



HAL
open science

Diffusion equations: convergence of the functional scheme derived from the Binomial tree with local volatility for non smooth payoff functions.

Julien Baptiste, Emmanuel Lépinette

► To cite this version:

Julien Baptiste, Emmanuel Lépinette. Diffusion equations: convergence of the functional scheme derived from the Binomial tree with local volatility for non smooth payoff functions.. Applied Mathematical Finance, 2018, 25 (5-6), 10.1080/1350486X.2018.1513806 . hal-01507267v3

HAL Id: hal-01507267

<https://hal.science/hal-01507267v3>

Submitted on 29 Dec 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Diffusion equations: convergence of the functional scheme derived from the Binomial tree with local volatility for non smooth payoff functions.

Julien Baptiste,¹
Emmanuel LEPINETTE,¹

¹ *Paris Dauphine university, PSL research university, Ceremade, CNRS, UMR, Place du Maréchal De Lattre De Tassigny, 75775 Paris cedex 16, France.*

Email: emmanuel.lepinette@ceremade.dauphine.fr

Abstract: The function solution to the functional scheme derived from the Binomial tree financial model with local volatility converges to the solution of a diffusion equation of type $h_t(t, x) + \frac{x^2 \sigma^2(t, x)}{2} h_{xx}(t, x) = 0$ as the number of discrete dates $n \rightarrow \infty$. Contrarily to classical numerical methods, in particular finite difference methods, the principle is only based on a discretization in time. We establish the uniform convergence in time of the scheme and provide the rate of convergence when the payoff function is not necessarily smooth as in finance. We illustrate the convergence result and compare its performance to the finite difference and finite element methods by numerical examples.

Keywords and phrases: Binomial tree model, European option pricing, Diffusion partial differential equations, Finite difference scheme, Finite element scheme.

2000 MSC: 60G44, G11-G13.

1. Introduction

There are three major methods to simulate a diffusion equation of type

$$h_t(t, x) + \frac{x^2 \sigma^2(t, x)}{2} h_{xx}(t, x) = 0, \quad t \in [0, T], \quad (1.1)$$

with the boundary condition $h(T, x) = g(x)$. The first one is to use the Monte Carlo methods as the solution h admits a probabilistic representation, see [8]. The second one is to numerically compute the solution directly from the PDE. In particular, the very well known finite difference (FD) technique is based

on approximations of the successive derivatives. It requires a discretization, both in time and in the space variable, of some compact subset $[0, T] \times [a, b]$, $a \leq b$. Moreover, as we need to fix extra boundary conditions, when $x = a$ or $x = b$, that are a priori approximations of h , a second type of approximation errors may appear. For instance, if b is large enough, we generally set the condition $h(t, b) = \lim_{x \rightarrow \infty} h(t, x)$ if the latter may be estimated while the condition $h(t, 0) = g(0)$ is chosen if $a = 0$ under some mild conditions on σ . There is a lot of articles in the literature focusing on this technique, see for example [2] and [16]. The finite element (FE) method is more sophisticated, see [3], but it is based on a discretization of the space variable as well as other techniques as finite volume [14] or spectral methods [13], [1].

The third one is based on binomial trees. Such a scheme is proposed by Milstein [11] where a convergence theorem is proven when the terminal condition, i.e. the payoff function in finance, is smooth. In particular, it is supposed that the successive derivatives of the solution h of the P.D.E. are uniformly bounded. In finance, e.g. for the Call payoff function $g(x) = (x - K)^+$, this is not the case. On the contrary the successive derivatives of the P.D.E. solution explode at the horizon date [10]. Motivated by pricing in finance, we consider non smooth payoff functions and we study the convergence of this very simple alternative scheme derived on the binomial tree with local volatility financial model, which is only based on a discretization in time of the interval $[0, T]$ and allows to simulate $t \mapsto h(t, x)$ for a fixed x . The advantage of this method is the ease of implementation of the functional scheme which is basically defined. The numerical scheme we study is defined by two functions depending on the number n of discretization dates. We focus here on a particular choice of such functions which leads to a uniform approximation in time of the diffusion equation under some mild conditions. A convergence theorem which generalizes [11] to non smooth terminal conditions is proved. We present some examples where the functional scheme appears to outperform the finite difference method and, moreover, it is well adapted to parallel computing. An open problem is to study the suggested functional scheme more generally.

Motivation and interpretation in finance. Let us consider a binomial model in finance with T steps defined as follows: we consider on a stochastic basis $(\Omega, (\mathcal{F}_t)_{t=0, \dots, T}, \mathbb{P})$ a price process $(S_t)_{t=0, \dots, T}$ where S_0 is given at time 0. Moreover, we suppose that $\mathbb{P}(S_{t+1} = S_t k_t^u(S_t) | \mathcal{F}_t) = \mathbb{P}(S_{t+1} = S_t k_t^d(S_t) | \mathcal{F}_t) = 1/2$ where $0 \leq k_t^d < 1 < k_t^u$ are two function-valued processes such that

$k^d + k^u = 2$. By the assumptions, we easily deduce that $(S_t)_{t=0, \dots, T}$ is a P-martingale.

Let us define a portfolio process V as an adapted process satisfying $\Delta V_t := V_t - V_{t-1} = \theta_{t-1} \Delta S_t$ for all $t = 1, \dots, T$, where θ_t is \mathcal{F}_t -measurable for all $t = 0, \dots, T - 1$. Moreover, we say that it replicates the contingent claim $g(S_T)$, for some function g , if $V_T = g(S_T)$. If $g(S_T)$ is integrable, we easily deduce that V is a martingale and, by induction, we get that $V_t = C(t, S_t)$ for some measurable function C . Using the martingale property $V_{t-1} = \mathbb{E}(V_t | \mathcal{F}_t)$, we finally deduce that

$$C(t-1, S_{t-1}) = \frac{1}{2}C(t, S_{t-1}k_{t-1}^d) + \frac{1}{2}C(t, S_{t-1}k_{t-1}^u), \quad t = 1, \dots, T. \quad (1.2)$$

The natural idea is then to consider this scheme as the number $n+1$ of dates $(Ti/n)_{i=0, \dots, n}$ tends to ∞ when we discretize the continuous-time interval $[0, T]$. This leads to the scheme we propose to study in the next section when $k_t^d(x) = 1 - \sigma(t, x)\sqrt{T/n}$ and $k_t^u(x) = 1 + \sigma(t, x)\sqrt{T/n}$. Actually, the asymptotic behaviour of such binomial model is well known, at least for the price process. Indeed, let us consider the continuous-time price process S^n of the binomial model defined by $S_t^n = S_{T(i-1)/n}^n$ if $t \in [T(i-1)/n, Ti/n)$ where $S_{T(i-1)/n}^n$ is defined from k^d and k^u when starting from S_0 at $t = 0$ as explained above. Then, by [12, Proposition 3.2.1] under mild conditions on σ , the sequence S^n weakly converges to the diffusion process S satisfying the stochastic differential equation

$$dS_t = S_t \sigma(t, S_t) dW_t, \quad (1.3)$$

where W is a standard Brownian motion. We also deduce the convergence of the price functions $(C^n)_{n \geq 1}$ given by (1.2) towards the limit price C of the continuous-time model given by (1.3), i.e. the solution to the diffusion equation (1.1). This is indeed the case when σ is constant for a large class of payoff functions g as proven in [2, Chapter 4] where it is also proven that the best rate we can get in general is $\frac{1}{\sqrt{n}}$, see also [18] and the papers [4], [15], [7], and [17] among others and, finally, [11] for smooth terminal conditions and local volatility function. Actually, our main contribution is to show that the result is still valid for local diffusion coefficients even if the solution h of the diffusion equation admits unbounded derivatives, e.g. when the terminal condition is not smooth. Moreover, the numerical experiments we present confirm the accuracy of the functional scheme in comparison to the (FD) and (FE) methods.

2. Functional scheme for diffusion equations

Let us consider a bounded diffusion function $\sigma : [0, T] \times \mathbb{R}_+$ and the associated backward parabolic equation

$$h_t(t, x) + \frac{x^2 \sigma^2(t, x)}{2} h_{xx}(t, x) = 0, \quad t \in [0, T), \quad h(T, x) = g(x), \quad (2.4)$$

where the terminal condition is defined by a Lipschitz function g with Lipschitz constant L_g and h_t and h_{xx} are respectively the first and the second derivatives of h with respect to time t and space variable $x \in \mathbb{R}$. This equation is very well known in physics but also in mathematical finance for models without friction when the risky asset S is driven by a standard Brownian motion W so that it satisfies the stochastic differential equation $dS_t = \sigma(t, S_t) S_t dW_t$ under a risk neutral probability measure P ¹. In that case, $h(t, S_t)$ is the value at time t of the unique self-financing portfolio process V composed of a fraction of the risk-less bond $B = 1$ and the risky asset S such that it replicates the European option payoff $g(S_T)$, i.e. $V_T = g(S_T)$. The quantity $\sigma(t, S_t)$ is then interpreted as a local volatility coefficient.

