ the probability of reaching jump level k at maturity crossing the border $\overline{k}$ at least once;

$Q_{2\overline{l}}(k)$ the probability of reaching jump level k at maturity crossing the border $-\overline{l}$ at least once.

As we have already remarked in the general case, $Q_2(\pm k) \leq \widetilde{Q}_2(k)$ for all $-2n \leq k \leq 2n$.

We underline that q_k, c_k, etc. are generally not the same as in the previous section, but we are using the same notation, since they will only be used in separate settings and there will be no risk of confusing the two sets.

For the difference $V_E - \widehat{V_E^T}$, articulated in the $N = 2$ case, the following inequality (due to Equation (1.108)) is true:

$$V_E - \widehat{V_E^T} \leq e^{(\alpha-r)\tau} S_0 \left(\sum_{k=\overline{k}+1}^{2n} e^{hk} \widetilde{Q}_2(k) + \sum_{k=\overline{l}+1}^{2n} e^{-hk} \widetilde{Q}_2(k) \right. \\ \left. \sum_{k=-\overline{l}}^{\overline{k}} e^{hk} (Q_2^{\overline{k}}(k) + Q_{2\overline{l}}(k)) \right). \tag{1.120}$$

We relate $Q_2^{\overline{k}}$ and $Q_{2\overline{l}}$ to $\widetilde{Q}_2$ via the following lemma.

Lemma 1.6.10.

$$Q_2^{\overline{k}}(k) \leq \widetilde{Q}_2(2\overline{k} - k + 2) + \widetilde{Q}_2(2\overline{k} - k + 4)$$

$$Q_{2\overline{l}}(k) \leq \widetilde{Q}_2(2\overline{l} + k + 2) + \widetilde{Q}_2(2\overline{l} + k + 4)$$

for all

$$-\overline{l} \leq k \leq \overline{k}$$

Proof. In order to prove the first inequality, let us consider a path which steps outside the upper boundary $\overline{k}$ but re-enters the allowed region reaching level $-\overline{l} \leq k \leq \overline{k}$ at maturity. We define its reflection as the path that is exactly the same as the original one up until the first moment this crosses the upper boundary, and afterwards it moves in the exact opposite way. Since $N = 2$, that is we are dealing with the possibility of ± 1 and ± 2 jumps, the first time a path trespasses the upper boundary it can reach level $\overline{k} + 1$ or level $\overline{k} + 2$; depending on this, its reflection will end up at level $2\overline{k} - k + 2$ or $2\overline{k} - k + 4$.

If the original path has a certain probability $q_2^a q_1^b q_{-2}^c q_{-1}^d q_0^{n-a-b-c-d}$, both the original path's and its reflection's probability will be less or equal then $\left(\frac{w_2}{n}\right)^{a+c} \left(\frac{w_1}{n}\right)^{b+d} q_0^{n-a-b-c-d}$, since reflecting it we exchange some of the $+2$ with -2 moves and vice versa, and $+1$ with -1 moves and vice versa.

Knowing that for each path which reaches k surpassing $\overline{k}$ there is one who reaches level $2\overline{k} - k + 2$ or $2\overline{k} - k + 4$, we have

$$Q_2^{\overline{k}}(k) \leq \widetilde{Q}_2(2\overline{k} - k + 2) + \widetilde{Q}_2(2\overline{k} - k + 4).$$

The same argument stands for $Q_{2\bar{l}}(k)$, where the reflections can reach either level $-2\bar{l}-k-2$ or $-2\bar{l}-k-4$.

$\square$

Lemma 1.6.10 allows us to modify Equation (1.120) into

$$V_E - \widehat{V_E^T} \leq e^{(\alpha-r)\tau} S_0 \left(\sum_{k=\bar{k}+1}^{2n} e^{hk} \widetilde{Q}_2(k) + \sum_{k=\bar{l}+1}^{2n} e^{-hk} \widetilde{Q}_2(k) + \right.$$
$$\left. + \sum_{k=-\bar{l}}^{\bar{k}} e^{hk} (\widetilde{Q}_2(2\bar{k}-k+2) + \widetilde{Q}_2(2\bar{k}-k+4) + \widetilde{Q}_2(2\bar{l}+k+2) + \widetilde{Q}_2(2\bar{l}+k+4)) \right) \tag{1.121}$$

$$\leq e^{(\alpha-r)\tau} S_0 \left(\sum_{k=\bar{k}+1}^{2n} e^{hk} \widetilde{Q}_2(k) + \sum_{k=\bar{l}+1}^{2n} e^{-hk} \widetilde{Q}_2(k) + \right.$$
$$+ e^{h\bar{k}} \sum_{k=-\bar{l}}^{\bar{k}} (\widetilde{Q}_2(2\bar{k}-k+2) + \widetilde{Q}_2(2\bar{k}-k+4)) +$$
$$\left. + e^{h\bar{k}} \sum_{k=-\bar{l}}^{\bar{k}} (\widetilde{Q}_2(2\bar{l}+k+2) + \widetilde{Q}_2(2\bar{l}+k+4)) \right)$$
$$\leq e^{(\alpha-r)\tau} S_0 \left(\sum_{k=\bar{k}+1}^{2n} e^{hk} \widetilde{Q}_2(k) + \sum_{k=\bar{l}+1}^{2n} e^{-hk} \widetilde{Q}_2(k) + \right.$$
$$\left. + e^{h\bar{k}} \sum_{s=\bar{k}+2}^{\min\{2\bar{k}+\bar{l}+2,2n\}} (\widetilde{Q}_2(s) + \widetilde{Q}_2(s+2)) + e^{h\bar{k}} \sum_{s=\bar{l}+2}^{\min\{2\bar{l}+\bar{k}+2,2n\}} (\widetilde{Q}_2(s) + \widetilde{Q}_2(s+2)) \right)$$
$$\leq e^{(\alpha-r)\tau} S_0 \left(\sum_{k=\bar{k}+1}^{2n} (e^{hk} + 2e^{h\bar{k}}) \widetilde{Q}_2(k) + \sum_{k=\bar{l}+1}^{2n} (e^{-hk} + 2e^{h\bar{k}}) \widetilde{Q}_2(k) \right) \tag{1.122}$$

As in the $N = 1$ case, we can compute Equation (1.122) with a $O(n^2)$ procedure, thus determining numerically the largest integers $\bar{l}$ and $\bar{k}$ such that the loss is inferior to an arbitrary ε,[3] but we can deepen our investigation and find theoretical closed-form formulas for $\bar{l}$ and $\bar{k}$.

[3] In the numerical computations for this work, we have considered the following less refined conditions: $\sum_{k=\bar{k}+1}^{2n} 3e^{hk}\widetilde{Q}_N(k)$ and $\sum_{k=\bar{l}+1}^{2n} (2e^{h\bar{k}}+1)\widetilde{Q}_N(k)$ We proceeded in the following way. Given ε, we considered $\eta = \frac{\varepsilon}{2e^{(\alpha-r)\tau}S_0}$ and we computed, starting from $i = Nn$, the

Proposition 1.6.11. *For k integer, $2W - 1 \le k \le 2n$,*

$$\widetilde{Q}_2(2k) \le 2(1 + w_1^2)e^W \frac{W^k}{k!} \tag{1.123}$$

$$\widetilde{Q}_2(2k+1) \le 2w_1(1 + W)e^W \frac{W^k}{k!} \tag{1.124}$$

thus for k integer, $2\lceil 2W - 1\rceil \le k \le 2n$,

$$\widetilde{Q}_2(k) \le 2ce^W \frac{W^{\lfloor \frac{k}{2} \rfloor}}{\lfloor \frac{k}{2} \rfloor !} \tag{1.125}$$

where $c = \max\{1 + w_1^2, w_1(1 + W)\}$.

For $\overline{k}$ integer, $2\lceil 2W - 1\rceil \le \overline{k} \le 2n$

$$\sum_{k=\overline{k}}^{2n} \widetilde{Q}_2(k) \le 4Ce^W \frac{W^{\left\lfloor \frac{\overline{k}}{2} \right\rfloor}}{\left\lfloor \frac{\overline{k}}{2} \right\rfloor !} \tag{1.126}$$

where $C = 1 + w_1^2 + w_1(1 + W)$.

For $\overline{k}$ integer, $2\lceil 2e^{2h}W - 1\rceil \le \overline{k} \le 2n$

$$\sum_{k=\overline{k}}^{2n} e^{hk}\widetilde{Q}_2(k) \le 4C^+e^W \frac{(We^{2h})^{\left\lfloor \frac{\overline{k}}{2} \right\rfloor}}{\left\lfloor \frac{\overline{k}}{2} \right\rfloor !} \tag{1.127}$$

where $C^+ = 1 + w_1^2 + e^h w_1(1 + W)$.

For $\overline{k}$ integer, $2\lceil 2W - 1\rceil \le \overline{k} \le 2n$:

$$\sum_{k=\overline{k}}^{2n} e^{-hk}\widetilde{Q}_2(-k) \le 4C^-e^W \frac{(We^{-2h})^{\left\lfloor \frac{\overline{k}}{2} \right\rfloor}}{\left\lfloor \frac{\overline{k}}{2} \right\rfloor !} \tag{1.128}$$

where $C^- = 1 + w_1^2 + e^{-h} w_1(1 + W)$.

Proof. In order to write an explicit upper bound for the probability of reaching a certain level in the $N = 2$ setting, for clarity purposes we distinguish between the even and the odd levels (which we indicate by $2k$ and $2k+1$ respectively), and for each of them we consider the possibility of achieving an even or an odd value for the total of the down moves, the "negative total".

---

sums $S_i^e = 3\sum_{k=i}^{2n} e^{hk}\widetilde{Q}_N(k)$ and $S_i = \sum_{k=i}^{2n} \widetilde{Q}_N(k)$. While $S_i^e < \eta$, we keep decreasing i. The first i we encounter such that $S_i^e \ge \eta$ is our $\overline{k}$. While $S_i < \eta$, we keep decreasing i. The first i we encounter such that $S_i < \frac{\eta}{2e^{h\overline{k}}+1}$ is our $\overline{l}$.

Let $k, z \geq 0$ be integers. We focus on the case of ending on an even level $2k$.

Let the negative total be $-2z$. If it is not entirely due to -2 moves, then there is a residual (that must be an even number, $2i$) that needs to be covered with -1 moves. In the same way, if the positive total $2k+2z$ is not entirely due to $k+z$ moves of the kind $+2$ there is a residual (that must be an even number, $2l$) that needs to be covered with $+1$ moves.

Once the quantities of $+1$, $+2$, -1, -2 moves are fixed, the number of moves where a jump does not occur is automatically determined. But the same number of $+1$, $+2$, -1, -2 jumps can be obtained with many paths, differing in the order of the jumps. To consider all possible permutations of a multiset of n elements: a jumps $+2$, b jumps $+1$, c jumps -2, d jumps -1 and the rest of no-jump moves, we need to use the multinomial coefficient

$$B(n,a,b,c,d) := \binom{n}{n-a-b-c-d, a, b, c, d} = \frac{n!}{(n-a-b-c-d)!a!b!c!d!}.$$

Notation $B(n,a,b,c,d)$ is not standard, we introduce it for brevity.

Considering all combinations of reaching a negative total of $-2z$ with a mixture of -1 and -2 moves, and at the same time adding up to $2k+2z$ with a mixture of $+2$ and $+1$ moves, in at most n steps, we obtain that the probability of ending at level $2k$ with a negative total of $-2z$ is

$$\sum_{l=0}^{\overline{l_E}} \sum_{i=0}^{\overline{i_E}} \frac{n!}{(n-k-l-2z-i)!(k+z-l)!(2l)!(z-i)!(2i)!} q_2^{k+z-l} q_1^{2l} q_{-2}^{z-i} q_{-1}^{2i} q_0^{n-k-l-2z-i} =$$
$$= \sum_{l=0}^{\overline{l_E}} \sum_{i=0}^{\overline{i_E}} B(n, k+z-l, 2l, z-i, 2i) q_2^{k+z-l} q_1^{2l} q_{-2}^{z-i} q_{-1}^{2i} q_0^{n-k-l-2z-i}$$

with $\overline{l_E} = \min\{k+z, n-k-2z\}$ and $\overline{i_E} = \min\{z, n-k-2z-l\}$.

Likewise the probability of ending at level $2k$ with a negative total of $-2z-1$ is

$$\sum_{l=0}^{\overline{l_O}} \sum_{i=0}^{\overline{i_O}} B(n, k+z-l, 2l+1, z-i, 2i+1) q_2^{k+z-l} q_1^{2l+1} q_{-2}^{z-i} q_{-1}^{2i+1} q_0^{n-k-l-2z-i-2}$$

with $\overline{l_O} = \min\{k+z, n-k-2z-2\}$ and $\overline{i_O} = \min\{z, n-k-2z-l-2\}$; here we used that both the positive and the negative totals are odd, therefore there must be an odd number of -1 and $+1$ moves.

The possible values for the negative total depends of course on the number of steps and on the level k: z can be at most $\left\lfloor\frac{n-k}{2}\right\rfloor$ in the former case, and $\left\lfloor\frac{n-k-2}{2}\right\rfloor$ in the latter.

Thus:

$$\begin{aligned}Q_2(2k) = &\sum_{z=0}^{\left\lfloor\frac{n-k}{2}\right\rfloor}\sum_{l=0}^{\overline{l_E}}\sum_{i=0}^{\overline{i_E}} B(n,k+z-l,2l,z-i,2i)q_2^{k+z-l}q_1^{2l}q_{-2}^{z-i}q_{-1}^{2i}q_0^{n-k-l-2z-i}+\\ +&\sum_{z=0}^{\left\lfloor\frac{n-k-2}{2}\right\rfloor}\sum_{l=0}^{\overline{l_O}}\sum_{i=0}^{\overline{i_O}} B(n,k+z-l,2l+1,z-i,2i+1)q_2^{k+z-l}q_1^{2l+1}q_{-2}^{z-i}q_{-1}^{2i+1}q_0^{n-k-l-2z-i-2}.\end{aligned} \tag{1.129}$$

We proceed similarly for level $2k+1$:

$$\begin{aligned}&Q_2(2k+1) =\\ &= \sum_{z=0}^{\left\lfloor\frac{n-k-1}{2}\right\rfloor}\sum_{l=0}^{\overline{l'_E}}\sum_{i=0}^{\overline{i'_E}} B(n,k+z-l,2l+1,z-i,2i)q_2^{k+z-l}q_1^{2l+1}q_{-2}^{z-i}q_{-1}^{2i}q_0^{n-k-l-2z-i-1}+\\ &+ \sum_{z=0}^{\left\lfloor\frac{n-k-2}{2}\right\rfloor}\sum_{l=0}^{\overline{l'_O}}\sum_{i=0}^{\overline{i_O}} B(n,k+z+1-l,2l,z-i,2i+1)q_2^{k+z+1-l}q_1^{2l}q_{-2}^{z-i}q_{-1}^{2i+1}q_0^{n-k-l-2z-i-2}\end{aligned} \tag{1.130}$$

with $\overline{l'_E} = \min\{k+z, n-k-2z-1\}$, $\overline{l'_O} = \min\{k+z+1, n-k-2z-2\}$ and $\overline{i'_E} = \min\{z, n-k-2z-l-1\}$.

From Equations (1.129)-(1.130) we can write

$$\begin{aligned}&\widetilde{Q}_2(2k) =\\ &= \sum_{z=0}^{\left\lfloor\frac{n-k}{2}\right\rfloor}\sum_{l=0}^{\overline{l_E}}\sum_{i=0}^{\overline{i_E}} B(n,k+z-l,2l,z-i,2i)\left(\frac{w_2}{n}\right)^{k+2z-l-i}\left(\frac{w_1}{n}\right)^{2l+2i}q_0^{n-k-l-2z-i}\\ &+ \sum_{z=0}^{\left\lfloor\frac{n-k-2}{2}\right\rfloor}\sum_{l=0}^{\overline{l_O}}\sum_{i=0}^{\overline{i_O}} B(n,k+z-l,2l+1,z-i,2i+1)\left(\frac{w_2}{n}\right)^{k+2z-l-i}\left(\frac{w_1}{n}\right)^{2l+2i+2}q_0^{n-k-l-2z-i-2}\\ &\widetilde{Q}_2(2k+1) =\\ &= \sum_{z=0}^{\left\lfloor\frac{n-k-1}{2}\right\rfloor}\sum_{l=0}^{\overline{l'_E}}\sum_{i=0}^{\overline{i'_E}} B(n,k+z-l,2l+1,z-i,2i)\left(\frac{w_2}{n}\right)^{k+2z-l-i}\left(\frac{w_1}{n}\right)^{2l+2i+1}q_0^{n-k-l-2z-i-1}\end{aligned}$$

$$+\sum_{z=0}^{\lfloor\frac{n-k-2}{2}\rfloor}\sum_{l=0}^{\overline{l'_O}}\sum_{i=0}^{\overline{i_O}} B(n,k+z+1-l,2l,z-i,2i+1)\left(\frac{w_2}{n}\right)^{k+2z+1-l-i}\left(\frac{w_1}{n}\right)^{2l+2i+1} q_0^{n-k-l-2z-i-2}.$$

Let us start with $\widetilde{Q}_2(2k)$.

First of all, $q_0<1$. Then we use the fact that $\frac{n!}{(n-t)!n^t}\leq 1$ for any t integer $0\leq t\leq n$, which implies:

$$\frac{B(n,a,b,c,d)}{n^{a-b-c-d}}\leq\frac{1}{a!b!c!d!}.$$

$$\begin{aligned}
&\widetilde{Q}_2(2k)<\\
&<\sum_{z=0}^{\lfloor\frac{n-k}{2}\rfloor}\sum_{l=0}^{\overline{l_E}}\sum_{i=0}^{\overline{i_E}} B(n,k+z-l,2l,z-i,2i)\left(\frac{w_2}{n}\right)^{k+2z-l-i}\left(\frac{w_1}{n}\right)^{2l+2i}+\\
&\quad+\sum_{z=0}^{\lfloor\frac{n-k-2}{2}\rfloor}\sum_{l=0}^{\overline{l_O}}\sum_{i=0}^{\overline{i_O}} B(n,k+z-l,2l+1,z-i,2i+1)\left(\frac{w_2}{n}\right)^{k+2z-l-i}\left(\frac{w_1}{n}\right)^{2l+2i+2}\\
&<\sum_{z=0}^{\lfloor\frac{n-k}{2}\rfloor}\sum_{l=0}^{\overline{l_E}}\sum_{i=0}^{\overline{i_E}}\frac{w_2^{k+2z-l-i}w_1^{2l+2i}}{(k+z-l)!(2l)!(2i)!(z-i)!}\\
&\quad+\sum_{z=0}^{\lfloor\frac{n-k}{2}\rfloor-1}\sum_{l=0}^{\overline{l_O}}\sum_{i=0}^{\overline{i_O}}\frac{w_2^{k+z-l}w_1^{2l+1}w_2^{z-i}w_1^{2i+1}}{(k+z-l)!(2l+1)!(2i+1)!(z-i)!}\\
&=\sum_{z=0}^{\lfloor\frac{n-k}{2}\rfloor}\sum_{l=0}^{\overline{l_E}}\frac{w_2^{k+z-l}w_1^{2l}}{(k-l+z)!(2l)!}\sum_{i=0}^{\overline{i_E}}\frac{w_2^{z-i}w_1^{2i}}{(z-i)!(2i)!}+\\
&\quad+\sum_{z=0}^{\lfloor\frac{n-k}{2}\rfloor-1}\sum_{l=0}^{\overline{l_O}}\frac{w_2^{k+z-l}w_1^{2l+1}}{(k-l+z)!(2l+1)!}\sum_{i=0}^{\overline{i_O}}\frac{w_2^{z-i}w_1^{2i+1}}{(z-i)!(2i+1)!}.
\end{aligned}$$

We start from the inner summations (indexes i) and then proceed outwards.

Since $\overline{i_E},\overline{i_O}\leq z$, and $\frac{1}{(2i+1)!}\leq\frac{1}{(2i)!}\leq\frac{1}{i!}$, we have:

$$\sum_{i=0}^{\overline{i_E}}\frac{w_2^{z-i}w_1^{2i}}{(z-i)!(2i)!}\leq\sum_{i=0}^{z}\frac{w_2^{z-i}(w_1^2)^i}{(z-i)!i!}\leq\sum_{i=0}^{z}\binom{z}{i}\frac{w_2^{z-i}(w_1^2)^i}{z!}=\frac{(w_2+w_1^2)^z}{z!}$$

and

$$\sum_{i=0}^{\overline{i_O}} \frac{w_2^{z-i} w_1^{2i+1}}{(2i+1)!(z-i)!} \leq w_1 \sum_{i=0}^{z} \binom{z}{i} \frac{w_2^{z-i}(w_1^2)^i}{z!} = w_1 \frac{(w_2+w_1^2)^z}{z!}.$$

Remembering that $\overline{l_E}, \overline{l_O} \leq k+z$ and applying the same argument to the sum over l, we get:

$$\sum_{l=0}^{\overline{l_E}} \frac{w_2^{k+z-l} w_1^{2l}}{(k+z-l)!(2l)!} \leq \frac{(w_2+w_1^2)^{k+z}}{(k+z)!}$$

and

$$\sum_{l=0}^{\overline{l_O}} \frac{w_2^{k+z-l} w_1^{2l+1}}{(k+z-l)!(2l+1)!} \leq w_1 \frac{(w_2+w_1^2)^{k+z}}{(k+z)!}.$$

Therefore, taking $W = w_2 + w_1^2$,

$$\begin{aligned}
\widetilde{Q}_2(2k) < & \sum_{z=0}^{\lfloor \frac{n-k}{2} \rfloor} \frac{W^{k+z}}{(k+z)!} \frac{W^z}{z!} + w_1^2 \sum_{z=0}^{\lfloor \frac{n-k}{2} \rfloor -1} \frac{W^{k+z}}{(k+z)!} \frac{W^z}{z!} \\
< & (1+w_1^2) \sum_{z=0}^{\lfloor \frac{n-k}{2} \rfloor} \frac{W^{k+z}}{(k+z)!} \frac{W^z}{z!} \\
< & (1+w_1^2) \left(\sum_{z=0}^{\lfloor \frac{n-k}{2} \rfloor} \frac{W^z}{z!} \right) \left(\sum_{z=0}^{\lfloor \frac{n-k}{2} \rfloor} \frac{W^{k+z}}{(k+z)!} \right) \\
< & e^W (1+w_1^2) \left(\sum_{t=k}^{\lfloor \frac{n+k}{2} \rfloor} \frac{W^t}{t!} \right).
\end{aligned}$$

By Lemma 1.6.3 with $j=2$, for $k \geq 2W-1$ we have

$$\widetilde{Q}_2(2k) \leq 2e^W (1+w_1^2) \frac{W^k}{k!}. \tag{1.131}$$

For a net balance $2k+1$ we act similarly and find:

$$\widetilde{Q}_2(2k+1) < \sum_{z=0}^{\lfloor \frac{n-k-1}{2} \rfloor} \sum_{l=0}^{\overline{l'_E}} \sum_{i=0}^{\overline{i'_E}} \frac{w_2^{k+2z-l-i} w_1^{2l+2i+1}}{(k+z-l)!(2l+1)!(2i)!(z-i)!}$$

$$
\begin{aligned}
&+\sum_{z=0}^{\left\lfloor\frac{n-k}{2}\right\rfloor-1}\sum_{l=0}^{\overline{l'_O}}\sum_{i=0}^{\overline{i_O}}\frac{w_2^{k+2z-l-i+1}w_1^{2l+2i+1}}{(k+z+1-l)!(2l)!(2i+1)!(z-i)!}\\
<&\sum_{z=0}^{\left\lfloor\frac{n-k-1}{2}\right\rfloor}\sum_{l=0}^{\overline{l'_E}}\frac{w_2^{k+z-l}w_1^{2l+1}}{(k+z-l)!(2l+1)!}\sum_{i=0}^{\overline{i'_E}}\frac{w_2^{z-i}w_1^{2i}}{(z-i)!(2i)!}+\\
&+\sum_{z=0}^{\left\lfloor\frac{n-k}{2}\right\rfloor-1}\sum_{l=0}^{\overline{l'_O}}\frac{w_2^{k+z+1-l}w_1^{2l}}{(k+z+1-l)!(2l)!}\sum_{i=0}^{\overline{i_O}}\frac{w_2^{z-i}w_1^{2i+1}}{(z-i)!(2i+1)!}\\
<&\sum_{z=0}^{\left\lfloor\frac{n-k-1}{2}\right\rfloor}\sum_{l=0}^{k+z}w_1\frac{w_2^{k+z-l}(w_1^2)l}{(k+z-l)!l!}\sum_{i=0}^{z}\frac{w_2^{z-i}(w_1^2)^i}{(z-i)!i!}+\\
&+\sum_{z=0}^{\left\lfloor\frac{n-k}{2}\right\rfloor-1}\sum_{l=0}^{k+z+1}\frac{w_2^{k+z+1-l}(w_1^2)^l}{(k+z+1-l)!l!}w_1\sum_{i=0}^{z}\frac{w_2^{z-i}(w_1^2)^i}{(z-i)!i!}\\
<&w_1\sum_{z=0}^{\left\lfloor\frac{n-k-1}{2}\right\rfloor}\frac{W^{k+z}}{(k+z)!}\frac{W^z}{z!}+w_1\sum_{z=0}^{\left\lfloor\frac{n-k}{2}\right\rfloor-1}\frac{W^{k+z+1}}{(k+z+1)!}\frac{W^z}{z!}\\
<&w_1\sum_{z=0}^{\left\lfloor\frac{n-k-1}{2}\right\rfloor}\frac{W^{k+z}}{(k+z)!}\frac{W^z}{z!}+w_1W\sum_{z=0}^{\left\lfloor\frac{n-k}{2}\right\rfloor-1}\frac{W^{k+z}}{(k+z)!}\frac{W^z}{z!}\\
<&w_1(1+W)\sum_{z=0}^{\left\lfloor\frac{n-k-1}{2}\right\rfloor}\frac{W^{k+z}}{(k+z)!}\frac{W^z}{z!}\\
<&w_1(1+W)\left(\sum_{z=0}^{\left\lfloor\frac{n-k-1}{2}\right\rfloor}\frac{W^z}{z!}\right)\left(\sum_{z=0}^{\left\lfloor\frac{n-k-1}{2}\right\rfloor}\frac{W^{k+z}}{(k+z)!}\right)\\
<&w_1(1+W)e^W\left(\sum_{t=k}^{\left\lfloor\frac{n+k}{2}\right\rfloor}\frac{W^t}{t!}\right)
\end{aligned}
$$

and applying Lemma 1.6.3 with $j=2$ for $k\geq 2W-1$, we get

$$\widetilde{Q}_2(2k+1)\leq 2w_1(1+W)e^W\frac{W^k}{k!}. \tag{1.132}$$

Juxtaposing Equations (1.131) and (1.132) we obtain

$$\widetilde{Q}_2(k)\leq 2\max\{1+w_1^2,w_1(1+W)\}e^W\frac{W^{\left\lfloor\frac{k}{2}\right\rfloor}}{\left\lfloor\frac{k}{2}\right\rfloor!}$$

for k integer, $2\lceil 2W-1\rceil \leq k \leq 2n$, which is Equation (1.125).

Let us now consider the summation of all $\widetilde{Q}_2(k)$ from a certain $\overline{k}$ upwards. It is convenient to separate its even and odd terms.

$$\begin{aligned}
\sum_{k=\overline{k}}^{2n} \widetilde{Q}_2(k) = & \sum_{\substack{k=\overline{k}\\ \text{even } k}}^{2n} \widetilde{Q}_2(k) + \sum_{\substack{k=\overline{k}\\ \text{odd } k}}^{2n} \widetilde{Q}_2(k) \\
& \leq \sum_{k=\left\lfloor\frac{\overline{k}}{2}\right\rfloor}^{2n} \widetilde{Q}_2(2k) + \sum_{k=\left\lfloor\frac{\overline{k}}{2}\right\rfloor}^{2n} \widetilde{Q}_2(2k+1) \\
& \leq \sum_{k=\left\lfloor\frac{\overline{k}}{2}\right\rfloor}^{2n} 2e^W(1+w_1^2)\frac{W^k}{k!} + \sum_{k=\left\lfloor\frac{\overline{k}}{2}\right\rfloor}^{2n} 2w_1(1+W)e^W\frac{W^k}{k!} \\
& \leq 2(1+w_1^2+w_1(1+W))e^W \sum_{k=\left\lfloor\frac{\overline{k}}{2}\right\rfloor}^{2n} 2w_1(1+W)e^W\frac{W^k}{k!}.
\end{aligned}$$

With a further application of Lemma 1.6.3 we obtain:

$$\sum_{k=\overline{k}}^{2n} \widetilde{Q}_2(k) \leq 4(1+w_1^2+w_1(1+W))e^W \frac{W^{\left\lfloor\frac{\overline{k}}{2}\right\rfloor}}{\left\lfloor\frac{\overline{k}}{2}\right\rfloor!}$$

for k integer, $2\lceil 2W-1\rceil \leq k \leq 2n$, which is Equation (1.126).

