

This is an ACCEPTED VERSION of the following published document:

J. G. López-Salas y C. Vázquez, «PDE formulation of some SABR/LIBOR market models and its numerical solution with a sparse grid combination technique», *Computers & Mathematics with Applications*, vol. 75, n.o 5, pp. 1616-1634, mar. 2018, doi: 10.1016/j.camwa.2017.11.024.

Link to published version: <https://doi.org/10.1016/j.camwa.2017.11.024>

General rights:

© 2018 Elsevier. This manuscript version is made available under the CC-BY-NC-ND 4.0 license <https://creativecommons.org/licenses/by-nc-nd/4.0/>. This version of the article has been accepted for publication in *Computers & Mathematics with Applications*. The Version of Record is available online at <https://doi.org/10.1016/j.camwa.2017.11.024>.

PDE formulation of some SABR/LIBOR market models and its numerical solution with a sparse grid combination technique[☆]

J. G. López-Salas^a, C. Vázquez^{a,*}

^a*Department of Mathematics, Faculty of Informatics, Campus Elviña s/n, 15071-A
Coruña (Spain)*

Abstract

SABR models have been used to incorporate stochastic volatility to LIBOR market models (LMM) in order to describe interest rate dynamics and price interest rate derivatives. From the numerical point of view, the pricing of derivatives with SABR/LIBOR market models (SABR/LMMs) is mainly carried out with Monte Carlo simulation. However, this approach could involve excessively long computational times. For first time in the literature, in the present paper we propose an alternative pricing based on partial differential equations (PDEs). Thus, we pose original PDE formulations associated to the SABR/LMMs proposed by Hagan [22], Mercurio & Morini [31] and Rebonato [35]. Moreover, as the PDEs associated to these SABR/LMMs are high dimensional in space, traditional full grid methods (like standard finite differences or finite elements) are not able to price derivatives over more than three or four underlying interest rates. In order to overcome this curse of dimensionality, a sparse grid combination technique is proposed. A comparison between Monte Carlo simulation results and the ones obtained with the sparse grid technique illustrates the performance of the method.

Keywords:

Stochastic volatility models, SABR/LIBOR market models, High

[☆]Partially financed by Spanish Grant MTM2013-47800-C2-1-P and by Xunta de Galicia (Grant CN2014/044). First author has also been founded by a FPU Spanish Grant.

*Corresponding author. Tel.: +34 981167000; fax: +34 981167160.

Email addresses: jose.lsalas@udc.es (J. G. López-Salas), carlosv@udc.es (C. Vázquez)

1. Introduction

The LMM [6, 25, 32] has become the most popular interest rate model. The main reason is the agreement between this model and Black's formulas [7]. The standard LIBOR market model considers constant volatilities for the forward rates. However, this is a very limited hypothesis since it is impossible to reproduce market volatility smiles.

Among the different stochastic volatility models offered in the literature, the SABR model proposed by Hagan, Kumar, Lesniewski and Woodward [21] in the year 2002 stands out for becoming the market standard to reproduce the price of European options. SABR is the acronym for Stochastic, Alpha, Beta and Rho, three of the four model parameters. The SABR model can not be used to price derivatives whose payoff depends on several forward rates. In fact, SABR model works in the terminal measure, under which both the forward rate and its volatility are martingales. This can always be done if we work with one forward rate in isolation at a time. Under this same measure, however, the process for another forward rate and for its volatility would not be driftless.

In order to allow LMM to fit market volatility smiles, different extensions of the LMM that incorporate the volatility smile by means of the SABR model were proposed. These models are known as SABR/LIBOR market models (SABR/LMMs). In this article we will deal with the models proposed by Hagan [22], Mercurio and Morini [31] and Rebonato [35].

While Monte Carlo [15] simulation remains the common choice for pricing interest rate derivatives within SABR/LMM setting, several difficulties motivate to address alternative approaches based on PDE formulations. The first issue is that the convergence of Monte Carlo methods, although it depends only very weakly on the dimension of the problem, is very slow. Indeed, if the standard deviation of the result using a single simulation is ϵ then the standard deviation of the error after N simulations is ϵ/\sqrt{N} . Therefore, to improve the accuracy of the solution by a factor of 10, 100 times as many simulations must be performed. The second drawback of Monte Carlo methods is the valuation of options with early-exercise, like in the case of the American options, due to the so-called "Monte Carlo on Monte Carlo" effect. Available Monte Carlo methods for American options are also quite costly, see [29]

for example. In contrast, the modification of the PDE to a linear complementarity problem is usually straightforward. Finally, the weakest point of Monte Carlo methods appears to be the computation of the sensitivities of the solution with respect to the underlyings, the so-called “Greeks”, which are very used by traders, and are directly given by the partial derivatives of the PDE solution. Besides, path-dependent options, like barrier options, can be easily priced in the PDE context where only the boundary conditions need to be changed, in contrast to Monte Carlo methods, where Brownian bridge techniques [16] must be applied.

In view of previous arguments, in the present paper we pose equivalent PDE formulations for the three above mentioned SABR/LMMs. As far as we now, this is the first time in the literature that these PDE formulations are posed. From the numerical point of view, one main difficulty in these PDE formulations lies in their high dimensionality in space-like variables. In order to cope with this so-called *curse of dimensionality* several methods are available in the literature, see [14, 4] for example, which can be put into three categories. The first group uses the Karhunen-Loeve transformation to reduce the stochastic differential equation to a lower dimensional equation, therefore this results in a lower dimensional PDE associated to the previously reduced SDE. The second category gathers those methods which try to reduce the dimension of the PDE itself, like for example dimension-wise decomposition algorithms. Finally, the third category groups the methods which reduce the complexity of the problem in the discretization layer, like for example the method of sparse grids, which we use in the present article.

The sparse grid method was originally developed by Smolyak [40], who used it for numerical integration. It is mainly based on a hierarchical basis [41, 42], a representation of a discrete function space which is equivalent to the conventional nodal basis, and a sparse tensor product construction. Zenger [44] and Bungartz and Griebel [8] extended this idea and applied sparse grids to solve PDEs with finite elements, finite volumes and finite differences methods. Besides working directly in the hierarchical basis, the sparse grid can also be computed using the combination technique [19] by linearly combining solutions on traditional Cartesian grids with different mesh widths. This is the approach we follow in this article. Recently, this technique has been used for a financial application related to the pricing of basket options in [24, 27, 37].

The paper is organized as follows. In Section 2 some basic concepts related to interest rate derivatives and the corresponding terminology and

notation are introduced. In Section 3 we pose the PDE formulations for the SABR/LMMs. In Section 4 we describe the use of a full grid finite differences scheme for the Mercurio and Morini model, the application of which is analogous for the other two SABR/LMMs. Numerical results show the limitations of the full grid method when the number of forward rates increases. Therefore, in Section 5 we describe the sparse grid combination technique applied to the SABR/LMM and show numerical results that illustrate the behaviour of the method when the number of forward rates increases. For this purpose, a comparison with Monte Carlo simulation results is used when analytic expressions of the solution are not available, as it happens in most of the cases. Note that Monte Carlo techniques are the usual alternative to price with SABR/LMM.

2. Interest rate derivatives. Caplets and swaptions

This section provides a brief introduction to the interest rate derivatives we deal with in the present article, for a deeper study we refer the reader to [7]. Interest rate derivatives consist of financial contracts that depend on some interest rates.

A zero coupon *bond* with maturity at time T is a contract that pays its holder one unit of currency at time T . The value of this product at time $t < T$ is denoted by $P(t, T)$, and is called the discount factor from time T to time t . Note that $P(T, T) = 1$ for all T .

A *tenor structure* is a set of ordered payment dates $\{T_i, i = 0, \dots, N\}$, such that

$$T_0 < T_1 < \dots < T_{N-1} < T_N.$$

The time between the payment dates is denoted by $\tau_i = T_{i+1} - T_i$. In terms of the corresponding discount factor, a payment of x units at time T_i is worth $xP(t, T_i)$ at time $t < T_i$.

A *forward* interest rate $F_i(t)$ is an interest rate we can contract in order to borrow or lend money during the future time period $[T_i, T_{i+1}]$, and can be expressed in terms of discount factors in the form:

$$F_i(t) = F(t; T_i, T_{i+1}) = \frac{1}{\tau_i} \left(\frac{P(t, T_i)}{P(t, T_{i+1})} - 1 \right) \text{ where } t \leq T_i.$$

Conversely, the price of a bond at time T_i that matures at T_j , $P(T_i, T_j)$,

can be expressed in terms of forward LIBOR rates as follows:

$$P(T_i, T_j) = \prod_{k=i}^{j-1} \frac{1}{1 + \tau_k F_k(T_i)}.$$

A *caplet* is a European call option on a forward rate. If the caplet expires at time T_{i+1} , at that time we will receive the payoff $\tau_i(F_i(T_i) - K)^+$, so that its discounted payoff at time $t < T_{i+1}$ is given by

$$P(t, T_{i+1})\tau_i(F_i(T_i) - K)^+,$$

where $(\cdot)^+$ denotes the function $\max(\cdot, 0)$ and K is the strike of the contract, which is given by a fixed interest rate in the contract. If constant volatilities are assumed, the caplet pricing can be computed with Black's formula [7]

$$P(t, T_{i+1})\tau_i \text{Bl}\left(K, F_i(t), \sigma_{Black}\sqrt{T_i - t}\right),$$

where

$$\text{Bl}(K, F, \nu) = F\Phi(d_1(K, F, \nu)) - K\Phi(d_2(K, F, \nu)), \quad (1)$$

$$d_1(K, F, \nu) = \frac{\ln(F/K) + \nu^2/2}{\nu},$$

$$d_2(K, F, \nu) = \frac{\ln(F/K) - \nu^2/2}{\nu},$$

and σ_{Black} is the constant volatility of the forward rate which can be retrieved from market quotes.

An interest rate *swap* (IRS) is a contract to exchange interest payments at future fixed dates. At every time instant T_{i+1} in a prescribed set of dates T_{a+1}, \dots, T_b the contract holder pays a fixed interest rate K and receives a floating interest rate at the LIBOR rate $F_i(T_i)$ fixed at time T_i . The discounted payoff at time $t < T_a$ of this swap can be expressed as

$$\text{IRS}(t; T_a, \dots, T_b) = \sum_{i=a}^{b-1} P(t, T_{i+1})\tau_i(F_i(T_i) - K). \quad (2)$$

A European $T_a \times (T_b - T_a)$ *swaption* is an option giving the right (and not the obligation) to enter a swap at the future time T_a , called the swaption maturity. The underlying swap length $T_b - T_a$ is referred as the tenor of the

swaption. Therefore, the discounted swaption payoff to the current time t is equal to

$$P(t, T_a)(\text{IRS}(T_a; T_a, \dots, T_b))^+.$$

As indicated in the introduction, the main objective of this article is the pricing of the previously described interest rate derivatives in the framework of SABR/LMM which incorporates the stochastic volatility by means of PDE formulations. Note that in most cases there are no analytical formulas for the solution. Thus, the new models are posed and solved with suitable numerical methods that overcome the high dimension in space of the equations. Moreover, the proposed sparse grid technique is parallelized to make the approach computationally efficient. Although along the paper we concentrate on Mercurio and Morini SABR/LMM, the proposed methodology can be analogously applied to Rebonato and Hagan models.

3. Derivation of the PDE from the stochastic processes

In [12] the authors analyzed the three SABR/LIBOR market models proposed by Hagan, Mercurio & Morini and Rebonato using Monte Carlo simulation and their implementation on GPUs in order to price several interest rate derivatives. They have concluded that the Mercurio & Morini model is the one with the best performance: it is the easiest to calibrate, it achieves the best fit to swaption market prices and it results the fastest one in the pricing with Monte Carlo simulation. Taking into account these reasons, in the present article we mainly choose this model to pose the PDE formulation and develop its numerical solution with the proposed methods. Nevertheless, at the end of this section we also pose the PDEs for the models of Hagan and Rebonato.