In the following, we consider the uniform grid on $[0, T]$ given by $t_i^n = (T/n)i$, $i = 0, \dots, n$ where $n \geq 1$. We then define the following functions:

$$\begin{aligned} k_t^{n+}(x) &:= 1 + \sigma(t, x) \sqrt{\frac{T}{n}}, & k_t^{n-}(x) &:= 1 - \sigma(t, x) \sqrt{\frac{T}{n}}, & (2.5) \\ \lambda_t^n(x) &:= \frac{1 - k_t^{n-}(x)}{k_t^{n+}(x) - k_t^{n-}(x)} = \frac{1}{2}, & \mu_t^n(x) &:= 1 - \lambda_t^n(x) = \frac{1}{2}. \end{aligned}$$

Let h^n be the piecewise constant function defined on $[0, T]$ by $h^n(t, x) = h^n(t_{i-1}, x)$ if $t \in [t_{i-1}^n, t_i^n]$, $i \leq n$, where the sequence $(h^n(t_i, x))_{i \leq n}$ is defined recursively by the following functional scheme:

$$\begin{aligned} h^n(t_{i-1}, x) &= \lambda_{t_{i-1}}^n(x) h^n(t_i, k_t^{n-}(x)x) + \mu_{t_{i-1}}^n(x) h^n(t_i, k_t^{n+}(x)x), \\ h^n(T, x) &= g(x), \quad i \leq n. \end{aligned} \quad (2.6)$$

Our main result states that the sequence of functions $(h^n)_{n \in \mathbb{N}}$ uniformly converges to h as $n \rightarrow \infty$ under some mild conditions (called Condition

¹If we suppose that the risk-free interest rate is $r = 0$

D below) satisfied by the successive derivatives of h , solution to (2.4). Of course, the functional scheme (2.6) may also be considered for other choices of functions k^{n-} and k^{n+} . In this paper, we restrict ourselves to the functions defined by (2.5), i.e. when the coefficients defining the binomial model are symmetric, but it is an interesting problem to study such a scheme in its full generality, i.e. for more general coefficients k^{n-} and k^{n+} .

Recall that $h(t, x) = \mathbb{E}g(S_{x,t}(T))$ where, for $t \leq T$, $S_{x,t}$ is the unique solution to the stochastic differential equation (6.14) on the interval $[t, T]$ with initial condition $S_{x,t}(t) = x$. We deduce that $h(t, 0) = g(0)$. In the following, we suppose that σ is locally Lipschitz and bounded so that existence and uniqueness holds for (6.14), see for instance [8, Theorem 2.2, p104]. The conditions below hold in particular when the payoff function g is continuous with linear growth at infinity, e.g. $g(x) = (x - K)^+$, $K \in \mathbb{R}$, and the diffusion coefficient σ is smooth enough and positive as proven in [10, Theorem 4.1] for a more general case ², see also [6, Section 5]. Observe that we do not suppose that the successive derivatives are bounded contrarily to [11].

Condition D:

There exists a constant $C > 0$ independent of n such that:

$$\begin{aligned} |h_x(t, x)| &\leq C, & |h_{xx}(t, x)| &\leq \frac{C}{|x|\sqrt{T-t}}, \\ |h_{xxx}(t, x)| &\leq \frac{C}{x^2(T-t)}, & |h_{xt}(t, x)| &\leq \frac{C}{T-t}, \\ |h_{xxt}(t, x)| &\leq \frac{C}{|x|(T-t)^{3/2}}, & |h_{tt}(t, x)| &\leq \frac{C|x|}{(T-t)^{3/2}}, \\ |h_{ttt}(t, x)| &\leq \frac{C|x|}{(T-t)^{5/2}}, & |h_{xtt}(t, x)| &\leq \frac{C}{(T-t)^2}, \end{aligned}$$

Theorem 2.1. *Suppose that σ is a bounded Lipschitz function and assume that the solution to Equation (2.4) satisfies Condition (D). Then, there exists a constant $C(x) > 0$ depending on x , g and σ such that*

$$\sup_{t \in [0, T]} |h^n(t, x) - h(t, x)| \leq \frac{C|x|}{\sqrt{n}}, \quad n > 0.$$

²Here, the coefficient γ_n of [10] appears to be a constant independent of n .

The proof is deduced from the lemmas and the corollary of Section 5. Note that, by choosing specific discretization dates, see [5], we should improve the convergence rate of the functional scheme. This is left for future research.

3. Numerical examples

In the following, we compare the performance of the functional scheme to the Crank-Nicolson finite difference method one on some examples.

3.1. The case where the diffusion coefficient only depends on time.

Suppose that $\sigma(t, x) = \sigma(t)$, $t \in [0, T]$, only depends on time. In that case, we may show (see [9]) that

$$h(t, x) = \int_{-\infty}^{\infty} g\left(xe^{\rho_t y - \frac{\rho_t^2}{2}}\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy,$$

where $\rho_t \geq 0$ is defined by $\rho_t^2 = \int_t^T \sigma_u^2 du$. Moreover, for some particular functions g , e.g. $g(x) = (x - K)^+$, it is trivial to derive an explicit formula of $h(t, x)$. As k^{n-} and k^{n+} do not depend on the space variable, we also deduce an explicit expression of h^n , solution to (2.6):

$$h(t_{n-i}^n, x) = \frac{1}{2^i} \sum_{z \in \mathbb{E}_{n-i}} g(xz), \quad i = 0, \dots, n, \quad (3.7)$$

where the sets $(E_i)_{i=0, \dots, n}$ does not depend on x and are recursively defined as $\mathbb{E}_n = \{1\}$ and $\mathbb{E}_{n-i-1} = \left(k_{t_{n-i-1}^n}^{n-} \mathbb{E}_{n-i}\right) \cup \left(k_{t_{n-i-1}^n}^{n+} \mathbb{E}_{n-i}\right)$ for all $i = 0, \dots, n-1$.

Regarding the implementation, we may see each E_i as a vector with 2^{n-i} components. Once these vectors computed for all, we may simulate the scheme (2.6) for any x and any arbitrary terminal condition g . Indeed, just compute the vectors G_i whose components are $(g(xz))_{z \in \mathbb{E}_i}$ and deduce $h(t_{n-i}^n, x)$ as the scalar product of G_i with $\mathbf{1} = (1, \dots, 1) \in \mathbb{R}^{2^{n-i}}$.

Let us now consider the case $\sigma(t) = \sigma(2 + \cos(t))$. Then,

$$\rho_t^2 = 4\sigma^2 [(T + \sin T) - (t + \sin t)] + \frac{\sigma^2}{2} \left[(T + \frac{1}{2} \sin(2T)) - (t + \frac{1}{2} \sin(2t)) \right]$$

and for $g(x) = (x - K)^+$, $h(t, x)$ is given explicitly by

$$\begin{aligned}
h(t, x) &= \int_{-\infty}^{\infty} g\left(xe^{\rho_t y - \frac{\rho_t^2}{2}}\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\
&= \int_{-\infty}^{\infty} \left(xe^{\rho_t y - \frac{\rho_t^2}{2}} - K\right)^+ \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\
&= \int_{\rho_t^{-1}\left(\ln\left(\frac{K}{x}\right) + \frac{\rho_t^2}{2}\right)}^{\infty} \left(xe^{\rho_t y - \frac{\rho_t^2}{2}} - K\right) \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \\
&= \int_{\rho_t^{-1}\left(\ln\left(\frac{K}{x}\right) + \frac{\rho_t^2}{2}\right)}^{\infty} x \frac{e^{-\frac{(y-\rho_t)^2}{2}}}{\sqrt{2\pi}} dy - K \left(1 - \Phi\left(\frac{\ln\left(\frac{K}{x}\right) + \frac{\rho_t^2}{2}}{\rho_t}\right)\right) \\
&= x \left(1 - \Phi\left(\frac{\ln\left(\frac{K}{x}\right) + \frac{\rho_t^2}{2}}{\rho_t} - \rho_t\right)\right) - K \Phi\left(\frac{\ln\left(\frac{x}{K}\right) - \frac{\rho_t^2}{2}}{\rho_t}\right) \\
&= x \Phi\left(\frac{\ln\left(\frac{x}{K}\right) + \frac{\rho_t^2}{2}}{\rho_t}\right) - K \Phi\left(\frac{\ln\left(\frac{x}{K}\right) - \frac{\rho_t^2}{2}}{\rho_t}\right).
\end{aligned}$$

Therefore, we are in position to numerically compare the convergence error between the explicit solution given above and the functional scheme as well as the finite difference approximations. We first consider the parameters $\sigma = 0.5$, $T = 10$, $K = 100$, $x = 120$ and 30 discrete dates. As we may observe in Figure 1, with a space discretization of 400 points, the finite difference method does not provide a good approximation and, moreover, the monotonicity we expect to is not satisfied.

With the parameters $\sigma = 1.5$, $T = 1$, $K = 115$, $x = 95$ and the same discretization in time and in space (see Figure 22), the finite difference method provides an approximating function which explodes as time tends to 0.

At last, with the parameters $\sigma = 0.1$, $T = 5$, $K = 115$, we consider 25 discrete dates and 400 points for the discretization in time. We observe that the approximating function given by the finite difference method is not very efficient as well.