Let us consider now Equation (1.127). Again we separate the even and odd terms.

$$\begin{aligned}
\sum_{k=\overline{k}}^{2n} e^{hk}\widetilde{Q}_2(k) = & \sum_{\substack{k=\overline{k}\\ \text{even } k}}^{2n} e^{kh}\widetilde{Q}_2(k) + \sum_{\substack{k=\overline{k}\\ \text{odd } k}}^{2n} e^{kh}\widetilde{Q}_2(k) \leq \\
& \leq \sum_{k=\left\lfloor\frac{\overline{k}}{2}\right\rfloor}^{2n} \left(e^{2kh}\widetilde{Q}_2(2k) + e^{(2k+1)h}\widetilde{Q}_2(2k+1)\right) \leq \\
& \leq \sum_{k=\left\lfloor\frac{\overline{k}}{2}\right\rfloor}^{2n} \left(e^{2kh}2e^W(1+w_1^2)\frac{W^k}{k!} + e^{(2k+1)h}2w_1(1+W)e^W\frac{W^k}{k!}\right) \leq \\
& \leq 2e^W(1+w_1^2+w_1e^h(1+W)) \sum_{k=\left\lfloor\frac{\overline{k}}{2}\right\rfloor}^{2n} \frac{(We^{2h})^k}{k!}.
\end{aligned}$$

Let us call $C^+ = 1 + w_1^2 + w_1 e^h (1+W)$; we obtain

$$\sum_{k=\overline{k}}^{2n} e^{hk}\widetilde{Q}_2(k) \le 2C^+ e^W \sum_{k=\left\lfloor \frac{\overline{k}}{2}\right\rfloor}^{2n} \frac{(We^{2h})^k}{k!} \le 4e^W C^+ \frac{(We^{2h})^{\left\lfloor \frac{\overline{k}}{2}\right\rfloor}}{\left\lfloor \frac{\overline{k}}{2}\right\rfloor !}$$

for $\left\lfloor \frac{\overline{k}}{2}\right\rfloor \ge 2We^{2h} - 1$, i.e. $\overline{k} \ge 2\left\lceil 2We^{2h} - 1\right\rceil$.

Similarly

$$\begin{aligned}
\sum_{k=\overline{k}}^{2n} e^{-hk}\widetilde{Q}_2(k) &= \sum_{\substack{k=\overline{k}\\ \text{even } k}}^{2n} e^{-kh}\widetilde{Q}_2(k) + \sum_{\substack{k=\overline{k}\\ \text{odd } k}}^{2n} e^{-kh}\widetilde{Q}_2(k)\\
&\le \sum_{k=\left\lfloor \frac{\overline{k}}{2}\right\rfloor}^{2n} \left(e^{-2kh}\widetilde{Q}_2(2k) + e^{-(2k+1)h}\widetilde{Q}_2(2k+1)\right).
\end{aligned}$$

For $\overline{k} \ge 2\left\lceil 2We^{-2h} - 1\right\rceil$ we can apply Lemma 1.6.3, therefore

$$\begin{aligned}
\sum_{k=\overline{k}}^{2n} e^{-hk}\widetilde{Q}_2(k) &\le \sum_{k=\left\lfloor \frac{\overline{k}}{2}\right\rfloor}^{2n} \left(e^{-2kh} 2e^W (1+w_1^2)\frac{W^k}{k!} + e^{-(2k+1)h} 2w_1(1+W)e^W \frac{W^k}{k!}\right)\\
&\le 2e^W (1 + w_1^2 + e^{-h} w_1 (1+W)) \sum_{k=\left\lfloor \frac{\overline{k}}{2}\right\rfloor}^{2n} \frac{(We^{-2h})^k}{k!}\\
&\le 4C^- e^W \frac{(We^{-2h})^{\left\lfloor \frac{\overline{k}}{2}\right\rfloor}}{\left\lfloor \frac{\overline{k}}{2}\right\rfloor !}
\end{aligned}$$

where $C^- = 1 + w_1^2 + w_1 e^{-h}(1+W)$, for $\overline{k} \ge 2\left\lceil 2W - 1\right\rceil$.

$\square$

Let us now indicate with $V(0,0,0)$ and $V^T(0,0,0)$ the European call values we obtain via a backward procedure using Equation (1.96) with $N = 2$ in the HS and in the truncated case (cf. Equation (1.103)) respectively.

Theorem 1.6.12. *Given $\varepsilon > 0$, taking*

$$\begin{aligned}
C^+ &= 1 + w_1^2 + w_1 e^h (1+W)\\
C^- &= 1 + w_1^2 + w_1 e^{-h} (1+W)\\
g^+ &= e^{2h+1} W + (\alpha - r)\tau + W + \log(8S_0) + \log(C^+ + (1+e^{2h})e^{4h} C^-)
\end{aligned}$$

$$g^{-} = e^{-2h+1}W + (\alpha - r)\tau + W + \log(8S_0) + \log(C^{-} + (1+e^{-2h})C^{+})$$

for

$$\overline{k} \geq \max\{2\lceil g^{+} - \log\varepsilon\rceil, 2\left\lceil 2We^{2h} - 1\right\rceil - 1\}$$

and

$$\bar{l} \geq \max\{2\lceil g^{-} - \log\varepsilon\rceil, 2\left\lceil 2We^{2h} - 1\right\rceil - 1\},$$

we have

$$\left|V(0,0,0) - V^{T}(0,0,0)\right| < \varepsilon.$$

Proof. We apply Proposition 1.6.11 to Equation (1.121).

$$\begin{aligned}
V_E - \widehat{V_E^T} \leq & e^{(\alpha-r)\tau} S_0 \left(\sum_{k=\overline{k}+1}^{2n} e^{hk}\widetilde{Q}_2(k) + \sum_{k=\bar{l}+1}^{2n} e^{-hk}\widetilde{Q}_2(k) + \right. \\
& + \sum_{s=\overline{k}+2}^{\min\{2\overline{k}+\bar{l}+2,2n\}} e^{h(2\overline{k}+2-s)}\widetilde{Q}_2(s) + \sum_{s=\overline{k}+4}^{\min\{2\overline{k}+\bar{l}+4,2n\}} e^{h(2\overline{k}+4-s)}\widetilde{Q}_2(s) \\
& \left. + \sum_{s=\bar{l}+2}^{\min\{2\bar{l}+\overline{k}+2,2n\}} e^{h(-2\bar{l}-2+s)}\widetilde{Q}_2(s) + \sum_{s=\bar{l}+4}^{\min\{2\bar{l}+\overline{k}+4,2n\}} e^{h(-2\bar{l}-4+s)}\widetilde{Q}_2(s) \right) \\
\leq & e^{(\alpha-r)\tau} S_0 \left(\sum_{k=\overline{k}+1}^{2n} e^{hk}\widetilde{Q}_2(k) + \sum_{k=\bar{l}+1}^{2n} e^{-hk}\widetilde{Q}_2(k) + \right. \\
& + e^{h(2\overline{k}+2)} \sum_{s=\overline{k}+2}^{\min\{2\overline{k}+\bar{l}+2,2n\}} e^{-hs}\widetilde{Q}_2(s) + e^{h(2\overline{k}+4)} \sum_{s=\overline{k}+4}^{\min\{2\overline{k}+\bar{l}+4,2n\}} e^{-hs}\widetilde{Q}_2(s) \\
& \left. + e^{h(-2\bar{l}-2)} \sum_{s=\bar{l}+2}^{\min\{2\bar{l}+\overline{k}+2,2n\}} e^{hs}\widetilde{Q}_2(s) + e^{h(-2\bar{l}-4)} \sum_{s=\bar{l}+4}^{\min\{2\bar{l}+\overline{k}+4,2n\}} e^{hs}\widetilde{Q}_2(s) \right) \\
\leq & e^{(\alpha-r)\tau} S_0 \left(\sum_{k=\overline{k}+1}^{2n} e^{hk}\widetilde{Q}_2(k) + \sum_{k=\bar{l}+1}^{2n} e^{-hk}\widetilde{Q}_2(k) + \right. \\
& + (1+e^{2h}) e^{h(2\overline{k}+2)} \sum_{s=\overline{k}+1}^{\min\{2\overline{k}+\bar{l}+2,2n\}} e^{-hs}\widetilde{Q}_2(s)
\end{aligned}$$

$$
\begin{aligned}
&+(1+e^{-2h})e^{h(-2\bar{l}-2)}\sum_{s=\bar{l}+1}^{\min\{2\bar{l}+\bar{k}+2,2n\}} e^{hs}\widetilde{Q}_2(s)\Bigg) \\
&\le 4e^{W}e^{(\alpha-r)\tau}S_0\left(C^+\frac{(We^{2h})^{\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor}}{\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor!}+C^-\frac{(We^{-2h})^{\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor}}{\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor!}+\right.\\
&+(1+e^{2h})e^{h(2\bar{k}+2)}C^-\frac{(We^{-2h})^{\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor}}{\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor!}\\
&\left.+(1+e^{-2h})e^{h(-2\bar{l}-2)}C^+\frac{(We^{2h})^{\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor}}{\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor!}\right)
\end{aligned}
$$

for $\bar{k},\bar{l}\ge 2\left\lceil 2We^{2h}-1\right\rceil-1$.

Since $\bar{k}+1\le 2\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor+1$ and $-\bar{l}-1\le -2\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor$ we can write:

$$
\begin{aligned}
V_E-\widehat{V_E^T}\le & 4e^{(\alpha-r)\tau+W}S_0\left(C^+\frac{(We^{2h})^{\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor}}{\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor!}+C^-\frac{(We^{-2h})^{\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor}}{\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor!}+\right.\\
&+(1+e^{2h})e^{4h}C^-\frac{(We^{-2h}e^{4h})^{\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor}}{\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor!}+\\
&\left.+(1+e^{-2h})C^+\frac{(We^{2h}e^{-4h})^{\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor}}{\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor!}\right) \qquad (1.133)\\
\le & 4e^{(\alpha-r)\tau+W}S_0\left((C^++(1+e^{2h})e^{4h}C^-)\frac{(We^{2h})^{\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor}}{\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor!}+\right.\\
&\left.+(C^-+(1+e^{-2h})C^+)\frac{(We^{-2h})^{\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor}}{\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor!}\right).
\end{aligned}
$$

If we want an error less than ε to be equally shared between the elimination of the upper and lower regions of the tree on the whole period, retracing our steps in Theorem 1.6.7, we want to determine under which conditions on $\bar{k}$ and $\bar{l}$ we have:

$$(C^+ + (1+e^{2h})e^{4h}C^-)\frac{(We^{2h})^{\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor}}{\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor!} \leq \frac{\varepsilon}{8e^{(\alpha-r)\tau+W}S_0}$$

and

$$(C^- + (1+e^{-2h})C^+)\frac{(We^{-2h})^{\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor}}{\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor!} \leq \frac{\varepsilon}{8e^{(\alpha-r)\tau+W}S_0}.$$

By Lemma 1.6.4, in order for the inequality

$$\frac{(We^{2h})^{\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor}}{\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor!} \leq \frac{\varepsilon}{8e^{(\alpha-r)\tau+W}S_0(C^+ + (1+e^{2h})e^{4h}C^-)}$$

to stand, it suffices to ask

$$-\left\lfloor\frac{\bar{k}+1}{2}\right\rfloor + e^{2h+1}W \leq \log\frac{\varepsilon}{8e^{(\alpha-r)\tau+W}S_0(C^+ + (1+e^{2h})e^{4h}C^-)}$$

i. e.

$$\bar{k} \geq 2\left\lceil e^{2h+1}W - \log\varepsilon + (\alpha-r)\tau + W + \log(8S_0) + \log(C^+ + (1+e^{2h})e^{4h}C^-)\right\rceil - 1.$$

In the same way, inequality

$$\frac{(We^{-2h})^{\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor}}{\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor!} \leq \frac{\varepsilon}{8e^{(\alpha-r)\tau+W}S_0(C^- + (1+e^{-2h})C^+)}$$

is guaranteed for

$$-\left\lfloor\frac{\bar{l}+1}{2}\right\rfloor + e^{-2h+1}W \leq \log\frac{\varepsilon}{8e^{(\alpha-r)\tau+W}S_0(C^- + (1+e^{-2h})C^+)}$$

i. e.

$$\bar{l} \geq 2\left\lceil e^{-2h+1}W - \log\varepsilon + (\alpha-r)\tau + W + \log(8S_0) + \log(C^- + (1+e^{-2h})C^+)\right\rceil - 1.$$

Collecting all the requirements on $\bar{k}$ and $\bar{l}$, we have the thesis.

□

If for some reasons we wish for a symmetric cut, since h is positive, hence $e^{-h} < 1 < e^h$ and $C^- < C^+$, we can take the more stringent condition:

$$\bar{l} = \bar{k} \geq \max\{2\left\lceil 2We^{2h} - 1\right\rceil - 1,$$
$$2\left\lceil e^{2h+1}W - \log\varepsilon + (\alpha - r)\tau + W + \log(8S_0) + \log(C^+(1 + e^{4h} + e^{6h}))\right\rceil - 1\}.$$

The previous theorem allows to extend to $N = 2$ the result of Theorem 1.6.8 and Theorem 1.6.9, which guarantee that for appropriately chosen $\bar{k}$ and $\bar{l}$, the value V_E of the European call option (or the value $V(0,0,0)$ obtained via the backward procedure, that is the same) can be approximated by

$$\widehat{V_E^T} = e^{-r\tau} \sum_{k=-\bar{l}}^{\bar{k}} \sum_{j=0}^{n} (S_0 e^{(-n+2j)\sigma\sqrt{\Delta t} + hk} - K_0)^+ P(j)\widehat{Q}_1(k)$$

and with the value $V^T(0,0,0)$ obtained through the backward truncated procedure.

Theorem 1.6.13. *Given $\varepsilon = \frac{1}{n} > 0$, and $\bar{k}$ and $\bar{l}$ the smallest integers as in Theorem 1.6.12, the proposed procedure for $\widehat{V_E^T}$ converges to the HS price and its computational complexity is $O(n \log n)$.*

Theorem 1.6.14. *Given $\varepsilon = \frac{1}{n} > 0$, and $\bar{k}$ and $\bar{l}$ the smallest integers as in Theorem 1.6.12, the backward procedure described above for $V^T(0,0,0)$ converges to the HS price and its computational complexity is $O(n^2 \log n)$.*

The proofs of Theorems 1.6.13 and 1.6.14 coincide with those of Theorems 1.6.8 and 1.6.9.

European call option, arbitrary N

In this Section the results are generalised to an arbitrary N. Similarly to what we have previously done, we will denote:

q_k the probability of a jump of amplitude kh for k integer, $-N \leq k \leq N$, in a Δt time interval;

c_k as the constant $q_k \cdot n$ for $k \neq 0$, $-N \leq k \leq N$;

w_k as the maximum between c_{+k} and c_{-k} for $1 \leq k \leq N$.

We also define recursively

$$
\begin{aligned}
W_1 &= w_1 \\
W_{i+1} &= w_{i+1} + W_i^{\frac{i+1}{i}}
\end{aligned}
\tag{1.134}
$$

and

$$M_i = \max\{W_i, W_i^{\frac{-i+1}{i}}\}. \tag{1.135}$$

Recall that $Q_N(k)$ will indicate the probability of reaching level k with the jumps, while $\widetilde{Q}_N(k)$ is the "enlarged probability" with q_{+i} and q_{-i} both substituted by their maximum, and $\widehat{Q}_N(k)$ is the probability of reaching level k computed recursively forwards without taking into account the nodes (i, j, k') with $k' < -\bar{l}$ or $k' > \overline{k}$.

We also recall the limitation (1.108) we had obtained on the difference between the Hilliard and Schwartz European call option price V_E and the price $\widehat{V_E^T}$ obtained cutting the tree at levels $-\bar{l}$ and $\overline{k}$ (cf. Equation (1.102)).

$$V_E - \widehat{V_E^T} \le e^{(\alpha - r)\tau} S_0 \left(\sum_{k=\overline{k}+1}^{Nn} e^{hk} \widetilde{Q}_N(k) + \sum_{k=\bar{l}+1}^{Nn} e^{-hk} \widetilde{Q}_N(k) + \sum_{k=-\bar{l}}^{\bar{k}} e^{hk} (Q_N^{\overline{k}}(k) + Q_{N\bar{l}}(k)) \right)$$

The following Lemma relates $Q_N^{\overline{k}}$ and $Q_{N\bar{l}}$ with $\widetilde{Q}_N$.

Lemma 1.6.15.

$$Q_N^{\overline{k}}(k) \le \sum_{i=1}^{N} \widetilde{Q}_N(2\overline{k} - k + 2i)$$

$$Q_{N\bar{l}}(k) \le \sum_{i=1}^{N} \widetilde{Q}_N(2\bar{l} + k + 2i)$$

for all

$$-\bar{l} \le k \le \overline{k}$$

Proof. The proof is similar to that of Lemma 1.6.10: the original path, at the first moment it trespasses the $\overline{k}$ level, can reach level $\overline{k}+1, \overline{k}+2, \ldots, \overline{k}+N$, depending on the amplitude of the jump that steps over the boundary. Therefore the reflection of the path can end at level $2\overline{k}-k+2$, $2\overline{k}-k+4$, ..., $2\overline{k}-k+2N$. Likewise for the paths that cross the $-\bar{l}$ level. □

By Lemma 1.6.15 we have

$$V_E^T - \widehat{V_E^T} \le e^{(\alpha - r)\tau} S_0 \left(\sum_{k=-\bar{l}}^{\bar{k}} e^{hk} \sum_{i=1}^{N} \widetilde{Q}_N(2\overline{k} - k + 2i) + \sum_{k=-\bar{l}}^{\bar{k}} e^{hk} \sum_{i=1}^{N} \widetilde{Q}_N(2\bar{l} + k + 2i) \right)$$

$$\leq e^{(\alpha-r)\tau}S_0\left(\sum_{s=\overline{k}+2}^{\min\{2\overline{k}+\overline{l}+2,Nn\}} e^{h(2\overline{k}-s+2)}\sum_{i=0}^{N-1}\widetilde{Q}_N(s+2i)+\right.$$
$$\left.\sum_{s=\overline{l}+2}^{\min\{2\overline{l}+\overline{k}+2,Nn\}} e^{h(s-2\overline{l}-2)}\sum_{i=0}^{N-1}\widetilde{Q}_N(s+2i)\right). \tag{1.136}$$

Since the first sum is over $\overline{k}+2\leq s\leq 2\overline{k}+\overline{l}+2$, we have $e^{h(2\overline{k}-s+2)}\leq e^{h\overline{k}}$ and likewise for $\overline{l}+2\leq s\leq 2\overline{l}+\overline{k}+2$ we have $e^{h(s-2\overline{l}-2)}\leq e^{h\overline{k}}$.

$$V_E^T-\widehat{V_E^T}\leq$$
$$\leq e^{(\alpha-r)\tau}e^{h\overline{k}}S_0\left(\sum_{s=\overline{k}+2}^{\min\{2\overline{k}+\overline{l}+2,Nn\}}\sum_{i=0}^{N-1}\widetilde{Q}_N(s+2i)+\sum_{s=\overline{l}+2}^{\min\{2\overline{l}+\overline{k}+2,Nn\}}\sum_{i=0}^{N-1}\widetilde{Q}_N(s+2i)\right) \tag{1.137}$$
$$\leq e^{(\alpha-r)\tau}e^{h\overline{k}}S_0\left(\sum_{i=0}^{N-1}\sum_{s=\overline{k}+2}^{\min\{2\overline{k}+\overline{l}+2,Nn\}}\widetilde{Q}_N(s+2i)+\sum_{i=0}^{N-1}\sum_{s=\overline{l}+2}^{\min\{2\overline{l}+\overline{k}+2,Nn\}}\widetilde{Q}_N(s+2i)\right)$$
$$\leq e^{(\alpha-r)\tau}e^{h\overline{k}}S_0\left(N\sum_{s=\overline{k}+2}^{\min\{2\overline{k}+\overline{l}+2,Nn\}}\widetilde{Q}_N(s)+N\sum_{s=\overline{l}+2}^{\min\{2\overline{l}+\overline{k}+2,Nn\}}\widetilde{Q}_N(s)\right). \tag{1.138}$$

Combining Equations (1.106) and (1.138) we obtain:

$$V_E-\widehat{V_E^T}\leq e^{(\alpha-r)\tau}S_0\left(\sum_{k=\overline{k}+1}^{Nn}e^{hk}\widetilde{Q}_N(k)+\sum_{k=\overline{l}+1}^{Nn}e^{-hk}\widetilde{Q}_N(k)+\right.$$
$$\left.e^{h\overline{k}}N\sum_{k=\overline{k}+2}^{Nn}\widetilde{Q}_N(k)+e^{h\overline{k}}N\sum_{k=\overline{l}+2}^{Nn}\widetilde{Q}_N(k)\right). \tag{1.139}$$

As in the $N=1$ case, we can compute (1.139) with a $O(n^2)$ procedure, thus determining numerically the largest integers $\overline{l}$ and $\overline{k}$ such that the loss is inferior to an arbitrary ε, [4] but we are interested in obtaining a closed formula for $\overline{l}$ and $\overline{k}$.

[4] In the numerical computations for this work, we have considered the following less refined conditions: $\sum_{k=\overline{k}+1}^{Nn}(N+1)e^{hk}\widetilde{Q}_N(k)$ and $\sum_{k=\overline{l}+1}^{Nn}(Ne^{h\overline{k}}+1)\widetilde{Q}_N(k)$ We proceeded in the following way. Given ε, we considered $\eta=\frac{\varepsilon}{2e^{(\alpha-r)\tau}S_0}$ and we computed, starting from $i=Nn$, the sums $S_i^e=(N+1)\sum_{k=i}^{Nn}e^{hk}\widetilde{Q}_N(k)$ and $S_i=\sum_{k=i}^{Nn}\widetilde{Q}_N(k)$. While $S_i^e<\eta$, we keep decreasing i. The first i we encounter such that $S_i^e\geq\eta$ is our $\overline{k}$. While $S_i<\eta$, we keep decreasing i. The first i we encounter such that $S_i<\frac{\eta}{Ne^{h\overline{k}}+1}$ is our $\overline{l}$.

Let us denote $G = 2N \max\{W_N, 1\} e^{W_N} \prod_{i=1}^{N-1} M_i^2$.

Proposition 1.6.16. *For $k \geq N\lceil 2W_N - 1 \rceil$*

$$\widetilde{Q}_N(k) \leq G \frac{W_N^{\lfloor \frac{k}{N} \rfloor}}{\lfloor \frac{k}{N} \rfloor !}. \tag{1.140}$$

For $\overline{k} \geq N\lceil 2W_N - 1 \rceil$

$$\sum_{k=\overline{k}}^{Nn} \widetilde{Q}_N(k) \leq 2GN \frac{W_N^{\left\lfloor \frac{\overline{k}}{N} \right\rfloor}}{\left\lfloor \frac{\overline{k}}{N} \right\rfloor !}. \tag{1.141}$$

For $\overline{k} \geq N\lceil 2e^{Nh} W_N - 1 \rceil$

$$\sum_{k=\overline{k}}^{Nn} e^{hk} \widetilde{Q}_N(k) \leq 2G \frac{(e^{hN} W_N)^{\left\lfloor \frac{\overline{k}}{N} \right\rfloor}}{\left\lfloor \frac{\overline{k}}{N} \right\rfloor !} \sum_{r=0}^{N-1} e^{hr}.$$

For $\overline{k} \geq N\lceil 2W_N - 1 \rceil$:

$$\sum_{k=\overline{k}}^{Nn} e^{-hk} \widetilde{Q}_N(-k) \leq 2G \frac{(e^{-hN} W_N)^{\left\lfloor \frac{\overline{k}}{N} \right\rfloor}}{\left\lfloor \frac{\overline{k}}{N} \right\rfloor !} \sum_{r=0}^{N-1} e^{-hr}.$$

Proof. We focus on the probability $Q_N(k)$ of reaching level $k \geq 0$ in the jump dynamics at maturity. Level k can be reached in n time steps with a variety of possible combinations of jumps, therefore, in order to consider all the possible paths, we distinguish between the positive and the negative jumps: if $k \geq 0$ is the total balance, when the sum of all negative jumps is $-l$, with $l \geq 0$, then the sum of all positive jumps must be $k + l$. In short, $Q_N(k)$ is the sum over all possible non negative l, subject to the condition of a total of n moves, of all probabilities of reaching balance level k with a negative balance of $-l$.[5]

For $k \geq 0$, the probability $Q_N(k)$ is given by:

$$Q_N(k) = \sum_l \sum_{e_N^+,\ldots,e_1^+} \sum_{e_N^-,\ldots,e_1^-} C(e_N^+, e_N^-, e_{N-1}^+, e_{N-1}^-, \ldots, e_1^+, e_1^-) q_{+N}^{e_N^+} \cdots q_{+1}^{e_1^+} q_{-N}^{e_N^-} \cdots q_{-1}^{e_1^-} q_0^{e_0}$$

where the indexes of the summations are non negative integers that individuate the number of jumps of each kind: e_i^+ is the number of $+i$ jumps and e_i^- is the

[5] We will be more precise with regard to this condition in the following.

number of $-i$ jumps, while e_0 is the number of no-jump moves. All summations are limited by the conditions

$$e_0 + \sum_{i=1}^{N} e_i^- + \sum_{i=1}^{N} e_i^+ = n$$

and

$$\sum_{i=1}^{N} i(e_i^+ - e_i^-) = k.$$

$C(e_N^+, e_N^-, e_{N-1}^+, e_{N-1}^-, \ldots, e_1^+, e_1^-)$ denotes the number of combinations of the n factors, once the exponents are fixed, and is equal to

$$C(e_N^+, e_N^-, e_{N-1}^+, e_{N-1}^-, \ldots, e_1^+, e_1^-) = \frac{n!}{e_N^+! e_N^-! e_{N-1}^+! e_{N-1}^-! \ldots e_1^+! e_1^-! e_0!}.$$

By Euclidean division we can write l as a multiple of N plus a remainder $0 \le r_N^- \le N-1$: $l = Nz + r_N^-$. This means that the negative balance $-l$ is due to at most z jumps of the $-N$ kind, and the difference between Nz and l shall be covered with smaller jumps.

Instead of summing over all possible l's, then, it will be easier in order to express the limitations for the indexes, to consider the summation over all possible z and $0 \le r_N^- \le N-1$.

For any fixed z and r_N^-, we will have at most z jumps of the $-N$ kind, therefore e_N^- needs to vary between 0 and z. The same idea is to be followed for the positive balance: given k, the values t and $0 \le r_N \le N-1$ such that $k = Nt + r_N$ are uniquely determined; therefore for any given pair of z and r_N^- the positive balance can be written as $N(t+z) + r_N + r_N^-$. This provides the limitation for e_N^+, which must be at most $t + z + \left\lfloor \frac{r_N + r_N^-}{N} \right\rfloor$. The choice of every e_i^+ (e_i^-) imposes further conditions on the possible values for e_{i-1}^+ (e_{i-1}^-, respectively).

A change in perspective in the summations will permit to better express the relationships and mutual limitations existing between the exponents.

For any fixed z, we have $0 \le e_N^- \le z$. Let us define $b_{N-1} = z - e_N^-$. The value b_{N-1} ($0 \le b_{N-1} \le z$) will represent the part of $-Nz$ that needs to be covered with jumps not greater than $N-1$. Of the negative balance $-(Nz + r_N^-)$, then, $-Ne_N^-$ will be covered by $-N$ jumps and the rest, $-(Nb_{N-1} + r_N^-)$, by jumps of smaller amplitude.

Once z, r_N^- and e_N^- are fixed, we have a negative balance of $-(Nb_{N-1} + r_N^-)$ to cover with negative jumps of amplitude at most $N-1$. In order to find the limitation for e_{N-1}^-, we compute the Euclidean division of $Nb_{N-1} + r_N^-$ by

$N-1$: the quotient $z_{N-1} = \lfloor \frac{Nb_{N-1}+r_N^-}{N-1} \rfloor$ is an upper bound for e_{N-1}^- (we shall consider as limitation for the summation the more stringent between this value and the condition of a total of n moves), and we call r_{N-1}^- the remainder. Once again, instead of summing over e_{N-1}^-, we sum over $b_{N-2} = z_{N-1} - e_{N-1}^-$.

We operate in the same way repeatedly using Euclidean division in order to find upper bounds for all e_j^-, and similarly for the positive jumps, where we introduce the a_j and r_j^+ values.

