Computers and Mathematics with Applications 169 (2024) 88–98

Contents lists available at ScienceDirect

Computers and Mathematics with Applications

journal homepage: www.elsevier.com/locate/camwa

PDEs for pricing interest rate derivatives under the new generalized Forward Market Model (FMM)

José Germán López-Salas [a], Soledad Pérez-Rodríguez [b], Carlos Vázquez [a,*]

[a] Department of Mathematics, Faculty of Informatics and CITIC, Campus Elviña s/n, 15071-A Coruña, Spain
[b] Departamento de Análisis Matemático and IMAULL, Universidad de La Laguna, 38208, La Laguna, Tenerife, Canary Islands, Spain

ARTICLE INFO

Keywords:
IBOR replacement
Generalized forward market model
Forward rates
Finite differences
AMFR-W methods

ABSTRACT

In this article we derive partial differential equations (PDEs) for pricing interest rate derivatives under the generalized Forward Market Model (FMM) recently presented by A. Lyashenko and F. Mercurio in [1] to model the dynamics of the Risk Free Rates (RFRs) that are replacing the traditional IBOR rates in the financial industry. Moreover, for the numerical solution of the proposed PDEs formulation, we develop some adaptations of the finite differences methods developed in [2] that are very suitable to treat the presence of spatial mixed derivatives. This work is the first article in the literature where PDE methods are used to value RFR derivatives. Additionally, Monte Carlo-based methods will be designed and the results are compared with those obtained by the numerical solution of PDEs.

1. Introduction

For decades, financial institutions have been using InterBank Offered Rates (IBORs) as reference rates to determine interest or as underlyings of derivatives in several currencies, perhaps the more popular example being the London InterBank Offered Rate (LIBOR). More than 350 trillion dollars in derivatives and other financial products are tied to these rates.

IBORs reflect the average unsecured short-term interest rate at which large global banks can borrow from each other. They are based on surveys to banks and not on real transactions.

At the beginning of the 21st century, several big banks manipulated the interest rate they reported that they could borrow at. Firstly, to allow their traders, who were taking derivative bets on where this IBOR would be, to make more money because the rate was artificially distorted. Later, in the depth of the 2008 financial crisis, banks again manipulated IBORs, this time not to make profits from derivatives trading, but to make themselves look financially stronger than they were [3–6].

In view of previous IBORs scandals, a few years ago financial authorities worldwide initiated the replacement of IBORs with alternative Risk Free Rates (RFRs). RFRs are reported to be robust because they rely on real transactions. The whole banking industry is adapting its products by offering them based on RFR for new trades (see [7,8], for example). This transition, known as IBOR transition, is complex for clients, dealers, and financial authorities.

In fact, on December 31st, 2021, some IBORs ceased to be published and became non-representative. This concerned all tenors of IBOR Japanese Yen, British Pounds, Swiss Francs, Euros, and two tenors of IBOR USD dollars. Recently, on June 30th, 2023, the remaining tenors of IBOR USD dollars also stopped being reported and became irrelevant. For the time being, all major economies have selected RFRs to replace their corresponding IBORs. For example, the United States of America adopted SOFR (Secured Overnight Financing Rate), the European Union selected ESTER (Euro Short-TErm Rate), the United Kingdom designed SONIA (Sterling OverNight Index Average), Switzerland took SARON (Swiss Average Rate OverNight), and Japan selected TONAR or TONA (Tokyo OverNight Average Rate).

The main general features of RFRs are the following. Firstly, RFRs are overnight rates and not term rates like IBORs (i.e. one week, one month, three months, ...). Secondly, by definition, RFRs are *backward-looking*, which means that the rate to be paid for the application period is calculated by reference to historical transaction data and set at the end of that time interval. Unlike RFRs, IBORs are *forward-looking* rates, meaning that the rate to be paid for the application period is set at the beginning of that time interval. Additionally, RFRs are risk-free since one-day credit risk can be neglected. On top of that, RFRs not only represent the interbank market; in fact they are rates for the entire market.

* Corresponding author.
E-mail addresses: jose.lsalas@udc.es (J.G. López-Salas), sperezr@ull.edu.es (S. Pérez-Rodríguez), carlosv@udc.es (C. Vázquez).

https://doi.org/10.1016/j.camwa.2024.06.010
Received 27 July 2023; Received in revised form 31 March 2024; Accepted 20 June 2024

The LIBOR Market Model (LMM) was widely used by financial institutions for the valuation of interest rate derivatives based on IBORs (see the seminal works [9,10] and the book [11], for example). The main reason behind its great success comes from the fact that LMM considers rate dynamics consistent with the well-established Black-Scholes market formulas for pricing caplets and swaptions. Since then, a lot of work has been devoted to the pricing of interest rate derivatives by using different methods, such as Monte Carlo or PDEs. Particularly, concerning PDEs formulations, the pricing of some complex derivatives with classical LMM has been addressed in [12–15], among others. Extensions of the LMM framework to incorporate stochastic volatility have been discussed in [16–19] and references therein, with a first formulation in terms of PDEs developed in [20]. Thus, in the setting of LMM, the formulation in terms of PDEs for pricing different interest rate derivatives has been widely studied. It is important to notice that the LMM contemplates only forward-looking rates. Therefore, it is no longer valid to price financial products based on the new RFRs, that are backward-looking.

Nowadays, the community working on quantitative finance is very active in proposing new mathematical models able to price the new derivatives based on RFRs. Having in mind that RFRs can be converted into compounded setting-in-arrears term rates, mathematical models for pricing RFR-based derivatives can mainly follow two different approaches. The first strategy is to directly simulate daily the underlying RFRs in their corresponding application periods. The second approach models term rates based on RFRs. The most promising interest rate model following the second strategy is the one proposed by Lyashenko and Mercurio in [1]. This model, referred to as the generalized Forward Market Model (FMM), is the main starting point in this work.

The FMM is a modeling framework that allows for the joint modeling of forward-looking and backward-looking rates inside the same parsimonious setup. Actually, FMM is an extension of the successful LMM. More precisely, in the post-LIBOR world, IBORs have to be replaced with some more general forward rates, and that is exactly what explains the term *generalized forward* market model that has been coined for this new setting.

FMM accommodates both forward-looking and backward-looking rates inside the same framework in a very natural way. Indeed, the FMM incorporates additional advantages over the LMM. One of them comes from the possibility of modeling rates directly under the classic risk-neutral money-market measure, something that was not possible with the LMM. In [21], Lyashenko and Mercurio explain how to complete the FMM in order to generate rates that are outside the given time grids that are initially assigned to the model. The authors build a general Heath–Jarrow–Morton (HJM) model that originates the FMM. Once this HJM model is constructed, since it is very general, one can create general rates and discount factors. This approach was not possible in the classical LMM. In fact, under LMM the approach is usually to complete the model by adding some interpolation method, which is typically called rate interpolation.

In order to price RFR-based derivatives under the FMM when the payoff depends on the joint distribution of several interest rates, numerical methods must be considered. For this purpose, expectation-based formulations or PDE-based formulations of the pricing problem can be mainly used. Although the standard approach uses Monte Carlo simulations for expectation-based formulations, it exhibits several drawbacks when pricing interest rate derivatives, as it has been pointed out in [20], for example. Although some disadvantages could be smoothed or even removed by enhanced Monte Carlo techniques applied to specific interest rate derivatives, the pricing of early exercise derivatives based on RFRs, like Bermudan swaptions, would require a highly computational demanding suitable adaption of Monte Carlo methodology. In the PDEs formulation setting, pricing these interest rate derivatives does not involve a huge increase of computational time with respect to analogous derivatives without early exercise opportunity.

These previous arguments motivate the interest in developing suitable PDE formulations for solving the pricing problem in the new recently established FMM model. In this work, we formulate the pricing of RFRs derivatives under the FMM in terms of PDEs. To the best of our knowledge, our presentation is the first in the literature.