In order to better compare the two methods, we have computed the needed time for several strikes to obtain a reasonable convergence error for the value of $h(0, x)$ with respect to the analytic expression of the later. Initially, the time discretization is done with 10 dates while the space is discretized in

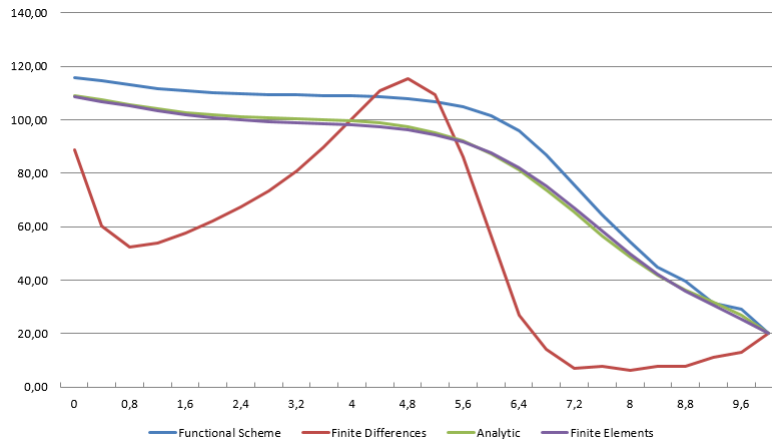


FIG 1. Numerical example with $\sigma = 0.5$, $T = 10$, $K = 100$, $x = 120$.

400 points. If the error is larger than 5%, we add as many discrete dates as necessary (see Tables 4, 5, 6 and 6).

Strike	Result	Error	Time Discretization	Computation time (s)
50	117.74	4.78%	23	2.36
60	117.04	4.91%	24	4.90
70	116.18	4.83%	26	19.62
80	115.66	4.99%	27	76.45
90	115.00	4.99%	28	152.14
100	114.19	4.80%	30	323.16
110	113.76	4.93%	31	1273.74

FIG 4. Values at $t = 0$ by the functional scheme with parameters $\sigma = 0.5$, $T = 10$, $x = 120$.

Clearly, the results are bad for the finite difference method, see Table 5.

Strike	Result	Error	Time Discretization	Computation time (s)
50	105.34	-6.26%	4000	> 1400
60	103.84	-6.93%	4000	> 1400
70	102.63	-7.40%	4000	> 1400
80	101.59	-7.78%	4000	> 1400
90	100.81	-7.97%	4000	> 1400
100	100.12	-8.11%	4000	> 1400
110	99.60	-8.12%	4000	> 1400

FIG 5. Values at $t = 0$ by finite differences with parameters $\sigma = 0.5$, $T = 10$, $x = 120$.

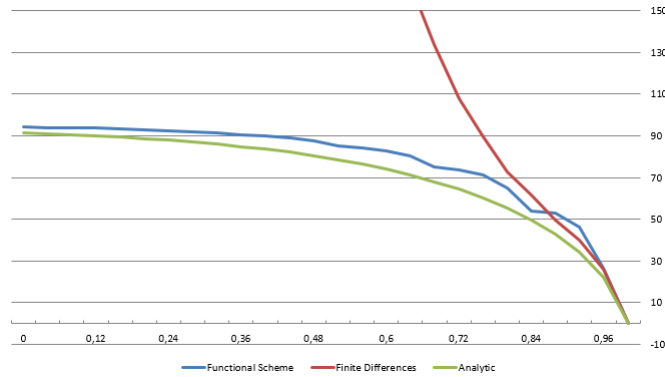


FIG 2. Numerical example with $\sigma = 1.5$, $T = 1$, $K = 115$, $x = 95$.

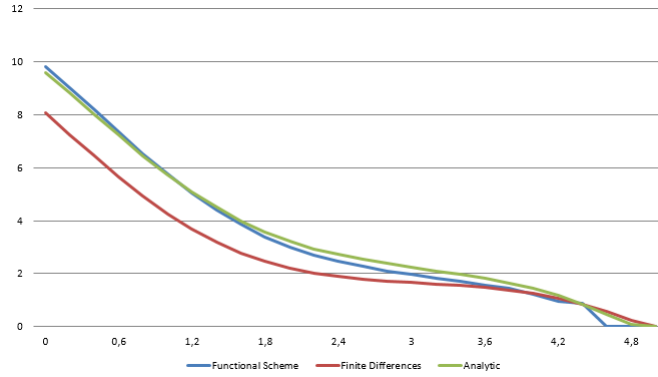


FIG 3. Numerical example with $\sigma = 0.1$, $T = 5$, $K = 115$, $x = 115$.

With the parameters $\sigma = 1.5$, $T = 1$, $x = 95$, we get the following:

Strike	Result	Error	Time Discretization	Computation time (s)
50	95.84	3.34%	20	0.29
60	95.84	3.58%	20	0.29
70	95.83	3.80%	20	0.29
80	95.83	4.01%	20	0.29
90	95.82	4.20%	20	0.29
100	95.81	4.38%	20	0.29
110	95.81	4.54%	20	0.29

FIG 6. Values at $t = 0$ by the functional scheme with parameters $\sigma = 1.5$, $T = 1$, $x = 95$.

Notice that the results given by the finite difference method are not ac-

ceptable, e.g. when K is too large:

Strike	Result	Error	Time Discretization	Computation time (s)
50	3876.89	> 4000%	100	> 20
60	3874.93	> 4000%	100	> 20
70	3873.15	> 4000%	100	> 20
80	3871.49	> 4000%	100	> 20
90	3869.98	> 4000%	100	> 20
100	3868.58	> 4000%	100	> 20
110	3867.27	> 4000%	100	> 20

FIG 7. Values at $t = 0$ by the functional scheme with parameters $\sigma = 1.5$, $T = 1$, $x = 95$.

3.2. The case where the diffusion coefficient only depends on space.

Let us consider the diffusion coefficient $\sigma(t, x) = \sigma(x) = \sigma * (1 + e^{-x^2})$. As $h(t, x)$ does not admit any analytic expression, we first evaluate it by using its probabilistic representation through a Monte-Carlo discretization of the associated diffusion process (6.14). The discretization in time is composed of 10000 dates with 50000 samples for the Monte-Carlo method and 30 dates for the two other methods while we discretize the space variable in 400 points. As we may observe in Figure 8, the finite difference method does not provide a non increasing function in time as expected and the approximation is not accurate.

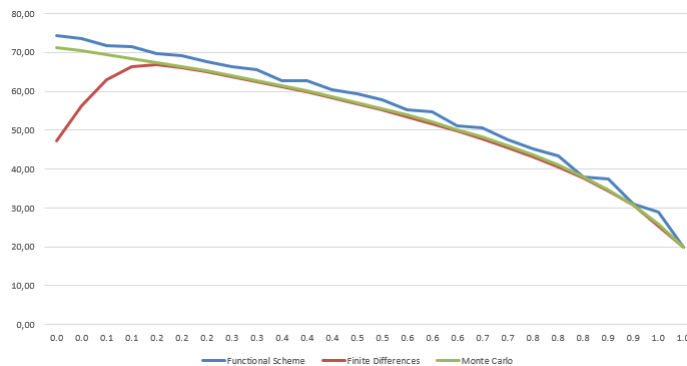


FIG 8. Numerical example with $\sigma = 0.5$, $T = 10$, $K = 100$, $x = 120$.

Strike	Result	Error	Time Discretization	Computation time (s)
50	91.25	4.93%	12	< 0.01
60	87.17	4.81%	13	< 0.01
70	83.51	4.68%	12	< 0.01
80	80.47	4.91%	10	< 0.01
90	76.94	4.13%	19	0.13
100	73.96	3.73%	19	0.13
110	72.16	4.72%	19	0.13

FIG 9. Values at $t = 0$ by the functional scheme with parameters $\sigma = 0.5$, $T = 10$, $K = 100$, $x = 120$.

Strike	Result	Error	Time Discretization	Computation time (s)
50	60.40	-30.56%	30	12.46
60	57.21	-31.22%	30	12.10
70	54.39	-31.82%	30	12.34
80	51.80	-32.46%	30	12.27
90	49.57	-32.92%	30	12.19
100	47.46	-33.44%	30	11.99
110	45.62	-33.80%	30	12.24

FIG 10. Values at $t = 0$ by the finite differences with parameters $\sigma = 0.5$, $T = 10$, $x = 120$.

In Table 11, we see that refining more the space variable grid does not significantly improve the results while the computation time increases.

Number of points	Result	Error	Time discretization	Computation time (s)
400	93.45	-13.78%	30	11.58
600	94.57	-12.75%	20	112.23
800	94.96	-12.39%	30	305.76
1000	95.12	-12.24%	30	573.46
1200	95.23	-12.14%	30	1071.47
1400	95.23	-12.10%	30	1913.85
1600	95.32	-12.05%	30	2938.44

FIG 11. Values at $t = 0$ by the finite differences with parameters $\sigma = 0.5$, $T = 10$, $x = 120$ and $K = 100$ when refining the space variable grid.