To recap, the indices a_i (b_i) are indicators of how much of the total positive (respectively, negative) balance is due to moves of amplitude at most i, and are related to the exponents in the following way:

$$
\begin{aligned}
e_N^- &= z - b_{N-1} \\
e_{N-1}^- &= \left\lfloor \frac{Nb_{N-1} + r_N^-}{N-1} \right\rfloor - b_{N-2} \\
&\dots \\
e_i^- &= \left\lfloor \frac{(i+1)b_i + r_{i+1}^-}{i} \right\rfloor - b_{i-1} \qquad \text{for } i < N \\
&\qquad \text{where } r_i^- \text{ is the remainder of } \frac{(i+1)b_i + r_{i+1}^-}{i} \qquad \text{for } i < N \\
&\dots \\
e_2^- &= \left\lfloor \frac{3b_2 + r_3^-}{2} \right\rfloor - b_1 \\
e_1^- &= 2b_1 + r_2^-
\end{aligned}
$$

$$
\begin{aligned}
e_N^+ &= t + z + \left\lfloor \frac{r_N + r_N^-}{N} \right\rfloor - a_{N-1} \\
e_{N-1}^+ &= \left\lfloor \frac{Na_{N-1} + r_N^+}{N-1} \right\rfloor - a_{N-2} \\
&\qquad \text{where } r_N^+ \text{ is the remainder of } \frac{r_N + r_N^-}{N} \\
&\dots \\
e_i^+ &= \left\lfloor \frac{(i+1)a_i + r_{i+1}^+}{i} \right\rfloor - a_{i-1} \qquad \text{for } i < N \\
&\qquad \text{where } r_i^+ \text{ is the remainder of } \frac{(i+1)a_i + r_{i+1}^+}{i} \qquad \text{for } i < N \\
&\dots \\
e_2^+ &= \left\lfloor \frac{3a_2 + r_3^+}{2} \right\rfloor - a_1
\end{aligned}
$$

$e_1^+ = 2a_1 + r_2^+$

The probability $Q_N(k)$ of reaching at maturity level $k \geq 0$ for the jump dynamics can then be written as:

$$Q_N(k) = \sum_{r_N^-=0}^{N-1} \sum_z \sum_{a_{N-1}} \cdots \sum_{a_1} \sum_{b_{N-1}} \cdots \sum_{b_1} C(e_N^+, e_N^-, e_{N-1}^+, e_{N-1}^-, \ldots, e_1^+, e_1^-)$$
$$q_{+N}^{e_N^+} \cdots q_{+1}^{e_1^+} q_{-N}^{e_N^-} \cdots q_{-1}^{e_1^-} q_0^{e_0}.$$

Substituting $c_{\pm i}$ with w_i, we obtain

$$\widetilde{Q}_N(k) = \sum_{r_N^-=0}^{N-1} \sum_z \sum_{a_{N-1}} \cdots \sum_{a_1} \sum_{b_{N-1}} \cdots \sum_{b_1} \frac{n!}{e_N^+! e_N^-! e_{N-1}^+! e_{N-1}^-! \ldots e_1^+! e_1^-! e_0!}$$
$$\frac{w_N^{e_N^+} \cdots w_1^{e_1^+} w_N^{e_N^-} \cdots w_1^{e_1^-}}{n^{\sum_{i=1}^N e_i^+ + \sum_{i=1}^N e_i^-}} q_0^{e_0}.$$

Since $q_0 \leq 1$ and $\frac{n!}{e_0! n^{\sum_{i=1}^N e_i^+ + \sum_{i=1}^N e_i^-}} \leq 1$:

$$\widetilde{Q}_N(k) \leq \sum_{r_N^-=0}^{N-1} \sum_z \sum_{a_{N-1}} \cdots \sum_{a_1} \sum_{b_{N-1}} \cdots \sum_{b_1} \frac{w_N^{e_N^+} \cdots w_1^{e_1^+} w_N^{e_N^-} \cdots w_1^{e_1^-}}{e_N^+! e_N^-! e_{N-1}^+! e_{N-1}^-! \ldots e_1^+! e_1^-!}.$$

We treat separately the positive and the negative parts, and we work from the inside outwards. The summation over the negative jumps is given by:

$$\sum_{b_{N-1}} \frac{w_N^{e_N^-}}{e_N^-!} \cdots \sum_{b_3} \frac{w_4^{e_4^-}}{e_4^-!} \sum_{b_2} \frac{w_3^{e_3^-}}{e_3^-!} \sum_{b_1} \frac{w_2^{e_2^-}}{e_2^-!} \frac{w_1^{e_1^-}}{e_1^-!} =$$

$$= \sum_{b_{N-1}} \frac{w_N^{e_N^-}}{e_N^-!} \cdots \sum_{b_3} \frac{w_4^{e_4^-}}{e_4^-!} \sum_{b_2} \frac{w_3^{e_3^-}}{e_3^-!} \sum_{b_1} \frac{w_2^{\left\lfloor \frac{3b_2+r_3^-}{2} \right\rfloor - b_1}}{\left(\left\lfloor \frac{3b_2+r_3^-}{2} \right\rfloor - b_1\right)!} \frac{w_1^{2b_2+r_2^-}}{(2b_1 + r_2^-)!}$$

$$\leq \sum_{b_{N-1}} \frac{w_N^{e_N^-}}{e_N^-!} \cdots \sum_{b_3} \frac{w_4^{e_4^-}}{e_4^-!} \sum_{b_2} \frac{w_3^{e_3^-}}{e_3^-!} w_1^{r_2^-} \frac{(w_2 + w_1^2)^{\left\lfloor \frac{3b_2+r_3^-}{2} \right\rfloor}}{\left\lfloor \frac{3b_2+r_3^-}{2} \right\rfloor !}.$$

Since r_2^- is the remainder of $\frac{3b_2+r_3^-}{2}$, it can only assume the values 0 or 1; therefore we can write:

$$\sum_{b_{N-1}} \frac{w_N^{e_N^-}}{e_N^-!} \cdots \sum_{b_3} \frac{w_4^{e_4^-}}{e_4^-!} \sum_{b_2} \frac{w_3^{e_3^-}}{e_3^-!} w_1^{r_2^-} \frac{(w_2 + w_1^2)^{\left\lfloor \frac{3b_2+r_3^-}{2} \right\rfloor}}{\left\lfloor \frac{3b_2+r_3^-}{2} \right\rfloor !} \leq$$

$$\leq \sum_{b_{N-1}} \frac{w_N^{e_N^-}}{e_N^-!} \cdots \sum_{b_3} \frac{w_4^{e_4^-}}{e_4^-!} \sum_{b_2} \frac{w_3^{\left\lfloor \frac{4b_3+r_4^-}{3} \right\rfloor - b_2}}{\left(\left\lfloor \frac{4b_3+r_4^-}{3} \right\rfloor - b_2\right)!} \max\{w_1, 1\} \frac{(w_2+w_1^2)^{\frac{3b_2+r_3^- - r_2^-}{2}}}{\frac{3b_2+r_3^- - r_2^-}{2}!}.$$

According to the definitions in 1.134, $\max\{w_1, 1\} = \max\{W_1^1, W_1^0\} = M_1$, and $w_2 + w_1^2 = W_2$. Therefore we can write the previous formula as:

$$M_1 \sum_{b_{N-1}} \frac{w_N^{e_N^-}}{e_N^-!} \cdots \sum_{b_3} \frac{w_4^{e_4^-}}{e_4^-!} \sum_{b_2} \frac{w_3^{\left\lfloor \frac{4b_3+r_4^-}{3} \right\rfloor - b_2}}{\left(\left\lfloor \frac{4b_3+r_4^-}{3} \right\rfloor - b_2\right)!} \frac{(W_2^{\frac{3}{2}})^{b_2}}{b_2!} \cdot W_2^{\frac{r_3^- - r_2^-}{2}}.$$

Since $-\frac{1}{2} \leq \frac{r_3^- - r_2^-}{2} \leq 1$ we have that $W_2^{\frac{r_3^- - r_2^-}{2}} \leq \max\{W_2^{-\frac{1}{2}}, W_2\} = M_2$, and

$$M_1 \sum_{b_{N-1}} \frac{w_N^{e_N^-}}{e_N^-!} \cdots \sum_{b_3} \frac{w_4^{e_4^-}}{e_4^-!} \sum_{b_2} \frac{w_3^{\left\lfloor \frac{4b_3+r_4^-}{3} \right\rfloor - b_2}}{(\left\lfloor \frac{4b_3+r_4^-}{3} \right\rfloor - b_2)!} \frac{(W_2^{\frac{3}{2}})^{b_2}}{b_2!} \cdot W_2^{\frac{r_3^- - r_2^-}{2}} \leq$$

$$\leq M_1 M_2 \sum_{b_{N-1}} \frac{w_N^{e_N^-}}{e_N^-!} \cdots \sum_{b_3} \frac{w_4^{e_4^-}}{e_4^-!} \frac{(w_3 + W_2^{\frac{3}{2}})^{\left\lfloor \frac{4b_3+r_4^-}{3} \right\rfloor}}{\left\lfloor \frac{4b_3+r_4^-}{3} \right\rfloor!} =$$

$$= M_1 M_2 \sum_{b_{N-1}} \frac{w_N^{e_N^-}}{e_N^-!} \cdots \sum_{b_3} \frac{w_4^{e_4^-}}{e_4^-!} \frac{W_3^{\left\lfloor \frac{4b_3+r_4^-}{3} \right\rfloor}}{\left\lfloor \frac{4b_3+r_4^-}{3} \right\rfloor!}.$$

In general, we take care of the sum over b_{i-1}, for $2 \leq i < N$, in the following way:

$$\prod_{j=1}^{i-2} M_j \sum_{b_{i-1}} \frac{w_i^{\left\lfloor \frac{(i+1)b_i+r_{i+1}^-}{i} \right\rfloor - b_{i-1}}}{\left(\left\lfloor \frac{(i+1)b_i+r_{i+1}^-}{i} \right\rfloor - b_{i-1}\right)!} \frac{W_{i-1}^{\left\lfloor \frac{ib_{i-1}+r_i^-}{i-1} \right\rfloor}}{\left\lfloor \frac{ib_{i-1}+r_i^-}{i-1} \right\rfloor!} =$$

$$= \prod_{j=1}^{i-2} M_j \sum_{b_{i-1}} \frac{w_i^{\left\lfloor \frac{(i+1)b_i+r_{i+1}^-}{i} \right\rfloor - b_{i-1}}}{\left(\left\lfloor \frac{(i+1)b_i+r_{i+1}^-}{i} \right\rfloor - b_{i-1}\right)!} \frac{W_{i-1}^{\frac{ib_{i-1}+r_i^- - r_{i-1}^-}{i-1}}}{\left\lfloor \frac{ib_{i-1}+r_i^-}{i-1} \right\rfloor!} \leq$$

$$\leq \prod_{j=1}^{i-2} M_j \sum_{b_{i-1}} \frac{w_i^{\left\lfloor \frac{(i+1)b_i+r_{i+1}^-}{i} \right\rfloor - b_{i-1}}}{\left(\left\lfloor \frac{(i+1)b_i+r_{i+1}^-}{i} \right\rfloor - b_{i-1}\right)!} \frac{(W_{i-1}^{\frac{i}{i-1}})^{b_{i-1}}}{b_{i-1}!} W_{i-1}^{\frac{r_i^- - r_{i-1}^-}{i-1}} \leq$$

$$\leq \prod_{j=1}^{i-2} M_j \sum_{b_{i-1}} \frac{w_i^{\left\lfloor \frac{(i+1)b_i+r^-_{i+1}}{i} \right\rfloor - b_{i-1}}}{\left(\left\lfloor \frac{(i+1)b_i+r^-_{i+1}}{i} \right\rfloor - b_{i-1}\right)!} \frac{(W_{i-1}^{\frac{i}{i-1}})^{b_{i-1}}}{b_{i-1}!} \max\{W_{i-1}, W_{i-1}^{-\frac{i-2}{i-1}}\} =$$

$$= \prod_{j=1}^{i-1} M_j \frac{(w_i + W_{i-1}^{\frac{i}{i-1}})^{\left\lfloor \frac{(i+1)b_i+r^-_{i+1}}{i} \right\rfloor}}{\left\lfloor \frac{(i+1)b_i+r^-_{i+1}}{i} \right\rfloor !} = \prod_{j=1}^{i-1} M_j \frac{W_i^{\left\lfloor \frac{(i+1)b_i+r^-_{i+1}}{i} \right\rfloor}}{\left\lfloor \frac{(i+1)b_i+r^-_{i+1}}{i} \right\rfloor !}.$$

For $i = N$ we apply the same reasoning:

$$\prod_{j=1}^{N-2} M_j \sum_{b_{N-1}} \frac{w_N^{z-b_{N-1}}}{(z-b_{N-1})!} \frac{W_{N-1}^{\left\lfloor \frac{Nb_{N-1}+r^-_N}{N-1} \right\rfloor}}{\left\lfloor \frac{Nb_{N-1}+r^-_N}{N-1} \right\rfloor !} =$$

$$\leq \prod_{j=1}^{N-2} M_j \sum_{b_{N-1}} \frac{w_N^{z-b_{N-1}}}{(z-b_{N-1})!} \frac{(W_{N-1}^{\frac{N}{N-1}})^{b_{N-1}}}{b_{N-1}!} W_{N-1}^{\frac{r^-_N - r^-_{N-1}}{N-1}} =$$

$$\leq \prod_{j=1}^{N-2} M_j \sum_{b_{N-1}} \frac{w_N^{z-b_{N-1}}}{(z-b_{N-1})!} \frac{(W_{N-1}^{\frac{N}{N-1}})^{b_{N-1}}}{b_{N-1}!} \max\{W_{N-1}, W_{N-1}^{-\frac{N-2}{N-1}}\} =$$

$$= \prod_{j=1}^{N-1} M_j \frac{(w_N + W_{N-1}^{\frac{N}{N-1}})^z}{z!}$$

$$= \prod_{j=1}^{N-1} M_j \frac{W_N^z}{z!}.$$

We can deal with the sum over a_{i-1}, for $2 \leq i \leq N$ in the same way. Bringing together the positive jump and negative jump parts, we obtain:

$$\begin{aligned}
\widetilde{Q}_N(k) &\leq \prod_{j=1}^{N-1} M_j^2 \sum_{r^-_N=0}^{N-1} \sum_{z} \frac{W_N^z}{z!} \frac{W_N^{t+z+\left\lfloor \frac{r_N+r^-_N}{N} \right\rfloor}}{\left(t+z+\left\lfloor \frac{r_N+r^-_N}{N} \right\rfloor\right)!} \\
&\leq \prod_{j=1}^{N-1} M_j^2 \sum_{z} \frac{W_N^z}{z!} \frac{W_N^{t+z}}{(t+z)!} \sum_{r^-_N=0}^{N-1} W_N^{\left\lfloor \frac{r_N+r^-_N}{N} \right\rfloor} \\
&\leq \prod_{j=2}^{N-1} M_j^2 \sum_{z} \frac{W_N^z}{z!} \sum_{z} \frac{W_N^{t+z}}{(t+z)!} \sum_{r^-_N=0}^{N-1} \max\{W_N, 1\}
\end{aligned}$$

$$\le N \max\{W_N, 1\} \prod_{j=2}^{N-1} M_j^2 e^{W_N} \cdot 2\frac{W_N^t}{t!}$$

for $t \ge 2W_N - 1$. Denoting $G = 2N \max\{W_N, 1\} \prod_{j=2}^{N-1} M_j^2 e^{W_N}$ we have Equation (1.140) for $k \ge N\lceil 2W_N - 1\rceil$.

Now we apply the Equation (1.140) to the summation $\sum_{k=\overline{k}}^{Nn} \widetilde{Q}_N(k)$ for $k \ge N\lceil 2W_N - 1\rceil$.

$$\begin{aligned}
\sum_{k=\overline{k}}^{Nn} \widetilde{Q}_N(k) &\le \sum_{k=\overline{k}}^{Nn} G\frac{W_N^{\left\lfloor \frac{k}{N}\right\rfloor}}{\left\lfloor \frac{k}{N}\right\rfloor!} \\
&\le 2GN\frac{W_N^{\left\lfloor \frac{\overline{k}}{N}\right\rfloor}}{\left\lfloor \frac{\overline{k}}{N}\right\rfloor!}
\end{aligned}$$

provided that $\overline{k} \ge N\lceil 2W_N - 1\rceil$.

Applying Equation (1.140) to the summation $\sum_{k=\overline{k}}^{Nn} e^{hk}\widetilde{Q}_N(k)$ for $k \ge N\lceil 2W_N - 1\rceil$, instead, we get:

$$\begin{aligned}
\sum_{k=\overline{k}}^{Nn} e^{hk}\widetilde{Q}_N(k) &\le \sum_{k=\overline{k}}^{Nn} e^{hk} G\frac{W_N^{\left\lfloor \frac{k}{N}\right\rfloor}}{\left\lfloor \frac{k}{N}\right\rfloor!} \\
&\le G \sum_{t=\left\lfloor \frac{\overline{k}}{N}\right\rfloor}^{Nn} \sum_{r=0}^{N-1} e^{hNt+hr}\frac{W_N^t}{t!} \\
&\le G \sum_{r=0}^{N-1} e^{hr} \sum_{t=\left\lfloor \frac{\overline{k}}{N}\right\rfloor}^{Nn} \frac{(e^{hN}W_N)^t}{t!} \\
&\le 2G\frac{(e^{hN}W_N)^{\left\lfloor \frac{\overline{k}}{N}\right\rfloor}}{\left\lfloor \frac{\overline{k}}{N}\right\rfloor!} \sum_{r=0}^{N-1} e^{hr}
\end{aligned}$$

for $\overline{k} \ge N\lceil 2e^{Nh}W_N - 1\rceil$.

Similarly, we obtain the analogous inequality for $\sum_{k=\overline{k}}^{Nn} e^{-hk}\widetilde{Q}_N(-k)$ with $\overline{k} \ge N\lceil 2W_N - 1\rceil$. □

Now let us consider the value $V(0,0,0)$ we obtain via backward recursion from Equation (1.96) with a fixed, arbitrary N, and $V^T(0,0,0)$ the value we obtain via backward recursion from Equation (1.103) for some $\overline{k}$, $\overline{l}$.

We can extend the result from Theorem 1.6.7 to the difference between $V(0,0,0)$ and $V^T(0,0,0)$.

Let W_i, M_i be as defined in Equations (1.134) and (1.135), and recall

$$G = 2N \max\{W_N, 1\} e^{W_N} \prod_{i=1}^{N-1} M_i^2.$$

Theorem 1.6.17. *Given $\varepsilon > 0$, taking*

$$\begin{aligned}
C^+ &= \sum_{r=0}^{N-1} e^{hr} + N \max\{W_N^2, 1\} e^{2hN} \sum_{r=0}^{N-1} e^{-hr} \\
C^- &= \sum_{r=0}^{N-1} e^{-hr} + N \max\{W_N^2, 1\} \sum_{r=0}^{N-1} e^{hr} \\
g^+ &= e^{hN+1} W_N + \log(4S_0 G) + (\alpha - r)\tau + \log C^+ \\
g^- &= e^{-hN+1} W_N + \log(4S_0 G) + (\alpha - r)\tau + \log C^-
\end{aligned}$$

and the value $V^T(0,0,0)$ obtained via truncation of the tree at levels $\overline{k}$ and $-\overline{l}$, where $\overline{k}$ and $\overline{l}$ are the smallest integers which satisfy:

$$\overline{k} \geq \max\{N \lceil g^+ - \log\varepsilon \rceil - 1, N \lceil 2e^{hN} W_N - 1 \rceil - 1\} \tag{1.142}$$

$$\overline{l} \geq \max\{N \lceil g^- - \log\varepsilon \rceil - 1, N \lceil 2e^{hN} W_N - 1 \rceil - 2, N \lceil 2W_N - 1 \rceil - 1\} \tag{1.143}$$

it holds true:

$$\left| V(0,0,0) - V^T(0,0,0) \right| < \varepsilon$$

Proof. Combining Equations (1.106) and (1.136) the difference between V_E and $\widehat{V_E^T}$ is less or equal than the sum of four discarded parts:

$$\begin{aligned}
V_E - \widehat{V_E^T} \leq e^{(\alpha-r)\tau} S_0 \Bigg(&\sum_{k=\overline{k}+1}^{Nn} e^{hk} \widetilde{Q}_N(k) + \sum_{k=\overline{l}+1}^{Nn} e^{-hk} \widetilde{Q}_N(k) + \\
&+ e^{h(2\overline{k}+2)} \sum_{s=\overline{k}+2}^{Nn} e^{-hs} \sum_{i=0}^{N-1} \widetilde{Q}_N(s+2i) +
\end{aligned}$$

$$+e^{h(-2\bar{l}-2)}\sum_{s=\bar{l}+2}^{Nn}e^{hs}\sum_{i=0}^{N-1}\widetilde{Q}_N(s+2i)\Bigg).$$

By Proposition 1.6.16:

$$V_E-\widehat{V_E^T}\le e^{(\alpha-r)\tau}S_0G\left(2\frac{(e^{hN}W_N)^{\lfloor\frac{\bar{k}+1}{N}\rfloor}}{\lfloor\frac{\bar{k}+1}{N}\rfloor!}\sum_{r=0}^{N-1}e^{hr}+2\frac{(e^{-hN}W_N)^{\lfloor\frac{\bar{l}+1}{N}\rfloor}}{\lfloor\frac{\bar{l}+1}{N}\rfloor!}\sum_{r=0}^{N-1}e^{-hr}+\right. \tag{1.144}$$

$$\left.+e^{h(2\bar{k}+2)}\sum_{s=\bar{k}+2}^{Nn}e^{-hs}\sum_{i=0}^{N-1}\frac{W_N^{\lfloor\frac{s+2j}{N}\rfloor}}{\lfloor\frac{s+2j}{N}\rfloor!}+e^{h(-2\bar{l}-2)}\sum_{s'=\bar{l}+2}^{Nn}e^{hs'}\sum_{i=0}^{N-1}\frac{W_N^{\lfloor\frac{s'+2j}{N}\rfloor}}{\lfloor\frac{s'+2j}{N}\rfloor!}\right) \tag{1.145}$$

where $G=2N\max\{W_N,1\}e^{W_N}\prod_{i=1}^{N-1}M_i^2$, for $\bar{k}\ge N\lceil 2e^{hN}W_N-1\rceil-1$ and $\bar{l}\ge N\lceil 2W_N-1\rceil-1$.

Since $\lfloor\frac{s}{N}\rfloor\le\lfloor\frac{s+2j}{N}\rfloor\le\lfloor\frac{s}{N}+2\rfloor$, we have that $\frac{W_N^{\lfloor\frac{s+2j}{N}\rfloor}}{\lfloor\frac{s+2j}{N}\rfloor!}\le\frac{W_N^{\lfloor\frac{s}{N}\rfloor}}{\lfloor\frac{s}{N}\rfloor!}\cdot\max\{W_N^2,1\}$:

$$\begin{aligned}
V_E-\widehat{V_E^T}\le{}& 2e^{(\alpha-r)\tau}S_0G\left(\frac{(e^{hN}W_N)^{\lfloor\frac{\bar{k}+1}{N}\rfloor}}{\lfloor\frac{\bar{k}+1}{N}\rfloor!}\sum_{r=0}^{N-1}e^{hr}+\frac{(e^{-hN}W_N)^{\lfloor\frac{\bar{l}+1}{N}\rfloor}}{\lfloor\frac{\bar{l}+1}{N}\rfloor!}\sum_{r=0}^{N-1}e^{-hr}\right)+\\
&+e^{(\alpha-r)\tau}S_0GN\max\{W_N^2,1\}\left(e^{h(2\bar{k}+2)}\sum_{s=\bar{k}+2}^{Nn}e^{-hs}\frac{W_N^{\lfloor\frac{s}{N}\rfloor}}{\lfloor\frac{s}{N}\rfloor!}+e^{h(-2\bar{l}-2)}\sum_{s=\bar{l}+2}^{Nn}e^{hs}\frac{W_N^{\lfloor\frac{s}{N}\rfloor}}{\lfloor\frac{s}{N}\rfloor!}\right)=\\
\le{}& 2e^{(\alpha-r)\tau}S_0G\left(\frac{(e^{hN}W_N)^{\lfloor\frac{\bar{k}+1}{N}\rfloor}}{\lfloor\frac{\bar{k}+1}{N}\rfloor!}\sum_{r=0}^{N-1}e^{hr}+\frac{(e^{-hN}W_N)^{\lfloor\frac{\bar{l}+1}{N}\rfloor}}{\lfloor\frac{\bar{l}+1}{N}\rfloor!}\sum_{r=0}^{N-1}e^{-hr}\right)+\\
&+e^{(\alpha-r)\tau}S_0GN\max\{W_N^2,1\}e^{h(2\bar{k}+2)}\sum_{r=0}^{N-1}e^{-hr}\sum_{t=\lfloor\frac{\bar{k}+2}{N}\rfloor}^{n}e^{-hNt}\frac{W_N^t}{t!}+\\
&+e^{(\alpha-r)\tau}S_0GN\max\{W_N^2,1\}e^{h(-2\bar{l}-2)}\sum_{r=0}^{N-1}e^{hr}\sum_{t=\lfloor\frac{\bar{l}+2}{N}\rfloor}^{n}e^{hNt}\frac{W_N^t}{t!}=\\
={}&2e^{(\alpha-r)\tau}S_0G\left(\frac{(e^{hN}W_N)^{\lfloor\frac{\bar{k}+1}{N}\rfloor}}{\lfloor\frac{\bar{k}+1}{N}\rfloor!}\sum_{r=0}^{N-1}e^{hr}+\frac{(e^{-hN}W_N)^{\lfloor\frac{\bar{l}+1}{N}\rfloor}}{\lfloor\frac{\bar{l}+1}{N}\rfloor!}\sum_{r=0}^{N-1}e^{-hr}\right)+
\end{aligned}$$

$$+ 2e^{(\alpha-r)\tau}S_0GN\max\{W_N^2,1\}\left(e^{h(2\bar{k}+2)}\frac{(e^{-hN}W_N)^{\left\lfloor\frac{\bar{k}+1}{N}\right\rfloor}}{\left\lfloor\frac{\bar{k}+1}{N}\right\rfloor!}\sum_{r=0}^{N-1}e^{-hr}\right.$$
$$\left.+e^{h(-2\bar{l}-2)}\frac{(e^{hs}W_N)^{\left\lfloor\frac{\bar{l}+1}{N}\right\rfloor}}{\left\lfloor\frac{\bar{l}+1}{N}\right\rfloor!}\sum_{r=0}^{N-1}e^{hr}\right)$$

for $\bar{k} \le N\lceil 2e^{hN}W_N - 1\rceil - 1$ and $\bar{l} \ge \max\{N\lceil 2W_N - 1\rceil - 1, N\lceil 2e^{hN}W_N - 1\rceil - 2\}$.

Since we also have $hs \le hN\left\lfloor\frac{s}{N}\right\rfloor + hN$ and $-hs \le -hN\left\lfloor\frac{s}{N}\right\rfloor$, we can write:

$$V_E - \widehat{V_E^T} \le 2e^{(\alpha-r)\tau}S_0G\left[\frac{(e^{hN}W_N)^{\left\lfloor\frac{\bar{k}+1}{N}\right\rfloor}}{\left\lfloor\frac{\bar{k}+1}{N}\right\rfloor!}\sum_{r=0}^{N-1}e^{hr} + \frac{(e^{-hN}W_N)^{\left\lfloor\frac{\bar{l}+1}{N}\right\rfloor}}{\left\lfloor\frac{\bar{l}+1}{N}\right\rfloor!}\sum_{r=0}^{N-1}e^{-hr}\right.$$
$$\left.+N\max\{W_N^2,1\}\left(e^{2hN}\frac{(e^{hN}W_N)^{\left\lfloor\frac{\bar{k}+1}{N}\right\rfloor}}{\left\lfloor\frac{\bar{k}+1}{N}\right\rfloor!}\sum_{r=0}^{N-1}e^{-hr} + \frac{(e^{-hN}W_N)^{\left\lfloor\frac{\bar{l}+1}{N}\right\rfloor}}{\left\lfloor\frac{\bar{l}+1}{N}\right\rfloor!}\sum_{r=0}^{N-1}e^{hr}\right)\right]$$
$$\le 2e^{(\alpha-r)\tau}S_0G\left[\frac{(e^{hN}W_N)^{\left\lfloor\frac{\bar{k}+1}{N}\right\rfloor}}{\left\lfloor\frac{\bar{k}+1}{N}\right\rfloor!}\left(\sum_{r=0}^{N-1}e^{hr} + N\max\{W_N^2,1\}e^{2hN}\sum_{r=0}^{N-1}e^{-hr}\right) +\right.$$
$$\left.+\frac{(e^{-hN}W_N)^{\left\lfloor\frac{\bar{l}+1}{N}\right\rfloor}}{\left\lfloor\frac{\bar{l}+1}{N}\right\rfloor!}\left(\sum_{r=0}^{N-1}e^{-hr} + N\max\{W_N^2,1\}\sum_{r=0}^{N-1}e^{hr}\right)\right].$$

In order to have the desired inequality, $V_E - \widehat{V_E^T} < \varepsilon$, we ask for:

$$\frac{(e^{hN}W_N)^{\left\lfloor\frac{\bar{k}+1}{N}\right\rfloor}}{\left\lfloor\frac{\bar{k}+1}{N}\right\rfloor!}\left(\sum_{r=0}^{N-1}e^{hr} + N\max\{W_N^2,1\}e^{2hN}\sum_{r=0}^{N-1}e^{-hr}\right) < \frac{\varepsilon}{4e^{(\alpha-r)\tau}S_0G}$$

and

$$\frac{(e^{-hN}W_N)^{\left\lfloor\frac{\bar{l}+1}{N}\right\rfloor}}{\left\lfloor\frac{\bar{l}+1}{N}\right\rfloor!}\left(\sum_{r=0}^{N-1}e^{-hr} + N\max\{W_N^2,1\}\sum_{r=0}^{N-1}e^{hr}\right) < \frac{\varepsilon}{4e^{(\alpha-r)\tau}S_0G}.$$

Let us call

$$C^+ = \sum_{r=0}^{N-1}e^{hr} + N\max\{W_N^2,1\}e^{2hN}\sum_{r=0}^{N-1}e^{-hr}$$

$$C^- = \sum_{r=0}^{N-1} e^{-hr} + N \max\{W_N^2, 1\} \sum_{r=0}^{N-1} e^{hr}.$$

With this notation our requests become:

$$\frac{(e^{hN} W_N)^{\left\lfloor \frac{\bar{k}+1}{N} \right\rfloor}}{\left\lfloor \frac{\bar{k}+1}{N} \right\rfloor !} < \frac{\varepsilon}{4e^{(\alpha-r)\tau} S_0 G C^+}$$

and

$$\frac{(e^{-hN} W_N)^{\left\lfloor \frac{\bar{l}+1}{N} \right\rfloor}}{\left\lfloor \frac{\bar{l}+1}{N} \right\rfloor !} < \frac{\varepsilon}{4e^{(\alpha-r)\tau} S_0 G C^-}.$$

Using Lemma 1.6.4 we impose:

$$e^{hN+1} W_N - \left\lfloor \frac{\bar{k}+1}{N} \right\rfloor \leq \log \varepsilon - \log(4S_0 G) - (\alpha - r)\tau - \log C^+$$

$$e^{-hN+1} W_N - \left\lfloor \frac{\bar{l}+1}{N} \right\rfloor \leq \log \varepsilon - \log(4S_0 G) - (\alpha - r)\tau - \log C^-$$

which means

$$\bar{k} \geq N \left\lceil g^+ - \log \varepsilon \right\rceil - 1$$

and

$$\bar{l} \geq N \left\lceil g^- - \log \varepsilon \right\rceil - 1$$

for

$$\begin{aligned} g^+ &= e^{hN+1} W_N + \log(4S_0 G) + (\alpha - r)\tau + \log C^+ \\ g^- &= e^{-hN+1} W_N + \log(4S_0 G) + (\alpha - r)\tau + \log C^- \end{aligned}$$

These conditions must be compared with those we needed in order to repeatedly apply Proposition 1.6.16, therefore:

$$\begin{aligned} \bar{k} &\geq \max\{N \left\lceil g^+ - \log \varepsilon \right\rceil - 1, N \left\lceil 2e^{hN} W_N - 1 \right\rceil - 1\} \\ \bar{l} &\geq \max\{N \left\lceil g^- - \log \varepsilon \right\rceil - 1, N \lceil 2W_N - 1 \rceil - 1, N \lceil 2e^{hN} W_N - 1 \rceil - 2\}. \end{aligned}$$

$\square$

The previous theorem allows to extend to arbitrary N the result of Theorem 1.6.8 and Theorem 1.6.9, which guarantee that for appropriately chosen $\overline{k}$ and $\overline{l}$, the value V_E of the European call option (or the value $V(0,0,0)$ obtained via the backward procedure, that is the same) can be approximated by

$$\widehat{V_E^T} = e^{-r\tau} \sum_{k=-\overline{l}}^{\overline{k}} \sum_{j=0}^{n} (S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk} - K_0)^+ P(j) \widehat{Q}_1(k)$$

and with the value $V^T(0,0,0)$ obtained through the backward truncated procedure.