Moreover, we present an efficient and recent numerical algorithm to approximate the solution of the proposed PDEs formulation by using finite differences and the AMFR-W1 method for the time integration, as already used in [2]. This method belongs to the class of AMFR-W-methods [22,23], which are very efficient when dealing with parabolic problems involving mixed derivatives, as they avoid computing explicitly the part of the Jacobian that includes the discretization of such mixed derivatives. In [2], a numerical strategy that combines appropriate finite differences schemes to deal with terms containing mixed derivatives with sparse grids technique has been successfully used for pricing interest rate derivatives when classical LMM for forward rates including stochastic volatility is considered.

However, in the present work, our aim will be different. More precisely, as the payoff function of the derivative that determines the dynamics of the PDE has differentiability issues near the strike values, we have explored the integration on non-uniform meshes, which contain many more points near the payoff non-differentiability area than in the rest of the domain. As we will see, the consideration of appropriate non-uniform meshes improves the accuracy and reliability of the approximation and we will obtain an approximation of higher quality than that provided by the Monte Carlo method, at least when the "spatial"[1] dimension of the PDE is less than or equal to 4. The application of sparse grids to solve the so-called curse of dimensionality in these PDEs will be left for future work.

The article is organized as follows. In Section 2 we review the extended zero-coupon bonds, which are the cornerstone of the recent FMM. A detailed description of how this concept of extended bonds allows to define not only the classical forward-looking forward rates but also the novel backward-looking forward rates, is also carried out. Besides, a thorough illustration of how to compute the extended discount factors from forward rates values is presented. In Section 3 the system of stochastic differential equations (SDEs) of the FMM is introduced, considering joint dynamics for the interest rates. In Section 4, we derive the PDEs for the FMM. Next, numerical methods to price derivatives under the FMM are designed in Section 5. In Section 6, numerical experiments are carried out to assess the behavior of the developed numerical methods. Finally, conclusions and future work are discussed in Section 7.

2. Main assumptions, definitions and notations

In this section, we present the basic notations and definitions that will be used throughout the article. A continuous-time financial market is considered. It has an instantaneous RFR whose value at time t is denoted by $r(t)$.

Definition 2.1 *(Bank account).* Let $B(t)$ be the value of the bank account at time $t \geq 0$. B is the classic process that satisfies the ordinary differential equation $\mathrm{d}B(t) = r(t)B(t)\,\mathrm{d}t$ with $B(0) = 1$, so that $B(t) = e^{\int_0^t r(u)\mathrm{d}u}$.

Moreover, we assume the existence of a risk-neutral measure Q, whose associated numeraire is the bank account B we have just defined. Besides, $\mathbb{E}$ will denote the expectation with respect to the risk-neutral

[1] When working with time-dependent PDEs such as the ones we consider in this work, it is common to use physics-like terminology where the word "spatial" refers to variables other than time. In our case, the "spatial" variables are the forward rates, so that the number of these rates is the "spatial dimension" of the PDE. In what follows, we will use this terminology.

measure, and $\mathcal{F}_t$ will be the σ-algebra generated by risk factors up to the evaluation time.

Definition 2.2 *(Zero-coupon bond price)*. A zero-coupon bond with maturity T is a very simple contract that pays its holder one unit of currency at time T, with no intermediate payments. For $t < T$, let $P(t,T)$ be the value at time t of this product. We have the following valuation formula, which is given by risk-neutral pricing:

$$P(t,T) = \mathbb{E}\left[e^{-\int_t^T r(u)\mathrm{d}u} \middle| \mathcal{F}_t \right]. \tag{1}$$

Note that $P(T,T) = 1$ for all T.

The formula (1) is clearly defined for valuation times t before the maturity T ($t \leq T$) since one wants to calculate the bond price before it expires. In the new FMM framework, the advantage is that it is mathematically possible to define $P(t,T)$ even for those times t after the maturity T ($t > T$).

Definition 2.3 *(Extended zero-coupon bond price)*. For $t > T$, Equation (1) reduces to

$$P(t,T) = \mathbb{E}\left[e^{\int_T^t r(u)\mathrm{d}u} \middle| \mathcal{F}_t \right] = e^{\int_T^t r(u)\mathrm{d}u} = \frac{B(t)}{B(T)}. \tag{2}$$

Note that $P(t,0) = B(t)$.

One must observe that the integral in equation (1) can be also defined for $t > T$, and it is equal to $e^{\int_T^t r(u)\mathrm{d}u}$, which is measurable with respect to $\mathcal{F}_t$. Then, using properties of conditional expectations, the conditioned expected value is equal to the term inside (see [24], for example). Finally, one just uses the definition of $B(t)$ to obtain that the price of the bond after its maturity is equal to the bank account at the valuation time t divided by the bank account at the maturity T.

This concept of extended zero coupon bonds is taken from [1]. Although it was already considered in [25] and [26] to define hybrid numeraires and measures, in [1] it is used for the first time to extend the definition from forward-looking to backward-looking rates, so that also forward rates can be appropriately extended after their maturity dates.

Definition 2.4 *(Extended T-forward measure)*. The extended T-forward measure, denoted by Q^T, is the martingale measure associated with the extended bond price $P(t,T)$. Note that the risk-neutral measure is a particular case of the extended T-forward measure where $T = 0$, i.e. $Q = Q^0$.

2.1. The compounded setting-in-arrears term rate

Financial derivatives written on RFRs consider as underlyings daily compounded setting-in-arrears term rates, which by definition are backward-looking in nature. Hereafter, $N \geq 1$ denotes the number of rates to be modeled. Let us define them in the next paragraphs.

We start with the tenor structure $0 = T_0 < T_1 < \ldots < T_N$. Let τ_k be the year fraction of the k-th time interval $[T_{k-1}, T_k)$. Next, we define the backward-looking spot rate.

Definition 2.5 *(Backward looking spot rate)*. The simple backward-looking spot rate is defined as

$$R(T_{k-1},T_k) = \frac{1}{\tau_k}\left[e^{\int_{T_{k-1}}^{T_k} r(u)\mathrm{d}u} - 1 \right] = \frac{1}{\tau_k}\left[\frac{B(T_k)}{B(T_{k-1})} - 1 \right]$$
$$= \frac{1}{\tau_k}\left[P(T_k, T_{k-1}) - 1 \right].$$

$R(T_{k-1},T_k)$ is the simple interest rate such that the investment of one unit of currency at time T_{k-1} yields $P(T_k,T_{k-1})$ units of currency at time T_k.

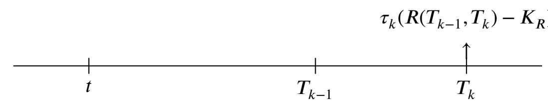

Fig. 1. Swaplet based on the backward-looking rate.

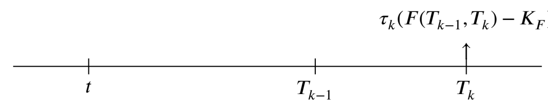

Fig. 2. Swaplet based on the forward-looking rate.

Additionally, we also need to consider forward-looking rates, which are the same as in the LMM (see [11]).

Definition 2.6 *(Forward-looking spot rate)*. The simple forward-looking spot rate is defined as

$$F(T_{k-1},T_k) = \frac{1}{\tau_k}\left[\frac{1}{P(T_{k-1},T_k)} - 1 \right].$$

$F(T_{k-1},T_k)$ is the simple interest rate such that the investment of $P(T_{k-1},T_k)$ units of currency at time T_{k-1} yields one unit of currency at time T_k.

2.2. Forward rates

Once we have defined spot rates, we move to the definition of forward rates.