In order to describe the SABR/LMM setting, we first consider a set of $N - 1$ LIBOR forward rates F_i , $1 \leq i \leq N - 1$, $\mathbf{F} = (F_1, \dots, F_{N-1})$ on the tenor structure $[T_0, T_1, \dots, T_{N-1}, T_N]$, the accruals being $\tau_i = T_{i+1} - T_i$. The Mercurio & Morini model is defined by the following system of stochastic differential equations [31]:

$$\begin{aligned} dF_i(t) &= \mu_i(t)F_i(t)^\beta dt + \alpha_i V(t)F_i(t)^\beta dW_i^{\mathcal{Q}}(t), & F_i(0) \text{ given,} \\ dV(t) &= \sigma V(t)dZ^{\mathcal{Q}}(t), & V(0) = \alpha, \end{aligned} \tag{3}$$

which are posed on a probability space $\{\Omega, \mathcal{F}, \mathcal{Q}\}$ with filtration $\{\mathcal{F}_t\}$, $t \in [T_0, T_N]$. In (3) μ_i is the drift of the i -th forward rate, $\beta \in [0, 1]$ is the local

volatility coefficient, α_i is a deterministic (constant) instantaneous volatility coefficient, $W_i^{\mathcal{Q}}$ are standard Brownian motions under the risk neutral measure \mathcal{Q} , ρ is the correlation matrix between the forward rates, i.e.

$$\langle dW_i^{\mathcal{Q}}(t), dW_j^{\mathcal{Q}}(t) \rangle = \rho_{ij} dt, \quad \forall i, j \in \{1, \dots, N-1\},$$

V is the stochastic volatility of the forward rates, $dZ^{\mathcal{Q}}$ is a standard Brownian motion correlated with the Brownian motions of the forward rates and ϕ is the correlation vector between the forward rates and the stochastic volatility, i.e.

$$\langle dW_i^{\mathcal{Q}}(t), dZ^{\mathcal{Q}}(t) \rangle = \phi_i dt, \quad \forall i \in \{1, \dots, N-1\}.$$

Due to the fact that the volatility process is lognormal, one can set the initial value of the volatility equal to one, i.e. $\alpha = 1$ with no loss of generality, since any different initial value can be embedded in the model by adjusting the deterministic coefficients α_i . This is the choice we adopt in the following.

The drifts of the forward rates are determined by the chosen numeraire. Under the terminal probability measure \mathcal{Q}^{T_N} associated with choosing the bond $P(t, T_N)$ as numeraire, the drifts of the forwards rates are given by

$$\mu_i(t) = \begin{cases} -\alpha_i V(t)^2 \sum_{j=i+1}^{N-1} \frac{\tau_j F_j(t)^\beta}{1 + \tau_j F_j(t)} \rho_{ij} \alpha_j & \text{if } i < N-1, \\ 0 & \text{if } i = N-1. \end{cases}$$

In order to price $T_a \times (T_b - T_a)$ swaptions we will consider the probability measure \mathcal{Q}^{T_a} associated with choosing the bond $P(t, T_a)$ as numeraire. In this case the drifts of the forward rates are given by

$$\mu_i(t) = \begin{cases} 0 & \text{if } i = a, \\ \alpha_i V(t)^2 \sum_{j=a+1}^i \frac{\tau_j F_j(t)^\beta}{1 + \tau_j F_j(t)} \rho_{ij} \alpha_j & \text{if } i > a. \end{cases}$$

Our model for the correlation structure is taken from Rebonato [34], who suggests the time independent function

$$\rho_{ij} = e^{-\lambda|T_i - T_j|}. \quad (4)$$

This function reflects the fact that the correlation increases as the time between the forward rates expiry decreases, so that two consecutive forward rates influence each other more than a forward rate in many years time.

A European option is characterized by its payoff function G , which determines the amount $G(T, \mathbf{F}(T), V(T))$ its holder receives at time $t = T$. The arbitrage-free value of the option relative to a numeraire \mathcal{N} is then given by

$$u(t, \mathbf{F}(t), V(t)) = \frac{U(t, \mathbf{F}(t), V(t))}{N(t)} = \mathbb{E}^{\mathcal{Q}} \left(\frac{G(T, \mathbf{F}(T), V(T))}{\mathcal{N}(T)} \middle| \mathcal{F}_t \right). \quad (5)$$

Closed-form solutions based on (5) are rarely available due to the multi-asset feature of most LIBOR derivatives. In the next paragraphs we sketch the derivation of the PDE formulation associated to the Mercurio & Morini model.

By using Itô's formula, see [39] for example, the stochastic differential equation for u is given by

$$\begin{aligned} du(t, \mathbf{F}(t), V(t)) &= \frac{\partial u}{\partial t} dt + \sum_{i=1}^{N-1} \frac{\partial u}{\partial F_i(t)} dF_i(t) + \frac{\partial u}{\partial V(t)} dV(t) + \\ &\quad \frac{1}{2} \sum_{i,j=1}^{N-1} \frac{\partial^2 u}{\partial F_i(t) \partial F_j(t)} dF_i(t) dF_j(t) + \frac{1}{2} \frac{\partial^2 u}{\partial V(t)^2} (dV(t))^2 + \\ &\quad \sum_{i=1}^{N-1} \frac{\partial^2 u}{\partial F_i(t) \partial V(t)} dF_i(t) dV(t), \end{aligned} \quad (6)$$

with box algebra [30]:

| | | | | |
|----------------------|------|----------------------|----------------------|--------------------|
| | dt | $dW_i^{\mathcal{Q}}$ | $dW_j^{\mathcal{Q}}$ | $dZ^{\mathcal{Q}}$ |
| dt | 0 | 0 | 0 | 0 |
| $dW_i^{\mathcal{Q}}$ | 0 | dt | $\rho_{ij} dt$ | $\phi_i dt$ |
| $dW_j^{\mathcal{Q}}$ | 0 | $\rho_{ij} dt$ | dt | $\phi_j dt$ |
| $dZ^{\mathcal{Q}}$ | 0 | $\phi_i dt$ | $\phi_j dt$ | dt |

The interpretation of the box algebra is the following. In an expansion to terms of order dt , as $dt \rightarrow 0$ higher order terms such as $(dt)^j$ are all negligible for $j > 0$. For example, $(dt)^2$ is of order 0 as $dt \rightarrow 0$, which is denoted as $(dt)(dt) \sim 0$. Similarly, cross terms such as $(dt)(dW_i^{\mathcal{Q}})$ are negligible because the increment $dW_i^{\mathcal{Q}}$ is normally distributed with mean 0 and standard deviation $(dt)^{1/2}$ and so $(dt)(dW_i^{\mathcal{Q}})$ has standard deviation $(dt)^{3/2}$ which tends to 0 as $dt \rightarrow 0$.

Substituting equations (3) in (6) and using the box algebra, we get

$$\begin{aligned}
du(t, \mathbf{F}(t), V(t)) = & \left(\frac{\partial u}{\partial t} + \sum_{i=1}^{N-1} \mu_i(t) F_i(t)^\beta \frac{\partial u}{\partial F_i(t)} + \right. \\
& \frac{1}{2} \sum_{i,j=1}^{N-1} \alpha_i \alpha_j V(t)^2 F_i(t)^\beta F_j(t)^\beta \rho_{ij} \frac{\partial^2 u}{\partial F_i(t) \partial F_j(t)} + \\
& \left. \frac{1}{2} \sigma^2 V(t)^2 \frac{\partial^2 u}{\partial V(t)^2} + \sum_{i=1}^{N-1} \sigma V(t)^2 \alpha_i F_i(t)^\beta \phi_i \frac{\partial^2 u}{\partial F_i(t) \partial V(t)} \right) dt + \\
& \sum_{i=1}^{N-1} \alpha_i V(t) F_i(t)^\beta \frac{\partial u}{\partial F_i(t)} dW_i^\mathcal{Q} + \sigma V(t) \frac{\partial u}{\partial V(t)} dZ^\mathcal{Q}. \quad (7)
\end{aligned}$$

In order to comply with the no-arbitrage conditions and (5), the process $du(t, \mathbf{F}, V)$ has to be martingale under the measure \mathcal{Q} . Thus, to satisfy this requirement, the drift term dt in (7) must be equal to zero. The same result could be directly obtained by applying Feynman-Kac theorem, see [39, 7]. The final parabolic PDE takes the following form:

$$\begin{aligned}
\frac{\partial u}{\partial t} + \frac{1}{2} \sigma^2 V^2 \frac{\partial^2 u}{\partial V^2} + \frac{1}{2} V^2 \sum_{i,j=1}^{N-1} \rho_{ij} \alpha_i \alpha_j F_i^\beta F_j^\beta \frac{\partial^2 u}{\partial F_i \partial F_j} + \\
\sigma V^2 \sum_{i=1}^{N-1} \phi_i \alpha_i F_i^\beta \frac{\partial^2 u}{\partial F_i \partial V} + \sum_{i=1}^{N-1} \mu_i(t) F_i^\beta \frac{\partial u}{\partial F_i} = 0, \quad (8)
\end{aligned}$$

with the terminal condition given by the derivative payoff,

$$u(T, \mathbf{F}, V) = g(T, \mathbf{F}, V),$$

on $\mathbb{R}^{N-1} \times \mathbb{R}$. For simplicity of notation, we have used the relative payoff $g(\cdot) = \frac{G(\cdot)}{\mathcal{N}(T)}$. The derivative price at time $t < T$ is given by $\mathcal{N}(t)u(t, \mathbf{F}(t), V(t))$.

Analytic solutions for (8) can be only found for suitable simple specifications of the functionals forms of the PDE and for straightforward boundary conditions (e.g. simple caplets without stochastic volatility, i.e. $\sigma = 0$, see Section 4.2).

Finally, we are going to present the PDE for Hagan model, which is

defined by the following system of stochastic differential equations [22]:

$$\begin{aligned} dF_i(t) &= \mu^{F_i}(t)F_i(t)^{\beta_i}dt + V_i(t)F_i(t)^{\beta_i}dW_i^{\mathcal{Q}}(t), \quad F_i(0) \text{ given,} \\ dV_i(t) &= \mu^{V_i}(t)V_i(t)dt + \sigma_iV_i(t)dZ_i^{\mathcal{Q}}(t), \quad V_i(0) = \alpha_i, \end{aligned} \quad (9)$$

with the associated correlations denoted by

$$\begin{aligned} \langle dW_i^{\mathcal{Q}}(t), dW_j^{\mathcal{Q}}(t) \rangle &= \rho_{ij}dt, \\ \langle dW_i^{\mathcal{Q}}(t), dZ_j^{\mathcal{Q}}(t) \rangle &= \phi_{ij}dt, \\ \langle dZ_i^{\mathcal{Q}}(t), dZ_j^{\mathcal{Q}}(t) \rangle &= \theta_{ij}dt. \end{aligned}$$

The PDE for this model is obtained in the same way as previously with the Mercurio & Morini model, thus obtaining:

$$\begin{aligned} \frac{\partial u}{\partial t} + \frac{1}{2} \sum_{i,j=1}^{N-1} \theta_{ij} \sigma_i V_i \sigma_j V_j \frac{\partial^2 u}{\partial V_i \partial V_j} + \frac{1}{2} \sum_{i,j=1}^{N-1} \rho_{ij} V_i F_i^{\beta_i} V_j F_j^{\beta_j} \frac{\partial^2 u}{\partial F_i \partial F_j} + \\ \sum_{i,j=1}^{N-1} \phi_{ij} V_i F_i^{\beta_i} \sigma_j V_j \frac{\partial^2 u}{\partial F_i \partial V_j} + \sum_{i=1}^{N-1} \mu^{F_i}(t) F_i^{\beta_i} \frac{\partial u}{\partial F_i} + \sum_{i=1}^{N-1} \mu^{V_i}(t) V_i \frac{\partial u}{\partial V_i} = 0. \end{aligned} \quad (10)$$

Rebonato model [35] is analogous to Hagan one, therefore its PDE will be also quite similar to (10).