Theorem 1.6.18. *Given $\varepsilon = \frac{1}{n} > 0$, and $\overline{k}$ and $\overline{l}$ the smallest integers as in Theorem 1.6.17, the proposed procedure for $\widehat{V_E^T}$ converges to the HS price and its computational complexity is $O(n \log n)$.*

Theorem 1.6.19. *Given $\varepsilon = \frac{1}{n} > 0$, and $\overline{k}$ and $\overline{l}$ the smallest integers as in Theorem 1.6.17, the backward procedure described above for $V^T(0,0,0)$ converges to the HS price and its computational complexity is $O(n^2 \log n)$.*

The proofs of Theorems 1.6.18 and 1.6.19 coincide with those of Theorems 1.6.8 and 1.6.9.

1.6.2 European put options

European put option, $N = 1$

With similar notation to the call case, we take:

$$P_E = e^{-r\tau} \sum_{k=-n}^{n} \sum_{j=0}^{n} (K_0 - S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk})^+ P(j) Q_1(k)$$

$$P_E^T = e^{-r\tau} \sum_{k=-\overline{l}}^{\overline{k}} \sum_{j=0}^{n} (K_0 - S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk})^+ P(j) Q_1(k)$$

$$\widehat{P_E^T} = e^{-r\tau} \sum_{k=-\overline{l}}^{\overline{k}} \sum_{j=0}^{n} (K_0 - S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk})^+ P(j) \widehat{Q}_1(k)$$

and we consider separately the differences $P_E - P_E^T$ and $P_E^T - \widehat{P_E^T}$.

Discarding the value of the underlying, we obtain:

$$\begin{aligned}
P_E - P_E^T =& e^{-r\tau} \sum_{k=-n}^{-\bar{l}-1} \sum_{j=0}^{n} (K_0 - S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk})^+ P(j) Q_1(k) + \\
&+ e^{-r\tau} \sum_{k=\bar{k}+1}^{n} \sum_{j=0}^{n} (K_0 - S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk})^+ P(j) Q_1(k) \\
&\le e^{-r\tau} K_0 \sum_{k=\bar{k}+1}^{n} Q_1(k) \sum_{j=0}^{n} P(j) + e^{-r\tau} K_0 \sum_{k=\bar{l}+1}^{n} Q_1(k) \sum_{j=0}^{n} P(j) \\
&\le e^{-r\tau} K_0 \left(\sum_{k=\bar{k}+1}^{n} \widetilde{Q}_1(k) + \sum_{k=\bar{l}+1}^{n} \widetilde{Q}_1(k) \right) \qquad (1.146)
\end{aligned}$$

while Lemma 1.6.5 gives

$$\begin{aligned}
P_E^T - \widehat{P_E^T} &\le e^{-r\tau} K_0 \left(\sum_{k=-\bar{l}}^{\bar{k}} \widetilde{Q}_1(2\bar{k} + 2 - k) + \sum_{k=-\bar{l}}^{\bar{k}} \widetilde{Q}_1(-2\bar{l} - 2 - k)) \right) \\
&\le e^{-r\tau} K_0 \left(\sum_{s=\bar{k}+2}^{\min\{2\bar{k}+\bar{l},n\}} \widetilde{Q}_1(s) + \sum_{s=\bar{l}+2}^{\min\{2\bar{l}+\bar{k},n\}} \widetilde{Q}_1(s) \right) \qquad (1.147)
\end{aligned}$$

Combining Equations (1.146) and (1.147) we obtain:

$$P_E - \widehat{P_E^T} \le e^{-r\tau} K_0 \left(\sum_{k=\bar{k}+1}^{n} \widetilde{Q}_1(k) + \sum_{k=\bar{l}+1}^{n} \widetilde{Q}_1(k) + \sum_{k=\bar{k}+2}^{n} \widetilde{Q}_1(k) + \sum_{k=\bar{l}+2}^{n} \widetilde{Q}_1(k) \right) \qquad (1.148)$$

which can be brought to

$$P_E - \widehat{P_E^T} \le 4e^{-r\tau} K_0 \sum_{k=\bar{k}+1}^{n} \widetilde{Q}_1(k)$$

taking $\bar{l} = \bar{k}$.

Applying Proposition 1.6.6 we obtain that, when $\bar{k} \ge 2w - 1$, provided that

$$4e^w \frac{w^{\bar{k}+1}}{(\bar{k}+1)!} \le \frac{\varepsilon}{4e^{-r\tau} K_0},$$

we have

$$4e^{-r\tau}K_0 \sum_{k=\overline{k}+1}^{n} \widetilde{Q}_1(k) \leq \varepsilon.$$

By Lemma 1.6.4 this means that $\overline{k}$ is the appropriate level for the cut when $-\overline{k}-1+ew \leq \log \frac{\varepsilon}{16e^{-r\tau+w}K_0}$ i.e.

$$\overline{k} \geq \max\{-1+ew-\log\varepsilon+\log(16K_0)-r\tau+w, 2w-1\}$$

European put option, arbitrary N

With the same notation as before,

$$P_E = e^{-r\tau} \sum_{k=-n\cdot N}^{n\cdot N} \sum_{j=0}^{n} (K_0 - S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk})^+ P(j) Q_N(k)$$

$$P_E^T = e^{-r\tau} \sum_{k=-\overline{l}}^{\overline{k}} \sum_{j=0}^{n} (K_0 - S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk})^+ P(j) Q_N(k).$$

and

$$\widehat{P_E^T} = e^{-r\tau} \sum_{k=-\overline{l}}^{\overline{k}} \sum_{j=0}^{n} (K_0 - S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk})^+ P(j) \widehat{Q}_N(k)$$

Discarding the value of the underlying, as we did in the $N=1$ case, we obtain:

$$P_E - P_E^T \leq e^{-r\tau} K_0 \left(\sum_{k=\overline{k}+1}^{Nn} \widetilde{Q}_N(k) + \sum_{k=\overline{l}+1}^{Nn} \widetilde{Q}_N(k) \right) \tag{1.149}$$

while recalling that $Q_1^{\overline{k}}(k)$ and $Q_{1\overline{l}}(k)$ are the probabilities of reaching at maturity the "jump level" k trespassing level $\overline{k}$ and $-\overline{l}$ respectively, we can write

$$P_E^T - \widehat{P_E^T} = \tag{1.150}$$

$$= e^{-r\tau} \sum_{k=-\overline{l}}^{\overline{k}} \sum_{j=0}^{n} (K_0 - S_0 e^{(-n+2j)\sigma\sqrt{\Delta t}+hk})^+ P(j) (Q_N(k) - \widehat{Q}_N(k)) \tag{1.151}$$

$$\leq e^{-r\tau} K_0 \sum_{k=-\overline{l}}^{\overline{k}} (Q_N(k) - \widehat{Q}_N(k)) \tag{1.152}$$

$$\leq e^{-r\tau} K_0 \sum_{k=-\bar{l}}^{\bar{k}} (Q_N^{\bar{k}}(k) + Q_{N\bar{l}}(k)) \tag{1.153}$$

$$\leq e^{-r\tau} K_0 \left(\sum_{k=-\bar{l}}^{\bar{k}} \sum_{i=1}^{N} \widetilde{Q}_N(2\overline{k} + 2i - k) + \sum_{k=-\bar{l}}^{\bar{k}} \sum_{i=1}^{N} \widetilde{Q}_N(2\bar{l} + 2i + k)) \right)$$

$$\leq e^{-r\tau} K_0 \left(\sum_{s=\bar{k}+2}^{\min\{2\bar{k}+\bar{l}+2, Nn\}} \sum_{i=0}^{N-1} \widetilde{Q}_N(s + 2i) + \sum_{s=\bar{l}+2}^{\min\{2\bar{l}+\bar{k}+2, Nn\}} \sum_{i=0}^{N-1} \widetilde{Q}_N(s + 2i) \right)$$

$$\leq e^{-r\tau} K_0 N \left(\sum_{s=\bar{k}+2}^{Nn} \widetilde{Q}_N(s) + \sum_{s=\bar{l}+2}^{Nn} \widetilde{Q}_N(s) \right). \tag{1.154}$$

Combining Equations (1.149) and (1.154) we obtain:

$$P_E - \widehat{P_E^T} \leq e^{-r\tau} K_0 (N+1) \left(\sum_{k=\bar{k}+1}^{Nn} \widetilde{Q}_N(k) + \sum_{k=\bar{l}+1}^{Nn} \widetilde{Q}_N(k) \right) \tag{1.155}$$

$$\leq 2e^{-r\tau} K_0 (N+1) \sum_{k=\bar{k}+1}^{Nn} \widetilde{Q}_N(k) \tag{1.156}$$

supposing we take $\bar{l} = \bar{k}$.

As in the call case, we can compute (1.155) with a $O(n^2)$ procedure, thus determining numerically the largest integers $\bar{l}$ and $\bar{k}$ such that the loss is inferior to an arbitrary ε.[6]

For the theoretical bounds we need the following result.

Theorem 1.6.20. *Given $\varepsilon > 0$, taking $G = 2N \max\{W_N, 1\} \prod_{i=1}^{N-1} M_i^2 e^{W_N}$, the value $\widehat{P_E^T}$ obtained via truncation of the tree at levels $\bar{k}$ and $-\bar{k}$, with $\bar{k}$ the smallest integer which satisfies:*

$\bar{k} \geq \max\{N\lceil 2W_N - 1\rceil - 1, N\lceil W_N e - \log\varepsilon - r\tau + \log(4N(N+1)K_0 G)\rceil - 1\}$, *we have*

$$\left| P_E - \widehat{P_E^T} \right| < \varepsilon$$

Proof. Applying Equation (1.141) to Equation (1.156) we obtain:

[6] Given ε, we consider $\eta = \frac{\varepsilon}{2e^{-r\tau} K_0(N+1)}$ and we compute, starting from $i = Nn$ and proceeding backwards, the sum $S_i = \sum_{k=i}^{Nn} \widetilde{Q}_N(k)$. While $S_i < \eta$, we keep decreasing i. The first i we encounter such that $S_i \geq \eta$ is our $\bar{k} = \bar{l}$.

$$
\begin{aligned}
P_E - \widehat{P_E^T} &\leq 2e^{-r\tau} K_0 (N+1) \sum_{k=\overline{k}+1}^{Nn} \widetilde{Q}_N(k) \\
&\leq 4e^{-r\tau} K_0 (N+1) GN \frac{W_N^{\left\lfloor \frac{\overline{k}+1}{N} \right\rfloor}}{\left\lfloor \frac{\overline{k}+1}{N} \right\rfloor !}
\end{aligned}
$$

for $\overline{k} \geq N\lceil 2W_N - 1 \rceil - 1$.

We ask $\overline{k} \geq N\lceil W_N e - \log\varepsilon - r\tau + \log(4N(N+1)KG) \rceil - 1$, in order to have

$$
4e^{-r\tau} K_0 (N+1) GN \frac{W_N^{\left\lfloor \frac{\overline{k}+1}{N} \right\rfloor}}{\left\lfloor \frac{\overline{k}+1}{N} \right\rfloor !} < \varepsilon.
$$

Collecting all the requirements on $\overline{k}$, we get that for

$\overline{k} \geq \max\{N\lceil 2W_N - 1 \rceil - 1, N\lceil W_N e - \log\varepsilon - r\tau + \log(4N(N+1)KG) \rceil - 1\}$ we have

$$
|P_E - \widehat{P_E^T}| < \varepsilon.
$$

□

We can then state the following results.

Theorem 1.6.21. *Given $\varepsilon = \frac{1}{n} > 0$, and $\overline{k}$ and $\overline{l}$ the smallest integers as in Theorem 1.6.20, the proposed procedure for $\widehat{P_E^T}$ converges to the HS price and its computational complexity is $O(n \log n)$.*

Theorem 1.6.22. *Given $\varepsilon = \frac{1}{n} > 0$, and $\overline{k}$ and $\overline{l}$ the smallest integers as in Theorem 1.6.20, the backward procedure described above for $P^T(0,0,0)$ converges to the HS price and its computational complexity is $O(n^2 \log n)$.*

The proofs of Theorems 1.6.21 and 1.6.22 coincide with those of Theorems 1.6.8 and 1.6.9.

1.6.3 American case

In the following, we extend the results on the backward procedure for the evaluation of the European derivatives to the American put option pricing, by showing that the truncation error for the American case must be less or equal then the error in the European case.

We define the value $V^K(i,j,k)$ obtained via backward procedure according to the following formula: $V^K(i,j,k) = e^{-r\Delta t}\sum_{l=-N}^{N}(V^K(i+1,j+1,k+l)p + V^K(i+1,j,k+l)(1-p))q_l$ if $k \in [-\bar{l},\bar{k}]$, K_0 otherwise; with initial data $V^K(n,j,k) = (S(n,j,k)-K_0)^+$, for j integer between 0 and n and k integer such that $-\bar{l} \le k \le \bar{k}$, $V^K(n,j,k) = K_0$ for j integer between 0 and n and k integer such that $-nN \le k \le -\bar{l}-1$ or $\bar{k}+1 \le k \le nN$.

We also consider value $\widehat{V_E^K}$, defined as

$$\widehat{V_E^K} = \widehat{V_E^T} + \sum_{\text{paths that reach } \tau \text{ and trespass}} \text{prob(path)} \cdot K_0 e^{-r\Delta t \mathrm{i(path)}} \tag{1.157}$$

where prob(path) identifies the probability of a single path and i(path) the time $0 < i \le n$ of the first exit of the path from the allowed zone.

For the put options, similarly we call $P^K(i,j,k)$ the value obtained via backward procedure according to the following formula:

$P^K(i,j,k) = e^{-r\Delta t}\sum_{l=-N}^{N}(P^K(i+1,j+1,k+l)p + P^K(i+1,j,k+l)(1-p))q_l$ if $k \in [-\bar{l},\bar{k}]$, K_0 otherwise; with initial data $P^K(n,j,k) = (K_0 - S(n,j,k))^+$, for j integer between 0 and n and k integer such that $-\bar{l} \le k \le \bar{k}$, $P^K(n,j,k) = K_0$ for j integer between 0 and n and k integer such that $-nN \le k \le -\bar{l}-1$ or $\bar{k}+1 \le k \le nN$, and $\widehat{P_E^K}$ the value

$$\widehat{P_E^K} = \widehat{P_E^T} + \sum_{\text{paths that reach } \tau \text{ and trespass}} \text{prob(path)} \cdot K_0 e^{-r\Delta t \mathrm{i(path)}}. \tag{1.158}$$

Lemma 1.6.23. *$V^K(0,0,0) = \widehat{V_E^K}$ and $P^K(0,0,0) = \widehat{P_E^K}$*

Proof. Since from now on we will focus on the put option, we will write the proof for the put case for induction on the number of steps n; for the call options the proof is analogous.

Let us start with $n = 1$. Our tree has only one step, which means that the values at maturity of the option are given by the $2(2N+1)$ children of $(0,0,0)$. $\Delta t = \tau$. Let $0 \le \bar{l},\bar{k} \le N$, that means that $(0,0,0)$ is surely in the allowed zone, while some of its children may be not. Since the value of the option on the nodes $(1,j,k)$ with $k \notin [-\bar{l},\bar{k}]$ is K_0, we can write:

$$P^K(0,0,0) = e^{-r\tau}\sum_{l=-N}^{N}(P^K(1,j+1,l)p + P^K(1,j,l)(1-p))q_l =$$

$$=e^{-r\tau}\sum_{l=-\bar{l}}^{\bar{k}}(P^K(1,j+1,l)pq_l + P^K(1,j,l)(1-p)q_l) + e^{-r\tau}\sum_{l=-N}^{-\bar{l}-1} K_0 + e^{-r\tau}\sum_{l=\bar{k}+1}^{N} K_0 =$$

$$=e^{-r\tau} \sum_{\text{paths that reach } \tau \text{ and do not trespass}} \text{prob(path)} \cdot \text{value(path)} +$$

$$+ \sum_{\text{paths that reach } \tau \text{ and trespass}} \text{prob(path)} \cdot K_0 e^{-r\Delta t \mathrm{i}(\text{path})}$$

$$=\widehat{P_E^T} +$$

$$+ \sum_{\text{paths that reach } \tau \text{ and trespass}} \text{prob(path)} \cdot K_0 e^{-r\Delta t \mathrm{i}(\text{path})}$$

where we take into account the fact that in a single step the paths that trespass are those that end outside the boundaries.

Let us now suppose the thesis is true for all trees in $n-1$ steps. Let us consider a tree of n steps. $\Delta t = \tau/n$. We focus on the first step and compute the value of $P^K(0,0,0)$ with the backward procedure: $P^K(0,0,0) = e^{-r\Delta t}\sum_{l=-N}^{N}(P^K(1,1,l)p + P^K(1,0,l)(1-p))q_l$.

If $l \notin [-\bar{l},\bar{k}]$, $P^K(1,1,l) = P^K(1,0,l) = K_0$. Otherwise, we can consider the $n-1$ tree that starts at $(1,j,l)$ for $j=0,1$ and $l \notin [-\bar{l},\bar{k}]$ and ends at maturity τ. We apply induction and write that the value $P^K(1,j,l)$ for this smaller tree is given by

$$P^K(1,j,l) =$$

$$=e^{-r\tau'} \sum_{\substack{\text{paths that go from } (1,j,l) \text{ to } \tau \\ \text{and do not trespass}}} \text{prob(path')} \cdot \text{value(path')} +$$

$$+ \sum_{\substack{\text{paths that go from } (1,j,l) \text{ to } \tau \\ \text{and trespass}}} \text{prob(path')} \cdot K_0 e^{-r\Delta t \mathrm{i}(\text{path'})}$$

where τ' indicates $\tau' = \tau - \Delta t$, $\Delta t' = \tau'/(n-1)$ and path' indicates a generic path going from $(1,j,l)$ to τ. Therefore

$$P^K(0,0,0) =$$

$$=e^{-r\Delta t}\sum_{\substack{l=-N \\ l\in[-\bar{l},\bar{k}]}}^{N}(P^K(1,1,l)p + P^K(1,0,l)(1-p))q_l + e^{-r\Delta t}\sum_{\substack{l=-N \\ l\notin[-\bar{l},\bar{k}]}}^{N} K_0 q_l =$$

$$
\begin{aligned}
=&e^{-r\tau}\sum_{l=-N}^{N}\Bigg(pq_l\sum_{\substack{\text{paths that go from }(1,1,l)\text{ to }\tau\\ \text{and do not trespass}}}\text{prob(path')}\cdot\text{value(path')}+\\
&+pq_l\sum_{\substack{\text{paths that go from }(1,1,l)\text{ to }\tau\\ \text{and trespass}}}\text{prob(path')}\cdot K_0e^{-r\Delta t\mathrm{i}(\text{path'})}+\\
&+(1-p)q_l\sum_{\substack{\text{paths that go from }(1,0,l)\text{ to }\tau\\ \text{and do not trespass}}}\text{prob(path')}\cdot\text{value(path')}+\\
&+(1-p)q_l\sum_{\substack{\text{paths that go from }(1,0,l)\text{ to }\tau\\ \text{and trespass}}}\text{prob(path')}\cdot K_0e^{-r\Delta t\mathrm{i}(\text{path'})}\Bigg)+\\
&+e^{-r\Delta t}\sum_{\substack{l=-N\\ l\notin[-\bar{l},\bar{k}]}}^{N}K_0q_l=\\
=&e^{-r\tau}\sum_{\substack{l=-N\\ l\in[-\bar{l},\bar{k}]}}^{N}\Big(pq_lP^T(1,j,l)+(1-p)q_lP^T(1,0,l)+\\
&+pq_l\sum_{\substack{\text{paths that go from }(1,1,l)\text{ to }\tau\\ \text{and trespass}}}\text{prob(path')}\cdot K_0e^{-r\Delta t\mathrm{i}(\text{path'})}+\\
&+(1-p)q_l\sum_{\substack{\text{paths that go from }(1,0,l)\text{ to }\tau\\ \text{and trespass}}}\text{prob(path')}\cdot K_0e^{-r\Delta t\mathrm{i}(\text{path'})}\Bigg)+\\
&+e^{-r\Delta t}\sum_{\substack{l=-N\\ l\notin[-\bar{l},\bar{k}]}}^{N}K_0q_l
\end{aligned}
$$

where we used Lemma 1.6.1 in order to use the values $P^T(1,j,l)$. Now we consider a path starting from the node $(0,0,0)$, visiting node $(1,j,l)$ and reaching maturity trespassing the barriers, following a path which we call path'. If

$j=0$ then prob(path)$=(1-p)q_l$ prob(path'), while if $j=1$ prob(path)$=pq_l$ prob(path'). If $l \notin [-\overline{l}, \overline{k}]$, then i(path)=1, otherwise i(path)=i(path')+1. This means we can write

$$\begin{aligned} P^K(0,0,0) &= P^T(0,0,0)+ \\ &+ \sum_{\substack{\text{paths that trespass} \\ \text{after first step}}} \text{prob(path)} \cdot K_0 e^{-r\Delta t \text{i(path)}} + \\ &+ \sum_{\substack{\text{paths that trespass} \\ \text{in the first step}}} \text{prob(path)} \cdot K_0 e^{-r\Delta t \text{i(path)}} = \\ &= \widehat{P_E^K} \end{aligned}$$

□

Substituting 0 with K_0 in the nodes above the barrier increases the value of the option, therefore we have $V^K(0,0,0) \geq V^T(0,0,0)$ and $P^K(0,0,0) \geq P^T(0,0,0)$. Since for the put options we also have $P(i,j,k) \leq K_0$ for every (i,j,k), we have that $P^K(0,0,0) \geq P(0,0,0) \geq P^T(0,0,0)$.

Therefore

$$|P^K(0,0,0) - P(0,0,0)| \leq |P^K(0,0,0) - P^T(0,0,0)| = |\widehat{P_E^K} - \widehat{P_E^T}|.$$

In order to control $\widehat{P_E^K} - P_E$ we only need to control $\widehat{P_E^K} - \widehat{P_E^T}$.

Theorem 1.6.24. *Given $\varepsilon > 0$, taking $G = 2N \max\{W_N, 1\} \prod_{i=1}^{N-1} M_i^2 e^{W_N}$, the values $P^K(0,0,0)$ and $P^T(0,0,0)$ obtained via truncation of the tree at levels $\overline{k}$ and $-\overline{k}$, with $\overline{k}$ the smallest integer which satisfies:*

$\overline{k} \geq \max\{N\lceil 2W_N - 1\rceil - 1, N\lceil W_N e - \log\varepsilon + \log(4N(N+1)K_0 G)\rceil - 1\}$, *we have*

$$\left|P^K(0,0,0) - P^T(0,0,0)\right| < \varepsilon$$

Proof.

$$P^K(0,0,0) - P^T(0,0,0) = \sum_{\text{paths that reach } \tau \text{ and trespass}} \text{prob(path)} \cdot K_0 e^{-r\Delta t \text{i(path)}}$$

Therefore:

$$P^K(0,0,0) - P^T(0,0,0) \le \sum_{\text{paths that exit } [-\bar{l},\bar{k}]} \text{prob(path)} \cdot K_0$$

$$\le K_0 \sum_{k=-Nn}^{Nn} \sum_{\substack{\text{paths that exit } [-\bar{l},\bar{k}] \\ \text{and end up in } k}} \text{prob(path)}$$

$$\le K_0 \sum_{k=-Nn}^{-\bar{l}-1} \sum_{\substack{\text{paths that exit } [-\bar{l},\bar{k}] \\ \text{and end up in } k}} \text{prob(path)} + K_0 \sum_{k=-\bar{l}}^{\bar{k}} \sum_{\substack{\text{paths that exit } [-\bar{l},\bar{k}] \\ \text{and end up in } k}} \text{prob(path)}+$$

$$+ K_0 \sum_{k=\bar{k}+1}^{Nn} \sum_{\substack{\text{paths that exit } [-\bar{l},\bar{k}] \\ \text{and end up in } k}} \text{prob(path)}$$

$$\le K_0 \sum_{k=-Nn}^{-\bar{l}-1} \sum_{\text{paths that end up in } k} \text{prob(path)} + K_0 \sum_{k=-\bar{l}}^{\bar{k}} \sum_{\substack{\text{paths that exit } [-\bar{l},\bar{k}] \\ \text{and end up in } k}} \text{prob(path)}+$$

$$+ K_0 \sum_{k=\bar{k}+1}^{Nn} \sum_{\text{paths that end up in } k} \text{prob(path)}$$

$$\le K_0 \sum_{k=-Nn}^{-\bar{l}-1} Q_N(k) + K_0 \sum_{k=\bar{k}+1}^{Nn} Q_N(k)+$$

$$+ K_0 \sum_{k=-\bar{l}}^{\bar{k}} \sum_{\substack{\text{paths that exit } -\bar{l} \\ \text{and end up in } k}} \text{prob(path)} + K_0 \sum_{k=-\bar{l}}^{\bar{k}} \sum_{\substack{\text{paths that exit } \bar{k} \\ \text{and end up in } k}} \text{prob(path)}$$

$$\le K_0 \sum_{k=\bar{l}+1}^{Nn} \widetilde{Q}_N(k) + K_0 \sum_{k=\bar{k}+1}^{Nn} \widetilde{Q}_N(k) + K_0 \sum_{k=-\bar{l}}^{\bar{k}} Q_{N\bar{l}}(k) + K_0 \sum_{k=-\bar{l}}^{\bar{k}} {Q_N}^{\bar{k}}(k).$$

Therefore we have

$$P^K(0,0,0) - P^T(0,0,0) \le K_0(N+1) \left(\sum_{k=\bar{l}+1}^{Nn} \widetilde{Q}_N(k) + \sum_{k=\bar{k}+1}^{Nn} \widetilde{Q}_N(k) \right)$$

$$\leq 2K_0(N+1)\sum_{k=\overline{k}+1}^{Nn}\widetilde{Q}_N(k)$$

$$\leq 4K_0(N+1)GN\frac{W_N^{\left\lfloor\frac{\overline{k}+1}{N}\right\rfloor}}{\left\lfloor\frac{\overline{k}+1}{N}\right\rfloor!}$$

for $\overline{l}=\overline{k}\geq N\lceil 2W_N-1\rceil-1$ and applying Equation (1.141).