Definition 2.7 *(Backward-looking forward rate)*. The simple compounded backward-looking forward rate prevailing at time t for the time interval $[T_{k-1}, T_k)$ is denoted by $R_k(t)$ and defined by

$$R_k(t) = \frac{1}{\tau_k}\left(\frac{P(t,T_{k-1})}{P(t,T_k)} - 1 \right). \tag{3}$$

It is the value of the fixed rate K_R in the swaplet paying $\tau_k(R(T_{k-1},T_k) - K_R)$ at time T_k, such that this product has zero value at time t (see Fig. 1). Note that the definition (3) is valid for all times t, even those times $t > T_k$. The rate $R_k(t)$ satisfies the following properties:

- $R_k(T_{k-1}) = F(T_{k-1},T_k)$, i.e., at time T_{k-1} it is equal to the forward-looking spot rate.
- $R_k(T_k) = R(T_{k-1},T_k)$, i.e., at time T_k it is equal to the backward-looking spot rate.
- For $t > T_k$, $R_k(t) = R(T_{k-1},T_k)$, i.e., after time T_k it stops evolving.

Definition 2.8 *(Forward-looking forward rate)*. The simple compounded forward-looking forward rate prevailing at time t for the time interval $[T_{k-1}, T_k)$ is denoted by $F_k(t)$ and defined by

$$F_k(t) = \begin{cases} R_k(t) & \text{if} \quad t \leq T_{k-1} \\ F(T_{k-1},T_k) & \text{if} \quad t > T_{k-1}. \end{cases} \tag{4}$$

It is the value of the fixed rate K_F in the swaplet paying $\tau_k(R(T_{k-1},T_k) - K_F)$ at time T_k such that this product has zero value at time t, see Fig. 2.

So we have defined two types of forwards: the forward of the backward-looking rate and the forward of the forward-looking rate. Nevertheless, for each $k = 1, \ldots, N$, the backward-looking forward rate R_k and the forward-looking forward rate F_k can be modeled by a single rate, the forward of the backward-looking rate R_k. In fact, before the beginning of the application interval $[T_{k-1}, T_k)$, the two forwards are the same and given by the process R_k. At time T_{k-1}, R_k sets at the forward-looking spot rate. Note that in the old LMM, R_k will end at time T_{k-1}. Instead, in the FMM the rate still exists and continues to evolve. In fact, it evolves until the time T_k, where it fixes to the backward-looking spot rate. After time T_k, it continues to exist, and it is a constant.

2.3. Computation of extended discount factors from forward rates values

In this subsection we summarize how to compute $P(T_i, T_j)$:

$$P(T_i,T_j)=\begin{cases}\prod_{k=i+1}^{j}\dfrac{1}{1+\tau_k R_k(T_i)} & \text{if } T_i<T_j,\\ 1 & \text{if } T_i=T_j,\\ \prod_{k=j+1}^{i}\left(1+\tau_k R_k(T_k)\right) & \text{if } T_i>T_j.\end{cases} \tag{5}$$

We can abridge these three cases in the following formula

$$P(T_i,T_j)=P(T_i,T_0)\prod_{k=1}^{j}\frac{1}{1+\tau_k R_k(T_i)},$$

with the equality holding for $T_i=T_j$, $T_i<T_j$ or $T_i>T_j$, and being valid for $j>0$. The case $j=0$ is just the bank account, i.e.,

$$B(T_i)=P(T_i,T_0)=\prod_{k=1}^{i}\left(1+\tau_k R_k(T_k)\right).$$

3. The generalized FMM

In order to create a proper market model, it is not enough to model just the evolution of a single rate. Indeed, we need to model the evolution of the forward rates jointly. This is the objective of this section: we model the evolution of the forward rates under a common probability measure.

Also in the FMM, we can specify the forward rate dynamics under the classic spot-LIBOR measure Q^d and the general T_k-forward measure Q^{T_k}. In fact, the FMM dynamics under Q^d and Q^{T_k} are the same as those of the corresponding LMM (see [11], for example).

Thanks to the Definition 2.3 of extended bond prices, the new FMM allows also for forward-rates dynamics under the risk-neutral measure Q. This is a great advantage, which was not possible in the old LMM world. In fact, as previously stated, the bank account is a zero maturity zero coupon bond, i.e. a bond that instantaneously matures and directly transforms into the bank account, $P(t,0)=B(t)$. So, the risk-neutral measure Q, associated with the bank account as numeraire, can be seen as a forward measure, i.e. the forward measure associated with the zero maturity bond. Therefore, one can derive Q dynamics for the forward rates.

From now on in this article, we will restrict ourselves to the risk-neutral measure Q. Each forward rate R_k, $k=1,\dots,N$, is modeled as a continuous time stochastic process $R_k(t)$. The dynamics of the forward processes are driven by a N-dimensional correlated Wiener processes $W_1^Q(t),\dots,W_N^Q(t)$ under measure Q. In order to streamline the notation along the article, we let $W_k(t)=W_k^Q(t)$, for $k=1,\dots,N$. Let ρ_{ij} denote the correlation coefficient between $W_i(t)$ and $W_j(t)$, i.e. $\mathbb{E}[\mathrm{d}W_i(t)\mathrm{d}W_j(t)]=\rho_{ij}\mathrm{d}t$, where $\mathrm{d}W_k(t)$ denotes the increment of the Wiener process $W_k(t)$ under the measure Q. The system of SDEs of the FMM takes the form

$$\mathrm{d}R_k(t)=\mu_k(t)\mathrm{d}t+v_k(t)\mathrm{d}W_k(t),\quad k=1,\dots,N, \tag{6}$$

where $\mu_k(t)$ and $v_k(t)$ are the drift and diffusion terms of the forward rate $R_k(t)$, respectively. The drift terms are determined by requiring a lack of arbitrage. The diffusion terms have to capture the fact that the process $R_k(t)$ will not be killed at $t=T_{k-1}$ as it happened in the classic LMM. In the FMM we need to define dynamics of the forward rates $R_k(t)$ inside their application periods $[T_{k-1},T_k)$. Besides, the volatility of $R_k(t)$ inside $[T_{k-1},T_k)$ goes down progressively to zero: it becomes smaller and smaller until reaching the value zero at T_k. In order to model this behavior, the system (6) is modified in the following way

$$\mathrm{d}R_k(t)=\mu_k(t)\mathrm{d}t+v_k(t)\gamma_k(t)\mathrm{d}W_k(t),\quad k=1,\dots,N, \tag{7}$$

where the new parameter $\gamma_k(t)$ incorporates the volatility decay in the model. Therefore, the volatility is decomposed in two components, one is the classic LMM volatility $v_k(t)$, while $\gamma_k(t)$ is a deterministic function to control the volatility decay. This function γ_k is equal to one up to the beginning of the interval $[T_{k-1},T_k)$ and is equal to zero at T_k to model the fact that the rate has expired. Moreover, γ_k should be differentiable and decrease down to zero (going from one to zero) inside the interval $[T_{k-1},T_k)$. For example, in the Ho-Lee model, the volatility decay is linear inside the application period, thus γ_k is defined as:

$$\gamma_k(t)=\begin{cases}1 & \text{if } t\le T_{k-1},\\ \dfrac{T_k-t}{T_k-T_{k-1}} & \text{if } t\in(T_{k-1},T_k),\\ 0 & \text{if } t\ge T_k.\end{cases} \tag{8}$$

Let $\sigma_k(\cdot)$ be a deterministic function of time t. Some standard volatility specifications are the following. In the so-called normal model, $v_k(t)=\sigma_k(t)$. For the lognormal model, $v_k(t)=\sigma_k(t)R_k(t)$, while for the shifted-lognormal model, which allows for negative rates, $v_k(t)=\sigma_k(t)(R_k(t)+\vartheta_k)$, where ϑ_k is a positive constant. For the CEV model $v_k(t)=\sigma_k(t)R_k(t)^{\beta_k}$, where $0\le\beta_k\le 1$. In the following, except where otherwise indicated, we will assume a general model specification, i.e. we let v_k a general adapted process.

Under the measure Q^k associated with the numeraire $P(t,T_k)$, R_k is a martingale, i.e. R_k is the driftless process $\mathrm{d}R_k(t)=v_k(t)\gamma_k(t)\mathrm{d}W_k^{Q^k}(t)$, where $W_k^{Q^k}$ denotes a Wiener process under measure Q^k. In order to model all forward rates R_k, $k=1,\dots,N$ jointly, the computation of the dynamics of each forward rate under a common probability measure is needed. In this work, we will consider as numeraire the bank account $B(t)=P(t,0)$. As previously stated, the probability measure associated with this numeraire is the risk-neutral measure Q.