With the parameters $\sigma = 1.5$, $T = 1$, $K = 115$, $x = 95$ and only 25 discrete dates, the function scheme provides a rather good approximation contrarily to the finite difference method:



FIG 12. Numerical example with $\sigma = 1.5$, $T = 1$, $K = 115$, $x = 95$.

Strike	Result	Error	Time Discretization	Computation time (s)
50	66.41	4.36%	11	< 0.01
60	62.89	4.52%	12	< 0.01
70	59.20	3.66%	12	< 0.01
80	56.66	4.24%	17	0.03
90	53.64	3.41%	17	0.03
100	51.98	4.77%	17	0.03
110	49.47	4.04%	20	0.26

FIG 13. Values at $t = 0$ by the functional scheme with parameters $\sigma = 1.5$, $T = 1$, $x = 95$.

At last, we consider the parameters $\sigma = 2$, $T = 5$, $K = 10$, $x = 95$ with only 20 discrete dates for the functional scheme. We get satisfactory results in less than 0.26 seconds while with 30 discrete dates for the finite difference methods, the computation time is large. In particular, even if the finite difference method provides acceptable results with only 10 discrete dates with an error less than 5% it takes more than 3.60 seconds to compute them.

3.3. The case where the diffusion coefficient depends on time and space.

We consider the volatility function $\sigma(t, x) = \sigma(1 + \frac{0.01x^2}{1+0.01x^2} + t)$. The number of discrete dates is 25 for the functional scheme as well as the finite difference method and 400 points for the space variable. The Monte Carlo discretization is based on 10000 discrete dates and 50000 trajectories. When the parameters

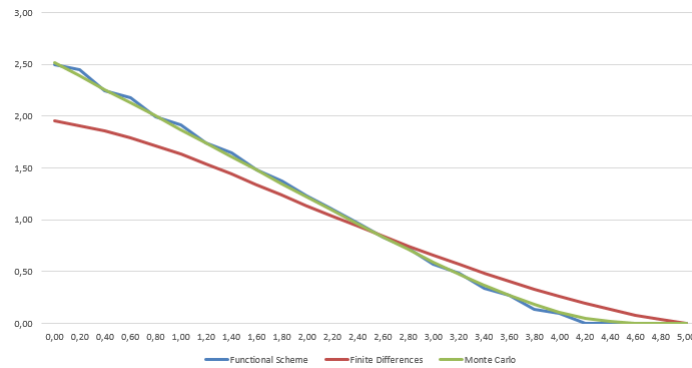


FIG 14. Numerical example with $\sigma = 2$, $T = 5$, $K = 10$, $x = 95$

are $\sigma = 0.5$, $T = 10$, $K = 100$, $x = 120$, the finite difference method provides bad results and it takes a long time to compute them as shown below.

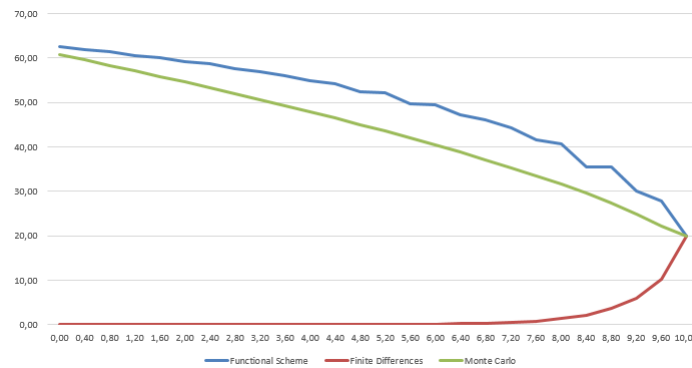


FIG 15. Numerical example with $\sigma = 0.5$, $T = 10$, $K = 100$, $x = 120$.

Strike	Result	Error	Time Discretization	Computation time (s)
50	83.21	4.36%	11	< 0.01
60	78.26	4.52%	11	< 0.01
70	73.64	3.66%	11	< 0.01
80	69.61	4.24%	11	< 0.01
90	66.18	3.41%	11	< 0.01
100	63.15	4.77%	11	< 0.01
110	60.32	4.04%	11	< 0.01

FIG 16. Values at $t = 0$ by the functional scheme with parameters $\sigma = 0.5$, $T = 1$, $x = 120$.

Strike	Result	Error	Time Discretization	Computation time (s)
50	0.02	-99.98%	30	11.42
60	0.01	-99.99%	30	13.00
70	0.00	-100.00%	30	11.82
80	0.00	-100.00%	30	12.23
90	0.00	-100.00%	30	11.77
100	0.00	-100.00%	30	12.42
110	0.00	-100.00%	30	11.77

FIG 17. Values at $t = 0$ by finite differences with parameters $\sigma = 0.5$, $T = 1$, $x = 120$.

For the parameters $\sigma = 0.1$, $T = 5$, $K = 110$, $x = 95$, the functional scheme over evaluates the Monte Carlo method while monotonicity is not satisfied for the finite difference method.

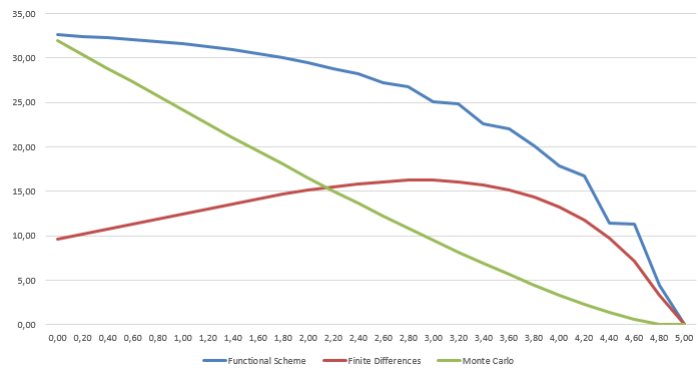


FIG 18. Numerical example with $\sigma = 0.1$, $T = 5$, $K = 110$, $x = 95$.

Strike	Result	Error	Time Discretization	Computation time (s)
50	56.13	4.36%	10	< 0.01
60	51.10	4.52%	10	< 0.01
70	46.69	3.66%	10	< 0.01
80	42.79	4.24%	10	< 0.01
90	39.33	3.41%	10	< 0.01
100	36.23	4.77%	10	< 0.01
110	33.46	4.04%	10	< 0.01

FIG 19. Values at $t = 0$ by the functional scheme with parameters $\sigma = 0.1$, $T = 5$, $x = 95$.

Strike	Result	Error	Time Discretization	Computation time (s)
50	27.21	-50.66%	30	11.69
60	22.88	-54.42%	30	11.58
70	19.37	-57.79%	30	11.55
80	16.38	-61.11%	30	11.56
90	14.03	-63.83%	30	11.66
100	12.04	-66.40%	30	11.60
110	10.41	-68.65%	30	11.59

FIG 20. Values at $t = 0$ for finite differences with parameters $\sigma = 0.1$, $T = 5$, $x = 95$.

Notice that it is possible to execute parallel computations by means of programming interfaces (e.g. OpenMP or MPI) i.e. we may use several calculus units simultaneously and reduce computation time. Using $n \geq 1$ processors, we should theoretically divide the needed time by n . In practice, that depends on how the executed parallel calculus are interconnected. In the finite difference method, when computing the value function at some point of the grid, we need to estimate some expressions depending on the adjacent nodes that may be calculated and used by other processors. This should increase the computation time, as it is necessary to wait for some computations to be executed before starting new ones. On the contrary, for the functional scheme, we do not face this problem as the nodes of the binomial tree at the same level are independent. This is confirmed by the realized time savings we observe to be very closed to the theoretical ones (e.g. 50% with two cores and 75% with 4 cores), see Table 21 with the example of the current subsection and the parameters $\sigma = 0.5$, $T = 10$, $K = 100$, $S_0 = 120$. Therefore, even with parallel computing, the functional scheme should outperform the finite difference method in the examples we have presented.

Number of cores	Computation time (s)	Theoretical time saving	Realized time saving
1	388.203	0%	0%
2	194.492	50%	49.90%
4	102.717	75%	73.54%

FIG 21. Time savings with the parallel functional scheme when $\sigma = 0.5$, $T = 10$, $K = 100$, $S_0 = 120$.

4. Comparison to the Finite Element method.

4.1. The case of a bounded payoff function.

Let us consider the payoff function $h(x) = (K - x)^+$. The case when $h(x) = (x - K)^+$ may be deduced by the Call-Put parity $(x - K)^+ - (K - x)^+ = x - K$.

When the volatility function only depends on time (see Section 3.1), we obtain very good approximations, see Figure 22 and Table 23, when the space variable interval is $[0, 10^4]$ discretised in 2000 points. This is confirmed by other sets of parameters.