4. Finite Differences Method with full grids

Hereafter, as we have motivated in the previous section, we are going to just focus on the PDE (8) of the Mercurio & Morini model. This backward parabolic PDE must be supplemented with a terminal condition, which describes the value of the variable u at the final time T . Moreover, appropriate boundary conditions are required, which prescribe how the function u , or its derivatives, behave at the boundaries of the necessarily bounded computational domain.

We are going to define a $(N+1)$ -dimensional mesh with the time sampled from today (time 0) to the final expiry of the option (time T) at $M+1$ points uniformly spaced by the time step $\Delta t = \frac{T}{M}$.

The variables representing the forward rates $\mathbf{F} = (F_1, \dots, F_{N-1})$ and their stochastic volatility V , often referred as the “space variables” will be sampled

at $M_i + 1$ ($i = 1, \dots, N - 1$) and $S + 1$ points spaced by $h_i = \frac{F_i^{max} - F_i^{min}}{M_i}$ and $h_v = \frac{V^{max} - V^{min}}{S}$, respectively.

Notice that while the choice of the range of the time variable is totally unambiguous, $[0, T]$, an a priori choice must be made about which values of the space variables are too high or too low to be of interest, so far we will denote them by $[F_i^{min}, F_i^{max}]$ and $[V^{min}, V^{max}]$. Selecting boundary values such that the option of interest is too deeply in or out-of-the money is a common and reasonable choice.

For a given mesh, each point is uniquely determined by the time level m ($m = 0, \dots, M$), the index vector of the $N - 1$ forward rates $\mathbf{f} = (f_1, \dots, f_i, \dots, f_{N-1})$ ($f_i = 0, \dots, M_i$) and the stochastic volatility level v ($v = 0, \dots, S$). We seek approximations of the solution at these mesh points, which will be denoted by

$$U_{\mathbf{f},v}^m \approx u(m\Delta t, (f_i h_i)_{1 \leq i \leq N-1}, v h_v).$$

It is natural for this PDE to be solved backwards in time. We approximate the time derivative by the time-forward approximation

$$\left. \frac{\partial u}{\partial t} \right|_{t=m\Delta t, \mathbf{F}=(f_i h_i)_{1 \leq i \leq N-1}, V=v h_v} = \left. \frac{\partial u}{\partial t} \right|_{m, \mathbf{f}, v} \approx \frac{U_{\mathbf{f},v}^{m+1} - U_{\mathbf{f},v}^m}{\Delta t}.$$

For the space derivatives we have chosen second-order approximations. We will write $\mathbf{f}_{i \pm 1}$ to mean the forward rates index vector $(f_1, \dots, f_i \pm 1, \dots, f_{N-1})$ which corresponds to the forward rates point $(f_1 h_1, \dots, (f_i \pm 1) h_i, \dots, f_{N-1} h_{N-1})$.

The first derivatives are approximated by central differences:

$$\left. \frac{\partial u}{\partial F_i} \right|_{m, \mathbf{f}, v} \approx \frac{U_{\mathbf{f}_{i+1},v}^m - U_{\mathbf{f}_{i-1},v}^m}{2h_i}.$$

The second derivatives are approximated by:

$$\begin{aligned} \bullet \left. \frac{\partial^2 u}{\partial F_i^2} \right|_{m, \mathbf{f}, v} &\approx \frac{U_{\mathbf{f}_{i+1},v}^m - 2U_{\mathbf{f},v}^m + U_{\mathbf{f}_{i-1},v}^m}{h_i^2}, \\ \bullet \left. \frac{\partial^2 u}{\partial V^2} \right|_{m, \mathbf{f}, v} &\approx \frac{U_{\mathbf{f},v+1}^m - 2U_{\mathbf{f},v}^m + U_{\mathbf{f},v-1}^m}{h_v^2}. \end{aligned}$$

The cross derivatives terms are approximated by:

- For $i \neq j$, $\left. \frac{\partial^2 u}{\partial F_i \partial F_j} \right|_{m, \mathbf{f}, v} \approx \frac{U_{\mathbf{f}_{i+1, j+1, v}}^m + U_{\mathbf{f}_{i-1, j-1, v}}^m - U_{\mathbf{f}_{i+1, j-1, v}}^m - U_{\mathbf{f}_{i-1, j+1, v}}^m}{4h_i h_j}$,
- $\left. \frac{\partial^2 u}{\partial F_i \partial V} \right|_{m, \mathbf{f}, v} \approx \frac{U_{\mathbf{f}_{i+1, v+1}}^m + U_{\mathbf{f}_{i-1, v-1}}^m - U_{\mathbf{f}_{i+1, v-1}}^m - U_{\mathbf{f}_{i-1, v+1}}^m}{4h_i h_v}$.

The finite differences solution under the so-called θ -scheme is:

$$\frac{U_{\mathbf{f}, v}^{m+1} - U_{\mathbf{f}, v}^m}{\Delta t} + \theta W_{\mathbf{f}, v}^m + (1 - \theta) W_{\mathbf{f}, v}^{m+1} = 0,$$

where $\theta \in [0, 1]$ and $W_{\mathbf{f}, v}^m$ is the discretization given by

$$\begin{aligned} W_{\mathbf{f}, v}^m &= \frac{1}{2} \sigma^2 V^2 \frac{U_{\mathbf{f}, v+1}^m - 2U_{\mathbf{f}, v}^m + U_{\mathbf{f}, v-1}^m}{h_v^2} + \\ &\frac{1}{2} V^2 \sum_{\substack{i, j=1 \\ i \neq j}}^{N-1} \rho_{ij} \alpha_i \alpha_j F_i^\beta F_j^\beta \frac{U_{\mathbf{f}_{i+1, j+1, v}}^m + U_{\mathbf{f}_{i-1, j-1, v}}^m - U_{\mathbf{f}_{i+1, j-1, v}}^m - U_{\mathbf{f}_{i-1, j+1, v}}^m}{4h_i h_j} + \\ &\frac{1}{2} V^2 \sum_{i=1}^{N-1} \alpha_i^2 F_i^{2\beta} \frac{U_{\mathbf{f}_{i+1, v}}^m - 2U_{\mathbf{f}_i, v}^m + U_{\mathbf{f}_{i-1, v}}^m}{h_i^2} + \\ &\sigma V^2 \sum_{i=1}^{N-1} \phi_i \alpha_i F_i^\beta \frac{U_{\mathbf{f}_{i+1, v+1}}^m + U_{\mathbf{f}_{i-1, v-1}}^m - U_{\mathbf{f}_{i+1, v-1}}^m - U_{\mathbf{f}_{i-1, v+1}}^m}{4h_i h_v} + \\ &\sum_{i=1}^{N-1} \mu_i (m \Delta t) F_i^\beta \frac{U_{\mathbf{f}_{i+1, v}}^m - U_{\mathbf{f}_{i-1, v}}^m}{2h_i}, \end{aligned} \quad (11)$$

and with terminal condition $U_{\mathbf{f}, v}^M = g(T, \mathbf{F}, V)$, where we have denoted $\mathbf{F} = (F_i = f_i h_i)_{1 \leq i \leq N-1}$ and $V = v h_v$.

Three different θ values represent three canonical discretization schemes, $\theta = 0$ is the explicit scheme, $\theta = 1$ the fully implicit scheme and $\theta = 0.5$ the Crank-Nicolson scheme. The fully implicit discretization is the best method with respect to stability, whereas the Crank-Nicolson timestepping provides the best convergence rate. Although the explicit method is the simplest to implement, it has the disadvantage of not being unconditionally stable.

We shall first discriminate explicit and implicit parts as follows:

$$\frac{U_{\mathbf{f},v}^m}{\Delta t} - \theta W_{\mathbf{f},v}^m = \frac{U_{\mathbf{f},v}^{m+1}}{\Delta t} + (1 - \theta)W_{\mathbf{f},v}^{m+1}. \quad (12)$$

As a result of such discretization we arrive to the linear system of equations $\mathbf{A}\mathbf{x} = \mathbf{b}$, where \mathbf{A} is the band matrix of known coefficients, \mathbf{x} is the vector of the unknown solutions $U_{\mathbf{f},v}^m$ and \mathbf{b} is the vector of known values corresponding to the right-hand side of (12).

Equation (12) can be rewritten as:

$$\begin{aligned} & d\theta U_{\mathbf{f},v-1}^m + d\theta U_{\mathbf{f},v+1}^m + \sum_{i=1}^{N-1} (b_i - r_i)\theta U_{\mathbf{f}_{i-1},v}^m + \sum_{i=1}^{N-1} (b_i + r_i)\theta U_{\mathbf{f}_{i+1},v}^m + \\ & \sum_{i=1}^{N-1} (a_i\theta U_{\mathbf{f}_{i-1},v-1}^m + a_i\theta U_{\mathbf{f}_{i+1},v+1}^m - a_i\theta U_{\mathbf{f}_{i-1},v+1}^m - a_i\theta U_{\mathbf{f}_{i+1},v-1}^m) + \\ & \sum_{ij \in C} (\psi_{ij}\theta U_{\mathbf{f}_{i-1},j-1,v}^m + \psi_{ij}\theta U_{\mathbf{f}_{i+1},j+1,v}^m - \psi_{ij}\theta U_{\mathbf{f}_{i-1},j+1,v}^m - \psi_{ij}\theta U_{\mathbf{f}_{i+1},j-1,v}^m) + \\ & \left(-1 - 2d\theta - 2\theta \sum_{i=1}^{N-1} b_i \right) U_{\mathbf{f},v}^m = \\ & - d\hat{\theta} U_{\mathbf{f},v-1}^{m+1} - d\hat{\theta} U_{\mathbf{f},v+1}^{m+1} - \sum_{i=1}^{N-1} (b_i - r_i)\hat{\theta} U_{\mathbf{f}_{i-1},v}^{m+1} - \sum_{i=1}^{N-1} (b_i + r_i)\hat{\theta} U_{\mathbf{f}_{i+1},v}^{m+1} \\ & - \sum_{i=1}^{N-1} (a_i\hat{\theta} U_{\mathbf{f}_{i-1},v-1}^{m+1} + a_i\hat{\theta} U_{\mathbf{f}_{i+1},v+1}^{m+1} - a_i\hat{\theta} U_{\mathbf{f}_{i-1},v+1}^{m+1} - a_i\hat{\theta} U_{\mathbf{f}_{i+1},v-1}^{m+1}) \\ & - \sum_{ij \in C} (\psi_{ij}\hat{\theta} U_{\mathbf{f}_{i-1},j-1,v}^{m+1} + \psi_{ij}\hat{\theta} U_{\mathbf{f}_{i+1},j+1,v}^{m+1} - \psi_{ij}\hat{\theta} U_{\mathbf{f}_{i-1},j+1,v}^{m+1} - \psi_{ij}\hat{\theta} U_{\mathbf{f}_{i+1},j-1,v}^{m+1}) + \\ & \left(-1 + 2d\hat{\theta} + 2\hat{\theta} \sum_{i=1}^{N-1} b_i \right) U_{\mathbf{f},v}^{m+1}, \end{aligned} \quad (13)$$

where $\hat{\theta} = (1 - \theta)$, C is the set containing the combinations of numbers $1, 2, \dots, N-1$ taken two at a time without repetition (the number of elements in C is $\binom{N-1}{2} = 2^{-1}(N-1)(N-2)$) and the known coefficients d, b_i, r_i, a_i

and ψ_{ij} are defined as

$$\begin{aligned} d &= \frac{\Delta t \sigma^2 V^2}{2h_v^2}, & b_i &= \frac{\Delta t V^2 \alpha_i^2 F_i^{2\beta}}{2h_i^2}, \\ r_i &= \frac{\Delta t \mu_i(t) F_i^\beta}{2h_i}, & a_i &= \frac{\Delta t \sigma V^2 \phi_i \alpha_i F_i^\beta}{4h_i h_v}, \\ \psi_{ij} &= \frac{\Delta t V^2 \rho_{ij} \alpha_i \alpha_j F_i^\beta F_j^\beta}{4h_i h_j}. \end{aligned}$$

4.1. Boundary conditions

In order to specify boundary conditions, a combination of mathematical, financial and heuristic reasoning allows us to find consistent and acceptable ones. There are several possibilities, see [11] for example.