Under the probability measure Q the price of the bonds $P(t,T_k)$ divided by the numeraire $B(t)=P(t,T_0)$ must be martingales. By using this condition, the drifts $\mu_k(t)$ for the forward rates can be computed starting from R_1 until R_N, thus obtaining (see [1] for details)

$$\mu_k(t)=v_k(t)\gamma_k(t)\sum_{i=1}^{k}\rho_{ik}\frac{\tau_i v_i(t)\gamma_i(t)}{1+\tau_i R_i(t)}. \tag{9}$$

Since $\gamma_k(t)=0$ for $t\ge T_k$, μ_k can be better expressed in terms of the index function

$$\eta(t)=\min\{j,T_j\ge t\}, \tag{10}$$

which provides the index of the element in the tenor structure being not smaller than t that is the nearest to time t. Therefore, we have

$$\mu_k(t)=v_k(t)\gamma_k(t)\sum_{i=\eta(t)}^{k}\rho_{ik}\frac{\tau_i v_i(t)\gamma_i(t)}{1+\tau_i R_i(t)}. \tag{11}$$

All in all, the dynamics of R_k under the measure Q satisfy the following system of SDEs:

$$\mathrm{d}R_k(t)=v_k(t)\gamma_k(t)\sum_{i=\eta(t)}^{k}\rho_{ik}\frac{\tau_i v_i(t)\gamma_i(t)}{1+\tau_i R_i(t)}\mathrm{d}t+v_k(t)\gamma_k(t)\mathrm{d}W_k(t),\ k=1,\dots,N. \tag{12}$$

4. PDE for the generalized FMM under the risk neutral measure

After the introduction of the generalized FMM in the previous section, where the dynamics of the forward rates R_k satisfy the system (12) under the risk-neutral measure $\mathcal{Q}$, in this section we derive the corresponding PDE formulation for the pricing of interest rate derivatives without early exercise opportunity (also referred to as European interest rate derivatives). As in the case of more classical models for interest rate derivatives, such as LMM, or any other derivatives, the statement of the PDEs formulation is based on the appropriate Feynman-Kàc theorem (see [11], for example). This theorem relates the formulation in terms of expectations with the one in terms of PDEs, so that the solution of the PDE represents the expectation of an appropriate process under a certain probability measure, in this case, we consider the risk-neutral one.

Proposition 4.1 *(FMM PDE). Let η be the index function defined in (10). Let $R^{min} \in \mathbb{R}$ be a potentially negative lower bound for the rates (to include the relevant cases of shifted-lognormal models). Let $v_k(t) = v_k(t, R_k(t))$ be a general instantaneous volatility for the forward rate $R_k(t)$. Under the risk-neutral measure $\mathcal{Q}$, the price of an interest rate derivative with maturity $T = T_k > T_0 = 0$ (for some $k = 1, \dots, N$), that depends on the fixing of the rates $R_1, \dots, R_N$, with payoff function $\varphi : [R^{min}, \infty)^N \to \mathbb{R}$, is given by*

$$V(t, R_1, \dots, R_N) = P(t, T_0)\Pi(t, R_1, \dots, R_N), \quad t \in [T_0, T],$$

where the relative price $\Pi : [T_0, T] \times [R^{min}, \infty)^N \to \mathbb{R}$ satisfies the PDE

$$\frac{\partial \Pi}{\partial t} + \sum_{k=1}^{N} \mu_k(t) \frac{\partial \Pi}{\partial R_k} + \frac{1}{2} \sum_{k,l=\eta(t)}^{N} \rho_{kl} v_k(t)\gamma_k(t)v_l(t)\gamma_l(t) \frac{\partial^2 \Pi}{\partial R_k \partial R_l} = 0, \quad t \in [T_0, T), \tag{13}$$

along with the terminal condition

$$\Pi(T, R_1, \dots, R_N) = \frac{\varphi(R_1, \dots, R_N)}{P(T, T_0)}, \quad R_1, \dots, R_N \geq R^{min}.$$

Proof. The derivative pays out $\varphi(R_1(T), \dots, R_N(T))$ at the maturity date T. Note that receiving an amount X of money at time T is equivalent to receiving an amount $\dfrac{X}{B(T)} = \dfrac{X}{P(T, T_0)}$ of T_0-bonds. Therefore, the payoff can be interpreted as a relative payoff, in the sense that it is a payoff in terms of an amount of extended zero-coupon bonds maturing at T_0. If $\Pi(t, R_1, \dots, R_N)$ denotes the time t relative price of such a derivative, standard pricing theory states that

$$\begin{aligned}\Pi(t, R_1, \dots, R_N) &= \frac{V(t, R_1, \dots, R_N)}{P(t, T_0)} \\ &= \mathbb{E}^{\mathcal{Q}} \left[\left. \frac{\varphi(R_1(T), \dots, R_N(T))}{P(T, T_0)} \right| \mathcal{F}_t \right],\end{aligned} \tag{14}$$

which allows to write the relative price of the interest rate derivative in terms of an expectation under the risk-neutral probability measure.

We refer the reader to Section 2.3 to check how to compute $P(T, T_0)$ in terms of $R_1(T), \dots, R_N(T)$. Having in mind that (12) describes the system of SDEs under the risk neutral measure $\mathcal{Q}$, the corresponding infinitesimal generator $\mathcal{L}_{R_1, \dots, R_N}$ of $R_1, \dots, R_N$ is thus given by (see [27], for example):

$$\mathcal{L}_{R_1, \dots, R_N} = \frac{\partial}{\partial t} + \sum_{k=1}^{N} \mu_k(t) \frac{\partial}{\partial R_k} + \frac{1}{2} \sum_{k,l=1}^{N} \rho_{kl} v_k(t)\gamma_k(t)v_l(t)\gamma_l(t) \frac{\partial^2}{\partial R_k \partial R_l}.$$

Since $\gamma_k(t)$ and $\gamma_l(t)$ are zero for $k, l < \eta(t)$ the infinitesimal generator can be written as

$$\mathcal{L}_{R_1, \dots, R_N} = \frac{\partial}{\partial t} + \sum_{k=1}^{N} \mu_k(t) \frac{\partial}{\partial R_k} + \frac{1}{2} \sum_{k,l=\eta(t)}^{N} \rho_{kl} v_k(t)\gamma_k(t)v_l(t)\gamma_l(t) \frac{\partial^2}{\partial R_k \partial R_l}.$$

Applying Feynman-Kac theorem, if Π satisfies the PDE

$$\mathcal{L}_{R_1, \dots, R_N} \Pi = 0, \tag{15}$$

$$\Pi(T, \cdot) = \frac{\varphi(\cdot)}{P(T, T_0)}, \tag{16}$$

then Π satisfies equation (14). □

PDE (13) diffuses a relative price, i.e., a price in terms of a bond. After having numerically solved the PDE and thereby having obtained the time t relative value function, the latter has to be multiplied by the time t bond price $P(t, T_0)$ to obtain the absolute value price (the price of the derivative itself). Note that if $t = T_0$, since $P(T_0, T_0) = 1$, then $V(T_0, R_1, \dots, R_N) = \Pi(T_0, R_1, \dots, R_N)$.

5. Numerical methods

In this section, we design both stochastic and deterministic numerical methods to price interest rate derivatives in the recently introduced FMM setting. Although there is a huge variety of products which could be priced, in this work we will consider swaptions. The main reason is that swaptions markets are one of the two main markets in the world of options on interest rates (along with caps and floors). However, the proposed methodology can be extended to a large variety of interest rate derivatives.

In this section and the forthcoming one about numerical results, for simplicity we will consider the lognormal model for volatilities, i.e., we assume that $v_k(t) = \sigma_k(t)R_k(t)$ (therefore $R^{min} = 0$). This model is one of the most popular, and it allows for a more straightforward presentation of the proposed numerical methodologies.