We ask $\overline{k}\geq N\lceil W_N e-\log\varepsilon+\log(4N(N+1)K_0G)\rceil-1$, in order to have

$$4K_0(N+1)GN\frac{W_N^{\left\lfloor\frac{\overline{k}+1}{N}\right\rfloor}}{\left\lfloor\frac{\overline{k}+1}{N}\right\rfloor!}<\varepsilon.$$

Collecting all the requirements on $\overline{k}$, we get that for $\overline{k}\geq\max\{N\lceil 2W_N-1\rceil-1, N\lceil W_N e-\log\varepsilon+\log(4N(N+1)K_0G)\rceil-1\}$ we have

$$\left|P^K(0,0,0)-P^T(0,0,0)\right|<\varepsilon.$$

□

Theorem 1.6.25. *Let $P_A=P_A(0,0,0)$ the binomial price, evaluated with the backward procedure, in the American put case. Fixed $\overline{k}$ and $\overline{l}$, let $P_A^K=P_A^K(0,0,0)$ the binomial price, evaluated with the backward procedure with the truncation and substitution of the value outside the allowed zone with the strike, in the American case. Let $\widehat{P_E^K}$ and P_E the European prices as above. One has:*

$$|P_A^K-P_A|\leq|\widehat{P_E^K}-P_E|.$$

Proof. We consider $P^K(i,j,l)$, $P_A^K(i,j,l)$, $j=0,...,i$, the backward procedures which have value K_0 for l under the $-\overline{l}$ and over the $\overline{k}$ barrier for all the time steps i, running from 0 to n. We consider $P(i,j,l)$, $P_A(i,j,l)$, $j=0,...,i$ the standard HS backward procedures.

We claim that

$$|P_A^K(i,j,l)-P_A(i,j,l)|\leq|P^K(i,j,l)-P(i,j,l)|$$

for all i,j,l.

Since the American put price is not smaller than the European price and it is not larger than K_0, out of the barriers the claim is true. Inside the boundaries, we prove it for induction on the step i.

Let $i = n$. On all the nodes at maturity the error is the same, since $P_A^K(i,j,l) = P^K(i,j,l)$ and $P_A(i,j,l) = P(i,j,l)$.

Consider now the case $i-1$.

We set the continuation value $C_A^K(i-1,j,l) = e^{-r\Delta t}\sum_{k=-N}^{N}(P_A^K(i,j+1,l+k)p + P_A^K(i,j,l+k)(1-p))q_k$ and $C_A(i-1,j,l)$, $C_E^K(i-1,j,l)$, $C_E(i-1,j,l)$, similarly. Consider the nodes inside the barriers. The truncation value $P_A^K(i-1,j,l)$ is then given by

$$P_A^K(i-1,j,l) = \max\left[C_A^K(i-1,j,l), K_0 - S(i-1,j,l)\right] \tag{1.159}$$

One has $C_A(i,j,l) \le C_A^K(i,j,l)$ for every i,j,l.

Only the following cases are possible:

- $C_A^K(i-1,j,l) \le K_0 - S(i-1,j,l)$

  This means $P_A^K(i-1,j,l) = K_0 - S(i-1,j,l) = P_A(i-1,j,l)$, and $|P_A^K(i-1,j,l) - P_A(i-1,j,l)| = 0$.

- $C_A(i-1,j,l) \le K_0 - S(i-1,j,l)$ and $C_A^K(i-1,j,l) \ge K_0 - S(i-1,j,l)$

  This means $P_A(i-1,j,l) = K_0 - S(i-1,j,l)$ and $P_A^K(i-1,j,l) = C_A^K(i-1,j,l)$, and

$$\begin{aligned}
&|P_A^K(i-1,j,l) - P_A(i-1,j,l)| = C_A^K(i-1,j,l) - (K_0 - S(i-1,j,l))\\
&\le C_A^K(i-1,j,l) - C_A(i-1,j,l)\\
\le& e^{-r\Delta t}\sum_{k=-N}^{N}(P_A^K(i,j+1,l+k)p + P_A^K(i,j,l+k)(1-p))q_k +\\
&- e^{-r\Delta t}\sum_{k=-N}^{N}(P_A(i,j+1,l+k)p + P_A(i,j,l+k)(1-p))q_k\\
=& e^{-r\Delta t}\sum_{k=-N}^{N}[(P_A^K(i,j+1,l+k) - P_A(i,j+1,l+k)]pq_k +\\
&+ e^{-r\Delta t}\sum_{k=-N}^{N}[P_A^K(i,j,l+k) - P_A(i,j,l+k)](1-p)q_k
\end{aligned}$$

  Either the nodes $(i,j+1,l+k)$ and $(i,j,l+k)$ are outside the boundaries, and then the claim is true, or we can use induction, therefore: $P_A^K(i,j+$

$1, l+k) - P_A(i, j+1, l+k) \leq P^K(i, j+1, l+k) - P(i, j+1, l+k)$, and $P_A^K(i, j, l+k) - P_A(i, j, l+k) \leq P^K(i, j, l+k) - P(i, j, l+k)$. Hence:

$$\begin{aligned}
&|P_A^K(i-1, j, l) - P_A(i-1, j, l)| \leq \\
\leq& e^{-r\Delta t} \sum_{k=-N}^{N} [(P^K(i, j+1, l+k) - P(i, j+1, l+k)]pq_k \\
&+ e^{-r\Delta t} \sum_{k=-N}^{N} [P^K(i, j, l+k) - P(i, j, l+k)](1-p)q_k \\
\leq& e^{-r\Delta t} \sum_{k=-N}^{N} (P^K(i, j+1, l+k)p + P(i, j, l+k)(1-p))q_k + \\
&- e^{-r\Delta t} \sum_{k=-N}^{N} (P^K(n, j+1, l+k)p + P(i, j, l+k)(1-p))q_k \\
\leq& P^K(i, j, l) - P(i, j, l).
\end{aligned}$$

- $C_A(i-1, j, l) \geq K_0 - S(i-1, j, l)$
  This means $P_A^K(i-1, j, l) = C_A^K(i-1, j, l)$ and $P_A(i-1, j, l) = C_A(i-1, j, l)$, and

$$|P_A(i-1, j, l) - P_A^K(i-1, j, l)| = C_A^K(i-1, j, l) - C_A(i-1, j, l)$$

which we have already considered in the previous case.

□

Hence, by Theorem 1.6.9 we can state the following analogous result in the American case.

Theorem 1.6.26. *Given $\varepsilon = \frac{1}{n} > 0$, and $\overline{k} = \bar{l}$ the smallest integer as in Theorem 1.6.24, the backward procedure described in Theorem 1.6.25 for $P_A^K(0,0,0)$ converges to the HS price and its computational complexity is $O(n^2 \log n)$.*

1.7 One-dimensional procedure

A further reduction of the computational complexity is possible, and in this Section we show how it can be obtained by modifying the previous procedure in order to make it one-dimensional. Our strategy is to start from a given $\varepsilon > 0$ (we took a $\varepsilon = 10^{-6}$ for the Tables provided in this work) and to specify a recursive backward step that only has a computational cost of order $O(n)$ - which in turn entails that the procedure as a whole will have a computational complexity of $O(n^2)$ - and that gives us a price, both in the European and the American case, with an error smaller than ε with respect to the price obtained via the Hilliard and Schwartz procedure described in Section 1.5.2.

In order to do this, we modify our tree by modelling both the Brownian moves and the jumps with only one variable. This feature has been proposed before - we have seen an example of it by Amin [5] - but with a very different strategy and a loss in precision which we want to avoid by the limitation of the error. We first address the particular case in which the jump nodes are nodes of the CRR tree (eventually extended). We will then show how to deal with the more general situation in which that is not the case. This section is based on Gaudenzi, Spangaro, Stucchi [34].

1.7.1 Jump is a multiple of the Brownian node distance

As a first step in transforming the bivariate tree in a tree with only one dimension for the random variable, we will suppose the minimal amplitude h of a jump to be a multiple of the distance between two Brownian nodes at the same height in the tree, that is, a multiple of $2\sigma\sqrt{\Delta t}$.

Let m be the (integer) ratio $\frac{h}{2\sigma\sqrt{\Delta t}}$, i. e. $h = 2m\sigma\sqrt{\Delta t}$.

We will then consider a tree where every node is labelled with a pair of indexes (i, j); the index i, for i integer between 0 and n, denotes the nodes at time $i\Delta t$, the index j, for j integer between 0 and $(2Nm + 1)i$, indicates the value $S_0 e^{(-i(2Nm+1)+2j)\sigma\sqrt{\Delta t}}$ for the underlying.

With the same notation as above, p will represent the probability of an "up" Brownian move (therefore $1 - p$ will be the probability of a "down" Brownian move), and q_k, for k integer between $-N$ and $+N$, will constitute the probability of a kh jump; since we assume that h is $2m$ times the Brownian move, q_k will be the probability of having a jump of amplitude $2km\sigma\sqrt{\Delta t}$.

Hence, from the node (i, j) we can reach the nodes:
$(i + 1, j)$ with probability $(1 - p)q_{-N}$,
$(i + 1, j + 1)$ with probability pq_{-N},

$(i + 1, j + m)$ with probability $(1 - p)q_{-(N-1)}$,
$(i + 1, j + 1 + m)$ with probability $pq_{-(N-1)}$,
$\vdots$
$(i + 1, j + 2Nm)$ with probability $(1 - p)q_N$,
$(i + 1, j + 1 + 2Nm)$ with probability pq_N.

In the European case, the backward procedure for the evaluation of the call option price is obtained via the recursion formula

$$V(i,j) = e^{-r\Delta t} \sum_{k=-N}^{N} (V(i+1,j+1+(k+N)m)p + V(i+1,j+(k+N)m)(1-p))q_k \tag{1.160}$$

with initial data $V(n,j) = (S(n,j) - K_0)^+$, for j integer between 0 and $n(2Nm+1)$.

In the American case we use the same initial data but replace the previous recursion formula by:

$$C_V(i,j) = \tag{1.161}$$

$$= \sum_{k=-N}^{N} e^{-r\Delta t}(V_A(i+1,j+1+(k+N)m)pq_k + V_A(i+1,j+(k+N)m)(1-p)q_k) \tag{1.162}$$

$$V_A(i,j) = \max\{S(i,j) - K_0, C_V(i,j)\}. \tag{1.163}$$

Similarly, we define $P(i,j)$, $C_P(i,j)$ and $P_A(i,j)$ for the put case. The highest node at time $i\Delta t$ is given by $j = (2Nm+1)i$, which we reach with a maximum amplitude jump and an up Brownian move at every time step up to time $i\Delta t$, which can be written as $i(2N\frac{h\sqrt{n}}{2\sigma\sqrt{\tau}} + 1)$, since $m = \frac{h}{2\sigma\sqrt{\Delta t}} = \frac{h\sqrt{n}}{2\sigma\sqrt{\tau}}$.

Considering all the possible nodes between the highest and the lowest, we have at most $i \cdot (2N \cdot \frac{h\sqrt{n}}{\sigma\sqrt{\tau}} + 1) + 1$ nodes at every time step (when $m > 2i$ this number also includes nodes that can never be reached). Therefore our tree has at most a total of

$$1 + \sum_{i=1}^{n} \left[i \cdot \left(2 \cdot N \cdot \frac{h\sqrt{n}}{\sigma\sqrt{\tau}} + 1 \right) + 1 \right] = 1 + \left(2 \cdot N \cdot \frac{h\sqrt{n}}{\sigma\sqrt{\tau}} + 1 \right) \sum_{i=1}^{n} i + n =$$

$$= 1 + \left(2 \cdot N \cdot \frac{h\sqrt{n}}{\sigma\sqrt{\tau}} + 1 \right) \left(\frac{n(n+1)}{2} \right) + n = 1 + \left(N \cdot \frac{h\sqrt{n}}{\sigma\sqrt{\tau}} + 2 \right) (n(n+1)) + n$$

nodes, and the computational complexity of the backward procedure is $O(n^{2.5})$.

1.7.2 Jump not a multiple of twice a Brownian

When the value $\sqrt{\gamma'^2 + \delta^2}$ that we usually take as the minimal amplitude of a jump is not a multiple of twice the amplitude of the Brownian step, we can recover the previous reasoning by introducing a transformation in the value b which intervenes in the definition of the jump amplitude $h = b\sqrt{\gamma'^2 + \delta^2}$ (see Equation (1.93)).

Instead of taking $b = 1$, we will define a b_n which depends on the number of steps n, and converges to $b = 1$ as n goes to infinity, such that $h_n = b_n\sqrt{\gamma'^2 + \delta^2}$ is a multiple of $2\sigma\sqrt{\Delta t}$. The value of the option is then computed as in the previous section. In order to show that this procedure is sound, we prove some convergence results of this modified process to Y_τ.

At first let us suppose $\sqrt{\gamma'^2 + \delta^2} > \frac{3}{2}\sigma\sqrt{\Delta t}$. Since γ' and δ are parameters of the process and do not depend on n, this is going to be true if we take big enough values for n; however, for completeness, later we will see how we can treat the case $\sqrt{\gamma'^2 + \delta^2} \leq \frac{3}{2}\sigma\sqrt{\Delta t}$.

Given n, let us define k as the nearest integer of the ratio between $\sqrt{\gamma'^2 + \delta^2}$ and $2\sigma\sqrt{\Delta t}$:

$$k = \left\| \frac{\sqrt{\gamma'^2 + \delta^2}}{2\sigma\sqrt{\Delta t}} \right\|$$

This implies $(k-1)\cdot 2\sigma\sqrt{\Delta t} \leq h \leq (k+1)\cdot 2\sigma\sqrt{\Delta t}$. Since we are considering $\sqrt{\gamma'^2 + \delta^2} > \frac{3}{2}\sigma\sqrt{\Delta t}$, k must be at least 1.

Let us call b_n the positive value such that $b_n\sqrt{\gamma'^2 + \delta^2} = k \cdot 2\sigma\sqrt{\Delta t}$, that is

$$b_n = k\frac{2\sigma\sqrt{\Delta t}}{\sqrt{\gamma'^2 + \delta^2}}.$$

We are not guaranteed the above value of b_n respects the HS prescribed boundaries of being between 0 and 1, but we are going to focus on its behaviour in the limit for $n \to +\infty$.

For $n \to +\infty$, we have $b_n = \left\| \frac{\sqrt{\gamma'^2+\delta^2}}{2\sigma\sqrt{\Delta t}} \right\| \frac{2\sigma\sqrt{\Delta t}}{\sqrt{\gamma'^2+\delta^2}} \to 1$. In fact,

$$\begin{aligned}
|b_n - 1| < \varepsilon \Leftrightarrow 1 - \varepsilon &< \left\| \frac{\sqrt{\gamma'^2 + \delta^2}}{2\sigma\sqrt{\Delta t}} \right\| \frac{2\sigma\sqrt{\Delta t}}{\sqrt{\gamma'^2 + \delta^2}} < 1 + \varepsilon \\
\Leftrightarrow (1-\varepsilon)\frac{\sqrt{\gamma'^2 + \delta^2}}{2\sigma\sqrt{\Delta t}} &< \left\| \frac{\sqrt{\gamma'^2 + \delta^2}}{2\sigma\sqrt{\Delta t}} \right\| < (1+\varepsilon)\frac{\sqrt{\gamma'^2 + \delta^2}}{2\sigma\sqrt{\Delta t}} \\
\Leftrightarrow (1-\varepsilon)\frac{\sqrt{\gamma'^2 + \delta^2}\sqrt{n}}{2\sigma\sqrt{\tau}} &< \left\| \frac{\sqrt{\gamma'^2 + \delta^2}\sqrt{n}}{2\sigma\sqrt{\tau}} \right\| < (1+\varepsilon)\frac{\sqrt{\gamma'^2 + \delta^2}\sqrt{n}}{2\sigma\sqrt{\tau}}.
\end{aligned}$$

Since

$$\frac{\sqrt{\gamma'^2+\delta^2}\sqrt{n}}{2\sigma\sqrt{\tau}}-\frac{1}{2}<\left\|\frac{\sqrt{\gamma'^2+\delta^2}\sqrt{n}}{2\sigma\sqrt{\tau}}\right\|\leq\frac{\sqrt{\gamma'^2+\delta^2}\sqrt{n}}{2\sigma\sqrt{\tau}}+\frac{1}{2}$$

if, given $\varepsilon>0$, we want $|b_n-1|<\varepsilon$,
we only need

$$\frac{\sqrt{\gamma'^2+\delta^2}\sqrt{n}}{2\sigma\sqrt{\tau}}+\frac{1}{2}<(1+\varepsilon)\frac{\sqrt{\gamma'^2+\delta^2}\sqrt{n}}{2\sigma\sqrt{\tau}}$$

and

$$(1-\varepsilon)\frac{\sqrt{\gamma'^2+\delta^2}\sqrt{n}}{2\sigma\sqrt{\tau}}<\frac{\sqrt{\gamma'^2+\delta^2}\sqrt{n}}{2\sigma\sqrt{\tau}}-\frac{1}{2},$$

which are both guaranteed by $n>\dfrac{\sigma^2\tau}{\varepsilon^2(\gamma'^2+\delta^2)}$.

In analogy to the two dimensional case, we will still call Y_n the discrete process we obtain from the sum of n random variables $Y_\Delta^{(n)}$ obtained with this variation from the original process from Hilliard and Schwartz:

$$Y_\Delta^{(n)}:=\begin{cases}kb_n\sqrt{\gamma'^2+\delta^2} & \text{with probability } q_k^{(n)} \text{ for } -N\leq k\leq N,\\ 0 & \text{with probability } 1-\lambda\Delta t.\end{cases}\tag{1.164}$$

We will call h_n the minimal amplitude $b_n\sqrt{\gamma'^2+\delta^2}$ of a single jump. Once the value h_n is fixed, imposing the matching of the moments for the discrete and the continuous process, as in Equation (1.94), we obtain the probabilities $q_k^{(n)}$ for $-N\leq k\leq N$.

Even when using $Y_\Delta^{(n)}$ instead of Y_Δ, the same results of convergence hold: there is weak convergence of Y_n to Y_τ in the case $\delta=0$ and $\gamma'>0$ (as in Hilliard and Schwartz) and convergence of the first $2N$ cumulants of Y_n to those of Y_τ in the general case.

These results, heavily based on those by Hilliard and Schwartz, are described below.

Theorem 1.7.1. *If the amplitude of the jump is fixed ($\delta=0$), the discrete process Y_n, sum of n i.i.d. $Y_\Delta^{(n)}$ defined as in Eq.(1.164), weakly converges to the continuous process Y_τ.*

Proof. We are going to prove that the characteristic functions of Y_n converge pointwise to the characteristic function of Y_τ. Therefore, by Lévy's continuity theorem, Y_n converges in distribution to Y_τ.

Since we take $\delta = 0$ and $\gamma' > 0$, the only possible values for $Y_\Delta^{(n)}$ are $b_n\gamma'$ (with probability $\lambda\Delta t$) and 0 (with probability $1 - \lambda\Delta t$).

The characteristic function of $Y_\Delta^{(n)}$ is therefore given by

$$\varphi_{Y_\Delta^{(n)}}(x) = \lambda\Delta t e^{ib_n|\gamma'|x} + (1-\lambda)\Delta t e^0.$$

Since Y_n is the sum of n i.i.d. random variables $Y_\Delta^{(n)}$, the characteristic function of the discrete process Y_n can be computed as the n^{th} power of the characteristic function of $Y_\Delta^{(n)}$:

$$\varphi_{Y_n}(x) = \left(\varphi_{Y_\Delta^{(n)}}(x)\right)^n = \left(1 + \frac{\lambda\tau\left(e^{ib_n\gamma' x} - 1\right)}{n}\right)^n.$$

Recalling Eq.(1.9), the characteristic function of the continuous process Y_τ is given by:

$$\varphi_{Y_\tau}(x) = e^{\lambda\tau\left(e^{ix\gamma' - \frac{\delta^2 x^2}{2}} - 1\right)} \tag{1.165}$$

which, for $\delta = 0$, gives:

$$\varphi_{Y_\tau}(x) = e^{\lambda\tau\left(e^{ix\gamma'} - 1\right)}.$$

We need to show that for $n \to +\infty$, $\varphi_{Y_n}(x) \to \varphi_{Y_\tau}(x)$.

Define $z_n = \lambda\tau\left(e^{ib_n\gamma' x} - 1\right)$. For $n \to +\infty$ we have $b_n \to 1$ and $z_n \to z = \lambda\tau\left(e^{i\gamma' x} - 1\right)$ by continuity of the function $f(t) = \lambda\tau\left(e^{i\gamma' xt} - 1\right)$. This also means that there is a value $c \in \mathbb{R}$ such that $|z_n| < c$ for all $n \in \mathbb{N}$.

Moreover, $e^{z_n} \to e^z$.

In order to see that $\left(1 + \frac{z_n}{n}\right)^n$ tends to e^z as n goes to $+\infty$, we use the triangle inequality:

$$\begin{aligned}
\left|\left(1 + \frac{z_n}{n}\right)^n - e^z\right| &= \left|\left(1 + \frac{z_n}{n}\right)^n - e^{z_n} + e^{z_n} - e^z\right| \\
&\leq \left|\left(1 + \frac{z_n}{n}\right)^n - e^{z_n}\right| + |e^{z_n} - e^z| \\
&\leq \left|\sum_{k=0}^{n}\binom{n}{k}\frac{z_n^k}{n^k} - \sum_{k=0}^{+\infty}\frac{z_n^k}{k!}\right| + |e^{z_n} - e^z| \\
&\leq \left|\sum_{k=0}^{n}\left(\frac{n!}{(n-k)!n^k} - 1\right)\frac{z_n^k}{k!} - \sum_{k=n+1}^{+\infty}\frac{z_n^k}{k!}\right| + |e^{z_n} - e^z|
\end{aligned}$$

$$\leq \left| \sum_{k=0}^{n} \left(\frac{n!}{(n-k)!n^k} - 1 \right) \frac{z_n^k}{k!} \right| + \left| \sum_{k=n+1}^{+\infty} \frac{z_n^k}{k!} \right| + |e^{z_n} - e^z|.$$

Now we choose M such that $\displaystyle\sum_{k=M+1}^{+\infty} \frac{c^k}{k!} < \frac{\varepsilon}{3}$. Then, for all $n > M$ we have that

$$\left| \left(1 + \frac{z_n}{n}\right)^n - e^z \right| \leq$$

$$\leq \sum_{k=0}^{M} \left| \frac{n!}{(n-k)!n^k} - 1 \right| \frac{c^k}{k!} + \sum_{k=M+1}^{n} \left| \frac{n!}{(n-k)!n^k} - 1 \right| \frac{c^k}{k!} +$$
$$+ \sum_{k=n+1}^{+\infty} \frac{c^k}{k!} + |e^{z_n} - e^z|$$
$$\leq \sum_{k=0}^{M} \left| \frac{n!}{(n-k)!n^k} - 1 \right| \frac{c^k}{k!} + \sum_{k=M+1}^{n} \frac{c^k}{k!} + \sum_{k=n+1}^{+\infty} \frac{c^k}{k!} + |e^{z_n} - e^z|$$
$$\leq \sum_{k=0}^{M} \left| \frac{n!}{(n-k)!n^k} - 1 \right| \frac{c^k}{k!} + \frac{\varepsilon}{3} + |e^{z_n} - e^z|$$

since $0 < \frac{n!}{(n-k)!n^k} \leq 1$, hence $\left| \frac{n!}{(n-k)!n^k} - 1 \right| \leq 1$.

Choosing $M' \geq M$ such that $|e^{z_n} - e^z| < \dfrac{\varepsilon}{3}$ and $\left| \dfrac{n!}{(n-k)!n^k} - 1 \right| < \dfrac{\varepsilon}{3e^c}$ for all $n > M'$ and $k \leq M$, we obtain $\left|\left(1 + \frac{z_n}{n}\right)^n - e^z\right| < \varepsilon$.

Therefore, the characteristic function of Y_n converges to the characteristic function of Y_τ.

□

This, joined with the weak convergence of X_n to X_τ, which is a standard result, is sufficient to guarantee the convergence of European prices.

Theorem 1.7.2. *The first $2N$ cumulants of the discrete process Y_n converge to the respective cumulants of the continuous process Y_τ.*

Since h_n and $q_k^{(n)}$ for $k = -N, \ldots, +N$ are constructed in order to have the matching of the first $2N$ moments, the proof is exactly the same as in Hilliard and Schwartz [38] and will be omitted.

Jump smaller than twice a Brownian

When $\sqrt{\gamma'^2+\delta^2} \le \frac{3}{2}\sigma\sqrt{\Delta t}$ for a certain n, let us define k as the nearest integer of the ratio $\frac{\sigma\sqrt{\Delta t}}{\sqrt{\gamma'^2+\delta^2}}$:

$$k = \left\| \frac{\sigma\sqrt{\Delta t}}{\sqrt{\gamma'^2+\delta^2}} \right\|$$

Let us call b_n the positive value such that $b_n\sqrt{\gamma'^2+\delta^2} = \frac{1}{k}\sigma\sqrt{\Delta t}$, that is

$$b_n = \frac{1}{k}\frac{\sigma\sqrt{\Delta t}}{\sqrt{\gamma'^2+\delta^2}}.$$

Again, we are not guaranteed the above value of b_n respects the prescribed boundaries of being between 0 and 1, but this case is only considered when we have a small n, while for a larger number of steps we will use the procedure described previously.

Defining $h_n = b_n\sqrt{\gamma'^2+\delta^2}$ we obtain that the Brownian step is a multiple of the minimal amplitude of the jump.

This means that the jumps are going to fit in the original grid, not expanding it but enriching it with extra nodes between the Brownian ones, if k is not greater than N.

1.7.3 Univariate cut

European case

Without loss of generality, we will suppose $h = 2m\sigma\sqrt{\Delta t}$ with $m \in \mathbb{N}$.

Here we consider the backward procedure introduced at page 112, where the European call option value is obtained via the recursion formula (1.160) with initial data $V(n,j) = (S(n,j)-K_0)^+$, for j integer between 0 and $n(2Nm+1)$.

Applying the truncation method, the procedure is modified as follows:

$$\begin{aligned} &V^T(i,j) = \\ &= e^{-r\Delta t}\sum_{k=-N}^{N}(V^T(i+1,j+1+(k+N)m)p + V^T(i+1,j+(k+N)m)(1-p))q_k \end{aligned} \tag{1.166}$$

for $0 \le i \le n$ and $\max\{0, -\bar{l}m+Nmi\} \le j \le \min\{\bar{k}m+i(Nm+1),(2Nm+1)i\}$, with initial data $V^T(n,j) = (S(n,j)-K_0)^+$, for j integer between $-\bar{l}m+$

$Nmn \leq j \leq \bar{k}m + n(Nm+1)$, while $V^T(n,j) = 0$ for $0 \leq j < -\bar{l}m + Nmn$ or $\bar{k}m + n(Nm+1) < j \leq (2Nm+1)n$, and $V^T(i,j) = 0$ for $0 \leq j < -\bar{l}m + Nmi$ and $\bar{k}m + i(Nm+1) < j \leq i(2Nm+1)$.

Similarly, we define $P^T(i,j)$ for the put option.

Therefore, in our procedure we only need to consider the nodes (i,j) with j between $\max\{0, -\bar{l}m + Nmi\}$ and $\min\{\bar{k}m + i(Nm+1), (2Nm+1)i\}$, for a number of nodes (at time i) inferior to $(\bar{k} + \bar{l})m + i$.

Theorem 1.7.3. *Given $\varepsilon = \frac{1}{n} > 0$, and $\bar{k}$ and $\bar{l}$ the smallest integers as in Theorem 1.6.17, in the one-dimensional case the backward procedure described above converges to the HS price and its computational complexity is $O(n^2)$.*

Proof. We focus on the call option, for the put option the proof is analogous. We are interested in how the value $V^T(0,0)$ obtained via the backward truncated procedure differs from $V^T(0,0,0)$. The value $V^T(0,0)$ is not equal to $V^T(0,0,0)$ since by substituting all the values outside the boundaries $\max\{0, -\bar{l}m + Nmi\}$ and $\min\{\bar{k}m + i(Nm+1), (2Nm+1)i\}$ with zero we are deleting the contribution of paths which at some point in time have surely made $\bar{k}$ up jumps or $\bar{l}$ down jumps, but we are not necessarily excluding all such paths. The following is true:

$$V^T(0,0,0) \leq V^T(0,0) \leq V_E.$$

Therefore, $|V^T(0,0) - V_E| \leq |V^T(0,0,0) - V_E| \leq \frac{1}{n}$ for $\bar{k}$ and $\bar{l}$ as in Theorem 1.6.17.

From the definition of $\bar{k}$ and $\bar{l}$, the number of nodes at maturity is at most proportional to $(\log n)\sqrt{n} + n$, while at any time $i \leq n$ it is proportional to $(\log n)\sqrt{n} + i$; therefore, by summing on all the time steps, we obtain that the computational complexity of the procedure is $O(n^2)$.

□

As we did in the bivariate case, we can also compute the European call price via the discounted expected payoff at maturity, substituting the jump probabilities with the truncated ones, obtained with a forward procedure.