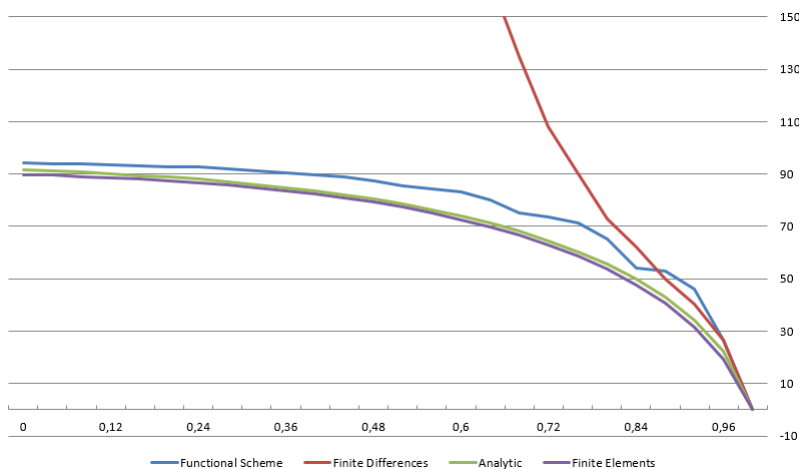


FIG 22. Numerical example with $\sigma = 1.5$, $T = 1$, $K = 115$, $x = 95$.

Strike	Result	Error	Time Discretization	Computation time (s)
50	90.94	-1.95%	10	0.005
60	90.62	-2.06%	10	0.005
70	90.32	-2.17%	10	0.005
80	90.04	-2.27%	10	0.005
90	89.78	-2.37%	10	0.005
100	89.53	-2.47%	10	0.004
110	89.29	-2.57%	10	0.005

FIG 23. Values at $t = 0$ by the FEM with parameters $\sigma = 1.5$, $T = 1$, $K = 115$, $x = 95$.

4.2. The case of an unbounded payoff function.

Let us consider the payoff function $h(x) = (x-K)^8$ and let us fix the following parameters: $\sigma = 0.25$, $T = 1$ and $K = 0.5$. We then compute $t \mapsto h(t, x)$ for $x = 1$. We compare the obtained results to the approach based on the probabilistic representation of $h(t, x)$, i.e. by implementing the Monte Carlo (MC) technique with a time discretization of 50000 dates and a sample of 100000 trajectories.

First, we observe that the (FE) method is very performant when $\sigma(t, x) = \sigma$ is a constant function and the interval space $[0, 10]$ is discretised in 500 points. To see it, we estimate the relative error of the (FE) approximation with respect to the (MC) method, see Table 24.

number of dates	(FE) method	Time (s)	Error
5	1.1201	0.001	215.32%
10	0.5969	0.001	68.02%
20	0.4508	0.002	26.91%
50	0.3855	0.005	8.52%
100	0.3666	0.009	3.19%
200	0.3576	0.017	0.66%
300	0.3547	0.025	0.16%

FIG 24. Approximation of $h(0, x)$ with the (FE) method.

In this example, the functional scheme seems to be more unstable even if it only takes 3 m.s. to get an error less than 5%. Nevertheless it takes too much time to obtain an error less than 1%, see Table 4.2.

We now consider the volatility function

$$\sigma(t, x) = \sigma \left(1 + \frac{x^2}{100 \left(1 + \frac{x^2}{100} \right)} + t \right).$$

In that case, it is difficult to obtain the desired convergence as the result is very sensitive to the chosen space variable and its discretization as well. It appears that we reduce the error by increasing the discretization in time but it is not possible to lower the error as much as desired, see Table 26.

Surprisingly, if we reduce the number of discrete points of the space variable to 100 points, the error is smaller as observed in Table 27. Nevertheless, even by increasing the number of discrete dates to 10^7 and more, it is not possible to reduce as much as desired the convergence error. If we enlarge the space

Number of dates	Functional Scheme	Time (s)	Error
5	0.1831	0.000	93.96%
6	0.3331	0.000	6.66%
7	0.3344	0.000	6.23%
8	0.2280	0.000	55.81%
9	0.2382	0.000	49.15%
10	0.3378	0.001	5.16%
11	0.2543	0.001	39.70%
12	0.2608	0.001	36.23%
13	0.3402	0.003	4.41%
14	0.3409	0.006	4.21%
15	0.3415	0.013	4.03%
16	0.2799	0.012	26.91%
17	0.2835	0.024	25.29%
18	0.2868	0.049	23.87%
19	0.3433	0.197	3.48%
20	0.2925	0.196	21.47%
21	0.2949	0.400	20.44%
22	0.3443	1.571	3.17%
23	0.3446	3.135	3.08%
24	0.3449	6.285	3.00%
25	0.3031	6.237	17.20%
26	0.3454	24.917	2.86%
27	0.3456	51.146	2.80%
28	0.3458	107.718	2.74%
29	0.3460	213.635	2.68%

FIG 25. Approximation of $h(0, x)$ with the (FS) method and constant volatility.

nbIterations	Finite Elements Method	Time (s)	Error
5	110.8036	0.001	644.37%
10	63.7537	0.001	328.29%
20	39.5068	0.003	165.40%
50	26.7455	0.006	79.67%
100	23.0440	0.011	54.81%
200	21.3239	0.022	43.25%
300	20.7713	0.034	39.54%

FIG 26. Approximation of $h(0, x)$ with the (FE) method and non constant volatility

variable to $[0, 25]$ the error is 837% for 2000 discrete dates so that we need to increase more the number of discrete dates.

On the contrary, the functional scheme appears to be efficient, see Table 28. In particular, it is possible to reduce as much as desired the convergence error. It takes about 145 seconds to obtain an error less than 5%. Moreover, a code parallelisation should significantly reduce the needed computation time.

5. Proof

From a line to the next one, we shall use the same notation C for some distinct constants independent of n . Also, notice that there is a constant $M > 0$ such that $|k_t^{n-}| + |k_t^{n+}| \leq M$ for all n and $t \in [0, T]$. We also use the notation $\Delta t_i^n = t_i^n - t_i^{n-1} = T/n$, $i \leq n$. At last, observe that the results below are trivial when $x = 0$ hence we only provide the proofs when $x \neq 0$.

nbIterations	Finite Elements Method	Time (s)	Error
5	103,3949	0,000	594,60%
10	56,7661	0,000	281,35%
20	33,6140	0,001	125,81%
50	21,9113	0,001	47,20%
100	18,6310	0,002	25,16%
200	17,1327	0,005	15,10%
300	16,6556	0,008	11,89%
500	16,2822	0,012	9,38%
1000	16,0070	0,03	7,53%
2000	15,8710	0,049	6,62%
1000000	15,7363	25,581	5,71%
10000000	15,7360	254,609	5,71%

FIG 27. Approximation of $h(0, x)$ with the (FE) method and non constant volatility for 100 discrete points.

nbIterations	Functional Scheme	Time (s)	Error
5	1,4608	0,000	919,00%
6	6,9680	0,000	113,63%
7	7,6439	0,000	94,74%
8	3,0234	0,000	392,35%
9	3,5389	0,001	320,63%
10	9,3410	0,001	59,36%
11	4,5235	0,000	229,07%
12	4,9881	0,001	198,43%
13	10,6514	0,01	39,75%
14	11,0229	0,008	35,04%
15	11,3674	0,017	30,95%
16	6,6509	0,017	123,81%
17	7,0200	0,032	112,04%
18	7,3720	0,066	101,92%
19	12,5264	0,261	18,83%
20	8,0281	0,261	85,42%
21	8,3339	0,527	78,61%
22	13,2189	2,098	12,61%
23	13,4236	4,167	10,89%
24	13,6171	8,331	9,32%
25	9,4280	8,436	57,89%
26	13,9739	34,286	6,52%
27	14,1387	71,794	5,28%
28	14,2953	144,877	4,13%
29	14,4445	291,819	3,05%

FIG 28. Approximation of $h(0, x)$ with the (FS) method and non constant volatility.

Lemma 5.1. *Suppose that the conditions of Theorem (2.1) hold. Then, there exists a constant $C > 0$ which does not depend on x such that*

$$\sup_{t \in [t_{n-1}^n, T]} |h^n(t, x) - h(t, x)| \leq \frac{C|x|}{\sqrt{n}}.$$

Proof. Let us define $\delta_t^n = |h(t, x) - h^n(t, x)|$. Recall that $h(t, x) = \mathbb{E}g(S_{x,t}(T))$ where $S_{x,t}$ is the solution to (6.14). We have:

$$\begin{aligned}
\delta_t^n &= |h(t, x) - \frac{1}{2}g(k_{t_{n-1}}^{n-} x) - \frac{1}{2}g(k_{t_{n-1}}^{n+} x)| \\
&\leq |h(t, x) - g(x)| + \frac{1}{2}|g(k_{t_{n-1}}^{n-} x) - g(x)| + \frac{1}{2}|g(k_{t_{n-1}}^{n+} x) - g(x)| \\
&\leq \sup_{t_{n-1}^n \leq t \leq T} |h(t, x) - g(x)| + \frac{\mathbb{L}g}{2}(1 - k_{t_{n-1}}^d)|x| + \frac{\mathbb{L}g}{2}(k_{t_{n-1}}^u - 1)|x| \\
&\leq \mathbb{L}g \sup_{t_{n-1}^n \leq t \leq T} \mathbb{E}|S_{x,t}(T) - x| + \frac{\mathbb{L}g}{2}\sigma(t, x)\sqrt{\frac{T}{n}}|x| + \frac{\mathbb{L}g}{2}\sigma(t, x)\sqrt{\frac{T}{n}}|x| \\
&\leq C|x|\sqrt{\frac{T}{n}} + \frac{\mathbb{L}g}{2}\sigma^*\sqrt{\frac{T}{n}}|x|,
\end{aligned}$$

where the last inequality is deduced from Lemma 6.1 and $\sigma^* = \sup_{t \in [0, T], x \in \mathbb{R}} |\sigma(t, x)|$.