5.1. RFR swaptions

Let us start defining interest rate swaps (IRS). An IRS is a contract with a counterparty to exchange payments indexed to interest rates at future fixed dates. At each time T_i, $i = a+1, \dots, b$, the contract holder pays a fixed interest rate K and receives the floating interest rate $R(T_{i-1}, T_i) = R_i(T_i)$. Therefore, at time T_a the value of the swap is given by

$$\mathrm{IRS}(T_a; T_a, \dots, T_b) = \sum_{i=a+1}^{b} P(T_a, T_i)\tau_i(R_i(T_a) - K). \tag{17}$$

As the underlying rates are RFRs, we can refer to it more precisely as an RFR swap.

Let us now define an option over an IRS with maturity at T_a and where the length of the underlying IRS is $(T_b - T_a)$. We denote such a derivate as a swaption $T_a \times (T_b - T_a)$. A swaption is an option giving the right to enter a swap at the swaption's maturity date T_a. More precisely, the RFR swaption payoff at time T_a is given by

$$\max(\mathrm{IRS}(T_a; T_a, \dots, T_b), 0).$$

Since backward-looking and forward-looking forward rates with the same application period are equal at every time before the start of the period, backward-looking and forward-looking swaptions have the same value.

5.2. Monte Carlo pricing of swaptions with the FMM

In the next paragraphs, we describe how to price RFR swaptions under the FMM using Monte Carlo simulation. The results obtained by this method will serve as benchmark prices for the forthcoming PDE numerical solutions.

In terms of expectations, according to expression (14), the relative price of the swaption $T_a \times (T_b - T_a)$ at time T_0 is given by

$$\mathbb{E}^{\mathcal{Q}} \left[\frac{\max(\mathrm{IRS}(T_a; T_a, \dots, T_b), 0)}{P(T_a, T_0)} \right]. \tag{18}$$

Next, taking into account (17) and the computation of the discount factors in Section 2.3, the expectation (18) depends on the joint distribution of the forward rates R_j, with $1 \leq j \leq b$, at time T_a, i.e. $R_1(T_a), \dots, R_a(T_a), R_{a+1}(T_a), \dots, R_b(T_a)$. According to the dynamics (7), we need to generate several simulations of such rates at time T_a. Finally, we evaluate the relative payoff in each simulation and average.

Since the dynamics (7) does not lead to a process with a known distribution, we perform a time discretization with the small time step Δt. Moreover, we introduce logarithmic interest rates in the FMM setting defined by (7), so that by using Ito's formula we get the following equivalent formulation of the FMM:

$$\mathrm{d}\ln R_k(t) = \left(\mu_k(t)\frac{1}{R_k(t)} - \frac{1}{2}\gamma_k^2(t)\sigma_k^2(t)\right)\mathrm{d}t + \sigma_k(t)\gamma_k(t)\mathrm{d}W_k(t).$$

This formulation has the advantage that the diffusion coefficient $\sigma_k\gamma_k(t)$ is deterministic. Therefore, Euler and Milstein schemes coincide. Consequently, the time discretization

$$R_k(t+\Delta t) = R_k(t)\exp\left(\mu_k(t)\frac{1}{R_k(t)}\Delta t - \frac{1}{2}\gamma_k^2(t)\sigma_k^2(t)\Delta t + \sigma_k(t)\gamma_k(t)(W_k(t+\Delta t) - W_k(t))\right),$$

leads to an approximation of the exact process. Note that the Brownian motion increments are normally distributed with mean 0 and standard deviation $\sqrt{\Delta t}$, and correlated with correlation matrix $\rho = (\rho_{kl})_{k,l=1,\dots,N}$.

5.3. Numerical integration of the multi-dimensional PDE

In this section we design an appropriate and efficient numerical scheme for solving the PDE stated in Section 4, when considering the lognormal model for volatilities.

Taking into account the notation of the tenor structure for the swap, let us focus on the pricing of the swaption $T_1 \times (T_N - T_1)$, with payoff function

$$\varphi(R_1,\dots,R_N) = \max\left\{\sum_{k=2}^{N} P(T_1,T_k)\tau_k(R_k(T_1) - K),\, 0\right\},$$

so the relative price $\Pi(t, R_1,\dots,R_N)$ satisfies the PDE (13) for $t \in [T_0, T_1]$ and the final condition

$$\Pi(T_1, R_1,\dots,R_N) = \frac{1}{P(T_1,T_0)}\max\left\{\sum_{k=2}^{N} P(T_1,T_k)\tau_k(R_k - K),\, 0\right\}.$$

In order to solve numerically this PDE problem, for simplicity we consider the following change of time variable and notations:

$$u(t,x_1,\dots,x_N) = \Pi(T_1 - t, R_1,\dots,R_N), \quad t\in[0, T_1 - T_0] = [0,\tau_1],\ x_k = R_k \geq 0. \tag{19}$$

Note that we are allowing a certain abuse of notation: we maintain the notation t, which initially represented physical time and hereafter represents the remaining time up to T_1. In this way the final condition at physical time T_1 becomes an initial condition in the PDE formulation in the new time variable. Numerical results in the next section will always refer to the physical time.

After the previously indicated change of time variable and notations, the PDE problem can be expressed as follows, find the function u, such that:

$$\frac{\partial u}{\partial t} = \sum_{k=1}^{N}\mu_k(t,x_1,\dots,x_k)\frac{\partial u}{\partial x_k} + \sum_{k=1}^{N}\delta_k(t,x_k)\frac{\partial^2 u}{\partial x_k^2} + \sum_{k=1}^{N-1}\sum_{l=k+1}^{N}\varrho_{kl}(t,x_k,x_l)\frac{\partial^2 u}{\partial x_k\,\partial x_l}, \tag{20}$$

for all $x_k > 0$, where

$$\mu_k(t,x_1,\dots,x_k) = \lambda_k(t)x_k\sum_{j=1}^{k}\rho_{kj}\lambda_j(t)\frac{\tau_j}{1+\tau_j x_j}x_j,$$
$$\delta_k(t,x_k) = \tfrac{1}{2}\lambda_k^2(t)x_k^2,$$
$$\varrho_{kl}(t,x_k,x_l) = \rho_{kl}\lambda_k(t)\lambda_l(t)x_k x_l,$$
$$\lambda_k(t) = \sigma_k(T_1 - t)\gamma_k(T_1 - t) = \begin{cases}\sigma_1(T_1 - t)\dfrac{t}{\tau_1}, & \text{if } k=1\\ \sigma_k(T_1 - t), & \text{if } k\geq 2,\end{cases} \quad t\in[0,\tau_1], \tag{21}$$

with initial condition

$$u(0,x_1,\dots,x_N) = u_0(x_1,\dots,x_N) := \max\{g(x_1,\dots,x_N), 0\},$$
$$g(x_1,\dots,x_N) = \sum_{k=2}^{N}\left(\prod_{l=1}^{k}\frac{1}{1+\tau_l x_l}\right)\tau_k(x_k - K). \tag{22}$$

Note that the expression of λ_k in (21) is valid for $t\in[0,\tau_1]$.

As usual, in financial problems initially posed in unbounded spatial domains, for the numerical integration of this PDE, the spatial domain must be restricted to a rectangle $(x_1,\dots,x_N)\in\Omega = (0, R_1^{max})\times\cdots\times(0, R_N^{max})$, by selecting the values $\{R_k^{max}\}_{k=1}^{N}$ large enough. On the upper boundaries, a linear behavior of the solution is assumed by considering the conditions

$$\frac{\partial^2 u}{\partial x_k^2}(t,x_1,\dots,x_N) = 0 \quad \text{if } x_k = R_k^{max}, \quad k=1,\dots,N, \tag{23}$$

while, due to the degeneracy of the PDE at the boundaries $x_k = 0$, no conditions are required at these boundaries.