We assume that forward rates and their stochastic volatility are non negative and hence take values in the range zero to infinity. We first truncate the unbounded interval to a bounded one and then we must specify conditions at the new boundary. Thus we will consider the truncated domain $[F_i^{min}, F_i^{max}] \times [V^{min}, V^{max}]$, with $F_i^{min} = 0$ and $V^{min} = 0$.

For the forward rates we consider Dirichlet boundary conditions. Particularly, the terminal condition holds on the forward rates boundaries, i.e.

$$U_{\{\mathbf{f}|\exists f_i=0\},v}^m = U_{\mathbf{f},v}^M, \quad \forall m = 0, \dots, M-1,$$

$$U_{\{\mathbf{f}|\exists f_i=M_i\},v}^m = U_{\mathbf{f},v}^M, \quad \forall m = 0, \dots, M-1.$$

At the stochastic volatility boundaries we consider the following conditions:

$$\frac{\partial u}{\partial t} + \sum_{i=1}^{N-1} \mu_i(t) F_i^\beta \frac{\partial u}{\partial F_i} = 0, \quad V = 0, \quad (14)$$

$$\frac{\partial u}{\partial V} = 0, \quad V = V_{max}. \quad (15)$$

Thus, when $V = 0$ we require that the PDE itself must be satisfied on this boundary. When V approaches to infinity, the price of the derivative becomes independent of V . This is reflected by using Neumann conditions instead of the Dirichlet ones used for the forward rates boundaries.

At the boundary $V = 0$, after discretizing the boundary condition (14) we obtain (note that the coefficients d , b_i , a_i and ψ_{ij} of equation (13) are zero):

$$-\sum_{i=1}^{N-1} r_i \theta U_{\mathbf{f}_{i-1},0}^m + \sum_{i=1}^{N-1} r_i \theta U_{\mathbf{f}_{i+1},0}^m - U_{\mathbf{f},0}^m = \sum_{i=1}^{N-1} r_i \hat{\theta} U_{\mathbf{f}_{i-1},0}^{m+1} - \sum_{i=1}^{N-1} r_i \hat{\theta} U_{\mathbf{f}_{i+1},0}^{m+1} + U_{\mathbf{f},0}^{m+1}.$$

For the boundary $V = V_{max}$ in order to maintain the second order accuracy in the discretization of the first derivative the ghost point method is considered. The ghost grid points $U_{\mathbf{f},S+1}$ are added. Then, the finite differences scheme of equation (13) can also be applied at the points $U_{\mathbf{f},S}$. However, we now have more unknowns than equations. The additional equations come from the central finite differences discretization of the Neumann boundary condition (15):

$$\frac{U_{\mathbf{f},S+1} - U_{\mathbf{f},S-1}}{2h_v} = 0,$$

which yields $U_{\mathbf{f},S+1} = U_{\mathbf{f},S-1}$. Inserting this into the finite differences equation at $V = V_{max}$ we achieve

$$\begin{aligned} & \hat{d}\theta U_{\mathbf{f},S-1}^m + \sum_{i=1}^{N-1} (b_i - r_i) \theta U_{\mathbf{f}_{i-1},S}^m + \sum_{i=1}^{N-1} (b_i + r_i) \theta U_{\mathbf{f}_{i+1},S}^m + \\ & \sum_{ij \in C} (\psi_{ij} \theta U_{\mathbf{f}_{i-1,j-1},S}^m + \psi_{ij} \theta U_{\mathbf{f}_{i+1,j+1},S}^m - \psi_{ij} \theta U_{\mathbf{f}_{i-1,j+1},S}^m - \psi_{ij} \theta U_{\mathbf{f}_{i+1,j-1},S}^m) + \\ & \left(-1 - \hat{d}\theta - 2\theta \sum_{i=1}^{N-1} b_i \right) U_{\mathbf{f},S}^m = \\ & - \hat{d}\hat{\theta} U_{\mathbf{f},S-1}^{m+1} - \sum_{i=1}^{N-1} (b_i - r_i) \hat{\theta} U_{\mathbf{f}_{i-1},S}^{m+1} - \sum_{i=1}^{N-1} (b_i + r_i) \hat{\theta} U_{\mathbf{f}_{i+1},S}^{m+1} + \\ & - \sum_{ij \in C} (\psi_{ij} \hat{\theta} U_{\mathbf{f}_{i-1,j-1},S}^{m+1} + \psi_{ij} \hat{\theta} U_{\mathbf{f}_{i+1,j+1},S}^{m+1} - \psi_{ij} \hat{\theta} U_{\mathbf{f}_{i-1,j+1},S}^{m+1} - \psi_{ij} \hat{\theta} U_{\mathbf{f}_{i+1,j-1},S}^{m+1}) + \\ & \left(-1 + \hat{d}\hat{\theta} + 2\hat{\theta} \sum_{i=1}^{N-1} b_i \right) U_{\mathbf{f},S}^{m+1}, \end{aligned}$$

where $\hat{d} = 2d = \frac{\Delta t \sigma^2 V_{max}^2}{h_v^2}$.

4.2. Numerical results

It is not clear where to place F_i^{max} and V^{max} . On one hand, it is advantageous to place them far away of the initial forward rates. This reduces the error of the artificial boundary conditions. On the other hand a large computational domain requires a large discretization width. This increases the error of the approximation of the derivatives. In our experiments we will consider $F_i^{max} = 0.1$ and $V^{max} = 3.5$.

Some specifications of the financial product are given in Table 1 and the employed market data, taken from [5], are shown in Table 2. We will consider $\lambda = 0.1$ in the model for the correlation structure (4). Besides, the Crank-Nicolson scheme will be used in (12). For solving the system (13) the Gauss-Seidel iterative solver has been employed using a tolerance of 10^{-6} .

The numerical experiments have been performed with the following hardware and software configurations: two recent multicore Intel Xeon CPUs E5-2620 v2 clocked at 2.10 GHz (6 cores per socket) with 62 GBytes of RAM, CentOS Linux, GNU C++ compiler 4.8.2.

| | |
|-----------|---------|
| Currency | EUR |
| Index | EURIBOR |
| Day Count | e30/360 |
| Strike | 5.5% |

Table 1: Specification of the interest rate model.

We are going to value $T_a \times (T_b - T_a)$ European swaptions under the Mercurio & Morini model.

First of all, the results when pricing a 1×1 European swaption are discussed. The value of this particular swaption is the same as the price of the corresponding caplet, and so it only depends on F_1 . Hence, in one dimension a closed form expression for the price of a European swaption can be found by using Black's formula (1) and is given by

$$P(T_0, T_2)\tau_1 \text{Bl}(K, F_1(T_0), \sigma_{Black}\sqrt{T_1 - T_0}).$$

This value is equal to 0.659096 basis points (one basis point is one hundredth of one percent, $\frac{1\%}{100} = 10^{-4}$). As Black's formula for caplets considers constant volatility σ_{Black} , in this first test the volatility of the volatility parameter of Mercurio & Morini model is considered equal to zero, i.e., $\sigma = 0$, therefore a

| | Start date | End date | LIBOR Rate (%) | Volatility (%) |
|-------|------------|----------|----------------|----------------|
| T_0 | 29-07-04 | 29-07-05 | 2.423306 | 0 |
| T_1 | 29-07-05 | 29-07-06 | 3.281384 | 24.73 |
| T_2 | 29-07-06 | 29-07-07 | 3.931690 | 22.45 |
| T_3 | 29-07-07 | 29-07-08 | 4.364818 | 19.36 |
| T_4 | 29-07-08 | 29-07-09 | 4.680236 | 17.43 |
| T_5 | 29-07-09 | 29-07-10 | 4.933085 | 16.15 |
| T_6 | 29-07-10 | 29-07-11 | 5.135066 | 15.02 |
| T_7 | 29-07-11 | 29-07-12 | 5.273314 | 14.24 |
| T_8 | 29-07-12 | 29-07-13 | 5.376115 | 13.42 |

Table 2: Market data used in pricing. Data taken from 27th July 2004.

standard LIBOR market model is used. The solution was found on several levels and Table 3 shows the convergence of the model when using 256 time steps. In all tables of this article, *Level* refers to the refinement level n , i.e., the mesh size is $h_i = 2^{-n} \cdot c_i$ in each coordinate direction, where c_i denotes the computational domain length in direction i , which is F_i^{max} in the case of the forward rates and V^{max} in the case of the stochastic volatility. Besides, the solution and the error with respect to the exact solution are also shown in basis points. Additionally, the execution time is measured in seconds and the column labeled as *Grid points* shows the number of grid points employed in the full grid used by the finite differences method without taking into account the time coordinate. The time discretization error decreases as increasing the number of time steps. As an example, in Table 4 we show the achieved solutions and the associated errors using only 12 time steps for the higher resolution levels ($n = 9, 10$), these results are to be compared with those ones of Table 3. From the results in Tables 3 and 4, we trust that the time step $\Delta t = 1/256$ is small enough so that the reported errors are dominated by the spatial discretization, at least up to level 10. Therefore, in the following numerical tests we will consider 256 time steps.

When the volatility of the volatility σ of the model is non zero or when the length of the underlying swap of the swaption being considered is greater than one, no closed form solutions are available. However, an estimate can be obtained from Monte Carlo simulations. On Table 5 Monte Carlo values for the 1×1 European swaption with $\sigma = 0$ are shown for several numbers of

paths ($\#Paths$). More details about Monte Carlo simulation of SABR/LMMs can be found in the article [12].

| Level | Solution | Error | Time | Grid points |
|-------|----------|----------|----------|-------------|
| 3 | 2.529194 | 1.870098 | 0.01 | 81 |
| 4 | 1.205248 | 0.546151 | 0.03 | 289 |
| 5 | 0.802744 | 0.143647 | 0.21 | 1089 |
| 6 | 0.674394 | 0.015297 | 1.95 | 4225 |
| 7 | 0.663856 | 0.004760 | 24.42 | 16641 |
| 8 | 0.659937 | 0.000840 | 340.57 | 66049 |
| 9 | 0.659381 | 0.000285 | 4775.27 | 263169 |
| 10 | 0.659163 | 0.000066 | 65013.50 | 1050625 |

Table 3: Convergence of the full grid finite differences solution in basis points for 1 LIBOR and stochastic volatility, $\sigma = 0$, $V(0) = 1$, $\beta = 1$, 256 time steps. Exact solution, 0.659096 basis points.

| Level | Solution | Error | Time | Grid points |
|-------|----------|----------|----------|-------------|
| 9 | 0.658014 | 0.001082 | 1884.24 | 263169 |
| 10 | 0.657795 | 0.001301 | 23436.59 | 1050625 |

Table 4: Convergence of the full grid finite differences solution in basis points for 1 LIBOR and stochastic volatility, $\sigma = 0$, $V(0) = 1$, $\beta = 1$, 12 time steps. Exact solution, 0.659096 basis points.

| $\#Paths$ | CI |
|-----------|----------------------|
| 10^5 | [0.608003, 0.703098] |
| 10^7 | [0.654232, 0.663567] |
| 10^9 | [0.658629, 0.659562] |

Table 5: 95% confidence intervals (CI) with Monte Carlo solution in basis points for 1 LIBOR and stochastic volatility, $\sigma = 0$, $V(0) = 1$, $\beta = 1$, 256 time steps. Exact solution, 0.659096 basis points.