Given $\varepsilon = \frac{1}{n} > 0$, and $\bar{k}$ and $\bar{l}$ as in Theorem 1.6.17, the procedure for the computation of the truncated probabilities is $O(n \log n)$, and so is the correct attribution of the probabilities to the nodes, therefore also in the one-dimensional case the truncated procedure for the computation of the European price as the discounted expected payoff at maturity as described above converges to the HS price and its computational complexity is $O(n \log n)$.

American case

Again, we will suppose $h = 2m\sigma\sqrt{\Delta t}$ with $m \in \mathbb{N}$.

Here we consider the backward procedure introduced at page 112, where the American put option value $P_A(i,j)$ is obtained via the recursion formula (1.161) with initial data $P_A(n,j) = (K_0 - S(n,j))^+$, for j integer between 0 and $n(2Nm+1)$, and its difference with the corresponding truncated procedure:

$$C_A^K(i,j) = \tag{1.167}$$

$$= \sum_{k=-N}^{N} e^{-r\Delta t}(P_A^K(i+1, j+1+(k+N)m)pq_l + P_A^K(i+1, j+(k+N)m)(1-p)q_k) \tag{1.168}$$

$$P_A^K(i,j) = \max\{K_0 - S(i,j), C_A^K(i,j)\} \tag{1.169}$$

for $0 \leq i \leq n$ and $-\bar{l}m + Nmi \leq j \leq \bar{k}m + i(Nm+1)$, with initial data $P_A^K(n,j) = (K_0 - S(n,j))^+$, for j integer between $-\bar{l}m + Nmn \leq j \leq \bar{k}m + n(Nm+1)$, while $P_A^K(n,j) = K_0$ for $0 \leq j < -\bar{l}m + Nmn$ or $\bar{k}m + n(Nm+1) < j \leq (2Nm+1)n$, and $P_A^K(i,j) = K_0$ for $0 \leq j < -\bar{l}m + Nmi$ and $\bar{k}m + i(Nm+1) < j \leq i(2Nm+1)$.

Since procedure $P^K(i,j)$ fixes to K_0 less nodes than those fixed to K_0 by the corresponding bivariate $P^K(i,j,k)$, we have:

$$P^T(0,0,0) \leq P^T(0,0) \leq P_E \leq P^K(0,0) \leq P^K(0,0,0)$$

therefore

$$P^K(0,0) - P_E \leq P^K(0,0,0) - P^T(0,0,0).$$

Retracing the argument in Theorem 1.6.25, one can prove that

$$|P_A^K(i,j) - P_A(i,j)| \leq |P^K(i,j) - P(i,j)|$$

for every $0 \leq i \leq n$, $0 \leq j \leq (2Nm+1)n$.

This gives the following Theorem.

Theorem 1.7.4. *Given $\varepsilon = \frac{1}{n} > 0$, and $\bar{k}$ and $\bar{l}$ the smallest integers as in Theorem 1.6.24, in the one-dimensional case the backward procedure described above converges to the HS price and its computational complexity is $O(n^2)$.*

1.8 Tables

We compare our results with the ones obtained by the procedures described by Hilliard and Schwartz [38], Amin [5], and Dai *et al.* [21], and closed formula by Merton, reporting the calculation times and the values $\overline{k}$ and $\bar{l}$ relative to the cutting of the tree with respect to the number of steps considered, in order to highlight the advantage provided by our procedures.

Table 1 represents European put prices for an option on an underlying with current value $S_0 = 40$, time to maturity $\tau = 1$ year, Poisson parameter $\lambda = 5.0$. In Table 2 we fix $\gamma = 0$ and $\delta^2 = 0.05$, with $n = 400$, but we allow τ to vary.

We may note that Dai *et al.* method is sensitive to the value of σ, while our method is sensitive to the increasing time to maturity.

In Tables 3 and 4 we compare our results in the bivariate (HScutB) and univariate (HScutU) cases with those of Simonato [68], Amin [5], Hilliard and Schwartz [38] and Dai *et al.* [21], for European and American call options. The benchmark for the American case in Table 4 is taken from Chiarella and Ziogas [15]. We see that the numerical results for the American call options present the same precision as in the case of put option, even in the absence of a theoretical result.

Table 1

Strike	Steps	European puts				
		Amin	Dai	HS	HScut	Merton
Panel A: $\gamma = 0$, $\delta^2 = 0.05$, $\sigma^2 = 0.05$						
	200	2.6253 (0.22)	2.6207 (0,47)	2.6215 (1.01)	2.6215 (0.08)	
30	400	2.6233 (1.70)	2.6209 (2.70)	2.6217 (8.04)	2.6217 (0.28)	
	800	2.6223 (13.5)	2.6210 (14.1)	2.6213 (63.9)	2.6213 (1.05)	2.6211
	200	6.7102 (0.22)	6.6972 (0.46)	6.6982 (1.01)	6.6982 (0.09)	
40	400	6.7029 (1.70)	6.6976 (2.70)	6.6070 (7.98)	6.6970 (0.29)	
	800	6.6995 (13.5)	6.6964 (14.1)	6.6968 (63.8)	6.6968 (1.05)	6.6970
	200	12.5486 (0.23)	12.5247 (0.45)	12.5260 (1.00)	12.5260 (0.09)	
50	400	12.5360 (1.69)	12.5243 (2.66)	12.5249 (7.88)	12.5249 (0.29)	
	800	12.5301 (13.3)	12.5241 (13.9)	12.5247 (61.2)	12.5247 (1.03)	12.5238
Panel B: $\gamma = 0$, $\delta^2 = 0.09$, $\sigma^2 = 0.01$						
	200	3.7542 (0.20)	3.9151 (1.24)	3.9154 (0.98)	3.9154 (0.08)	
30	400	3.9086 (1.79)	3.9138 (7.66)	3.9141 (8.03)	3.9141 (0.30)	
	800	3.9220 (13.8)	3.9131 (41.2)	3.9132 (62.4)	3.9132 (1.09)	3.9184
	200	8.3061 (0.22)	8.4652 (1.23)	8.4654 (0.98)	8,4654 (0.09)	
40	400	8.4547 (1.69)	8.4620 (7.59)	8.4621 (8.01)	8.4621 (0.30)	
	800	8.4648 (13.3)	8.4603 (41.3)	8.4604 (61.8)	8.4604 (1.10)	8.4578
	200	14.3182 (0.22)	14.4825 (1.24)	14.4831(0.96)	14.4831 (0.09)	
50	400	14.4621 (1.69)	14,4793 (7.61)	14.4795 (7.88)	14.4795 (0.30)	
	800	14,4697 (13.3)	14,4778 (41.4)	14.4778 (61.9)	14.4778 (1.11)	14.4604
Panel C: $\gamma = 0$, $\delta^2 = 0.05$, $\sigma^2 = 0.0025$						
	200	1,4498 (0.23)	2.1887 (1.82)	2.1888 (0.96)	2.1888 (0.09)	
30	400	1.9766 (1.68)	2.1883 (11.3)	2.1884 (7,94)	2.1884 (0.29)	
	800	2,1502 (13.3)	2.1881 (60.6)	2.1881 (62.0)	2.1881 (1.05)	2.1720
	200	5.2298 (0.24)	6.0039 (1.83)	6.0040 (0.97)	6.0040 (0.10)	
40	400	5.7905 (1.68)	6.0014 (11.3)	6.0015 (7.94)	6.0015 (0.29)	
	800	5.9625 (13.3)	6.0014 (60.6)	6.0002 (62.2)	6.0002 (1.03)	5.9800
	200	11.0203 (0.23)	11.7862 (1.82)	11.7866 (0.98)	11.7866 (0.09)	
50	400	11.5728 (1.68)	11.7839 (11.3)	11.7841 (8.01)	11.7841 (0.28)	
	800	11.7414 (13.3)	11.7828 (60.6)	11.7829 (62.1)	11.7829 (1.05)	11.7556

Table 2

Strike	European puts			
	Amin	Dai	HScut	Merton
Panel A: Maturity $\tau =$ one year, $\sigma^2 = 0.05$				
30	2.6233 (1.70)	2.6209 (2.70)	2.6217 (0.28)	2.6211
40	6.7029 (1.70)	6.6976 (2.70)	6.6970 (0.29)	6.6970
50	12.5360 (1.69)	12.5243 (2.66)	12.5249 (0.29)	12.5238
Panel B: Maturity $\tau =$ one year, $\sigma^2 = 0.01$				
30	2.2486 (1.68)	2.2448 (5.72)	2.2451 (0.27)	2.2436
40	6.1124 (1.68)	6.1029 (5.72)	6.1032 (0.27)	6.0995
50	11.9013 (1.68)	11.8860 (5.72)	11.8864 (0.27)	11.8819
Panel C: Maturity $\tau = 5$ years, $\sigma^2 = 0.05$				
30	5.6850 (1.68)	5.6178 (1.28)	5.6200 (0.50)	5.6013
40	9.5178 (1.68)	9.4120 (1.28)	9.4143 (0.50)	9.3861
50	13.8861 (1.68)	13.7415 (1.27)	13.7446 (0.50)	13.7055
Panel D: Maturity $\tau = 5$ years, $\sigma^2 = 0.01$				
30	5.0466 (1.68)	4.9361 (2.64)	4.9374 (0.50)	4.9198
40	8.6917 (1.68)	8.5266 (2.64)	8.5281 (0.50)	8.5003
50	12.9203 (1.68)	12.7024 (2.64)	12.7042 (0.50)	12.6657
Panel E: Maturity $\tau = 10$ years, $\sigma^2 = 0.05$				
30	5.4517 (1.68)	5.2829 (0.93)	5.2857 (0.73)	5.2834
40	8.3314 (1.68)	8.0925 (0.93)	8.0972 (0.73)	8.0927
50	11.4521 (1.68)	11.1450 (0.94)	11.1495 (0.73)	11.1450
Panel F: Maturity $\tau = 10$ years, $\sigma^2 = 0.01$				
30	4.9085 (1.68)	4.6494 (1.89)	4.6516 (0.73)	4.6491
40	7.6451 (1.68)	7.2843 (1.89)	7.2874 (0.73)	7.2832
50	10.6468 (1.68)	10.1889 (1.89)	10.1926 (0.73)	10.1872

Table 3

Maturity	European calls $\gamma' = -0.02$, $\delta^2 = 0.01$, $r = 0.05$, $d = 0$, $\sigma^2 = 0.04$, $\lambda = 5$, $n = 150$					
	Simonato	Amin	Dai	HScutB	HScutU	Merton
Panel A: $K_0 = 45$						
30/365	5.4304	5.4429	5.4430	5.4435	5.4436	5.4582
90/365	6.4372	6.4263	6.4367	6.4389	6.4393	6.4607
270/365	8.8432	8.7390	8.8323	8.8362	8.8369	8.8668
Panel B: $K_0 = 50$						
30/365	1.7306	1.6952	1.6960	1.6961	1.6958	1.7038
90/365	3.2149	3.1879	3.1952	3.1964	3.1954	3.2119
270/365	5.9859	5.8932	5.9731	5.9773	5.9730	6.0041
Panel C: $K_0 = 55$						
30/365	0.3030	0.3026	0.3023	0.3031	0.3031	0.2936
90/365	1.3251	1.3111	1.3152	1.3176	1.3172	1.3147
270/365	3.8720	3.7975	3.8632	3.8682	3.8705	3.8850

Table 4

Stock	American calls $\tau = 0.5$ year, $r = 0.05$, $d = 0.03$, $\sigma^2 = 0.16$, $\lambda = 1$, $n = 150$, $K_0 = 100$					
	Simonato	Dai	HS	HScutB	HScutU	Benchmark
Panel A: $\gamma' = 0.0000$, $\delta = 0.1980$,						
80	4.0966	4.0839	4.0956	5.0956	4.0940	4.0500
100	12.7026	12.6936	12.6912	12.6912	12.6862	12.6800
120	26.2072	26.2035	26.2015	26.2015	26.1978	26.2200
Panel B: $\gamma' = 0.0488$, $\delta = 0.1888$,						
80	4.2107	4.1867	4.1983	4.1983	4.1763	4.1200
100	12.7409	12.7344	12.7312	12.7312	12.7265	12.6800
120	26.1668	26.1624	26.1591	26.1591	26.1565	26.1400
Panel C: $\gamma' = -0.0513$, $\delta = 0.2082$,						
80	4.0685	4.0722	4.0836	4.0836	4.0828	4.0700
100	12.8002	12.7887	12.7868	12.7868	12.7813	12.8300
120	26.3915	26.3809	26.3794	26.3794	26.3755	26.4600

Chapter 2

Performance evaluation

In this chapter we will discuss performance evaluation. The main issue in determining a performance measure is connected to the problem of measuring risk. The somewhat intuitively clear concept of risk in finance - namely, the possibility of losing some of the initial investment - can be expressed formally in a variety of ways, and therefore quantified taking into account different risk factors.

We will then focus on different ranking methods, analysing their characteristics and comparing their performances in an ex post evaluation of asset class indices. We will first describe some predecessors of the performance measures currently used, and then we will concentrate on the traditional Sharpe index, followed by its modifications which take into account the Value at Risk and the CVaR. We will also include in our analysis two of the methods which use higher and lower partial moments with the aim of an even better estimation of the risk, the Omega ratio and the Sortino ratio. The last in the series is the Rachev ratio, which instead of modifying the risk evaluation intervenes on the profitability measure.

We will show the performances of three macro asset class indices (12 indices) in the period 2003-2015 for these six different rankings, and how strongly these are correlated. This brings to the idea that the information the more refined systems introduce is not capital to the analysis of the behaviour of the asset class indices.

2.1 Ranking criteria for portfolio selection and performance evaluation

The idea of a ranking criterium for portfolios spurs from two separate necessities or points of view. The first is the *ex ante* point of view: the main purpose is to find an objective way to select the best portfolio for an investment, which involves trying to predict the future behaviour of a portfolio and also trying to predict how the actual realisation of the investment is likely to differ from our forecasts (based on historical information); this practice is usually called risk adjustment, and it is also important for allocation of capital and setting position limits.

The choice is trivial when one of the alternatives dominates the others: higher return for the same or lower level of uncertainty, or lower uncertainty for the same or higher level of return; it becomes trickier when the alternative that provides higher expected returns allows for higher uncertainty.

The other reason for such a measurement is performance evaluation: the necessity to assess *ex post* the performance of a portfolio, for example in order to judge the portfolio management and appropriately reward virtuous traders.

In answer to these problems, many different measures and ranking criteria for risk adjustment and performance evaluation have been proposed. In this section we are going to recall the first historical contributions, describe some of the choices that have been favoured in time, and approach the measures that are currently used.

2.1.1 First proposals: Roy's ratio and the Treynor Index.

In 1952 the economist Arthur Roy, in an attempt to give a theoretical basis to analyse the aggregate market behaviour, opposes the view of maximum expected gain and states the principle of Safety First: an individual, in financial decision making, has the main aim of reducing the chance of what they perceive as catastrophic loss. In his own words:

> Decisions taken in practice are less concerned with whether a little more of this or of that will yield the largest net increase in satisfaction than with avoiding known rocks of uncertain position or with deploying forces so that, if there is an ambush round the next corner, total disaster is avoided. If economic survival is always taken for granted, the rules of behaviour applicable in an uncertain and ruthless world cannot be discovered.

That the rational investor would follow this principle gives, according to the author, a better explanation of the observed behaviour, particularly of the customary practice of diversification of resources among a wide range of assets.

Roy proposes then to choose one's assets in a way as to minimise the (upper bound of the) probability that the asset reaches the "disaster level" d, which is shown to be equivalent to maximise the ratio

$$Roy = \frac{E(R_P) - d}{s_P}$$

where $E(R_P)$ is the expected gross return of the portfolio and s_P its standard error (that is, the standard deviation of the sampling distribution of the mean). These are supposed to be known (while in reality they must be estimated analysing time series) and also to be the only information we can infer on every possible action we can choose, which is, as Roy himself pointed out, rather optimistical, since it relates to a normal distribution of the return (recall Chapter 1.4).

Roy's ratio therefore provides a measure of the performance of an asset by relating its excess return and its variability. The level d is, nevertheless, arbitrary, and shall be chosen by the investor accordingly to their preferences (which also means, Roy highlights, that a bigger excess return may make us overlook a more negative loss). In the following we will see that the concept introduced above of a minimum level will be then recovered by other authors, for example in the Minimal Acceptable Return, or MAR, by Sortino [71].

In the same year as Roy's paper, an article by Markowitz [48] changed the paradigm of economic science building the mean-variance framework, or, as it has been baptised, the Modern Portfolio theory.

In his paper *How to Rate Management of Investment Funds*, James Treynor ponders the problem of a satisfactory way to measure the performance of a fund manager, considering that the performance of the fund itself is influenced by two types of risk: the risk due to fluctuations of the market in general and that due to fluctuations of the specific securities held by the fund. While the latter kind of risk tends to average out with an appropriate diversification of the fund, there is no averaging out for the former kind of risk.

The sought measure needs to take into account risk-aversion, but also to be constant even if there are market fluctuations, as long as the management performance is constant, which is why once again it is not feasible to concentrate only on the average return.

For the above reasons, he introduces what he calls the Slope Angle. First of all he plots the percent annual rate of return of the fund against the percent annual rate of return for a general market average (in his paper, the Dow-Jones Industrial Average), for a decade. The resulting points tend to distribute in a straight line, which Treynor calls the characteristic line and can be fitted statistically: the slope of the characteristic line is a measure of the fund volatility. Excess deviations from this line either show an insufficient diversification of the securities held by the fund or are a symptom of an alteration in the fund volatility; a shift in the characteristic line instead is an indication for a drastic change in fund performance without any variation in its volatility.

A fund F can also be represented as a point in an expected rate of return vs volatility diagram. In such a diagram a risk-free asset (e.g. a government bond) B will be represented by a point on the vertical axis, and the straight line going through points B and F is the set of all possible combinations for an investor that sets a portfolio with the riskless asset and the fund. This line is called the portfolio-possibility line, and the measure Treynor proposes for a fund performance is the tangent of the angle α the portfolio-possibility line forms with the horizontal axis.

The slope angle value is given by:

$$SA = \tan\alpha = \frac{\mu_F - \mu^*}{\beta_F}$$

where μ_F is the expected fund rate of return at a particular market rate of return, μ^* is the expected rate of return of the riskless security, and β_F is the volatility of the fund (slope of the characteristic line).

2.1.2 CAPM and Sharpe ratio

In 1966 William Sharpe [64] builds on Treynor's ideas with the reward-to-variability ratio, a performance measure for an investment based on the ratio between reward - represented by the excess return - and risk - expressed as the standard deviation of return.

Applying the developing theory to mutual funds, Sharpe remarks on the impossibility for the manager of a mutual fund to detect the expected return risk combination the investor prefers, and therefore the necessity for the mutual fund manager to select a certain attitude in advance and then draw in investors who share that preference; this gives further reason for a shared definition of performance measure.

The excess return of an investment with respect to the return of a risk-free asset is interpreted as the reward for bearing risk.

In constructing the predictory, *ex ante* version of the ratio, Sharpe uses the two measures of the expected rate of return E_P and the standard deviation σ_P. Referring to its work on market equilibrium [63], he states that if information and prediction about future performance of the securities is agreed upon by all investors, and all investors are able to invest or borrow funds at a common risk-free rate, then for all efficient portfolios it holds true that

$$E_P = \mu^* + b \cdot \sigma_P$$

where the value $b > 0$ represents the risk premium.

On the (E, σ) plane, every possible combination of positions in the riskless security and in the portfolio P is represented by the straight line

$$E = \mu^* + \left(\frac{E(R_P) - \mu^*}{\sigma(R_P)} \right) \sigma.$$

The Sharpe ratio, defined as the slope of this line, allows to find the efficient portfolio P which gives the best boundary of (E, σ) combinations.

$$Sharpe_{ex\_ante}(P) = \frac{E(R_P) - \mu^*}{\sigma(R_P)}$$

Higher expected returns or lower standard deviations determine a higher Sharpe ratio, while lower expected returns or higher standard deviations cause the Sharpe ratio to be lower; therefore the preferable portfolio is the one that gives the greatest Sharpe ratio. In case there is more than one efficient portfolio, all such portfolios must belong to the same line in the (E, σ) plane and therefore share the same Sharpe ratio.

Predicting the performance of a fund needs the evaluation of $E(R_P)$ and $\sigma(R_P)$, which can only be done through estimation.

Changing perspective, the *ex post* Sharpe ratio is initially defined [64] as derived from the previous coefficient by substituting the expected rate of return with the average $A(R_P)$ over n historically observed rate of returns of the portfolio $r_P(i)$ in a given period of time, and the standard deviation of the rate of return with its observed counterpart[1], which Sharpe calls variability, $V(R_P)$.

$$A(R_P) = \frac{1}{n} \sum_{i=1}^{n} r_P(i) \tag{2.1}$$

$$V(R_P) = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (r_P(i) - A(R_P))^2} \tag{2.2}$$

The Sharpe ratio with these modifications becomes:

$$Sharpe_{ex\_post}(P) = \frac{A(R_P) - \mu^*}{V(R_P)} \tag{2.3}$$

which can be interpreted as the reward per unit of variability, hence the name Sharpe initially gave to the measure.

[1]for the estimator of the standard deviation of an aleatory variable see for example [80]

Sharpe argued that this index was preferable to Treynor's in the ex post analysis of a portfolio performance because, by considering variability instead of volatility, it also captured the unsystematic risk which is due to the lack of diversification of the securities in the fund. At the same time, if one is concerned about the usefulness of the indexes as predictors for future performance, the focus of the Treynor index on the volatility, which is the systematic risk, alone, can mean that its ranking of a fund is not distorted by transitory effects and can give a better decision criterium, provided the fund is sufficiently diversified.

Though Harry Markowitz, William Sharpe and Merton Miller received the Nobel prize in Economic Sciences in 1990, "for their pioneering work in the theory of financial economics", in particular respectively for the theory of portfolio choice, the CAPM and the contribution to the theory of corporate finance, there has been a lot of criticism in literature for the very restrictive hypothesis at the foundation of these theories.

In 1994, as an answer to some of the criticism his first measure received (as we will see in the following), Sharpe generalises the Sharpe Ratio, introducing the idea of differential return, which is the difference between the return of the fund we are considering and the return of a chosen benchmark. The updated Sharpe ratio in the *ex ante* version is then the ratio between the expected differential return and its standard deviation.

$$Sharpe'_{ex\_ante}(P) = \frac{E(R_P - R_B)}{\sigma(R_P - R_B)} \tag{2.4}$$

The selection of the benchmark shall be done by the investor accordingly to the aim of the evaluation. In literature, many different proposals for benchmark have been studied: a combination of riskless securities and the market portfolio [77], funds with specified factor loadings [61], or with an investment style similar to that of the considered fund [65].

If we take as a benchmark the risk-free asset, which has variability equal to zero, this definition coincides with the original one.

In the same way, the generalised *ex post* Sharpe ratio is the ratio between the average over n historically observed differential returns and its standard deviation.

$$Sharpe'_{ex\_post}(P) = \frac{A(R_P - R_B)}{V(R_P - R_B)} \tag{2.5}$$

William Sharpe himself underlined the limitations of the use of Sharpe ratio, since it is very simple, it doesn't convey information about, for example, the correlation of the fund with other assets.

The pros of adopting such a measure for investment performance are nevertheless many: it can be interpreted as the evaluation of a zero-investment strategy, with R_P being the acquired asset financed through the short position in R_B, which makes it scale independent; a simple formula for the change of the time period is available; the Sharpe ratio alone provides a means of strategy selection when choosing which fund to add to an up-to-that time riskless portfolio. Even when we are dealing with a risky portfolio, where choosing the fund with the greatest ratio is not sufficient to provide an optimal strategy on its own, Sharpe [66] shows that if the funds we are comparing have similar correlations with the other assets of our investment, we still shall select the fund with the highest Sharpe ratio.

2.1.3 Refining risk evaluation: Value at Risk

Just as the ratios we have seen above were created in order to enlarge the valuation of the desirability of an investment, from the focus on the return, to a more risk-averse mindset, other authors tried to come to a better definition which could benefit from more information on the behaviour of the considered assets.

This line of reasoning spurred the development of several different performance ratios, which proposed ways of including additional pieces of information in the risk evaluation part of the performance measure.

The problem of correlations has been addressed by Dowd [24], who investigated the situation of an investor who needs to decide whether or not to add an investment to their already acquired portfolio. In his work, a correct adjustment for risk of the expected return requires a generalised Sharpe ratio, defined on the different portfolios achievable with the choices feasible to the investor.

If an investor needs to choose whether to buy an additional asset to include in their portfolio, they will compare the Sharpe ratio of the old portfolio and the Sharpe ratio of the portfolio with the additional asset included (for a given proportion a of the total portfolio), and choose the solution with the higher Sharpe ratio. With this stratagem, Dowd builds on the validity of the traditional Sharpe ratio in the instance of non correlation between the feasible options and the rest of the investor's portfolio, simply because the feasible options - defined as seen above - do include all of the portfolio.

Initially, in his discussion, Dowd refers to the original Sharpe ratio, where the benchmark is supposed to be risk-free. The comparison between the old and the new portfolio Sharpe ratio will compel us to invest in the new asset if

$$\frac{E(R_P - \mu^*)}{\sigma_P} \leq \frac{E(aR_A + (1-a)R_P - \mu^*)}{\sigma_{new}} \tag{2.6}$$

where R_P and σ_P are the (aleatory) return and standard deviation of the portfolio already acquired by the investor, R_A is the return of the new asset and σ_{new} is the standard deviation of the new portfolio. This means there is a threshold the reward of the new asset has to surpass in order for the investor to be profitable to include it in the portfolio, and this threshold depends on the new asset's contribution to the overall portfolio risk:

$$E(R_A) \geq E(R_P) + \left(\frac{\sigma_{new}}{\sigma_P} - 1\right)\frac{E(R_P)}{a} \tag{2.7}$$

This decision rule quantifies the idea that an addition to our portfolio is worthwhile if it brings diversification: we will accept lower returns with the acquisition of the new asset if it helps lowering the risks, while if this acquisition carries with it a higher risk (e.g. when the new asset is positively correlated with those in our portfolio) we need a higher expected return as an incentive to buy it.

Dowd remarks that using the traditional Sharpe ratio instead of the generalised one brings miscalculations in the threshold, which will be underestimated in the case of positive correlation and overestimated in the case of negative correlation between the new asset and the old portfolio.

The rule for the optimal portfolio choice can be expressed in terms of the VaR (Value at Risk).

Given a fixed (for the purpose of it, low) probability $(1-c)$, the VaR_{1-c} represents the opposite of the level of return of an investment such that there is a $(1-c)$ probability the random return realizations of the investment will fall under it. We can write this with the following formula:

$$VaR_{1-c}(P) = -F_R^{-1}(1-c) \tag{2.8}$$

Assuming that the return of the portfolio P is normally distributed, the VaR_{1-c} of the portfolio is equal to

$$VaR_{1-c}(P) = -\alpha_{1-c}\sigma_P W \tag{2.9}$$

where α_{1-c} is the confidence parameter associated with the confidence level c and W is a scale parameter reflecting the size of the portfolio.

Going back to Equations (2.6) and (2.7), assuming that the returns of the portfolio with and without the new asset are both normally distributed, given that the scale parameter W and the confidence parameter α_{1-c} are the same for both portfolios, we can substitute the ratio $\frac{\sigma_{new}}{\sigma_P}$ with $\frac{VaR_{1-c}(P_{new})}{VaR_{1-c}(P)}$. This also means the generalised Sharpe ratio in the normal distributed case can be written as:

$$GSharpe(P) = \frac{E(R_P - \mu^*)}{VaR_{1-c}(P)} \tag{2.10}$$

and in order to apply it to the problem of adding a new investment to our portfolio we need to compare $GSharpe(P)$ and $GSharpe(P_{new})$.

In his paper Dowd also considered the more general case of a noncash benchmark, for example the opportunity cost of funds. Equations (2.6)-(2.10) are modified accordingly:

$$\frac{E(R_P - R_B)}{\sigma'_P} \leq \frac{E(aR_A + (1-a)R_P - R_B)}{\sigma'_{new}}$$
$$E(R_A - R_B) \geq E(R_P - R_B) + \left(\frac{\sigma'_{new}}{\sigma'_P} - 1\right)\frac{E(R_P - R_B)}{a}$$
$$GSharpe'(P) = \frac{E(R_P - R_B)}{BVaR_{1-c}(P)}$$

where R_B is the return of the benchmark, σ'_P and σ'_{new} are the standard deviation of $R_P - R_B$ and of $aR_A + (1-a)R_P - R_B$ respectively and $BVaR_{1-c}(P)$ is the benchmark-VaR as defined by Dembo [23].

Dowd remarks the difficulties that intervene in dealing with this generalised version of the Sharpe ratio: the choice of the benchmark is crucial in the determination of the optimal strategy: choosing the benchmark incorrectly not only can give different values for the measure, but also different rankings of the choices.

Also in these cases, the ex ante risk adjustment ratio has an ex post counterpart, where the expectations are substituted with the statistical estimators. Alexander and Baptista investigated the use of this measure for portfolio evaluation, and they baptised it the reward to VaR ratio. In their own words, it represents

> the additional average [periodic] rate of return that investors would have earned if they had borne an additional percentage point of VaR by moving a fraction of wealth from the risk-free security to the portfolio of risky securities that they have selected.[4]

The previous hypothesis that the portfolio characteristics can be comprised in terms of mean and variance, (which is justifiable only with normally or, more in general, elliptically distributed returns) encountered strong criticism, and with their paper Alexander and Baptista motivate the validity of the reward to VaR ratio showing its better performance in the case of heavy-tailed returns.