Next, we propose a finite difference method to approximate the solution of the PDE problem (20)-(21)-(22)-(23). Firstly, a discretization of the spatial derivatives $u_{x_k}, u_{x_k x_k}, u_{x_k x_j}$ must be done. In order to build a spatial mesh, it must be taken into account that the payoff (22) lacks differentiability at the points of Ω such that

$$\tilde{g}(x_2,\dots,x_N) := \sum_{k=2}^{N}\left(\prod_{l=2}^{k}\frac{1}{1+\tau_l x_l}\right)\tau_k(x_k - K) = 0, \tag{24}$$

so, as it is recommended in [28, Chap. 4] and the references therein, a non-uniform spatial mesh on each x_k-direction, for $k\geq 2$, results more convenient. Therefore, given N positive integers $\{M_1,\dots,M_N\}$ we consider the spatial grid

$$\begin{aligned}
&\text{for } x_1: \ x_{1,j_1} = j_1 h_1, \quad 0\leq j_1\leq M_1, \quad h_1 = \frac{R_1^{max}}{M_1},\\
&\text{for } x_k,\ k\geq 2: \ x_{k,j_k} = K + L_k\sinh\xi_{k,j_k},\ 0\leq j_k\leq M_k,\\
&\qquad h_{k,j_k} = x_{k,j_k} - x_{k,j_k-1},\\
&\qquad \xi_{k,j_k} = \xi_k^{min} + j_k\Delta\xi_k, \quad \Delta\xi_k = \frac{\xi_k^{max} - \xi_k^{min}}{M_k},\\
&\qquad \xi_k^{min} = \sinh^{-1}(-K/L_k),\\
&\qquad \xi_k^{max} = \sinh^{-1}\left((R_k^{max} - K)/L_k\right),
\end{aligned} \tag{25}$$

where the parameters L_k measure the fraction of grid points that are closer to K. In this case, we will consider $L_k = K/10$, $\forall k\geq 2$.

In addition, a cell averaging technique is applied to smooth the payoff at the grid points close to the hyperplane defined by (24), by adapting the technique proposed in [28, Chap. 4] for 1D-PDEs. In this case, for each subset of indices $(j_3,\dots,j_N)$, $0\leq j_k\leq M_k$, $k\geq 3$, we compute the value

$$\tilde{x}_2 := K - \frac{1}{\tau_2}\sum_{k=3}^{N}\left(\prod_{l=3}^{k}\frac{1}{1+\tau_l x_{l,j_l}}\right)\tau_k(x_{k,j_k} - K),$$

($\tilde{x}_2 = K$ for $N = 2$) and look for the index j_{ind}, with $0 \le j_{ind} \le M_2$, such that $|x_{2,j_{ind}} - \tilde{x}_2| = \min_{0\le j_2\le M_2} |x_{2,j_2} - \tilde{x}_2|$. Then, the cell $[x_2^-, x_2^+]$ is considered, where

$$x_2^- = \frac{x_{2,j_{ind}-1} + x_{2,j_{ind}}}{2}, \quad x_2^+ = \frac{x_{2,j_{ind}} + x_{2,j_{ind}+1}}{2}, \quad \tilde{h}_2 = x_2^+ - x_2^-,$$

and the initial condition at the points $(x_{1,j_1}, x_{2,j_{ind}}, x_{3,j_3}, \dots, x_{N,j_N})$, for all $j_1 = 0, \dots, M_1$, is taken as the average over this cell $[x_2^-, x_2^+]$, that is,

$$\begin{aligned} & u(0, x_{1,j_1}, x_{2,j_{ind}}, x_{3,j_3}, \dots, x_{N,j_N}) \\ &= \frac{1}{\tilde{h}_2} \int_{x_2^-}^{x_2^+} u_0(x_{1,j_1}, x_2, x_{3,j_3}, \dots, x_{N,j_N})\, dx_2 \\ &= \frac{1}{\tilde{h}_2(1+\tau_1 x_{1,j_1})} \left(x_2^+ - \tilde{x}_2 - \left(K + \frac{1-H_3}{\tau_2} \right) \log\left(\frac{1+\tau_2 x_2^+}{1+\tau_2 \tilde{x}_2} \right) \right), \end{aligned}$$

with

$$H_3 = \begin{cases} 0, & \text{if } N = 2, \\ \displaystyle\sum_{k=3}^{N} \left(\prod_{l=3}^{k} \frac{1}{1+\tau_l x_{l,j_l}} \right) \tau_k (x_{k,j_k} - K) & \text{if } N \ge 3. \end{cases}$$

On the other mesh points, the payoff function given in (22) is applied.

Following the notation given in [2, Sec. 3], the grid points are rearranged as

$$\begin{aligned} \overline{\Omega}_h = \{ & \vec{x}_J = (x_{1,j_1}, \dots, x_{N,j_N}) \,:\, J = \vartheta_0(j_1, \dots, j_N), \\ & 0 \le j_k \le M_k,\ 1 \le k \le N \}, \end{aligned}$$

where the bijection $\vartheta_0 : \mathcal{I}_N^{(0)} \to \{0, 1, \dots, M-1\}$, with $M = \displaystyle\prod_{k=1}^{N} (M_k + 1)$ is defined as

$$\vartheta_0(x_{1,j_1}, \dots, x_{N,j_N}) = \sum_{k=1}^{N} j_k E_k, \quad \text{with } E_1 = 1,\ E_k = \prod_{l=1}^{k-1} (M_l + 1),\ k \ge 2.$$

Because of the boundary conditions, the solution of the PDE on all the nodes (including the boundary points) of the spatial grid is unknown. Then, by approximating the spatial derivatives in the PDE by second-order central finite differences schemes on all the nodes of the spatial grid, for each $J = 0, 1, \dots, M-1$, $(j_1, \dots, j_N) = \vartheta_0^{-1}(J)$, we get

$$\begin{aligned} Y_J' = & \sum_{k=1}^{N} \mu_k(t, x_{1,j_1}, \dots, x_{k,j_k}) \nabla_J^{(k)} + \sum_{k=1}^{N} \delta_k(t, x_{k,j_k}) \Delta_J^{(k)} \\ & + \sum_{k=1}^{N-1} \sum_{l=k+1}^{N} \varrho_{kl}(t, x_{k,j_k}, x_{l,j_l}) \Delta_J^{(kl)}, \end{aligned} \tag{26}$$

where the involved discrete operators are

$$\begin{aligned} \nabla_J^{(1)} &= \frac{Y_{J+E_1} - Y_{J-E_1}}{2h_1}, \\ \nabla_J^{(k)} &= \beta_{j_k,-1}^{(k)} Y_{J-E_k} + \beta_{j_k,0}^{(k)} Y_J + \beta_{j_k,1}^{(k)} Y_{J+E_k},\ k \ge 2, \\ \Delta_J^{(1)} &= \frac{Y_{J+E_1} - 2Y_J + Y_{J-E_1}}{h_1^2}, \\ \Delta_J^{(k)} &= \eta_{j_k,-1}^{(k)} Y_{J-E_k} + \eta_{j_k,0}^{(k)} Y_J + \eta_{j_k,1}^{(k)} Y_{J+E_k},\ k \ge 2, \\ \Delta_J^{(kl)} &= \beta_{j_l,-1}^{(l)} \nabla_{J-E_l}^{(k)} + \beta_{j_l,0}^{(l)} \nabla_J^{(k)} + \beta_{j_l,1}^{(l)} \nabla_{J+E_l}^{(k)},\ l \ge k+1, \end{aligned} \tag{27}$$

with

$$\begin{aligned} \beta_{j_k,-1}^{(k)} &= \frac{-h_{k,j_k+1}}{h_{k,j_k}(h_{k,j_k} + h_{k,j_k+1})}, & \eta_{j_k,-1}^{(k)} &= \frac{2}{h_{k,j_k}(h_{k,j_k} + h_{k,j_k+1})}, \\ \beta_{j_k,0}^{(k)} &= \frac{h_{k,j_k+1} - h_{k,j_k}}{h_{k,j_k} h_{k,j_k+1}}, & \eta_{j_k,0}^{(k)} &= \frac{-2}{h_{k,j_k} h_{k,j_k+1}}, \\ \beta_{j_k,1}^{(k)} &= \frac{h_{k,j_k}}{h_{k,j_k+1}(h_{k,j_k} + h_{k,j_k+1})}, & \eta_{j_k,1}^{(k)} &= \frac{2}{h_{k,j_k+1}(h_{k,j_k} + h_{k,j_k+1})}. \end{aligned} \tag{28}$$