□

Lemma 5.2. *Suppose that the conditions of Theorem (2.1) hold. Then, there exists a constant $C > 0$ such that*

$$|h^n(t_{n-2}^n, x) - h(t_{n-2}^n, x)| \leq \frac{C|x|}{\sqrt{n}}.$$

Proof. Let us introduce $\delta_{t-2}^n = |h^n(t_{n-2}^n, x) - h(t_{n-2}^n, x)|$. We have

$$\begin{aligned}
\delta_{t-2}^n &= \left| \frac{1}{2}h^n(t_{n-1}^n, k_{t_{n-2}^n}^{n-} x) + \frac{1}{2}h^n(t_{n-1}^n, k_{t_{n-2}^n}^{n+} x) - h(t_{n-2}^n, x) \right| \\
&\leq \frac{1}{2}|h^n(t_{n-1}^n, k_{t_{n-2}^n}^{n-} x) - h(t_{n-2}^n, x)| + \frac{1}{2}|h^n(t_{n-1}^n, k_{t_{n-2}^n}^{n+} x) - h(t_{n-2}^n, x)| \\
&\leq \frac{1}{2}|h^n(t_{n-1}^n, k_{t_{n-2}^n}^{n-} x) - h(t_{n-1}^n, k_{t_{n-2}^n}^{n-} x)| + \frac{1}{2}|h(t_{n-1}^n, k_{t_{n-2}^n}^{n-} x) - h(t_{n-2}^n, k_{t_{n-2}^n}^{n-} x)| \\
&\quad + \frac{1}{2}|h(t_{n-2}^n, k_{t_{n-2}^n}^{n-} x) - h(t_{n-2}^n, x)| + \frac{1}{2}(|h^n(t_{n-1}^n, k_{t_{n-2}^n}^{n+} x) - h(t_{n-1}^n, k_{t_{n-2}^n}^{n+} x)| \\
&\quad + \frac{1}{2}|h(t_{n-1}^n, k_{t_{n-2}^n}^{n+} x) - h(t_{n-2}^n, k_{t_{n-2}^n}^{n+} x)| + \frac{1}{2}|h(t_{n-2}^n, k_{t_{n-2}^n}^{n+} x) - h(t_{n-2}^n, x)|).
\end{aligned}$$

Under Condition D, by the mean value theorem, we deduce that

$$\begin{aligned}
|h(t_{n-2}^n, x) - h(t_{n-2}^n, k_{t_{n-2}^n}^{n-} x)| &\leq C |x - k_{t_{n-2}^n}^{n-} x| \leq C |x| \cdot |1 - k_{t_{n-2}^n}^{n-}| \\
&\leq C |x| \sqrt{\frac{T}{n}}, \\
|h(t_{n-2}^n, x) - h(t_{n-2}^n, k_{t_{n-2}^n}^{n+} x)| &\leq C |x - k_{t_{n-2}^n}^{n+} x| \leq C |x| \cdot |1 - k_{t_{n-2}^n}^{n+}| \\
&\leq C |x| \sqrt{\frac{T}{n}}.
\end{aligned}$$

Note that, under Condition D, we have

$$|h_t(t, x)| = \frac{x^2 \sigma^2(t, x)}{2} |h_{xx}(t, x)| \leq \frac{C |x|}{\sqrt{T-t}}. \quad (5.8)$$

Therefore, by the mean value theorem, since $k_{t_{n-2}^n}^{n-}, k_{t_{n-2}^n}^{n+} \leq M$,

$$\begin{aligned}
|h(t_{n-1}^n, k_{t_{n-2}^n}^{n-} x) - h(t_{n-2}^n, k_{t_{n-2}^n}^{n-} x)| &\leq \frac{CM |x|}{\sqrt{T-t_{n-1}^n}} \Delta t_{n-1}^n \leq CM |x| \sqrt{\frac{T}{n}}, \\
|h(t_{n-1}^n, k_{t_{n-2}^n}^{n+} x) - h(t_{n-2}^n, k_{t_{n-2}^n}^{n+} x)| &\leq \frac{CM |x|}{\sqrt{T-t_{n-1}^n}} \Delta t_{n-1}^n \leq CM |x| \sqrt{\frac{T}{n}}.
\end{aligned}$$

At last, by Lemma 5.1,

$$\begin{aligned}
\gamma_-^n(t_{n-1}^n, x) &:= |h^n(t_{n-1}^n, k_{t_{n-2}^n}^{n-} x) - h(t_{n-1}^n, k_{t_{n-2}^n}^{n-} x)| \leq C k_{t_{n-2}^n}^{n-} |x| \sqrt{\frac{T}{n}}, \\
\gamma_+^n(t_{n-1}^n, x) &:= |h^n(t_{n-1}^n, k_{t_{n-2}^n}^{n+} x) - h(t_{n-1}^n, k_{t_{n-2}^n}^{n+} x)| \leq C k_{t_{n-2}^n}^{n+} |x| \sqrt{\frac{T}{n}}.
\end{aligned}$$

As $k_{t_{n-2}^n}^{n-} + k_{t_{n-2}^n}^{n+} = 2$, we deduce that

$$\frac{1}{2} \gamma_-^n(t_{n-1}^n, x) + \frac{1}{2} \gamma_+^n(t_{n-1}^n, x) \leq C |x| \sqrt{\frac{T}{n}}.$$

The conclusion follows. \square

Lemma 5.3. *Suppose that the conditions of Theorem (2.1) hold. Then, there exists a constant $C > 0$ such that*

$$|h^n(t_i^n, x) - h(t_i^n, x)| \leq \frac{C|x|}{\sqrt{n}}, \quad \text{for all } i \leq n-2.$$

Proof. We first prove the result for $i = n - 2$ and, then, we generalize the result by induction. Let us introduce the function

$$F_i(x) = \frac{1}{2}h(t_i^n, k_{t_{i-1}^n}^{n-})x + \frac{1}{2}h(t_i^n, k_{t_{i-1}^n}^{n+}x).$$

By the Taylor formula, we get that

$$\begin{aligned} F_i(x) &= \frac{1}{2} [h(t_{i-1}^n, x) + R_i^{11}(x) + R_i^{12}(x) + R_i^{13}(x)] \\ &+ \frac{1}{2} [h(t_{i-1}^n, x) + R_i^{21}(x) + R_i^{22}(x) + R_i^{23}(x)]. \end{aligned}$$

where, for some constants α_i , $i = 1, \dots, 4$, and $\tilde{t}_{i-1}^n \in [t_{i-1}^n, t_i^n]$, $\tilde{x} = x + \tilde{\theta}x(k_{t_{i-1}^n}^{n-} - 1)$, $\tilde{\theta} \in [0, 1]$, $\hat{t}_{i-1}^n \in [t_{i-1}^n, t_i^n]$, $\hat{x} = x + \hat{\theta}x(k_{t_{i-1}^n}^{n+} - 1)$, $\hat{\theta} \in [0, 1]$, we have

$$\begin{aligned} R_i^{11}(x) &= h_t(t_{i-1}^n, x)\frac{T}{n} - h_x(t_{i-1}^n, x)x\sigma(t_{i-1}^n, x)\sqrt{\frac{T}{n}}, \\ R_i^{12}(x) &= -h_{tx}(t_{i-1}^n, x)x\sigma(t_{i-1}^n, x)\left(\frac{T}{n}\right)^{\frac{3}{2}} + \frac{1}{2}h_{xx}(t_{i-1}^n, x)x^2\sigma^2(t_{i-1}^n, x)\frac{T}{n} \\ &+ \frac{1}{2}h_{tt}(t_{i-1}^n, x)\left(\frac{T}{n}\right)^2, \\ R_i^{13}(x) &= \alpha_1h_{ttt}(\tilde{t}_{i-1}^n, \tilde{x})\left(\frac{T}{n}\right)^3 - \alpha_2h_{tx}(\tilde{t}_{i-1}^n, \tilde{x})\sigma(t_{i-1}^n, x)x\left(\frac{T}{n}\right)^{5/2} \\ &+ \alpha_3h_{txx}(\tilde{t}_{i-1}^n, \tilde{x})\sigma^2(t_{i-1}^n, x)x^2\left(\frac{T}{n}\right)^2 - \alpha_4h_{xxx}(\tilde{t}_{i-1}^n, \tilde{x})\sigma^3(t_{i-1}^n, x)x^3\left(\frac{T}{n}\right)^{3/2}, \\ R_i^{21}(x) &= h_t(t_{i-1}^n, x)\frac{T}{n} + h_x(t_{i-1}^n, x)x\sigma(t_{i-1}^n, x)\sqrt{\frac{T}{n}}, \\ R_i^{22}(x) &= h_{tx}(t_{i-1}^n, x)x\sigma(t_{i-1}^n, x)\left(\frac{T}{n}\right)^{\frac{3}{2}} + \frac{1}{2}h_{xx}(t_{i-1}^n, x)x^2\sigma^2(t_{i-1}^n, x)\frac{T}{n} \\ &+ \frac{1}{2}h_{tt}(t_{i-1}^n, x)\left(\frac{T}{n}\right)^2, \\ R_i^{13}(x) &= \alpha_1h_{ttt}(\hat{t}_{i-1}^n, \hat{x})\left(\frac{T}{n}\right)^3 + \alpha_2h_{tx}(\hat{t}_{i-1}^n, \hat{x})\sigma(t_{i-1}^n, x)x\left(\frac{T}{n}\right)^{5/2} \\ &+ \alpha_3h_{txx}(\hat{t}_{i-1}^n, \hat{x})\sigma^2(t_{i-1}^n, x)x^2\left(\frac{T}{n}\right)^2 + \alpha_4h_{xxx}(\hat{t}_{i-1}^n, \hat{x})\sigma^3(t_{i-1}^n, x)x^3\left(\frac{T}{n}\right)^{3/2}. \end{aligned}$$