$VaR_{1-c}(P)$ can be easily computed from n historical data by ordering the set of realizations and choosing the value corresponding to the $(1-c)n^{th}$ position.

With the previous notations, the Reward to VaR ratio associated with probability c is:

$$VaRR_{1-c}(P) = \frac{A(R_P) - \mu^*}{VaR_{1-c}(P)} \tag{2.11}$$

Alexander and Baptista remark that the difference in ranking of a portfolio between the traditional Sharpe evaluation and the reward to VaR evaluation can be taken as a distress signal of the non-normality of the distribution of the rate of return.

The numerical analysis in their paper stresses the similarities and differences between the Sharpe ratio and the reward to VaR ratio, highlighting how the latter one is a judge of the performance of a portfolio in the case of fat tails but also pointing out that different confidence levels of VaR give different measures and different rankings.

2.1.4 Diverting attention from the mean: Omega and Sortino

Alexander and Baptista were not the first to point out how restrictive the mean-variance framework is.

In 1991, Sortino and van der Meer [71] tried to devise a risk measure that could better take into account the asymmetric, skewed, more pointy and fat-tailed distribution the market data seemed to report.

In their paper, the authors appealed to the meaning of the word “risk” to criticise the traditional mean-variance framework. According to them, there is no point in considering risky the deviation the return may have from the mean in the positive direction: such a deviation is beneficial for the investor. The dispersion from the mean that is harmful to the investor is the one which brings the portfolio return to an unacceptably low level. This Minimal Acceptable Return should be fixed by the investor according to their established goals (e.g. the cost of capital in a corporate planning), and given the MAR the volatility shall be divided into a “good” volatility, which would be the dispersion above the MAR level, and “bad”volatility, the dispersion below. Therefore we are interested in measuring risk using the bad volatility. Comparing various such measures (the shortfall or downside probability, the shortfall magnitude by Baumol [8] and Downside Variance by Fishburn [29]), Sortino and van der Meer state the supremacy of the Downside Variance (which has been afterwards renamed downside deviation), defined as

$$DV(P) = \int_{-\infty}^{M} (M - r)^2 f(r) dr \tag{2.12}$$

where r represents the return of the portfolio, whose continuous probability density $f(r)$ needs to be estimated (for example by fitting a three-parameters lognormal distribution to the data, as Sortino himself proposes in [72], using the tenth, fiftieth and ninetieth percentile. For a discussion on fitting parametric distribution, see Chapter 1.4). The downside deviation had been incorporated by Van Harlow and Rao [79] in the CAPM model, and it also has the advantage of being consistent with second- and third-degree stochastic dominance and expected utility theory, and compatible with the Modern Portfolio Theory, while shortfall magnitude is not.

The authors propose an optimisation strategy that constructs the mean-downside deviation efficient frontier, and show that in several cases in different Countries (Netherlands, Canada and UK) this strategy allows for better results than mean-variance one, provided that the investor has "near perfect" forecasting ability, that is, they are able to forecast what will happen in each quarter, knowing the average historical risk-return characteristics for each asset.

The concept of downside risk becomes so popular that in a short time several variations of it are born, and Sortino and Price [72] feel the need to clarify that only in particular cases the downside deviation can be calculated with respect to the mean of each asset (and can then take the name of semivariance); more in general the performances of each asset must be computed with the same reference point (that can be the risk-free asset, the mean return of the index, or some other level according to the investor's preferences).

The definition of the downside deviation as a risk measure allows for the creation of another performance measure, the Sortino ratio:

$$Sortino(P) = \frac{E(R_P - MAR)}{DD(P)} \tag{2.13}$$

Sortino and Price [72] also cite a Fouse index, which can be defined calling on utility theory. Setting a V parameter that incorporates the degree of risk-aversion of the investor, the Fouse index is given by

$$Fouse(P) = E(R_P) - V \cdot DD^2(P) \tag{2.14}$$

The larger the value of V, the more risk-averse is the investor ($V = 1$ identifies a somewhat aggressive investor, who would require the risk premium of an equity to be greater than 200 basis points to prefer it to the riskless asset).

This index, defined in relation with the utility function $U(P) = R_P - V \cdot DD^2(P)$ studied by Fishburn [29], has the additional plus that it can capture the investor's degree of risk-aversion.

Sortino ratio can be generalised by considering the higher and lower partial moments of order n.

Given a level MAR below which the investor feels the investment is not worthwhile, the lower partial moment of order n of the return of a portfolio P with respect to the MAR is given by the n^{th} root of the expected value of the n^{th} power of the difference between the return of the portfolio and the Minimal Acceptable Return.

$$LPM_n(MAR, P) = \sqrt[n]{E(((MAR - R_P)^+)^n)} \tag{2.15}$$

For $n = 2$, this reduces to Sortino's downside deviation.

Similarly, we can define the higher partial moment:

$$HPM_n(MAR, P) = \sqrt[n]{E(((R_P - MAR)^+)^n)} \tag{2.16}$$

Both the previous expressions can be interpreted in an ex post situation by substituting the expected value with the average:

$$LPM_n(MAR, P) = \sqrt[n]{\frac{1}{N}\sum_{i=1}^{N}((MAR - r_i)^+)^n} \tag{2.17}$$

$$HPM_n(MAR, P) = \sqrt[n]{\frac{1}{N}\sum_{i=1}^{N}((r_i - MAR)^+)^n} \tag{2.18}$$

Since, from the point of view of an investor, realisations above the MAR are good and those below are bad, higher partial moments can be considered as profitability indicators and, on the contrary, lower partial moments can play the role of risk indicators.

Depending on the degree n we focus on for the computation of the higher and lower partial moments, a wide variety of performance ratios of the kind HPM/LPM are feasible. The order of the moments shall be chosen accordingly to the level of investor risk aversion.

The Omega ratio (see [62]) is the performance measure we obtain by setting $n = 1$

$$Omega(P) = \frac{HPM_1(MAR, P)}{LPM_1(MAR, P)} = \frac{E(R_P - MAR)}{LPM_1(MAR, P)} + 1. \tag{2.19}$$

2.1.5 Conditional Value at Risk as a risk measure

On the same wake, Agarwal and Naik [2] focus on hedge funds, whose nonlinear payoff resembles - according to the evidence provided by the authors - that of a short position in a put option. Fung and Hsieh [32] had proposed to abandon the mean-variance framework in the case of portfolio construction with hedge funds, unless a quadratic preference of the investor or a normal distribution of the portfolio return is presumed. The option-like behaviour of the hedge fund return means that, in considering the risk, we may want to beware of left-tail risk, and the selected tool for such an analysis is the Conditional Value at Risk, or CVaR, which not only considers the chance of a liability below a certain level, but also has the additional feature of considering also the size of possible extreme losses. Moreover, VaR doesn't have useful properties such as subadditivity, convexity or differentiability, while CVaR is more treatable in that respect.

The CVaR, or Expected Shortfall, at a given level $1 - c$ has been defined by Artzner et al. [6] as the opposite of the expected value of the return R_P of the portfolio, conditional to the values of return below VaR_{1-c}. This can be written as:

$$CVaR_{1-c}(P) = -E(R_P | R_P \leq -VaR_{1-c}(P)) = -\frac{\int_{-\infty}^{-VaR_{1-c}(P)} rf(r)dr}{1-c} \tag{2.20}$$

In other words, CVaR provides the mean restricted to the return worst values. According to the authors [2], opposite to what Sortino and van der Meer [71] observed, in this case obtaining $f(r)$ by fitting a parametrised distribution to the historical data doesn't provide a good estimation of the risk, for the approximation may not be trustworthy in the tails, therefore the empirical distribution is used instead.

Starting from historical data,$CVaR_{1-c}$ can be evaluated as the opposite of the arithmetic mean of the observed values smaller than VaR_{1-c}.

Following the example of Palmquist, Uryasev and Krokhmal [53], Agarwal and Naik construct the mean-CVaR efficient frontier and compare the average returns and the CVaR of those portfolios with those of the mean-variance efficient frontier. While the CVaR of a portfolio grows with its volatility and its level of confidence (which means setting the critical value more the left), the ratio between the CVaR of a portfolio on the mean-variance frontier and the CVar of a portfolio on the M-CVaR frontier grows with the level of confidence but is lower the higher the volatility, suggesting that in high volatility portfolios the estimation of the loss via the mean variance method is not far from the results of the more refined M-CVaR optimisation, which is instead of paramount importance in a low volatility situation.

The Conditional Value at Risk allows for a variation of the reward to VaR ratio which gives a new portfolio performance measure, the Conditional Sharpe ratio or STARR ratio [57]:

$$CVaRR_{1-c}(P) = \frac{E(R_P) - \mu^*}{CVaR_{1-c}(P)} \tag{2.21}$$

2.1.6 Refining reward: Conditional Value at Risk with a twist

Building on the idea the VaR is a boundary that divides the losses for the investor from the possible profits, Rachev [56] and Biglova *et al.* [57] propose to consider a "good $CVaR$", that is the expected value of return, conditional to the values of return greater than VaR_{1-c_1}, as profitability index:

$$Rachev(P) = \frac{E(R_P|R_P > VaR_{1-c_1}(P))}{CVaR_{1-c_1}(P)} \tag{2.22}$$

2.2 Empirical Results

Eling and Schuhmacher [28] compare the Sharpe ratio with twelve other performance measures in the risks assessment of hedge funds. Contrary to what we would expect, their findings support the thesis that the Sharpe ratio is adequate for the ranking of hedge funds, regardless the non normality of their returns. Indeed, they argue that there is virtually no difference in the ranking one would obtain with Sharpe ratio instead of other more refined measures. In order to further investigate this hypothesis in the case of asset class indexes, we focus on six performance measures: the Sharpe index, the VaRR and CVaRR, the Sortino, the Omega and the Rachev ratio.

2.2.1 Data and Main Statistics

In our analysis, we have considered three main asset classes: equities, fixed income and real estate, for a total of 12 indexes, relative to Europe, US, Russia and China. Monthly data of closing adjusted prices from the period 2003 - 2015 from Bloomberg[2] database have been used, exception made for the Russian real estate index, which are given only quarterly. As for the other indexes considered, in the database they are quoted daily, weekly, monthly and so on. Namely, the indexes involved in our examination are the following:

- STOXX Europe 600 Price Index EUR (Bloomberg ticker: SXXP Index): a derivation of the STOXX Europe Total Market Index that has 600 components and represents large-, mid- and small-capitalization companies across 18 European countries;
- S&P 500 Index (SPX): a capitalization-weighted index of 500 US stocks;
- MSCI Russia Index (MXRU): a float-weighted equity index that captures the performance of the large- and mid-capitalization segments of the Russian market, covering approximately 85% of the float-adjusted market capitalization in Russia;
- Hong Kong Hang Seng Index (HSI): a float-weighted equity index of a selection of companies from the Stock Exchange of Hong Kong;
- EUG5TR Index: a Bloomberg/EFFAS (European Federation of Financial Analysts' Societies) long term European government bond index;
- USG5TR Index: the analogous of the EUG5TR index for US;
- Russian Government Bond Index (RGBI): a weighted index of Russian government bonds;
- FGGYCN1 Index: a FTSE index of medium- and long-term Chinese government bonds;
- Bloomberg Europe 500 Real Estate Index (BEREALE): a capitalization-weighted index of all companies that are in the real estate sector of the Bloomberg Europe 500 Index;
- Bloomberg NA REITs (BBREIT): a weighted index of US Real Investment Trusts with capitalization not less than $15 millions;

[2]Bloomberg Finance L.P.

- Russia Housing Prices New Apartments (RUPHNRF): the index of changes in residential property prices, developed by the Russian Federal Service of Statistics;
- Hang Seng Property (HSP): a weighted index of Chinese Real Investment Trusts.

For each of these indices, we evaluated the logarithmic return. In the estimation of its moments, we obtained for skewness a minimum value equal to -2.59 and a maximum equal to 2.39, while for the excess kurtosis a minimum of 1.13 and a maximum of 10.75 (in this analysis we excluded Russian real estate, since it is quoted quarterly and presents excess kurtosis anomalies). High values in both skewness and kurtosis bring us to assume the observed return does not have a normal distribution.

In evaluating indices we have imposed $\mu^* = MAR = 0$. For VaR and $CVaR$ in $VaRR$, $CVaRR$ and Rachev ratio, a value of c equal to 99% has been used, while for the "good" $CVaR$ in the Rachev ratio c_1 is taken equal to 50%.

2.2.2 Ranking Correlation

In order to quantify the relation between the performance measures, we have evaluated the Spearman rank correlation coefficients year by year.

Table 1 and 2 report the Spearman rank correlation coefficients related to the year with the minimum (2015) and the maximum (2013). All performance measures, except for Rachev ratio, display good rank correlation with respect to the Sharpe ratio, between 0.028 (CVaR ratio, 2015) and 0.979 (Omega ratio, 2013). It should be noted that the $CVaR$ ratio in 2015 shows a low level of correlation with any of the other performance ratios, which is probably due to the scarce amount of historical data considered in the $CVaR$ evaluation. On average, the rank correlation of the Sharpe ratio in relation to the other examined performance measures over the years amounts to 0.642. In general, there is also a good correlation between the other pairs of performance ratios, except sometimes for the Rachev ratio and the VaR ratio. Our analysis partially confirm the results obtained by Eling and Schuhmacher [28], who find a very high rank correlation of the examined performance measures with respect to the Sharpe ratio and also in relation with each other. A more refined analysis, where monthly data are substituted by weekly or daily data would probably give more precise answers and results strictly close to those shown by Eling and Schuhmacher.

Figure 2.1: Table 1

2015	SR	VaRR	CVaRR	Omega	Sortino	Rachev
SR	1	0.811	0.028	0.804	0.741	0.783
VaRR		1	0.056	0.993	0.923	0.958
CVaRR			1	0.077	0.154	0.014
Omega				1	0.951	0.972
Sortino					1	0.951
Rachev						1

Figure 2.2: Table 2

2013	SR	VaRR	CVaRR	Omega	Sortino	Rachev
SR	1	0.951	0.951	0.972	0.979	0.951
VaRR		1	0.888	0.979	0.951	0.986
CVaRR			1	0.937	0.909	0.895
Omega				1	0.951	0.979
Sortino					1	0.930
Rachev						1

2.2.3 Visualising the performance of the indexes

The similarities we have found in all the rankings can be perhaps better explained with an analogy: the different performance indexes can be considered as thermometers calibrated with different degrees; the values may differ but they all give the same ranking. In order to give a visual representation of this result and a more immediate analysis, we have used Excel contour charts. Year by year, the white areas cover the two best performing asset classes (the top), the grey areas represent the five medium performing asset classes, while the five worst performing asset classes are painted in black. Hereafter we report the graphs for every performance indicator.

The graphs shown above are very different only at first glance, because there are interesting similarities that are worthy of note. For example, the "black snake" on the right (in Picture 2.9 the detail from the Sortino graph) is everywhere evident (even though less so in the $CVaR$ ratio). This shows that from 2012 to 2015 American, Russian and Chinese fixed income indices have been the worst performers, whatever performance measure we choose. Another shared feature, except for $CVaR$ ratio, is the "giraffe" on the bottom left, shown in Picture 2.10 . This can be interpreted as the supremacy of Chinese fixed income and European and American real estate in the years

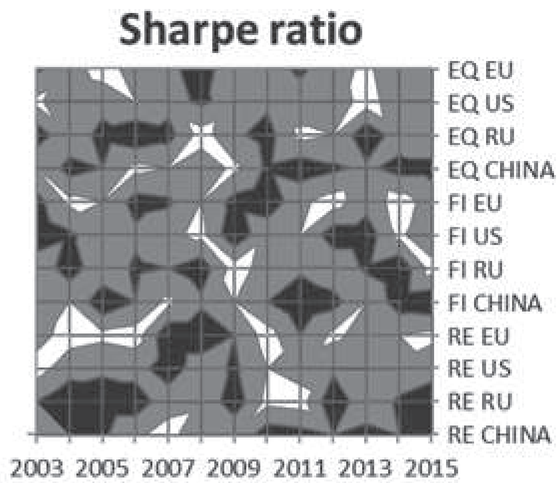


Figure 2.3: Sharpe

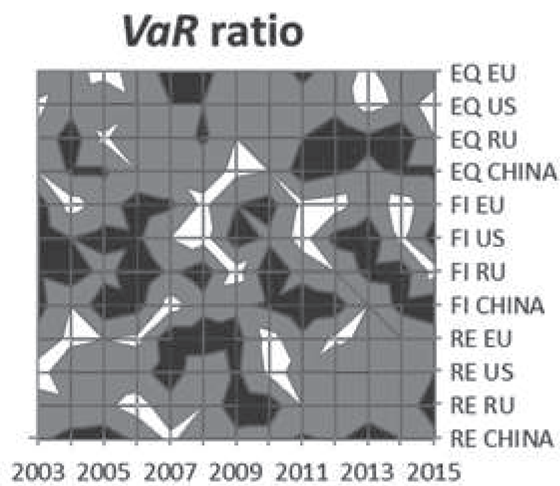


Figure 2.4: VaRR

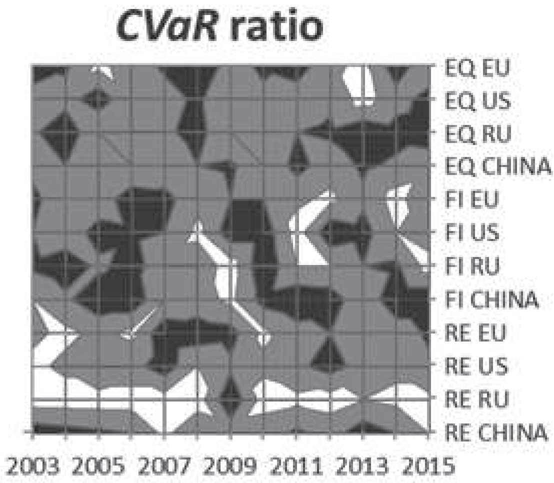


Figure 2.5: CVaRR

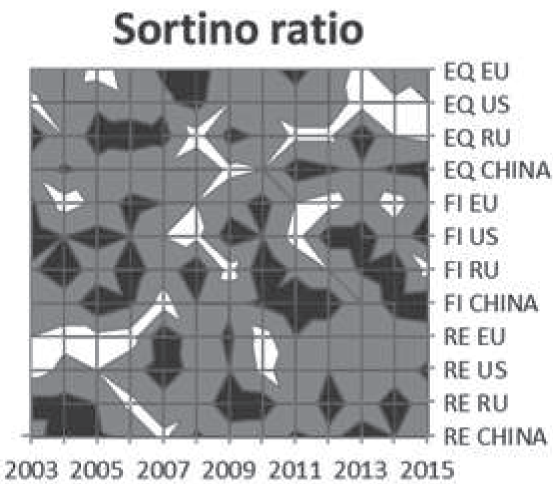


Figure 2.6: Sortino

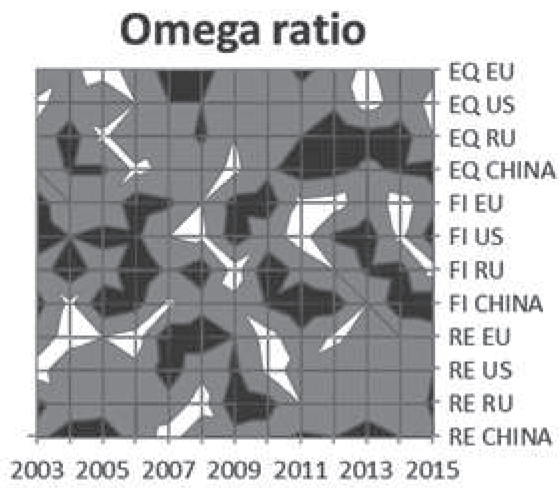


Figure 2.7: Omega

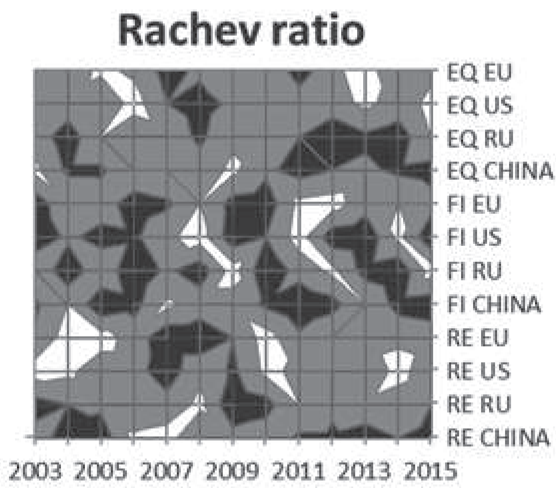


Figure 2.8: Rachev

2003-2008. Lastly, the black trapezoid in the upper part of almost all graphs (in Picture 2.11) corresponding to the year 2008 suggests US and European stock indices as the worst performers in that period, which we can easily relate to the financial crisis.

Figure 2.9: Detail from Sortino contour graph

Figure 2.10: Detail from Sortino contour graph

Figure 2.11: Detail from Sortino contour graph

Our analysis, both through Spearman rank correlation and visual means, suggest the opportunity to rely on the Sharpe ratio, despite the criticism its simplicity often encountered in literature. Indeed, this widely used ratio manages to capture the evolution in the financial period of 2003-2015 in Europe, US, Russia and China; and in doing so agrees with the more refined measures that have been described for a more precise valuation of risk. Moreover, this supports their use in *ex ante* risk assessment and portfolio selection, provided that an appropriate stochastic modelling of the return dynamics is available.

Chapter 3

Appendix

3.1 Formulas for the single jump probabilities

3.1.1 Probabilities in the $N = 2$ case

In the $N = 2$ case the possibilities for the amplitude of the jump are 0, $\pm h$ and $\pm 2h$; therefore we require five probabilities q_{-2}, q_{-1}, q_0, q_1, q_2. The values of the probabilities q_{-2}, q_{-1}, q_0, q_1, q_2 are obtained by matching the first 4 moments of the discrete single step random variable with the local moments (substituted by cumulants as already justified) of the continuous counterpart, and are thus the solutions of the following linear system:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 \\ -2 & -1 & 0 & 1 & 2 \\ 4 & 1 & 0 & 1 & 4 \\ -8 & -1 & 0 & 1 & 8 \\ 16 & 1 & 0 & 1 & 16 \end{pmatrix} \begin{pmatrix} q_{-2} \\ q_{-1} \\ q_0 \\ q_1 \\ q_2 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{k_1}{h} \\ \frac{k_2}{h^2} \\ \frac{k_3}{h^3} \\ \frac{k_4}{h^4} \end{pmatrix}$$

where k_1, k_2, k_3 and k_4 are the first four cumulants of the compound Poisson distribution (cf. page 150).

Therefore:

$$\begin{aligned} q_{-2} &= \frac{\alpha_j \lambda\tau}{12hn} - \frac{\lambda\tau}{24n} - \frac{(\alpha_j^3 + 3\alpha_j\delta^2)\lambda\tau}{12h^3n} + \frac{(\alpha_j^4 + 6\alpha_j^2\delta^2 + 3\delta^4)\lambda\tau}{24h^4n} =: \frac{c_{-2}}{n} \\ q_{-1} &= -\frac{2\alpha_j\lambda\tau}{3hn} + \frac{2\lambda\tau}{3n} + \frac{(\alpha_j^3 + 3\alpha_j\delta^2)\lambda\tau}{6h^3n} - \frac{(\alpha_j^4 + 6\alpha_j^2\delta^2 + 3\delta^4)\lambda\tau}{6h^4n} =: \frac{c_{-1}}{n} \\ q_0 &= 1 - \frac{5\lambda T}{4n} + \frac{\lambda T(\alpha_j^4 + 6\alpha_j^2\sigma_j^2 + 3\sigma_j^4)\}}{4h^4n} < 1 - \frac{\lambda T}{2n} \end{aligned}$$

$$q_1 = \frac{2\alpha_j\lambda\tau}{3hn} + \frac{2\lambda\tau}{3n} - \frac{(\alpha_j^3 + 3\alpha_j\delta^2)\lambda\tau}{6h^3n} - \frac{(\alpha_j^4 + 6\alpha_j^2\delta^2 + 3\delta^4)\lambda\tau}{6h^4n} =: \frac{c_1}{n}$$
$$q_2 = -\frac{\alpha_j\lambda\tau}{12hn} - \frac{\lambda\tau}{24n} + \frac{(\alpha_j^3 + 3\alpha_j\delta^2)\lambda\tau}{12h^3n} + \frac{(\alpha_j^4 + 6\alpha_j^2\delta^2 + 3\delta^4)\lambda\tau}{24h^4n} =: \frac{c_2}{n}$$

where we introduced the constants c_k in order to highlight the fact that the q_k's are inverse proportional to the number of steps n.

3.1.2 Probabilities in the $N = 3$ case

In the $N = 3$ the definition of seven probabilities requires the matching of the first six moments. The value of the probabilities $q_{-3}, q_{-2}, q_{-1}, q_0, q_1, q_2, q_3$ are thus the solutions of the following linear system:

$$\begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ -3 & -2 & -1 & 0 & 1 & 2 & 3 \\ 9 & 4 & 1 & 0 & 1 & 4 & 9 \\ -27 & -8 & -1 & 0 & 1 & 8 & 27 \\ 81 & 16 & 1 & 0 & 1 & 16 & 81 \\ -243 & -32 & -1 & 0 & 1 & 32 & 243 \\ 729 & 64 & 1 & 0 & 1 & 64 & 729 \end{pmatrix} \begin{pmatrix} q_{-3} \\ q_{-2} \\ q_{-1} \\ q_0 \\ q_1 \\ q_2 \\ q_3 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{k_1}{h} \\ \frac{k_2}{h^2} \\ \frac{k_3}{h^3} \\ \frac{k_4}{h^4} \\ \frac{k_5}{h^5} \\ \frac{k_6}{h^6} \end{pmatrix}$$

where k_1, k_2, k_3, k_4, k_5 and k_6 are the first six cumulants of the compound Poisson distribution (cf. page 150).