It must be observed that, with this procedure, the spatial discretization is expanded to the borders. On the one hand, at the points of the "right" borders the conditions (23) have to be involved, so when $j_k = M_k$, $\Delta_J^{(k)} = 0$ is imposed. Therefore, virtual points Y_{J+E_k} must be defined when $j_k = M_k$,

$$Y_{J+E_1} = 2Y_J - Y_{J-E_1}, \quad Y_{J+E_k} = -\frac{\eta_{j_k,-1}^{(k)}}{\eta_{j_k,1}^{(k)}} Y_{J-E_k} - \frac{\eta_{j_k,0}^{(k)}}{\eta_{j_k,1}^{(k)}} Y_J,\ k \ge 2.$$

On the other hand, for each $k = 1, \dots, N$, on the boundary $x_k = 0$, we get $\mu_k = \delta_k = 0$, $\forall k$, $\varrho_{kl} = 0$, $\forall l \ge k+1$. As a consequence, $\nabla_J^{(k)} = \Delta_J^{(k)} = 0$, $\forall k$, $\Delta_J^{(kl)} = 0$, $l \ge k+1$, when $j_k = 0$.

By grouping all the approximations in a vector $Y = (Y_J)_{J=0}^{M-1}$, these equations can be written as the initial value problem with a directional splitting

$$\begin{gathered} Y' = \mathcal{F}(t, Y) = \sum_{k=0}^{N} \mathcal{F}_k(t, Y), \quad Y(0) = Y_0, \\ \mathcal{F}_k(t, Y) = \mathcal{A}_k(t) Y, \quad k = 0, 1, \dots, N, \\ \mathcal{A}_1(t) = \lambda_1^2(t) \tilde{\mathcal{A}}_1, \quad \mathcal{A}_k(t) = \lambda_k^2(t) \tilde{\mathcal{A}}_k^{(1)} + \lambda_k(t) \mathcal{D}_k(t) \tilde{\mathcal{A}}_k^{(2)},\ k = 2, \dots, N, \end{gathered} \tag{29}$$

where each $\mathcal{F}_k(t, Y)$ stores the components of the discretization of the advection and diffusion terms in the x_k-direction, for $k = 1, \dots, N$, and $\mathcal{F}_0(t, Y)$ stores those of the discretization of the mixed derivatives. In this case, $\tilde{\mathcal{A}}_1$, $\{\tilde{\mathcal{A}}_k^{(1)}, \tilde{\mathcal{A}}_k^{(2)}\}_{k=2}^{N}$ are block tridiagonal constant matrices and $\mathcal{D}_k(t)$ is diagonal. For the sake of simplicity, the coefficients of these matrices are given in Appendix A.

Due to the increasing stiffness of (29) as the resolution of the spatial grid increases, explicit methods are not suitable for its time integration. On the other hand, fully implicit methods requiring the computation of the exact Jacobian of the derivative function are also unsuitable because of the complicated structure of the matrix $\mathcal{A}_0(t)$. Therefore, for the time integration of (29) a method from the class of AMFR-W-methods ([22, Sec. 4], [23]) is applied. In particular, we have selected the one-stage AMFR-W1 method. More precisely, given an approximation Y_n to the solution of (29) at the time $t = t_n$, this method approximates the solution at $t = t_{n+1} = t_n + \Delta t$ (with Δt being the constant step of the time discretization) by

$$\begin{aligned} K^{(0)} &= \Delta t\, \mathcal{F}(t_n, Y_n), \\ (I - \nu \Delta t\, \mathcal{A}_k(t_n)) K^{(k)} &= K^{(k-1)} + \nu (\Delta t)^2 \alpha_{k,n}, \quad k = 1, \dots, N, \\ \tilde{K}^{(0)} &= 2K^{(0)} + \theta (\Delta t)^2 G_n - (I - \theta \Delta t\, \mathcal{A}(t_n)) K^{(N)}, \\ (I - \nu \Delta t\, \mathcal{A}_k(t_n)) \tilde{K}^{(k)} &= \tilde{K}^{(k-1)} + \nu (\Delta t)^2 \alpha_{k,n}, \quad k = 1, \dots, N, \\ Y_{n+1} &= Y_n + \tilde{K}^{(N)}, \end{aligned} \tag{30}$$

where

$$\begin{aligned} & \mathcal{A}(t_n) = \frac{\partial \mathcal{F}}{\partial Y}(t_n, Y_n) = \sum_{k=0}^{N} \mathcal{A}_k(t_n), \\ & \alpha_{k,n} = \frac{\partial \mathcal{F}_k}{\partial t}(t_n, Y_n),\ k = 1, \dots, N, \quad G_n = \frac{\partial \mathcal{F}}{\partial t}(t_n, Y_n), \end{aligned}$$

with parameters $\theta = 1/2$ and $\nu = \theta$ for $N = 2, 3$ and $\nu = \kappa_N N \theta$ for $N \ge 4$, where the values of κ_N are given in [22, Table 2] and guarantee that the AMFR-W1 method is unconditionally stable on multi-dimensional linear constant coefficient PDEs with mixed derivatives.

Table 1
Hypothetical market data of RFRs for the numerical examples.

k	T_k	$R_k(0)$	$\sigma_k(t)$
1	0.25	0.01	0.2
2	0.5	0.013	0.15
3	0.75	0.014	0.25
4	1.0	0.015	0.26
5	1.25	0.016	0.27

It must be observed that, due to the block tridiagonal structure of the matrices $\mathcal{A}_k(t)$, $k = 1,\dots,N$, the linear systems with coefficient matrix $(I - \nu\Delta t\,\mathcal{A}_k(t_n))$ of dimension M in the method are decoupled in tridiagonal systems of dimension $M_k + 1$, which drastically reduces their computational cost (see details in Appendix A). In addition, it is not necessary to compute the full Jacobian $\mathcal{A}(t_n)$ in (30), which does not have a block structure that can reduce its computational cost, due to the presence of the discretization of the mixed derivatives. The product $\mathcal{A}(t_n)K^{(N)}$ in the right-hand side in the definition of $\tilde{K}^{(0)}$ in (30) is obtained by an additional evaluation of the derivative function, since

$$\mathcal{A}(t_n)K^{(N)} = \left(\sum_{k=0}^{N}\mathcal{A}_k(t_n)\right)K^{(N)} = \mathcal{F}(t_n, K^{(N)}).$$

6. Numerical experiments

In this section, we present some numerical results that correspond to several RFR swaptions to assess the correctness of the proposed numerical methods and the performance of the models. The selected swaptions aim to illustrate the behavior of the methodology when increasing the number of involved RFRs.

Some of the employed market data are shown in Table 1, where constant volatilities are taken into account. Also, we will consider that $T_0 = 0$ and that the pricing date is the physical time $t = T_0 = 0$ (which will correspond to the solution of the PDE at time $t = T_1$ in the new time variable with the previously mentioned abuse of notation, although we will always refer to the physical time $t = 0$).

Moreover, the constant correlation coefficients $\rho_{kl} = 0.5$ have been chosen, for all $k, l = 1,\dots,N$, with $k \neq l$.

In all forthcoming examples, we will start computing prices using Monte Carlo simulation, which will serve as benchmark RFR swaption prices for the corresponding PDE numerical solutions.

The numerical experiments have been performed with the following hardware and software configurations: an AMD Ryzen 9 5950X 16-Core Processor with 128 GBytes of RAM, CentOS Linux, and GNU C++ compiler. We have not used a numerical linear algebra software package, because the resulting linear systems of equations are tridiagonal, as previously stated. Thus, the codes developing the numerical methods have been implemented from scratch. Besides, parallel computing was not considered, since the AMFR-W1 time integrator is highly sequential.