We deduce that $F_i(x) = h(t_{i-1}^n, x) + R_i^1(x) + R_i^2(x) + R_i^3(x)$ where $R_i^1(x) = \frac{1}{2}(R_i^{11}(x) + R_i^{21}(x))$, $R_i^2(x) = \frac{1}{2}(R_i^{12}(x) + R_i^{22}(x))$ and $R_i^3(x) = \frac{1}{2}(R_i^{13}(x) + R_i^{23}(x))$. As h is the solution to (2.4), we may simplify $F_i(x)$ as

$$F_i(x) = h(t_{i-1}^n, x) + \frac{1}{2}h_{tt}(t_{i-1}^n, x) \left(\frac{T}{n}\right)^2 + R_i^3(x).$$

Under Condition D, there exists a constant C independent of i and n such that

$$\frac{1}{2} |h_{tt}(t_{i-1}^n, x)| \leq \frac{C |x|}{(T - t_{i-1}^n)^{\frac{3}{2}}}. \quad (5.9)$$

On the other hand, since $\tilde{t}_{i-1}^n, \hat{t}_{i-1}^n \in [t_{i-1}^n, t_i^n]$ and $|\hat{x}| + |\tilde{x}| \leq C|x|$ where $C > 0$ is independent of i and n , the residual error $R_i^3(x)$ is the sum of terms which may be dominated under Condition D using the following inequalities:

$$|\alpha_1 h_{ttt}(\hat{t}_{i-1}^n, \hat{x})| \leq \frac{C |x|}{(T - t_i^n)^{\frac{5}{2}}}. \quad (5.10)$$

$$|\alpha_2 h_{xtt}(\hat{t}_{i-1}^n, \hat{x}) \sigma(t_{i-1}^n, x) x| \leq \frac{C |x|}{(T - t_i^n)^2}. \quad (5.11)$$

$$|\alpha_3 h_{txx}(\hat{t}_{i-1}^n, \hat{x}) \sigma^2(t_{i-1}^n, x) x^2| \leq \frac{C |x|}{(T - t_i^n)^{\frac{3}{2}}}. \quad (5.12)$$

The inequalities above are also satisfied if we replace \hat{t}_{i-1}^n by \tilde{t}_{i-1}^n and \hat{x} by \tilde{x} . Recall that the constant C may denote distinct constants that change for a line to the next one but this constant does not depend on i and n . At last, by the Taylor formula, we rewrite the last term as

$$\begin{aligned} \epsilon_i &= \alpha_4 (h_{xxx}(\hat{t}_{i-1}^n, \hat{x}) \sigma^3(t_{i-1}^n, x) - h_{xxx}(\tilde{t}_{i-1}^n, \tilde{x}) \sigma^3(t_{i-1}^n, x)) x^3 \left(\frac{T}{n}\right)^{3/2}, \\ &= \alpha_4 h_{xxx}(\tilde{t}_{i-1}^n, \tilde{x}) x^3 (\hat{t}_{i-1}^n - \tilde{t}_{i-1}^n) \left(\frac{T}{n}\right)^{3/2} + \alpha_4 h_{xxxx}(\tilde{t}_{i-1}^n, \tilde{x}) x^3 (\hat{x} - \tilde{x}) \left(\frac{T}{n}\right)^{3/2}, \end{aligned}$$

where $\bar{t}_{i-1}^n \in [\tilde{t}_{i-1}^n, \hat{t}_{i-1}^n]$ and $\bar{x} = \alpha \tilde{x} + (1 - \alpha)\hat{x}$ for some $\alpha \in [0, 1]$. It is easily seen that $|\hat{t}_{i-1}^n - \tilde{t}_{i-1}^n| \leq CT/n$ for some constant C independent of i and n . Similarly, by definition of \tilde{x} and \hat{x} , there exists a constant C such that $|\hat{x} - \tilde{x}| \leq C|x|\sqrt{T/n}$. Therefore,

$$|\epsilon_i| \leq \frac{C|x|}{(T-t_i^n)^2} \left(\frac{T}{n}\right)^{5/2} + \frac{C|x|}{(T-t_i^n)^{3/2}} \left(\frac{T}{n}\right)^2. \quad (5.13)$$

Let us introduce $\delta_{t_{i-1}}^n(x) := |h^n(t_{i-1}^n, x) - h(t_{i-1}^n, x)|$, $i \leq n-2$. From above, we get that

$$\begin{aligned} \delta_{t_{i-1}}^n(x) &= \left| \frac{1}{2}h^n(t_i^n, k_{t_{i-1}}^{n-} x) + \frac{1}{2}h^n(t_i^n, k_{t_{i-1}}^{n+} x) - h(t_{i-1}^n, x) \right| \\ &= \left| \frac{1}{2}h^n(t_i^n, k_{t_{i-1}}^{n-} x) + \frac{1}{2}h^n(t_i^n, k_{t_{i-1}}^{n+} x) - F_i(x) + \frac{1}{2}h_{tt}(t_{i-1}^n, x) \left(\frac{T}{n}\right)^2 + R_i^3(x) \right| \\ &\leq \frac{1}{2}\delta_{t_i}^n(k_{t_{i-1}}^{n-} x) + \frac{1}{2}\delta_{t_i}^n(k_{t_{i-1}}^{n+} x) + \frac{1}{2} \left(\frac{T}{n}\right)^2 |h_{tt}(t_{i-1}^n, x)| + |R_i^3(x)|. \end{aligned}$$

Using the inequalities (5.9), (5.10), \dots , (5.13), we then deduce a constant $C > 0$ independent of i and n such that

$$\begin{aligned} \delta_{t_{i-1}}^n(x) &\leq \frac{1}{2}\delta_{t_i}^n(k_{t_{i-1}}^{n-} x) + \frac{1}{2}\delta_{t_i}^n(k_{t_{i-1}}^{n+} x) \\ &\quad + \frac{C|x|}{(T-t_i^n)^{1/2}} \left(\frac{T}{n}\right)^{3/2} + \frac{C|x|}{(T-t_i^n)^{3/2}} \left(\frac{T}{n}\right)^2 \\ &\quad + \frac{C|x|}{(T-t_i^n)^2} \left(\frac{T}{n}\right)^{5/2} + \frac{C|x|}{(T-t_i^n)^{5/2}} \left(\frac{T}{n}\right)^3. \end{aligned}$$

Recall that, by Lemma 5.2,

$$\frac{1}{2} \left(\delta_{t_{n-2}}^n(k_{t_{n-3}}^{n-} x) + \delta_{t_{n-2}}^n(k_{t_{n-3}}^{n+} x) \right) \leq \frac{1}{2}C|x|(k_{t_{n-3}}^{n-} + k_{t_{n-3}}^{n+})\sqrt{\frac{T}{n}} = C|x|\sqrt{\frac{T}{n}}.$$

We deduce that

$$\begin{aligned} |\delta_{t_{n-3}}^n(x)| &\leq C|x| \sqrt{\frac{T}{n}} + \frac{C|x|}{(T - t_{n-2}^n)^{\frac{1}{2}}} \left(\frac{T}{n}\right)^{\frac{3}{2}} + \frac{C|x|}{(T - t_{n-2}^n)^{\frac{3}{2}}} \left(\frac{T}{n}\right)^2 \\ &\quad + \frac{C|x|}{(T - t_{n-2}^n)^2} \left(\frac{T}{n}\right)^{\frac{5}{2}} + \frac{C|x|}{(T - t_{n-2}^n)^{\frac{5}{2}}} \left(\frac{T}{n}\right)^3. \end{aligned}$$