In Table 6 the pricing of the 1×1 European swaption with $\sigma = 0.3$ for different resolution levels n are shown. In Table 7 the results for the 1×2 swaption are given. Note that with this numerical method it was not feasible to price the swaption past refinement level $n = 8$ due to the huge number of

required grid points. In Table 8 the results for the 1×3 swaption are given. Full grid pricing is only possible on low grid levels. It is not achievable to obtain a solution for a level greater than 6 in reasonable computational time due to the high number of grid points.

| Level | Solution | Time | Grid points |
|-------|----------|----------|-------------|
| 3 | 3.440 | 0.01 | 81 |
| 4 | 2.168 | 0.03 | 289 |
| 5 | 1.800 | 0.21 | 1089 |
| 6 | 1.678 | 2.20 | 4225 |
| 7 | 1.670 | 27.57 | 16641 |
| 8 | 1.667 | 376.91 | 66049 |
| 9 | 1.665 | 5206.93 | 263169 |
| 10 | 1.663 | 70492.91 | 1050625 |

Table 6: Convergence of the full grid finite differences solution in basis points for 1 LIBOR and stochastic volatility, $\sigma = 0.3$, $\phi_i = 0.4$, $V(0) = 1$, $\beta = 1$, 256 time steps. 95% confidence interval with Monte Carlo simulation using 10^7 paths, $[1.652, 1.672]$ in basis points.

| Level | Solution | Time | Grid points |
|-------|----------|-----------|-------------|
| 3 | 7.162 | 0.14 | 729 |
| 4 | 5.565 | 1.84 | 4913 |
| 5 | 5.003 | 34.41 | 35937 |
| 6 | 4.865 | 806.02 | 274625 |
| 7 | 4.846 | 21903.33 | 2146689 |
| 8 | 4.824 | 611725.64 | 16974593 |

Table 7: Convergence of the full grid finite differences solution in basis points for 2 LIBORs and stochastic volatility, $\sigma = 0.3$, $\phi_i = 0.4$, $V(0) = 1$, $\beta = 1$, 256 time steps. 95% confidence interval with Monte Carlo simulation using 10^7 paths, $[4.800, 4.844]$ in basis points.

Theoretically, it is possible to solve the discrete system (13) for a general number of dimensions. However, in computational science, a major problem occurs when the number of dimensions increases. A natural way to diminish the discretization error is to decrease the mesh step in each coordinate direction. However, then the number of grid points in the resulting full grid grows exponentially with the dimension, i.e. the size of the discrete solution

| Level | Solution | Time | Grid points |
|-------|----------|-----------|-------------|
| 3 | 11.702 | 2.17 | 6561 |
| 4 | 9.497 | 73.90 | 83521 |
| 5 | 8.892 | 3033.58 | 1185921 |
| 6 | 8.771 | 157152.75 | 17850625 |

Table 8: Convergence of the full grid finite differences solution in basis points for 3 LIBORs and stochastic volatility, $\sigma = 0.3$, $\phi_i = 0.4$, $V(0) = 1$, $\beta = 1$, 256 time steps. 95% confidence interval with Monte Carlo simulation using 10^7 paths, [8.635, 8.700] in basis points.

increases drastically. This is called the *curse of dimensionality* [3]. Therefore, this procedure of improving the accuracy by decreasing the mesh step is mainly bounded by two factors, the storage and the computational complexity. Due to these limitations, using a full grid discretization method which achieves sufficiently accurate approximations is only possible for problems with up to three or four dimensions, even on the most powerful machines presently available [8].

5. Sparse grids and the combination technique

Two approaches to try to overcome the curse of dimensionality are increasing the order of accuracy of the applied numerical approximation scheme or reducing the dimension of the problem by choosing suitable coordinates. Both approaches are not always possible for every option pricing problem. In this article we will take advantage of the sparse grid combination technique first introduced by Zenger and co-workers [19] in order to try to overcome the curse of dimensionality and allow to use the PDE formulation of SABR/LMM for the pricing problem we are dealing with. The combination technique replicates the structure of a so-called sparse grid by linearly combining solutions on coarser grids of the same dimensionality. This technique reduces the computational effort and the storage space involved with the mentioned traditional finite differences discretization methods. The number of sub-problems to solve increases, while the computational time per problem decreases drastically. This method can be implemented in parallel as each sub-grid is independent of the others. In the next two subsections we give a brief introduction to sparse grids and the combination technique. For a detailed discussion we refer to [8].

5.1. Sparse grids

First, we introduce some notations and definitions. Let $\mathbf{l} = (l_1, l_2, \dots, l_d) \in \mathbb{N}_0^d$ denote a d -dimensional multi-index. Let $|\mathbf{l}|_1$ and $|\mathbf{l}|_\infty$ denote the discrete L_1 -norm and L_∞ -norm of the multi-index \mathbf{l} , respectively, that are defined as

$$|\mathbf{l}|_1 = \sum_{k=1}^d l_k \quad \text{and} \quad |\mathbf{l}|_\infty = \max_{1 \leq k \leq d} l_k.$$

We define the anisotropic grid $\Omega_{\mathbf{l}}$ with mesh size $\mathbf{h} = (h_1, h_2, \dots, h_d) = (2^{-l_1} c_1, 2^{-l_2} c_2, \dots, 2^{-l_d} c_d)$ with multi-index \mathbf{l} and grid length $\mathbf{c} = (c_1, c_2, \dots, c_d)$.

Then, the full grid at refinement level $n \in \mathbb{N}$ and mesh size $h_i = 2^{-n} \cdot c_i$ for all i can be defined via the sequence of subgrids

$$\Omega^n = \Omega_{(n, \dots, n)} = \bigcup_{|\mathbf{l}|_\infty \leq n} \Omega_{\mathbf{l}}.$$

Figure 1 visualizes two dimensional full grids for levels $n = 0, \dots, 4$.

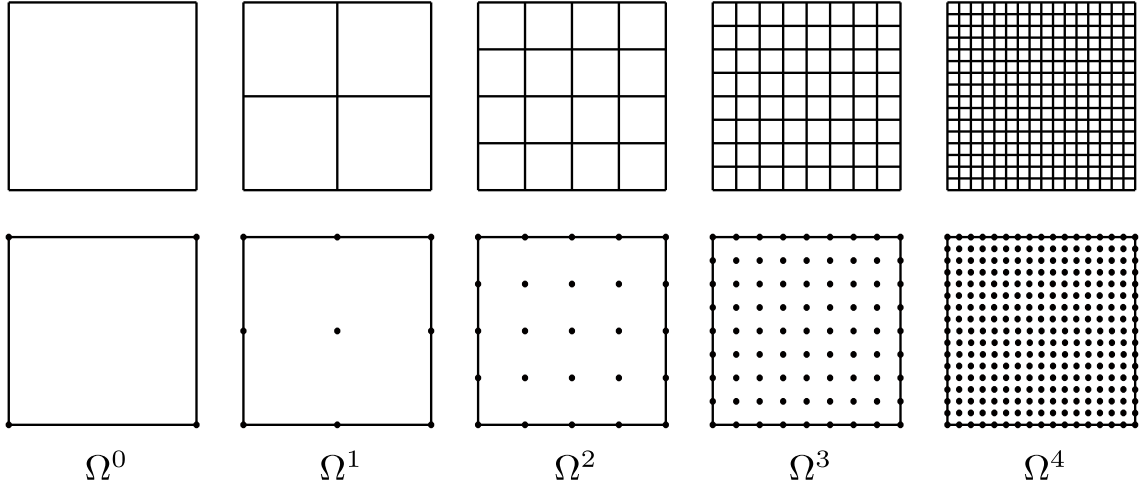


Figure 1: Two-dimensional full grid hierarchy up to level $n = 4$.

The number of grid points in each coordinate direction of the full grid is $2^n + 1$ and therefore the number of grid nodes in the full grid increases with $O(2^{n \cdot d})$, i.e. grows exponentially with the dimensionality d of the problem.

The sparse grid Ω_s^n at refinement level n consists of all anisotropic Cartesian grids $\Omega_{\mathbf{l}}$, where the total sum of all refinement factors l_k in each coordinate direction equals the resolution n . Then, the sparse grid Ω_s^n is given by

$$\Omega_s^n = \bigcup_{|\mathbf{l}|_1 \leq n} \Omega_{\mathbf{l}} = \bigcup_{|\mathbf{l}|_1 = n} \Omega_{\mathbf{l}}.$$

Figure 2 shows the two-dimensional grid hierarchy for levels $n = 0, \dots, 4$.

The total number of nodes in the grid $\Omega_{\mathbf{l}}$ is $\prod_{k=1}^d (2^{l_k} + 1) = O(2^{|\mathbf{l}|_1}) = O(2^n)$.

In addition, there exist exactly $\binom{n+d-1}{d-1}$ grids $\Omega_{\mathbf{l}}$ with $|\mathbf{l}|_1 = n$,

$$\begin{aligned} \binom{n+d-1}{d-1} &= \frac{(n+d-1)!}{(d-1)!n!} = \frac{(n+d-1) \cdot \dots \cdot (n+1)n!}{(d-1)!n!} \\ &= \frac{n+(d-1)}{d-1} \cdot \frac{n+(d-2)}{d-2} \cdot \dots \cdot \frac{n+(d-(d-1))}{d-(d-1)} \\ &= \left(1 + \frac{n}{d-1}\right) \cdot \left(1 + \frac{n}{d-2}\right) \cdot \dots \cdot \left(1 + \frac{n}{2}\right) \cdot \left(1 + \frac{n}{1}\right) \\ &\leq (1+n)^{d-1} = O(n^{d-1}). \end{aligned}$$

Thus, the total number of grid points of the sparse grid Ω_s^n grows according to

$$\binom{n+d-1}{d-1} \cdot \prod_{k=1}^d (2^{l_k} + 1) = O(n^{d-1})O(2^n) = O(n^{d-1}2^n), \quad (16)$$

which is far less the size of the corresponding full grid with $O(2^{nd})$ grid points. Let $h_n = 2^{-n}$, therefore the sparse grid employs $O(h_n^{-1} \cdot \log_2(h_n^{-1})^{d-1})$ grid points compared to $O(h_n^{-d})$ nodes in the full grid.

Bungartz and Griebel [8] show that the accuracy of the sparse grid using $O(h_n^{-1} \cdot \log_2(h_n^{-1})^{d-1})$ nodes is of order $O(h_n^2 \log_2(h_n^{-1})^{d-1})$ in the case of piecewise linear finite elements discretization and under the smoothness condition that the mixed derivatives are bounded. Thus, the accuracy of the sparse grid is only slightly deteriorated from the accuracy $O(h_n^2)$ of conventional full grid methods which need $O(h_n^{-d})$ grid points. Therefore, sparse grids need much less points than regular full grids to achieve a similar approximation quality.

However, the structure of a sparse grid is more complicated than the one of a full grid. Common PDE solvers usually manage only full grid solutions.

Existing sparse grid methods working directly in the hierarchical basis involve a challenging implementation [1, 43]. This handicap can be circumvented with the help of the sparse grid combination technique which not only exploits the economical structure of the sparse grids but also allows for the use of traditional full grid PDE solvers.

Concerning finite differences, in [9, 10] the authors obtain error bounds in terms of the Fourier transform coefficients when using a combination technique with a central difference scheme for the Laplace equation. The adaptive case with finite differences has been addressed in [18] for elliptic operators and the errors are obtained in L^2 and L^∞ -norms under the assumption that the $2d$ -th mixed derivatives are bounded. In [26], the smoothness condition is written in terms of Holder spaces, the continuity of the mixed derivatives and their associated semi-norm to be finite, also including more general finite differences schemes. More recently, in [36] a methodology to obtain error bounds for general finite differences schemes in any dimension is proposed. It is mainly based on an error correction scheme leading to an appropriate error expansion. The results are again based on the existence of bounded mixed derivatives of the solution in the L^∞ norm, which vanish at the boundary to avoid regularity problems. For this kind of functions, results about the expansion and interpolation are previously stated.

Concerning the smoothness of solution of the here treated PDEs, we note that the parabolic operator involves a smoothness effect (it is a kind of Black-Scholes operator in high spatial dimensions), so that the solution becomes C^∞ for $t < T$, and the usual payoff (final condition at $t = T$) is continuous. Note that sparse grids have been analyzed when applied to basket options in [37], for example.