The solutions are as follows, and again we define the coefficients c_k in order to explicitate the dependence of the q_k on n.

$$\begin{aligned} q_{-3} =& -\frac{\gamma'\,\lambda\,\tau}{60\,hn} + \frac{\left(\gamma'^3 + 3\,\delta^2\,\gamma'\right)\,\lambda\,\tau}{48\,h^3n} - \frac{\left(\gamma'^4 + 6\,\delta^2\,\gamma'^2 + 3\,\delta^4\right)\,\lambda\,\tau}{144\,h^4n} + \\ & - \frac{\left(\gamma'^5 + 10\,\delta^2\,\gamma'^3 + 15\,\delta^4\,\gamma'\right)\,\lambda\,\tau}{240\,h^5n} + \\ & + \frac{\left(\gamma'^6 + 15\,\delta^2\,\gamma'^4 + 45\,\delta^4\,\gamma'^2 + 15\,\delta^6\right)\,\lambda\,\tau}{720\,h^6n} + \frac{\lambda\,\tau}{180n} =: \frac{c_{-3}}{n} \\ q_{-2} =& \frac{3\,\gamma'\lambda\tau}{20\,hn} - \frac{\left(\gamma'^3 + 3\,\delta^2\,\gamma'\right)\lambda\tau}{6\,h^3n} + \frac{\left(\gamma'^4 + 6\,\delta^2\,\gamma'^2 + 3\,\delta^4\right)\lambda\tau}{12\,h^4n} + \\ & + \frac{\left(\gamma'^5 + 10\,\delta^2\,\gamma'^3 + 15\,\delta^4\,\gamma'\right)\lambda\tau}{60\,h^5n} + \\ & - \frac{\left(\gamma'^6 + 15\,\delta^2\,\gamma'^4 + 45\,\delta^4\,\gamma'^2 + 15\,\delta^6\right)\lambda\tau}{120\,h^6n} - \frac{3\lambda\tau}{40n} =: \frac{c_{-2}}{n} \end{aligned}$$

$$
\begin{aligned}
q_{-1} =& -\frac{3\,\gamma'\lambda\tau}{4\,hn} + \frac{13\,\left(\gamma'^3 + 3\,d^2\,\gamma'\right)\lambda\tau}{48\,h^3 n} - \frac{13\,\left(\gamma'^4 + 6\,\delta^2\,\gamma'^2 + 3\,\delta^4\right)\lambda\tau}{48\,h^4 n} + \\
& - \frac{\left(\gamma'^5 + 10\,\delta^2\,\gamma'^3 + 15\,\delta^4\,\gamma'\right)\lambda\tau}{48\,h^5 n} + \\
& + \frac{\left(\gamma'^6 + 15\,\delta^2\,\gamma'^4 + 45\,\delta^4\,\gamma'^2 + 15\,\delta^6\right)\lambda\tau}{48\,h^6 n} + \frac{3\lambda\tau}{4n} =: \frac{c_{-1}}{n} \\
q_0 =& \frac{7\,\left(\gamma'^4 + 6\,\delta^2\,\gamma'^2 + 3\,\delta^4\right)\lambda\tau}{18\,h^4 n} - \frac{\left(\gamma'^6 + 15\,\delta^2\,\gamma'^4 + 45\,\delta^4\,\gamma'^2 + 15\,\delta^6\right)\lambda\tau}{36\,h^6 n} - \frac{49\lambda\tau}{36n} + 1
\end{aligned}
$$

and we have

$$
\begin{aligned}
q_0 =& \frac{7\,\left(3\gamma'^4 + 6\,\delta^2\,\gamma'^2 + 3\,\delta^4 - 2\gamma'^4\right)\lambda\tau}{18\,h^4 n} + \\
& - \frac{\left(15\gamma'^6 + 45\,\delta^2\,\gamma'^4 + 45\,\delta^4\,\gamma'^2 + 15\,\delta^6 - 14\gamma'^6 - 30\,\delta^2\,\gamma'^4\right)\lambda\tau}{36\,h^6 n} - \frac{49\lambda\tau}{36n} + 1 = \\
=& 1 - \frac{49\lambda\tau}{36n} + \frac{7\cdot 3h^4\lambda\tau}{18\,h^4 n} - \frac{15h^6\lambda\tau}{36\,h^6 n} - \frac{7\cdot 2\gamma'^4\lambda\tau}{18\,h^4 n} + \frac{\left(14\gamma'^6 + 30\,\delta^2\,\gamma'^4\right)\lambda\tau}{36\,h^6 n} = \\
&= 1 - \frac{11\lambda\tau}{18n} - \frac{\gamma'^4\lambda\tau}{9\,h^4 n}\frac{\left(7\gamma'^2 - \delta^2\right)}{2h^2} \\
&= 1 - \frac{10\lambda\tau}{18n} - \frac{\lambda\tau}{18n}\frac{h^6 + \gamma'^4\left(7\gamma'^2 - \delta^2\right)}{2h^6} \\
&= 1 - \frac{5\lambda\tau}{9n} - \frac{\lambda\tau}{18n}\frac{h^6 + \gamma'^4\left(7\gamma'^2 - \delta^2\right)}{2h^6} < 1 - \frac{5\lambda\tau}{9n} < 1 - \frac{\lambda\tau}{2n} \\
q_1 =& \frac{3\,\gamma'\lambda\tau}{4\,hn} - \frac{13\,\left(\gamma'^3 + 3\,\delta^2\,\gamma'\right)\lambda\tau}{48\,h^3 n} - \frac{13\,\left(\gamma'^4 + 6\,\delta^2\,\gamma'^2 + 3\,\delta^4\right)\lambda\tau}{48\,h^4 n} + \\
& + \frac{\left(\gamma'^5 + 10\,\delta^2\,\gamma'^3 + 15\,\delta^4\,\gamma'\right)\lambda\tau}{48\,h^5 n} + \\
& + \frac{\left(\gamma'^6 + 15\,\delta^2\,\gamma'^4 + 45\,\delta^4\,\gamma'^2 + 15\,\delta^6\right)\lambda\tau}{48\,h^6 n} + \frac{3\lambda\tau}{4n} =: \frac{c_1}{n} \\
q_2 =& -\frac{3\,\gamma'\lambda\tau}{20\,hn} + \frac{\left(\gamma'^3 + 3\,\delta^2\,\gamma'\right)\lambda\tau}{6\,h^3 n} + \frac{\left(\gamma'^4 + 6\,\delta^2\,\gamma'^2 + 3\,\delta^4\right)\lambda\tau}{12\,h^4 n} + \\
& - \frac{\left(\gamma'^5 + 10\,\delta^2\,\gamma'^3 + 15\,\delta^4\,\gamma'\right)\lambda\tau}{60\,h^5 n} + \\
& - \frac{\left(\gamma'^6 + 15\,\delta^2\,\gamma'^4 + 45\,\delta^4\,\gamma'^2 + 15\,\delta^6\right)\lambda\tau}{120\,h^6 n} - \frac{3\lambda\tau}{40n} =: \frac{c_2}{n} \\
q_3 =& \frac{\gamma'\lambda\tau}{60\,hn} - \frac{\left(\gamma'^3 + 3\,\delta^2\,\gamma'\right)\lambda\tau}{48\,h^3 n} - \frac{\left(\gamma'^4 + 6\,\delta^2\,\gamma'^2 + 3\,\delta^4\right)\lambda\tau}{144\,h^4 n} + \\
& - \frac{\left(\gamma'^5 + 10\,\delta^2\,\gamma'^3 + 15\,\delta^4\,\gamma'\right)\lambda\tau}{240\,h^5 n} +
\end{aligned}
$$

$$+\frac{\left(\gamma'^6 + 15\,\delta^2\,\gamma'^4 + 45\,\delta^4\,\gamma'^2 + 15\,\delta^6\right)\lambda\tau}{720\,h^6 n} + \frac{\lambda\tau}{180n} =: \frac{c_3}{n}$$

3.2 Equalities concerning the hyperbolic functions

By straightforward calculation from the definition of $\sinh x$ and $\cosh x$ we have:

$$\sinh^2 x = \frac{1}{2}\left(\cosh(2x) - 1\right) \simeq \frac{e^{2x}}{4} \tag{3.1}$$

$$\sinh^3 x = \frac{1}{4}\left(\sinh(3x) - 3\sinh x\right) \simeq \frac{e^{3x}}{8} \tag{3.2}$$

$$\sinh^4 x = \frac{1}{8}\left(\cosh(4x) - 4\cosh 2x + 3\right) \simeq \frac{e^{4x}}{16} \tag{3.3}$$

$$\cosh^2 x = \frac{1}{2}\left(\cosh(2x) + 1\right) \simeq \frac{e^{2x}}{4} \tag{3.4}$$

$$\cosh^3 x = \frac{1}{4}\left(\cosh(3x) + 3\cosh x\right) \simeq \frac{e^{3x}}{8} \tag{3.5}$$

$$\cosh^4 x = \frac{1}{8}\left(\cosh(4x) + 4\cosh 2x + 3\right) \simeq \frac{e^{4x}}{16} \tag{3.6}$$

Combining the previous formulas:

$$\cosh(2x) = 2\sinh^2 x + 1 \tag{3.7}$$

$$\sinh(3x) = 4\sinh^3 x + 3\sinh x \tag{3.8}$$

$$\cosh(4x) = 8\sinh^4 x + 8\sinh^2 x + 1 \tag{3.9}$$

3.3 Cumulants of Y_t obtained via the characteristic function

The i^{th} cumulant of Y_t is the i^{th} coefficient of the Maclaurin expansion of $\ln(\varphi_{Y_t})(x) = \sum_{n=0} k_i \frac{(ix)^n}{n!}$. Therefore, in order to compute the first $2N$ coefficients, we need to evaluate the first $2N$ derivatives of $\ln\varphi_{Y_t}(x)$ in $x = 0$.

In this work we used the first 8 cumulants (in order to approximate the jump process with 4 up and 4 down jumps), therefore we needed up to the eight derivative of $\ln\varphi_{Y_t}(x)$.

$$f(x) = \ln\varphi_{Y_t}(x) = \lambda t\left(e^{ix\gamma' - \frac{\delta^2 x^2}{2}} - 1\right)$$

$$\begin{aligned}
f(0) =&0\\
f'(x) =&\lambda t(i\gamma' - \delta^2 x)\left(e^{ix\gamma' - \frac{\delta^2 x^2}{2}}\right) = \lambda t g'(x)\\
f^{(n)}(x) =&\lambda t g^{(n)}(x)
\end{aligned}$$

where we defined: $g(x) = e^{ix\gamma' - \frac{\delta^2 x^2}{2}}$, and of course $g(0) = 1$ and $g'(0) = i\gamma'$. We are then only interested in the derivatives of $g(x)$.

$$\begin{aligned}
g''(x) =&g(x)[(i\gamma' - \delta^2 x)^2 - \delta^2]\\
g''(0) =&-\gamma'^2 - \delta^2\\
g'''(x) =&g'(x)[(i\gamma' - \delta^2 x)^2 - \delta^2] + g(x)\cdot(-2)\delta^2(i\gamma' - \delta^2 x) = g'(x)[(i\gamma' - \delta^2 x)^2 - 3\delta^2]\\
g'''(0) =&-i\gamma'[\gamma'^2 + 3\delta^2]\\
g^{(4)}(x) =&g''(x)[(i\gamma' - \delta^2 x)^2 - 3\delta^2] + g'(x)\cdot(-2)\delta^2(i\gamma' - \delta^2 x) =\\
=&g(x)\{[(i\gamma' - \delta^2 x)^2 - \delta^2][(i\gamma' - \delta^2 x)^2 - 3\delta^2] - 2\delta^2[(i\gamma' - \delta^2 x)^2]\}\\
=&g(x)\{[(i\gamma' + \delta^2 x)^4 - 6\delta^2(i\gamma' + \delta^2 x)^2 + 3\delta^4]\} =\\
g^{(4)}(0) =&\gamma'^4 + 6\gamma'^2\delta^2 + 3\delta^4\\
g^{(5)}(x) =&g'(x)\{[(i\gamma' - \delta^2 x)^4 + 6\delta^2(i\gamma' - \delta^2 x)^2 + 3\delta^4]\}+\\
&+ g(x)\{-4\delta^2[(i\gamma' - \delta^2 x)^3 - 12\delta^4(i\gamma' - \delta^2 x)]\} =\\
=&g'(x)[(i\gamma' - \delta^2 x)^4 - 10\delta^2(i\gamma' - \delta^2 x)^2 + 15\delta^4]\\
g^{(5)}(0) =&i\gamma'(\gamma'^4 + 10\delta^2\gamma'^2 + 15\delta^4)\\
g^{(6)}(x) =&g''(x)[(i\gamma' - \delta^2 x)^4 - 10\delta^2(i\gamma' - \delta^2 x)^2 + 15\delta^4]+\\
&+ g'(x)[-4\delta^2(i\gamma' - \delta^2 x)^3 + 20\delta^4(i\gamma' - \delta^2 x)] =\\
=&g(x)[(i\gamma' - \delta^2 x)^2 - \delta^2][(i\gamma' - \delta^2 x)^4 - 10\delta^2(i\gamma' - \delta^2 x)^2 + 15\delta^4]+\\
&+ g(x)(i\gamma' - \delta^2 x)[-4\delta^2(i\gamma' - \delta^2 x)^3 + 20\delta^4(i\gamma' - \delta^2 x)] =\\
=&g(x)[(i\gamma' - \delta^2 x)^6 - 15(i\gamma' - \delta^2 x)^4\delta^2 + 45(i\gamma' - \delta^2 x)^2\delta^4 - 15\delta^6]\\
g^{(6)}(0) =&-\gamma'^6 - 15\gamma'^4\delta^2 - 45\gamma'^2\delta^4 - 15\delta^6\\
g^{(7)}(x) =&g'(x)[(i\gamma' - \delta^2 x)^6 - 15(i\gamma' - \delta^2 x)^4\delta^2 + 45(i\gamma' - \delta^2 x)^2\delta^4 - 15\delta^6]+\\
&+ g(x)[-6\delta^2(i\gamma' - \delta^2 x)^5 + 60\delta^4(i\gamma' - \delta^2 x)^3 - 90 - \delta^6(i\gamma' - \delta^2 x)] =\\
=&g'(x)[(i\gamma' - \delta^2 x)^6 - 21(i\gamma' - \delta^2 x)^4\delta^2 + 105(i\gamma' - \delta^2 x)^2\delta^4 - 105\delta^6]\\
g^{(7)}(0) =&i\gamma'(-\gamma'^6 - 21\gamma'^4\delta^2 - 105\gamma'^2\delta^4 - 105\delta^6)\\
g^{(8)}(x) =&g''(x)[(i\gamma' - \delta^2 x)^6 - 21(i\gamma' - \delta^2 x)^4\delta^2 + 105(i\gamma' - \delta^2 x)^2\delta^4 - 105\delta^6]+\\
&+ g'(x)[-6\delta^2(i\gamma' - \delta^2 x)^5 + 84\delta^4(i\gamma' - \delta^2 x)^3 - 210\delta^2(i\gamma' - \delta^2 x)] =\\
=&g(x)[(i\gamma' - \delta^2 x)^2 - \delta^2][(i\gamma' - \delta^2 x)^6 - 21(i\gamma' - \delta^2 x)^4\delta^2 + 105(i\gamma' - \delta^2 x)^2\delta^4 - 105\delta^6]+
\end{aligned}$$

$$
\begin{aligned}
&+ g(x)(i\gamma' - \delta^2 x)[-6\delta^2(i\gamma' - \delta^2 x)^5 + 84\delta^4(i\gamma' - \delta^2 x)^3 - 210\delta^2(i\gamma' - \delta^2 x)] = \\
&= g(x)[(i\gamma' - \delta^2 x)^8 - 28(i\gamma' - \delta^2 x)^6\delta^2 + 210(i\gamma' - \delta^2 x)^4\delta^4 - 420(i\gamma' - \delta^2 x)^2\delta^6 + 105\delta^8] \\
g^{(8)}(0) &= \gamma'^8 + 28\gamma'^6\delta^2 + 210\gamma'^4\delta^4 + 420\gamma'^2\delta^6 + 105\delta^8
\end{aligned}
$$

Since $\sum_{n=0} \frac{f^{(n)}(0)x^n}{n!} = \sum_{n=0} k_i \frac{(ix)^n}{n!}$, in order to compute k_i we only need to multiply $f^{(n)}(0)$ by $(-i)^n$.

The first 8 cumulants for Y_t are then:

$$
\begin{aligned}
k_1 &= -\, i \cdot f'(0) = -i \cdot i\gamma'\lambda t = \gamma'\lambda t \\
k_2 &= -\, f''(0) = \lambda t(\gamma'^2 + \delta^2) \\
k_3 &= i f'''(0) = \lambda t \gamma'(\gamma'^2 + 3\delta^2) \\
k_4 &= f^{(4)}(0) = \lambda t(\gamma'^4 + 6\gamma'^2\delta^2 + 3\delta^4) \\
k_5 &= -\, i f^{(5)}(0) = \lambda t \gamma'(\gamma'^4 + 10\gamma'^2\delta^2 + 15\delta^4) \\
k_6 &= -\, f^{(6)}(0) = \lambda t(\gamma'^6 + 15\gamma'^4\delta^2 + 45\gamma'^2\delta^4 + 15\delta^6) \\
k_7 &= i f^{(7)}(0) = \lambda t \gamma'(\gamma'^6 + 21\gamma'^4\delta^2 + 105\gamma'^2\delta^4 + 105\delta^6) \\
k_8 &= f^{(8)}(0) = \lambda t(\gamma'^8 + 28\gamma'^6\delta^2 + 210\gamma'^4\delta^4 + 420\gamma'^2\delta^6 + 105\delta^8)
\end{aligned}
\tag{3.10}
$$

3.4 Integration formulas

3.4.1 Newton-Cotes

The following formulas, necessary for Amin's algorithm, are taken from Abramowitz and Stegun ([1]), section 25.4.

closed, 5 points, also known as Boole's rule If the interval $[a, b]$ is equally divided in 4 segments, each long $h = \frac{b-a}{4}$, such that $a = x_0$ and $b = x_4$ then

$$\int_{x_0}^{x_4} f(x)dx = \frac{2h}{45}\left(7(f(x_0) + f(x_4)) + 32(f(x_1) + f(x_3)) + 12f(x_2)\right) - \frac{8f^{(6)}(\xi)h^7}{945}$$

closed, 10 points If the interval $[a, b]$ is equally divided in 9 segments, each long $h = \frac{b-a}{9}$, such that $a = x_0$ and $b = x_9$ then

$$\int_{x_0}^{x_9} f(x)dx = \frac{9h}{89600}(2857(f(x_0) + f(x_9)) +$$
$$+ 15741(f(x_1) + f(x_8)) + +1080(f(x_2) + f(x_7)) + 19344(f(x_3) + f(x_6)) +$$
$$+5778(f(x_4) + f(x_5))) - \frac{173f^{(10)}(\xi)h^{11}}{14620} \quad (3.11)$$

Bibliography

[1] Abramowitz, M., and Stegun, I. (1972) *Handbook of Mathematical Functions.* National Bureau of Standards. Applied Mathematics Series. 55

[2] Agarwal V., Naik N.Y. (2004). Risk and portfolio decisions involving hedge funds. *Review of Financial Studies* , vol.17, no.1, 63–98

[3] Aitchison J., Brown J.A.C. (1957). *The Lognormal Distribution.* Cambridge University Press.

[4] Alexander G.J., Baptista A.M. (2003). Portfolio Performance Evaluation Using Value at Risk. *The Journal of Portfolio Management*, Summer, pp. 93-102

[5] Amin K. I. (1993). Jump Diffusion Option Valuation in Discrete Time. *The Journal of Finance*, vol. XLVIII, no. 5, December, pp. 1833-1863.

[6] Artzner P., Delbaen F., Eber J., Heath D. (1999) Coherent Measures of Risk. *Mathematical Finance*, no. 9, pp. 203-228

[7] Baker G.A. (1934) Transformation of Non-Normal Frequency Distributions Into Normal Distributions. *The Annals of Mathematical Statistics* vol.5, no. 2, pp. 113-123.

[8] Baumol W.J. (1963) An Expected Gain Confidence Limit Criterion for Portfolio Selection. *Management Science*, October

[9] Biglova A., Ortobelli S., Rachev S.T., Stoyanov S. (2004). Different approaches to risk estimation in portfolio theory. *The Journal of Portfolio Management*, vol. 31, no.1, Fall, pp. 103-112

[10] Black F. and Scholes M. (1973). The Pricing of Options and Corporate Liabilities. *The Journal of Political Economy*, Vol. 81, No. 3, pp. 637-654

[11] Bowman K.O. and Shenton L.R. (1980). Evaluation of the parameters of S_U by rational fractions. *Communications in Statistics: Simulation and Computation*, Vol. B9, No. 2, pp. 127-132

[12] Boyle P.P. (1977). Options: a Monte Carlo Approach. *Journal of Financial Economics* Vol. 4, No. 3, pp. 323-338

[13] Charlier C.V.L. (1905) Über die Darstellung willkürlicher Funktionen. *Arkiv för matematik, astronomi och fysik*, vol. 2, no. 20

[14] Cheang G.H.L. and Chiarella C. (2011) A modern View on Merton's Jump Diffusion Model

[15] Chiarella C. and Ziogas A. (2005) Pricing American Options on Jump-Diffusion Processes Using Fourier Hermite Series Expansion *Quantitative Finance Research Centre* Research Paper 145

[16] Cont R., Tankov P. (2003). *Financial Modeling with Jump Processes.* CRC Press.

[17] Cont R., Voltchkova E. (2003) A finite difference scheme for option pricing in jump diffusion and exponential Lévy models *Rapport Interne CMAP* No. 513

[18] Costabile M., Leccadito A., Massabo' I., Russo E. (2014). Option Pricing Under Regime-Switching Jump-Diffusion Models. *Journal of Computational and Applied Mathematics*, no. 256, pp. 152-167

[19] Cox J.C., Ross S.A. (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics*. vol. 3 pp. 145-166

[20] Cox, J. C., Ross, S. A., Rubinstein, M. (1979). *Option Pricing: a Simplified Approach* Journal of Financial Economics, Vol. 7, pp. 229-263

[21] Dai, T.-S., Wang, C.-J., Lyuu, Y.-D., Liu, Y.-C. (2010). *An efficient and accurate lattice for pricing derivatives under a jump-diffusion process.* Applied Mathematics and Computation, vol. 217, pp. 3174-3189.

[22] Das S., Sundaram R. (1999) Of Smiles and smirks: a term structure perspective. *Journal of Financial Quantitative Analysis*, Vol. 34, pp. 13-32

[23] Dembo R.S. (1997). Value-At-Risk and Return *Net Exposure - The Electronic Journal of Financial Risk*, No. 1, October, pp. 1-12.

[24] Dowd K. (2000). Adjusting for risk: An improved Sharpe ratio. *International Review of Economics and Finance* vol. 9, no. 3, pp. 209–222

[25] Duffie D., Pan J., Singleton K. (2000) Transform Analysis and Asset Pricing for Affine Jump Diffusions *Econometrica*, Vol. 68, No. 6, pp. 1343-1376

[26] Edgeworth F.Y. (1898). On the Representation of Statistics by Mathematical Formulae (Part I.) *Journal of the Royal Statistical Society*, vol. 61, no. 4, pp. 670-700.

[27] Elderton W. P., and Johnson N.L. (1969) *Systems of Frequency Curves.* Cambridge University Press, London

[28] Eling M., Schuhmacher F. (2007). Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking & Finance*, vol.31, pp. 2632–2647

[29] Fishburn P.C. (1977). Mean-Risk Analysis with Risk Associated with Below Target Returns. *The American Economic Review*, March

[30] Feller, W. (1973). *An Introduction to Probability Theory and Its Applications*, Volume I, volume I. John Wiley and Sons.

[31] Figlewski, S., Gao, B. (1999) The adaptive mesh model: a new approach to efficient option pricing. *Journal of Financial Economics* no. 53 pp. 313-351

[32] Fung W., Hsieh D.A. (1999) Is Mean-Variance Analysis Applicable to Hedge Funds? *Economic Letters*, no. 62, pp. 53-58

[33] Gaudenzi M., Spangaro A., Stucchi P. Efficient European and American option pricing under a jump-diffusion process *(preprint)*

[34] Gaudenzi M., Spangaro A., Stucchi P. A Univariate Procedure for Efficient European and American option pricing under a jump-diffusion process *(preprint)*

[35] Heston, Steven L. (1993). A Closed-Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options. *The Review of Financial Studies.* Vol.6, No.2, pp. 327–343.

[36] Hill D. (1973) Algorithm AS 66: The Normal Integral *Applied Statistics* Vol. 22, No. 3, pp. 424-427.

[37] Hill I.D., Hill R., Holder R.L., (1976) Algorithm AS 99: Fitting Johnson curves by moments. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* vol. 25, pp. 180–189.

[38] Hilliard, J. E. & Schwartz , A. (2005). *Pricing European and American Derivatives under a Jump-Diffusion Process: A Bivariate Tree Approach.* Journal of Financial and Quantitative Analysis, vol.4, no. 3, pp. 1-21.

[39] Johnson N. L. (1949). Systems of Frequency Curves Generated by Methods of Translation. *Biometrika*, Vol. 36, No. 1/2, pp. 149-176.

[40] Johnson N. L. (1965). Tables to facilitate fitting S_U frequency curves *Biometrika*, Vol. 52, No. 3/4, pp. 547-558.

[41] Kapteyn J.C., Van Uven M.J. (1916). *Skew Frequency Curves in Biology and Statistics.* Groningen: Astronomical Laboratory.

[42] Kou S.G. (2002) A jump diffusion model for option pricing. *Management Science* vol. 48 pp- 1086-1101.

[43] Kou S.G., Wang H. (2004) Option Pricing Under a Double Exponential Jump Diffusion Model *Management Science*, Vol. 50, No. 9, pp.1178-1192

[44] Leslie D.C.M. (1959) Determination of parameters in the Johnson system of probability distributions. *Biometrika*, Vol. 46, pp. 229-231.

[45] Limpert E., Stahel W. A., Abbt M. (2001) *Log-normal Distributions across the Sciences: Keys and Clues* BioScience, Vol. 51, No. 5

[46] Longstaff F. A., Schwartz E.S. (2001) Valuing American Options by Simulation: A Simple Least-Squares Approach *The Review of Financial Studies* Vol. 14, No. 1, pp. 113-147

[47] Markowitz H.M., (1952) Portfolio Selection, *The Journal of Finance* vol. 7 no. 1, pp. 77-91

[48] Markowitz H.M., (1968) *Portfolio Selection: Efficient Diversification of Investments*, Cowles Foundation Monograph: No. 16

[49] McKean, H. P. Jr. (1969). *Stochastic Integrals.* Academic Press, New York.

[50] Merton R. C. (1976). Option pricing when underlying stocks returns are discontinuous. *Journal of Financial Economics*, vol. 3, pp. 125-144.

[51] Morgenstern, Oskar (1976). Some Reflections on Utility. In Andrew Schotter. Selected Economic Writings of Oskar Morgenstern. New York University Press. pp. 65–70.

[52] Nelson and Ramaswamy (1990). Simple Binomial Processes as Diffusion Approximations in Financial Models *The Review of Financial Studies* , vol.3. no. 3, pp. 393-430

[53] Palmquist J., Uryasev S. and Krokhmal P. (1999) Portfolio Optimization with Conditional Value-at-Risk Objective and Constraint. Research Report 99-14, University of FLorida, Gainesville.

[54] Pearson K. (1895). Contributions to the mathematical theory of evolution, II: Skew variation in homogeneous material. *Philosophical Transactions of the Royal Society.* vol. 186, pp. 343–414

[55] Pressacco, F., Stucchi, P. (2005) *Lezioni di matematica finanziaria*, Forum Edizioni

[56] Rachev S.T. Ed. (2003). Handbook of Heavy Tailed Distributions in Finance. *Elsevier*

[57] Rachev S.T., Martin D., Racheva-Iotova B., Stoyanov S. (2009). Stable ETL optimal portfolios and extreme risk management. In *Risk Assessment* Physika Verlag, pp. 235-262

[58] Rietz H.L. (1922) Frequency Distributions Obtained by Certain Transformations of Normally Distributed Variates *Annals of Mathematics*, vol. 23, no. 4, pp. 292-300

[59] Roy A.D. (1952). Safety First and the Holding of Assets. *Econometrica.* vol. 20, no. 3, pp. 431–450

[60] Rubinstein M. (1998). Edgeworth binomial trees. *Journal of Derivatives*, Spring, pp. 20-27.

[61] Rudd A., Clasing H.K., Modern Portfolio Theory: The Principles of Investment Management, Dow-Jones Irwin

[62] Shadwick W.F., Keating C. (2002). A universal performance measure. *Journal of Performance Measurement*, vol. 6, no. 3, pp. 59–84

[63] Sharpe W.F. (1964). Capital Asset Prices: a Theory of Market Equilibrium under Conditions of Risk. *The Journal of Finance*, vol. XIX, no. 3, pp. 425–442

[64] Sharpe W.F. (1966). Mutual fund performance. *Journal of Business*, vol. 39, no. 1, pp. 119–138

[65] Sharpe W.F. (1992) Asset allocation: Management Style and Performance Measurement, *The Journal of Portfolio Management*, vol. 18, no. 2 , pp. 7-19

[66] Sharpe W.F. (1994). The Sharpe Ratio. *Journal of Portfolio Management*, Fall, pp. 45-58

[67] Shenton L.R. and Bowman K.O. (1982). A Lagrange expansion for the parameters of Johnson's S_U. *Journal of Statistical Computation and Simulation*, Vol. 15, No. 2-3, pp. 89-95

[68] Simonato J.-G. (2011). Johnson binomial trees. *Quantitative Finance*, vol.11, no. 8, , pp. 1165-1176.

[69] Simonato J.-G. (2011). The Performance of Johnson Distributions for Computing Value at Risk and Expected Shortfall, *working paper*, May.

[70] Slifker J.F. and Shapiro S.S. (1980). The Johnson system: Selection and parameter estimation. *Technometrics*, Vol. 22, No. 2, pp. 239-246

[71] Sortino F.A., van der Meer R. (1991). Downside risk. *Journal of Portfolio Management.* vol. 17, no. 4, pp. 27-31.

[72] Sortino F.A., Price L.N. (1994). Performance measurement in a downside risk framework. *Journal of Investing.* vol. 3pp. 59-64.

[73] Spangaro A., Stucchi P. (2016). A Comparison of Ranking Criteria: an Application to Asset Class Indices of Europe, US, Russia and China, *Journal Transition Studies Review*, vol.23, pp. 3-10.

[74] Stuart and Ord (1987). *Kendall's Advanced Theory of Statistics*,Oxford University Press

[75] Tobin J. (1958). Liquidity Preference as Behavior towards Risk. *Review of Economic Studies*, vol. XXV, pp. 65-86

[76] Treynor J.L. (1965). How to Rate Management of Investment Funds. *Harvard Business Review*, vol. XLIII, pp. 63-75

[77] Treynor J.L., Black F. (1973) How to Use Security Analysis to Improve Portfolio Selection *The Journal of Business*,vol. 46, no. 1, pp. 66-86

[78] Tuenter H. J. H. (2001) An algorithm to determine the parameters of SU-curves in the Johnson system of probability distributions by moment matching. *The Journal of Statistical Computation and Simulation* vol.70, no.4, pp. 325–347

[79] Van Harlow, Rao (1989) An Expected Gain Confidence Limit Criterion for Portfolio Selection. *Management Science*, October

[80] Wasserman L. (2004) *All of Statistics. A Concise Course in Statistical Inference*, Springer-Verlag New York

[81] Wicksell S.D. (1917) On the genetic theory of frequency. *Arkiv för matematik, astronomi och fysik*, vol. 12, no. 20. pp. 1-56

[82] Wilmott P. (1998) *Derivatives. The Theory and Practice of Financial Engineering.* John Wiley & sons.

[83] Zelen and Severo (1964). Probability Functions *Handbook of Mathematical Functions* , Dover, New York, pp. 925-966

[84] Zhang X.L. (1997) Valuation of American Options in a Jump-Diffusion Model *Numerical methods in Finance*, Cambridge university Press, pp.93-114

[85] www.tradingtoday.com/black-scholes