Repeating the reasoning, given that $\frac{1}{2}(k_{t_{i-1}^n}^{n-} + k_{t_{i-1}^n}^{n+}) = 1$, we deduce by induction that, for every $i \leq n - 3$,

$$|\delta_{t_i}^n(x)| \leq C|x| \sqrt{\frac{T}{n}} + S_{i,n}^1 + S_{i,n}^2 + S_{i,n}^3 + S_{i,n}^4,$$

where, for $i \leq n - 3$,

$$\begin{aligned} S_{i,n}^1 &:= \sum_{j=i+1}^{n-2} \frac{C|x|}{(T - t_j^n)^{\frac{1}{2}}} \left(\frac{T}{n}\right)^{\frac{3}{2}}, & S_{i,n}^2 &= \sum_{j=i+1}^{n-2} \frac{C|x|}{(T - t_j^n)^{\frac{3}{2}}} \left(\frac{T}{n}\right)^2, \\ S_{i,n}^3 &= \sum_{j=i+1}^{n-2} \frac{C|x|}{(T - t_j^n)^2} \left(\frac{T}{n}\right)^{\frac{5}{2}}, & S_{i,n}^4 &= \sum_{j=i+1}^{n-2} \frac{C|x|}{(T - t_{n-2}^n)^{\frac{5}{2}}} \left(\frac{T}{n}\right)^3. \end{aligned}$$

Since $(T - u)^{-1} \geq (T - t_j^n)^{-1}$ if $u \in [t_j^n, t_{j+1}^n]$, we deduce that

$$\begin{aligned} S_{i,n}^1 &\leq C|x| \sqrt{\frac{T}{n}} \int_0^T \frac{1}{\sqrt{T-t}} dt \leq C|x| \sqrt{\frac{T}{n}}, \\ S_{i,n}^2 &\leq C|x| \frac{T}{n} \int_0^{T-T/n} \frac{1}{(T-t)^{3/2}} dt \leq C|x| \sqrt{\frac{T}{n}}, \\ S_{i,n}^3 &\leq C|x| \left(\frac{T}{n}\right)^{3/2} \int_0^{T-T/n} \frac{1}{(T-t)^2} dt \leq C|x| \sqrt{\frac{T}{n}}, \\ S_{i,n}^4 &\leq C|x| \left(\frac{T}{n}\right)^2 \int_0^{T-T/n} \frac{1}{(T-t)^{5/2}} dt \leq C|x| \sqrt{\frac{T}{n}}. \end{aligned}$$

Both with Lemma 5.1 and 5.2 and the inequalities above, we may conclude. \square

Corollary 5.4. *Suppose that the conditions of Theorem (2.1) hold. Then, there exists a constant $C > 0$ such that*

$$\sup_{t \in [0, T]} |h^n(t, x) - h(t, x)| \leq \frac{C|x|}{\sqrt{n}}.$$

Proof. Let $t \in [0, T]$ be such that $t \in [t_{i-1}^n, t_i^n)$ for some $i \geq 1$. By Lemma 5.1, we may suppose that $i \leq n - 2$. Then,

$$\begin{aligned} |h^n(t, x) - h(t, x)| &= |h^n(t_{i-1}^n, x) - h(t, x)| \\ &\leq |h^n(t_{i-1}^n, x) - h(t_{i-1}^n, x)| + |h(t_{i-1}^n, x) - h(t, x)|. \end{aligned}$$

By the mean value theorem and Inequality (5.8), since $T - t_i^n \geq T/n$, we deduce that

$$|h(t_{i-1}^n, x) - h(t, x)| \leq \frac{C|x|}{\sqrt{T - t_i^n}} \Delta t_i^n \leq \frac{C|x|}{\sqrt{n}}.$$

We then conclude by Lemma 5.3. \square

6. Appendix

Let us consider the unique solution $S_{x,t}$, $t \in [0, T]$, to the stochastic differential equation

$$dS_{t,x}(u) = S_{t,x}(u)\sigma(u, S_{t,x}(u))dW_u, \quad u \in [t, T], \quad S_{t,x}(t) = x \in \mathbb{R}, \quad (6.14)$$

where W is a standard Brownian motion.

Lemma 6.1. *Suppose that $t \in [0, T]$. Let $S_{x,t}$ be the solution to the stochastic differential equation (6.14). If σ is bounded by a constant $\sigma^* > 0$, there exists a constant C such that*

$$\mathbb{E} \sup_{t \leq u \leq T} S_{t,x}^2(u) \leq Cx^2, \quad \mathbb{E}(S_{t,x}(T) - x)^2 \leq Cx^2(T - t).$$

Proof. By the Doob's inequality, we obtain that for every $t \leq r \leq T$:

$$\phi(r) := \mathbb{E} \left| \sup_{t \leq u \leq r} S_{x,t}(u) \right|^2 \leq 4\mathbb{E} |S_{x,t}(r)|^2$$

As $S_{x,t}(r) = x + \int_t^r \sigma(u, S_{x,t}(u)) S_{x,t}(u) dW_u$, using the inequality $(a+b)^2 \leq 2(a^2 + b^2)$ and the Ito isometry, we get that :

$$\begin{aligned} \phi(r) &\leq 8x^2 + 8\mathbb{E} \left(\int_t^r \sigma(u, S_{x,t}(u)) S_{x,t}(u) dW_u \right)^2 \\ &\leq 8x^2 + 8\mathbb{E} \left(\int_t^r \sigma^2(u, S_{x,t}(u)) S_{x,t}^2(u) du \right) \\ &\leq 8x^2 + 8(\sigma^*)^2 \int_t^r \mathbb{E} S_{x,t}^2(u) du \\ &\leq 8x^2 + 8(\sigma^*)^2 \int_t^r \phi(u) du. \end{aligned}$$

Applying the Gronwall lemma, we deduce that:

$$\mathbb{E} | S_{x,t}(r) |^2 \leq 8x^2 \exp(8\sigma^*(T-t)) \leq Cx^2,$$

where C does not depend on x . By the Ito isometry, we then deduce that

$$\mathbb{E}(S_{x,t}(T) - x)^2 = \mathbb{E} \int_t^T \sigma^2(u, S_{x,t}(u)) S_{x,t}^2(u) du$$

Using the inequality above, we may conclude. \square

References

- [1] Bernardi C., Maday Y. Spectral methods. In Handbook of numerical analysis, Vol. V, Handb. Numer. Anal., V, 209-485. North-Holland, Amsterdam (1997).
- [2] Brenner P., Thome V. and Wahlbin Lars B. Besov spaces and applications to difference methods for initial value problems. Lecture Notes in Mathematics 434, Springer-Verlag, Berlin-Heidelberg-New York 1975.
- [3] Ciarlet P.G. The finite element method for elliptic problems. North-Holland, Amsterdam, 1978.
- [4] Cox J., Ross S.A., Rubinstein M. Option pricing: a simplified approach, Journal of Financial Economics, 7 (1979), 229-263.

- [5] Darses S., Lépinette E. Limit theorem for a modified Leland hedging strategy under constant transaction costs rate. Inspired by Finance, The Musiela Festschrift, Eds. Yu. Kabanov, M. Rutkowski, T. Zariphopoulou. Springer, 159-199 (2014), Springer.
- [6] Denis E. Marchés avec coûts de transaction: approximation de Leland et arbitrage. <http://indexation.univ-fcomte.fr/nuxeo/site/esupversions/8d5694b7-6409-413b-882a-37f2d7295a07>.
- [7] Diener F., Diener M. Asymptotics of the price oscillations of a European call option in a tree model. *Mathematical Finance* 14(2004), 271-293.
- [8] Friedman A. Stochastic differential equations and applications. Volume 1. *Academic Press* (1975).
- [9] Kabanov Y. and Denis E. Mean square error for the Leland-Lott hedging strategy: convex pay-off. *Finance and Stochastics*. 14 (2010),4, 626-667.
- [10] Lépinette E., Tran T. Approximate hedging in a local volatility model with proportional transaction costs. *Applied Mathematical Finance*, 21 (2014), 4, 313-341.
- [11] Milstein G.N. The Probability Approach to Numerical Solution of Non-linear Parabolic Equations. *Numerical methods for partial differential equations*, 18, 4 (2002), 490-522.
- [12] Prigent J.L. Weak convergence of financial markets. Springer-Verlag Berlin-Heidelberg (2003).
- [13] Quarteroni A. An introduction to spectral methods for partial differential equations. In *Advances in numerical analysis, I* (Lancaster, 1990), Oxford Sci. Publ., 96-146. Oxford Univ. Press, New York (1991).
- [14] Eymard R., Gallouet T., and Herbin R. Finite volume methods. In *Handbook of numerical analysis, Vol. VII, Handb. Numer. Anal.*, North-Holland, Amsterdam (2000), VII, 713-1020.
- [15] Hsia C-C. On binomial option pricing. *Journal of Financial Research*, 6 (1983), 41-5.
- [16] Richtmyer R.D., Morton K.W. Difference methods for initial value problems. Robert E. Krieger Publishing Co. Inc., Malabar, FL, second edition (1994).
- [17] Viitasaari L. Rate of convergence for discrete approximation of option prices. <https://arxiv.org/abs/1207.6756>.
- [18] Walsh J.B. The rate of convergence of the binomial tree scheme. *Finance and Stochastics* 7 (2003), 337-361.