Finally, two and three dimensional sparse grids for several resolution levels n are shown in Figures 3 and 4, respectively. Additionally, the growth of the grid points when increasing n can be observed.

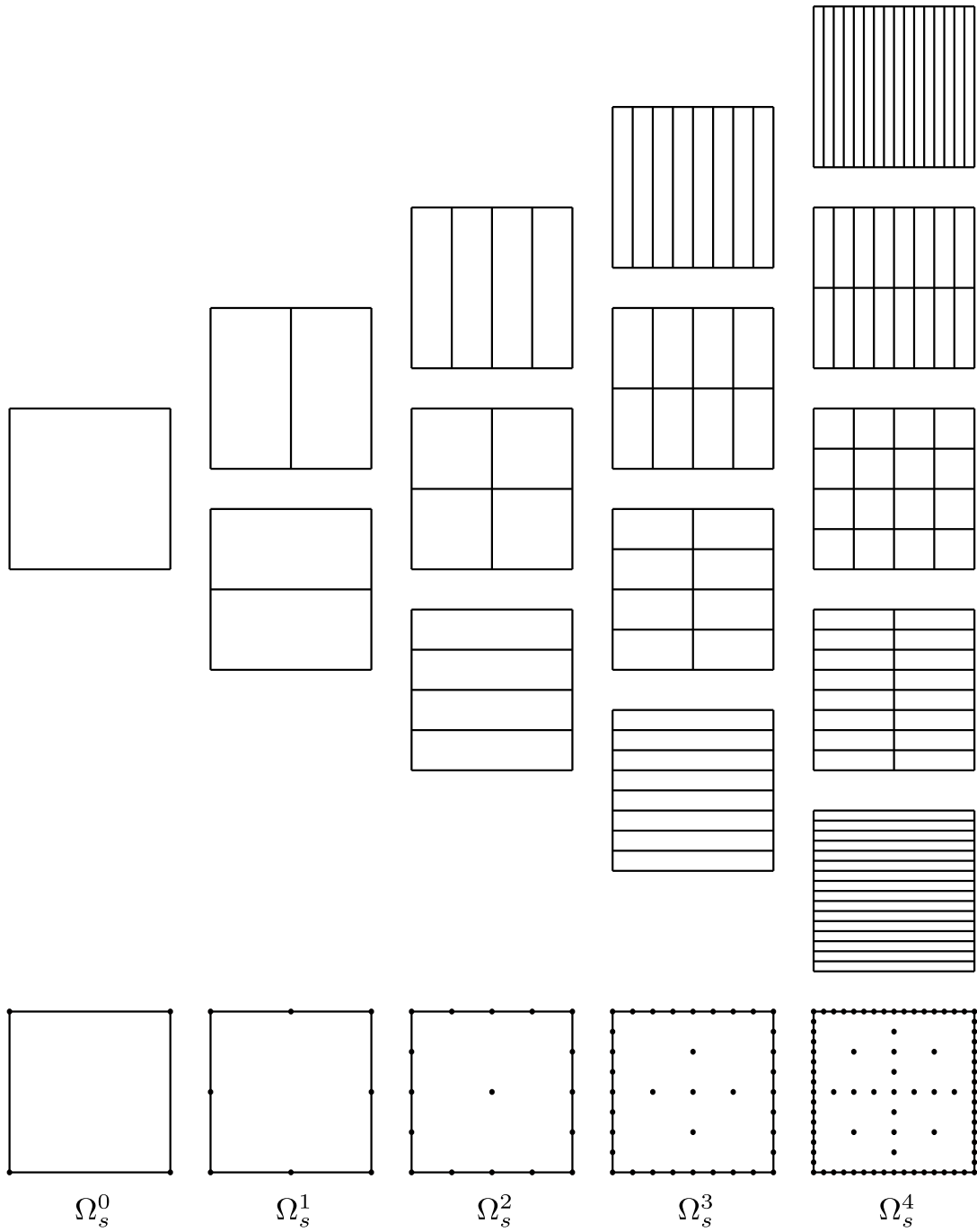
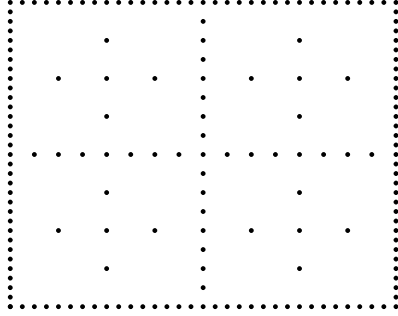
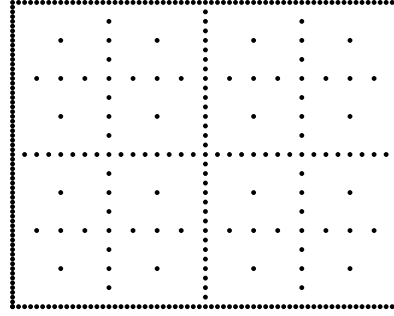


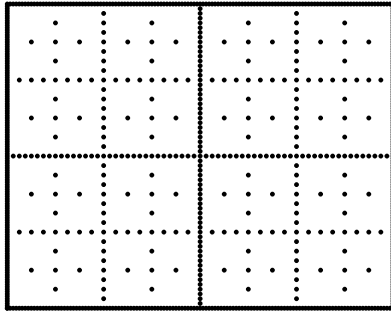
Figure 2: Two-dimensional sparse grid hierarchy up to level $n = 4$.



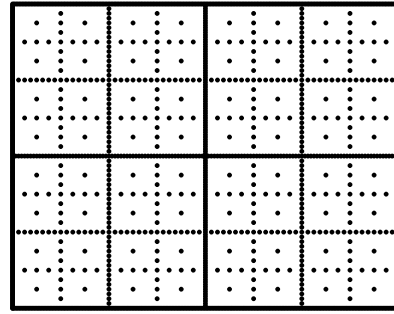
(a) Ω_s^5 , 177 grid points.



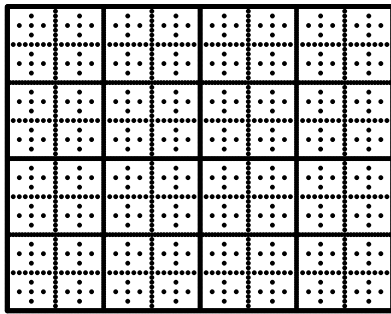
(b) Ω_s^6 , 385 grid points.



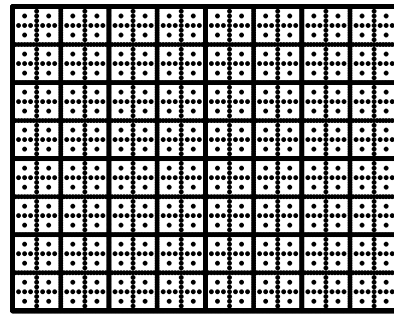
(c) Ω_s^7 , 833 grid points.



(d) Ω_s^8 , 1793 grid points.

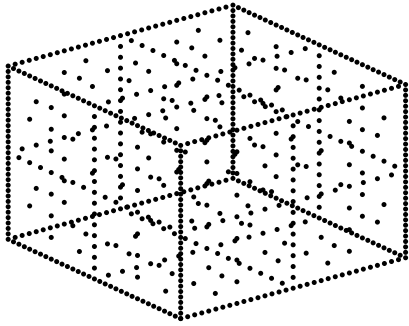


(e) Ω_s^9 , 3841 grid points.

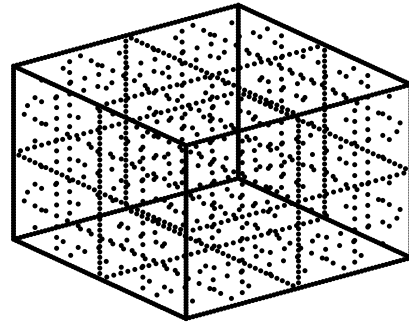


(f) Ω_s^{10} , 8193 grid points.

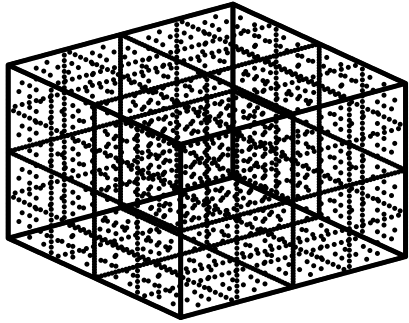
Figure 3: Two dimensional sparse grids for levels $n = 5, \dots, 10$.



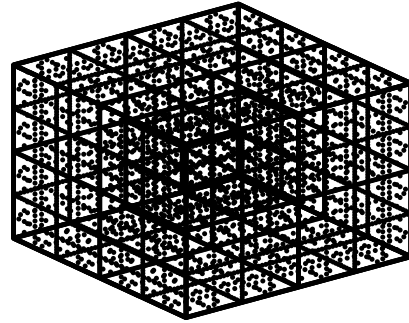
(a) Ω_s^5 , 705 grid points.



(b) Ω_s^6 , 1649 grid points.



(c) Ω_s^7 , 3809 grid points.



(d) Ω_s^8 , 8705 grid points.

Figure 4: Three dimensional sparse grids for levels $n = 5, 6, 7$ and 8.

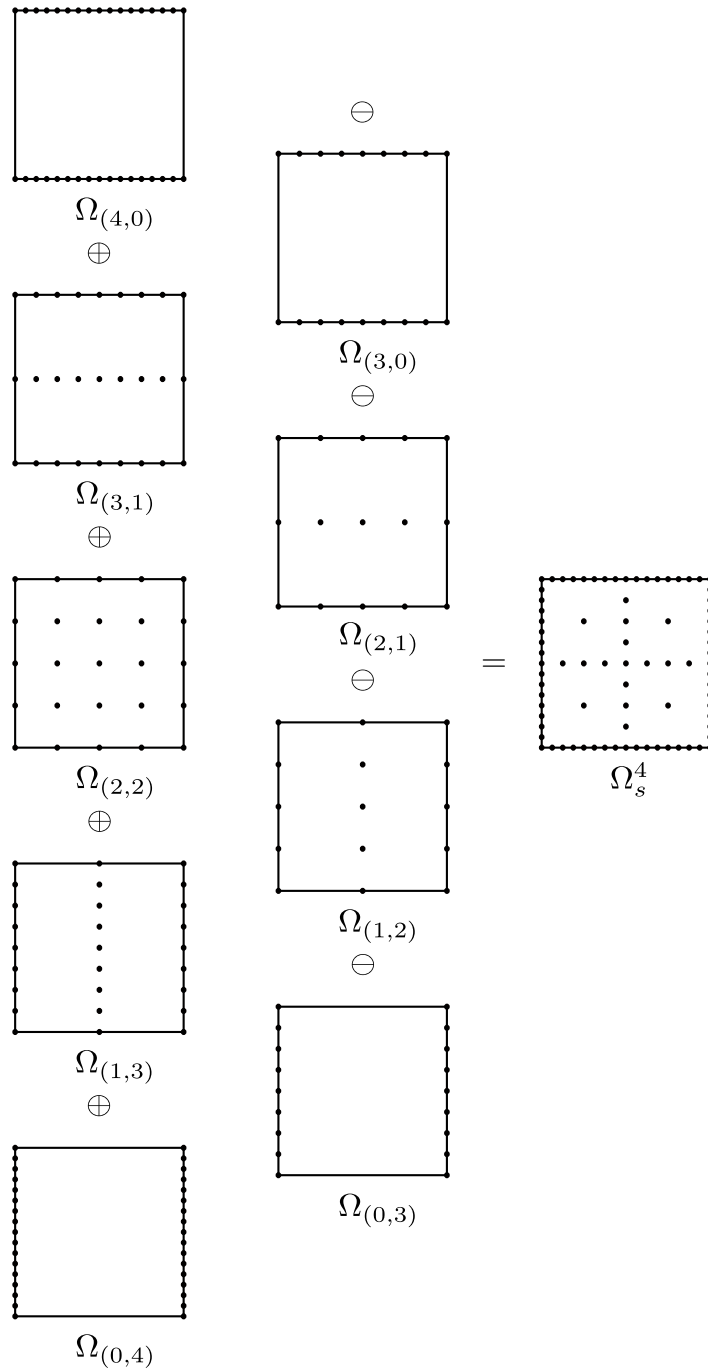


Figure 5: Combination technique with level $n = 4$ in two dimensions.

5.2. Combination technique

Similar to the Richardson extrapolation [38], the so-called combination technique linearly combines the numerical solution on the sequence of anisotropic grids Ω_1 where

$$|\mathbf{l}|_1 = n - q, \quad q = 0, \dots, d - 1.$$

The combination technique reads

$$U_s^n = \sum_{q=0}^{d-1} (-1)^q \cdot \binom{d-1}{q} \cdot \sum_{|\mathbf{l}|_1=n-q} U_1, \quad l_k \geq 0, \quad \forall k = 1, \dots, d, \quad (17)$$

where U_1 denotes the numerical solution on the grid Ω_1 and U_s^n the combined solution on the sparse grid Ω_s^n .

The grids employed by the combination technique of level $n = 4$ in two dimensions are shown in Figure 5.

The idea of this technique is that the leading order errors from the discretization on each grid cancel each other out in the combination solution.

The number of grid points involved in the approximation of U_s^n grows according to $O(n^{d-1} \cdot 2^n)$. In fact, from the formula (16) we have to solve $\binom{n+d-1}{d-1}$ problems with $O(2^n)$ unknowns, $\binom{n+d-2}{d-1}$ problems with $O(2^{n-1})$ unknowns, ... and $\binom{n}{d-1}$ problems with $O(2^{n-(d-1)})$ unknowns. This results in a total number of $O(n^{d-1} \cdot 2^n)$ grid points which is much less than the $O(2^{n \cdot d})$ grid nodes used by traditional full grid methods. Thus, the efficient use of sparse grids greatly reduces the computing time and the storage requirements which allows for the treatment of problems with ten variables and even more [8].

We have seen that the combination technique linearly combines the numerical solution on several traditional full grids. The solution can be calculated on each of these grids by using any existing PDE numerical method like finite differences, finite volume or finite elements. In addition, since all these sub-problems are independent the combination technique can be parallelized [17, 23].

The combination technique approach presumes the existence of a so-called error splitting. It requires for an associated numerical approximation method on the full grid Ω_1 an error splitting of the form

$$u(\mathbf{x}) - U_1(\mathbf{x}) = \sum_{k=1}^d \sum_{\substack{\{j_1, \dots, j_k\} \\ \subseteq \{1, \dots, d\}}} C_{j_1, \dots, j_k}(\mathbf{x}, h_{j_1}, \dots, h_{j_k}) \cdot h_{j_1}^p \cdot \dots \cdot h_{j_k}^p, \quad (18)$$

at each grid point $\mathbf{x} \in \Omega_1$. Here u denotes the exact solution of the partial differential equation under consideration, U_1 the numerical solution on the grid Ω_1 , $p > 0$ is the order of accuracy of the numerical approximation method with respect to each coordinate direction and the coefficient functions C_{j_1, \dots, j_k} of \mathbf{x} and the mesh sizes h_{j_k} , $k = 1, \dots, d$ are required to be bounded by a positive constant K such that

$$|C_{j_1, \dots, j_k}(\mathbf{x}, h_{j_1}, \dots, h_{j_k})| \leq K, \quad \forall k, 1 \leq k \leq d, \quad \forall \{j_1, \dots, j_m\} \subseteq \{1, \dots, d\}.$$

The existence of the expansion (18) is a crucial point to obtain the error bounds of the sparse grid recombination technique and usually requires the assumption of bounded mixed derivatives.

In [20] Griebel and Thurner showed that if the solution of the PDE is sufficiently smooth, the pointwise accuracy of the sparse grid combination technique is $O(n^{d-1} \cdot 2^{-n \cdot p}) = O([\log_2 h_n^{-1}]^{d-1} h_n^p)$, which is only slightly worse than $O(2^{-n \cdot p}) = O(h_n^p)$ obtained by the full grid solution.

The solution at points which do not belong to the sparse grid can be computed through interpolation. The applied interpolation method should provide at least the same order of accuracy of the numerical discretization scheme used to solve the PDE. Otherwise, the accuracy of the numerical scheme will be deteriorated.

Up to now we have assumed the existence of an error splitting of type (18). However, such an error splitting has to be proven for each problem. Nevertheless, proving the existence of this error splitting is usually very complex. Bungartz et al. [9, 10] showed the existence of such an error splitting for the finite differences discretization of the 2-d Laplace equation. Arciniega and Allen [2] proved the existence of this error splitting for the fully implicit as well as the Crank-Nicolson discretization scheme of the European call option. More recently, Reisinger [36] showed that such a splitting also holds for a wider class of linear PDEs, for example convection-diffusion equations. The author gives general conditions which need to be fulfilled to ensure the existence of the desired splitting structure: sufficiently smooth initial data and compatible boundary data, the existence of bounded mixed derivatives, a consistent numerical scheme which provides a truncation error of the desired splitting structure and stability of the discretization scheme. As a summary, we can say that the deduction of the error splitting formula is very complex and was until now only performed for some reference problems. However, we will see in the following Section 5.3 that the numerical results for the sparse

grid combination technique are promising, even for more complex financial products.

5.3. Numerical results

Taking advantage of the previously described sparse grid combination technique, in this section we are pricing the same interest rate derivatives that have been valued in the former Section 4.2 where traditional full grid finite differences methods were considered. In addition to those products, we are going to price interest rate derivatives with up to seven underlying LIBOR rates and their stochastic volatility, showing that the sparse grid combination technique is able to cope with the curse of dimensionality up to a certain extent. As in the previous Section 4.2, we will use Crank-Nicolson scheme, we will consider the Gauss-Seidel iterative solver and the same boundary conditions as in Section 4.1. In the present case, we are interested in the evaluation of the solution at a single point which corresponds with the value of the forward rates at time zero (see Table 2) and $V(0) = 1$. The numerical solution on each grid handled by the combination technique is interpolated at this point using multilinear interpolation and then added up with the appropriate weights.

The sparse grid combination technique has been implemented to run on multicore CPUs. The program was optimized and parallelized using OpenMP [45]. CPU times, measured in seconds, correspond to executions using 24 threads, so as to take advantage of Intel Hyperthreading. The speedups of the parallel version with respect to the pure sequential code are around 16. To the best of our knowledge, graphic processor units (GPUs) are not well-suited to parallelize the combination technique, due to the fact that the different grids employed by the combination technique involve memory accesses patterns totally different, therefore, it is not possible to access the device memory in a coalesced way [33], thus GPU global memory can not serve threads in parallel. In this scenario, the GPU code will be ill performing. In the work [13] the authors take advantage of GPUs to parallelize the solver of each full grid considered by the combination technique. However, they do not parallelize the combination technique itself.

In Table 9 a 1×1 European swaption is priced. The exact price of this derivative is 0.659096 basis points, as discussed in Section 4.2. These results are to be compared with those of Table 3, where it can be seen how the computational times and the grid points employed by the sparse grid combination technique have been substantially reduced.

| Level | Solution | Error | Time | Grid points |
|-------|----------|----------|-------|-------------|
| 3 | 6.864576 | 6.205480 | 0.04 | 37 |
| 4 | 2.207696 | 1.548600 | 0.04 | 81 |
| 5 | 1.107670 | 0.448573 | 0.05 | 177 |
| 6 | 0.788659 | 0.129562 | 0.05 | 385 |
| 7 | 0.668489 | 0.009393 | 0.06 | 833 |
| 8 | 0.662096 | 0.002999 | 0.12 | 1793 |
| 9 | 0.659715 | 0.000618 | 0.54 | 3841 |
| 10 | 0.659287 | 0.000191 | 2.68 | 8193 |
| 11 | 0.659127 | 0.000031 | 16.79 | 9217 |

Table 9: Convergence of the sparse grid finite differences solution in basis points for 1 LIBOR and stochastic volatility, $\sigma = 0$, $V(0) = 1$, $\beta = 1$, 256 time steps. Exact solution, 0.659096 basis points.

Next, in Table 10 a 1×1 European swaption is priced considering stochastic volatility. These results are to be compared with those of Table 6.

| Level | Solution | Time |
|-------|----------|-------|
| 3 | 6.243 | 0.03 |
| 4 | 3.653 | 0.04 |
| 5 | 2.199 | 0.04 |
| 6 | 2.069 | 0.08 |
| 7 | 1.779 | 0.08 |
| 8 | 1.720 | 0.39 |
| 9 | 1.681 | 1.13 |
| 10 | 1.668 | 6.91 |
| 11 | 1.662 | 43.03 |

Table 10: Convergence of the sparse grid finite differences solution in basis points for 1 LIBOR and stochastic volatility, $\sigma = 0.3$, $\phi_i = 0.4$, $V(0) = 1$, $\beta = 1$, 256 time steps. 95% confidence interval with Monte Carlo simulation using 10^7 paths, [1.652, 1.672] in basis points.

In the following Tables 11 and 12, the pricing of 1×2 and 1×3 European swaptions taking into account stochastic volatility is shown, as in Tables 7 and 8, respectively. For the higher resolution levels, the full grid method became very slow, while the sparse grid combination technique results much faster. Note that the combination technique is able to price successfully the 1×3 European swaption, this was not attainable in Table 7.

| Level | Solution | Time |
|-------|----------|--------|
| 5 | 9.560 | 0.12 |
| 6 | 7.416 | 0.12 |
| 7 | 5.896 | 0.17 |
| 8 | 5.318 | 0.46 |
| 9 | 5.007 | 1.05 |
| 10 | 4.826 | 5.34 |
| 11 | 4.833 | 31.42 |
| 12 | 4.820 | 197.90 |

Table 11: Convergence of the sparse grid finite differences solution in basis points for 2 LIBORs and stochastic volatility, $\sigma = 0.3$, $\phi_i = 0.4$, $V(0) = 1$, $\beta = 1$, 256 time steps. 95% confidence interval with Monte Carlo simulation using 10^7 paths, [4.800, 4.844] in basis points.

| Level | Solution | Time |
|-------|----------|---------|
| 7 | 5.872 | 0.34 |
| 8 | 13.279 | 0.88 |
| 9 | 7.466 | 2.29 |
| 10 | 8.642 | 8.37 |
| 11 | 9.809 | 29.74 |
| 12 | 9.686 | 156.75 |
| 13 | 8.694 | 895.75 |
| 14 | 8.671 | 5725.09 |

Table 12: Convergence of the sparse grid finite differences solution in basis points for 3 LIBORs and stochastic volatility, $\sigma = 0.3$, $\phi_i = 0.4$, $V(0) = 1$, $\beta = 1$, 256 time steps. 95% confidence interval with Monte Carlo simulation using 10^7 paths, [8.635, 8.700] in basis points.

Finally, in Tables from 13 to 16, 1×4 , ..., 1×7 European swaptions are priced considering stochastic volatility. The pricing of these interest rate derivatives was not viable with the full grid approach of Section 4.

In order to be able to price derivatives with more than 8 underlyings, the combination technique method should be executed on a cluster of processors. Furthermore, the use of a cluster would allow us to include higher resolution levels which would probably better illustrate some of the cases where the error at the considered point exhibits an oscillatory behaviour (Tables 13 and 15, for example), although all results are consistent with the corresponding Monte Carlo confidence intervals. In the Chapter 13 of the book [14] Philipp Schröder et al. discuss the parallelization of the combination technique using

| Level | Solution | Time |
|-------|----------|----------|
| 9 | 12.51 | 8.29 |
| 10 | 10.35 | 21.60 |
| 11 | 15.75 | 64.52 |
| 12 | 16.30 | 223.30 |
| 13 | 9.95 | 921.53 |
| 14 | 13.03 | 4504.83 |
| 15 | 13.24 | 25980.70 |

Table 13: Convergence of the sparse grid finite differences solution in basis points for 4 LIBORs and stochastic volatility, $\sigma = 0.3$, $\phi_i = 0.4$, $V(0) = 1$, $\beta = 1$, 256 time steps. 95% confidence interval with Monte Carlo simulation using 10^7 paths, [13.20, 13.29] in basis points.

| Level | Solution | Time |
|-------|----------|-----------|
| 11 | 23.71 | 305.46 |
| 12 | 9.45 | 855.16 |
| 13 | 20.38 | 2394.99 |
| 14 | 18.82 | 7584.20 |
| 15 | 18.60 | 29529.46 |
| 16 | 18.56 | 160027.61 |

Table 14: Convergence of the sparse grid finite differences solution in basis points for 5 LIBORs and stochastic volatility, $\sigma = 0.3$, $\phi_i = 0.4$, $V(0) = 1$, $\beta = 1$, 256 time steps. 95% confidence interval with Monte Carlo simulation using 10^7 paths, [18.51, 18.61] in basis points.

| Level | Solution | Time |
|-------|----------|----------|
| 13 | 18.30 | 627.97 |
| 14 | 15.55 | 2115.55 |
| 15 | 25.60 | 9892.70 |
| 16 | 24.49 | 50139.95 |

Table 15: Convergence of the sparse grid finite differences solution in basis points for 6 LIBORs and stochastic volatility, $\sigma = 0.3$, $\phi_i = 0.4$, $V(0) = 1$, $\beta = 1$, 12 time steps. 95% confidence interval with Monte Carlo simulation using 10^7 paths, [24.47, 24.59] in basis points.

MPI (*Message Passing Interface*) API. In [28] the authors parallelize the sparse grid combination technique taking advantage of a MapReduce framework, algorithms that are inherently fault tolerant.

In the present paper the obtention of an estimate for the sparse grid com-

| Level | Solution | Time |
|-------|----------|-----------|
| 15 | 19.56 | 24612.73 |
| 16 | 26.80 | 79463.11 |
| 17 | 30.93 | 324996.63 |

Table 16: Convergence of the sparse grid finite differences solution in basis points for 7 LIBORs and stochastic volatility, $\sigma = 0.3$, $\phi_i = 0.4$, $V(0) = 1$, $\beta = 1$, 2 time steps. 95% confidence interval with Monte Carlo simulation using 10^7 paths, $[30.85, 30.98]$ in basis points.

bination technique error in the solution of the SABR/LIBOR market model PDE remains open. In order to achieve this goal the possible strategies start from proving the existence of an expansion of the error and the bound of the associated coefficients. This is usually complex and has been specifically proven for each problem in which this analysis has been successfully addressed. In practice, we recommend to stop the refinement procedure as soon as the solution obtained by two consecutive refinements are close enough to each other.

6. Conclusion

In this work we have posed for first time in the literature the PDEs associated to the SABR/LIBOR market models proposed by Mercurio & Morini [31] and Hagan [22]. In order to price interest rate derivatives we have developed a traditional full grid finite difference method. This approach is able to successfully price derivatives up to two or three underlying forward rates in reasonable computational times. However, when the number of underlyings increases this scheme suffers from the well-known curse of dimensionality. In order to price derivatives over a moderately large number of forward rates we have proposed to use the sparse grid combination technique. Taking into account that this technique is embarrassingly parallel we have parallelized it so as to drastically reduce computational times. Finally, we have tested the proposed method in two different ways. On one hand, using the analytical solution when it exists. On the other hand, when the exact solution is not known, we have used as reference solution the one computed with the Monte Carlo method, thus ensuring the correctness of the developed scheme.

Acknowledgements: The authors would like to thank the anonymous referees for their remarks which helped to improve the paper.

References

- [1] S. Achatz. Higher order sparse grid methods for elliptic partial differential equations with variable coefficients. *Computing*, 71(1):1–15, 2003.
- [2] A. Arciniega and E. Allen. Extrapolation of difference methods in option valuation. *Applied Mathematics and Computation*, 135:165–186, 2004.
- [3] R. Bellmann. Adaptive control processes: A guided tour. *Princeton University Press*, 1961.
- [4] G. Beylkin and M. J. Mohlenkamp. Algorithms for numerical analysis in high dimensions. *SIAM Journal on Scientific Computing*, 26(6):2133–2159, 2005.
- [5] J. Blackham. Sparse grid solutions to the libor market model. Master’s thesis, Magdalen College, University of Oxford, 2004.
- [6] A. Brace, D. Gatarek, and M. Musiela. The Market model of interest rate dynamics. *Mathematical Finance*, 7(2):127–155, 1997.
- [7] D. Brigo and F. Mercurio. *Interest Rate Models - Theory and Practice. With Smile, Inflation and Credit*. Springer, second edition, 2007.
- [8] H. J. Bungartz and M. Griebel. Sparse grids. *Acta Numerica*, 13:147–269, 2004.
- [9] H. J. Bungartz, M. Griebel, D. Roschke, and C. Zenger. Pointwise convergence of the combination technique for the Laplace equation. *East-West Journal of Numerical Mathematics*, 2:21–45, 1994.
- [10] H. J. Bungartz, M. Griebel, D. Roschke, and C. Zenger. A proof of convergence for the combination technique for the Laplace equation using tools of symbolic computation. *Mathematics and Computers in Simulation*, 41:595–605, 1996.
- [11] D. J. Duffy. *Finite Difference methods in financial engineering. A Partial Differential Equation Approach*. Wiley Finance Series, 2006.
- [12] A. M. Ferreiro, J. A. García, J. G. López-Salas, and C. Vázquez. SABR/LIBOR market models: Pricing and calibration for some interest rate derivatives. *Applied Mathematics and Computation*, 242:65–89, 2014.

- [13] A. Gaikwad and I. M. Toke. GPU Based Sparse Grid Technique for Solving Multidimensional Options Pricing PDEs. In *Proceedings of the 2nd Workshop on High Performance Computational Finance*, WHPCF '09, pages 6:1–6:9, New York, NY, USA, 2009. ACM.
- [14] T. Gerstner and P. Kloeden, editors. *Recent Developments in Computational Finance. Foundations, Algorithms and Applications*. World Scientific Publishers, Interdisciplinary Mathematical Science, 2013.
- [15] P. Glasserman. *Monte Carlo Methods in Financial Engineering*. Springer-Verlag, New York, 2003.
- [16] E. Gobet. Advanced Monte Carlo methods for barrier and related exotic options. *Mathematical Modeling and Numerical Methods in Finance*, pages 497–528, 2009.
- [17] M. Griebel. The combination technique for sparse grids solution of PDEs on multiprocessor machines. *Parallel Processing Letters*, 3:66–71, 1992.
- [18] M. Griebel. Adaptive sparse grid multilevel methods for elliptic PDEs based on finite differences. *Computing*, 61:151–180, 1998.
- [19] M. Griebel, M. Schneider, and C. Zenger. A combination technique for the solution of sparse grid problems. In P. de Groen and R. Beauwens, editors, *Proceedings of the IMACS International Symposium on Iterative Methods in Linear Algebra*, pages 263–281, Brussels, April 1991. Elsevier, Amsterdam, 1992.
- [20] M. Griebel and V. Thurner. The efficient solution of fluid dynamics problems by the combination technique. *International Journal of Numerical Methods for Heat and Fluid Flow*, 5(3):251–269, 1995.
- [21] P. Hagan, D. Kumar, A. Lesniewski, and D. Woodward. Managing Smile Risk. *Wilmott Magazine*, pages 84–108, 2002.
- [22] P. Hagan and A. Lesniewski. LIBOR market model with SABR style stochastic volatility. *Working paper, available at <http://lesniewski.us/papers/working/SABRLMM.pdf>*, 2008.

- [23] M. Heene, Christoph Kowitz, and Dirk Pfluger. Higher order ADI schemes for parabolic equations in the combinations technique and application to finance. *Journal of Computational and Applied Mathematics*, 316:175–194, 2017.
- [24] C. Hendricks, C. Heuer, M. Ehrhardt, and M. Gunter. Higher order ADI schemes for parabolic equations in the combinations technique and application to finance. *Journal of Computational and Applied Mathematics*, 316:175–194, 2017.
- [25] F. Jamshidian. LIBOR and swap market models and measures. *Finance and Stochastic*, 1:293–330, 1997.
- [26] F. Koster. A proof of consistency of the finite differences technique on sparse grids. *Computing*, 65:247–261, 2000.
- [27] J. Kraus. Option pricing using the sparse grid combination technique. Master’s thesis, University of Waterloo, 2007.
- [28] J. W. Larson, M. Hegland, B. Harding, S. Roberts, L. Stals, A. P. Rendell, P. Strazdins, M. M. Ali, C. Kowitz, R. Nobes, J. Southern, N. Wilson, M. Li, and Y. Oishi. 2013 International Conference on Computational Science: Fault-Tolerant Grid-Based Solvers: Combining Concepts from Sparse Grids and MapReduce. *Procedia Computer Science*, 18:130–139, 2013.
- [29] F. A. Longstaff and E. S. Schwartz. Valuing american options by simulation: A simple least-squares approach. *The Review of Financial Studies*, 14(1):113–147, 2001.
- [30] D. L. McLeish. *Monte Carlo Simulation and Finance*. Wiley, 2005.
- [31] F. Mercurio and M. Morini. No-Arbitrage dynamics for a tractable SABR term structure Libor Model. *Modeling Interest Rates: Advances in Derivatives Pricing, Risk Books (2009)*, 2009.
- [32] K. Miltersen, K. Sandmann, and D. Sondermann. Closed-form solutions for term structure derivatives with lognormal interest rates. *Journal of Finance*, 52(1):409–430, 1997.
- [33] Nvidia Corporation. *CUDA C Programming guide*.

- [34] R. Rebonato. *Modern pricing of interest-rate derivatives: the Libor market model and beyond*. Princeton University Press, 2002.
- [35] R. Rebonato and R. White. Linking caplets and swaptions prices in the LMM-SABR model. *The Journal of Computational Finance*, 13(2):19–45, 2009.
- [36] C. Reisinger. Analysis of linear difference schemes in the sparse grid combination technique. *IMA Journal of Numerical Analysis*, 33(2):544–581, 2013.
- [37] C. Reisinger and G. Wittum. Efficient hierarchical approximations of high-dimensional option pricing problems. *SIAM Journal Scientific Computing*, 29:440–458, 2007.
- [38] L. F. Richardson. The approximate arithmetical solution by finite differences of physical problems including differential equations, with an application to the stresses in a masonry dam. *Philosophical Transactions of the Royal Society A*, 210(459-470):307–357, 1911.
- [39] S. E. Shreve. *Stochastic Calculus for Finance*. Springer, 2004.
- [40] S. Smolyak. Quadrature and interpolation formulas for tensor products of certain classes of functions. *Dokl. Akad. Nauk SSR*, 148:1042–1045, 1963.
- [41] H. Yserentant. On the multi-level splitting of finite element spaces. *Numerische Mathematik*, 49:379–412, 1986.
- [42] H. Yserentant. Hierarchical bases. In *Proceedings of the Second International Conference on Industrial and Applied Mathematics*, ICIAM 91, pages 256–276, Philadelphia, PA, USA, 1992. Society for Industrial and Applied Mathematics.
- [43] A. Zeiser. Fast Matrix-Vector Multiplication in the Sparse-Grid Galerkin Method. *Journal of Scientific Computing*, 47(3):328–346, 2010.
- [44] C. Zenger. Sparse grids. In W. Hackbusch, editor, *Parallel Algorithms for Partial Differential Equations, Proceedings of the Sixth GAMM-Seminar*, volume 31, pages 241–251, Kiel, Germany, 1990. Vieweg-Verlag, 1991.

[45] OpenMP web page: <http://www.openmp.org>.