6.1. 2-dimensional case ($N = 2$)

In this first example, we price, at time $t = T_0 = 0$, several RFR European swaptions $T_1 \times (T_2 - T_1)$ for the values of the spot forward rates $R_1(0)$ and $R_2(0)$ given on Table 1. More precisely, we value the swaption at-the-money (ATM) ($K = K_{ATM}$), a couple of swaptions out-of-the-money (OTM) ($K = 1.1K_{ATM}$, $K = 1.2K_{ATM}$), and two swaptions in-the-money (ITM) ($K = 0.8K_{ATM}$, $K = 0.9K_{ATM}$).

Monte Carlo and PDE results are shown in Table 2. Monte Carlo confidence intervals have been obtained with 10^7 simulations and 100 time steps for the Milstein discretization scheme. In order to check the performance of the PDE AMFR-W1 numerical method, firstly we have tested it on a spatial grid with $M_1 = M_2 = 1024$ and a small constant time step size $\Delta t = \tau_1/2^{11}$, so that the error associated to the time integration is negligible when compared to the one due to the spatial discretization. PDE prices for the values of Table 1 are obtained by using multi-linear interpolation since those points could not belong to the spatial mesh (25). Note that both Monte Carlo and PDE prices are consistent, thus validating both numerical techniques and also the well-posedness of the PDEs formulation and the associated boundary conditions. Besides, the implied volatilities corresponding to the PDE prices are also shown. Since the chosen dynamics are lognormal, implied volatilities are flat (not perfectly flat since the swap rate is not exactly lognormal).

Table 2
95% Monte Carlo confidence intervals for swaption prices, 10^7 simulations with a 100 time steps. PDE prices, computed on the non-uniform grids, and corresponding implied volatilities.

Swaption $T_1 \times (T_2 - T_1)$			
K	Monte Carlo Confidence Interval	PDE	Impl vol
$1.2\,K_{ATM}$	$[6.569174\times10^{-7}, 6.705475\times10^{-7}]$	6.610817×10^{-7}	0.150103
$1.1\,K_{ATM}$	$[1.229203\times10^{-5}, 1.235655\times10^{-5}]$	1.230812×10^{-5}	0.150014
K_{ATM}	$[9.663654\times10^{-5}, 9.681989\times10^{-5}]$	9.666517×10^{-5}	0.150003
$0.9\,K_{ATM}$	$[3.313149\times10^{-4}, 3.315975\times10^{-4}]$	3.314849×10^{-4}	0.150035
$0.8\,K_{ATM}$	$[6.460959\times10^{-4}, 6.463961\times10^{-4}]$	6.463699×10^{-4}	0.150143
Time	73.32 s	603.82 s, $M_1 = M_2 = 1024$	

Table 3
ATM Swaption $T_1 \times (T_2 - T_1)$: spatial errors and estimated orders on the non-uniform grids (25) with $L = M_1 = M_2$.

L	l_2-error	l_∞-error	l_2-order	l_∞-order	Time (s)
4	4.53×10^{-05}	1.02×10^{-04}	-	-	1.19×10^{-02}
8	3.76×10^{-06}	1.04×10^{-06}	3.59	3.29	3.53×10^{-02}
16	4.00×10^{-07}	1.07×10^{-06}	3.23	3.28	0.12
32	1.07×10^{-07}	2.71×10^{-07}	1.91	1.98	0.45
64	2.59×10^{-08}	6.45×10^{-08}	2.04	2.07	1.73
128	6.72×10^{-09}	1.68×10^{-08}	1.95	1.94	7.62
256	1.67×10^{-09}	4.15×10^{-09}	2.01	2.02	31.86
512	3.97×10^{-10}	9.87×10^{-10}	2.07	2.07	138.07
1024	7.93×10^{-11}	1.97×10^{-10}	2.32	2.32	603.82

For the sole purpose of illustrating the behavior of the PDE solution, in Fig. 3 the approximated values of u at $t = T_1$ are reported in the same case with $M_1 = M_2 = 1024$ mesh points. It must be emphasized again that, due to the time reversal (19), the values of u when $t = T_1$ are equivalent to that of Π when $t = T_0 = 0$. The solution on the computational domain and a zoom in the area of financial interest near the strike rate are shown.

In order to illustrate the convergence of the finite difference method with AMFR-W1 (30), we have applied it for several spatial meshes (25), with $M_1 = M_2 = L = 2^r$, $r = 2,3,\dots,10$, and the same constant time step size $\Delta t = \tau_1/2^{11}$ as above. In Table 3 the errors obtained for all cases are displayed, in both l_2- and l_∞-norms. There, the errors (l_2-error and l_∞-error) have been computed with respect to the numerical solution given by the same method with $M_1 = M_2 = 2^{11}$, and the estimated spatial orders (l_2-order and l_∞-order) gathered in the last two columns are calculated by the well-known formula

$$l_p - \text{order} = \frac{\log(\|\, l_p - \text{error}(2m)\,\| \,/\, \|\, l_p - \text{error}(m)\,\|)}{\log(2)},$$

for $m = 2^r$, $r = 2,3,\dots,9$, with $p = 2$ and $p = \infty$. The second order of convergence is reflected in Table 3.

In addition, in Table 4 the results obtained using a uniform spatial mesh $x_{k,j_k} = j_k h_k$, $0 \le j_k \le L$, with $h_k = R_k^{max}/L$, are shown. Although also achieving second-order convergence with uniform grids, the advantage of using non-uniform meshes is clearly illustrated when comparing Table 3 and Table 4, as it improves the accuracy achieved by the method by two orders of magnitude.

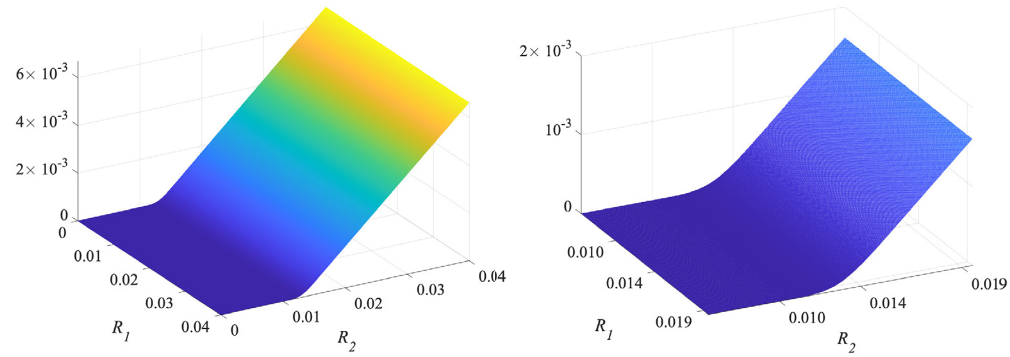

Fig. 3. PDE prices for the $T_1 \times (T_2 - T_1)$ swaption, with 1024×1024 grid. Left: full computational domain. Right: zoom in the area of interest $[0.5K_{ATM}, 1.5K_{ATM}]^2$.

Table 4
ATM Swaption $T_1 \times (T_2 - T_1)$: spatial errors and estimated orders on uniform grids with $L = M_1 = M_2$.

L	l_2-error	l_∞-error	l_2-order	l_∞-order	Time (s)
4	4.30×10^{-6}	8.33×10^{-6}	-	-	1.21×10^{-02}
8	5.53×10^{-6}	1.57×10^{-5}	-0.36	-0.91	3.59×10^{-02}
16	3.87×10^{-6}	1.23×10^{-5}	0.52	0.35	0.12
32	1.13×10^{-6}	5.00×10^{-6}	1.77	1.30	0.45
64	3.21×10^{-7}	1.95×10^{-6}	1.82	1.36	1.72
128	6.91×10^{-8}	3.03×10^{-7}	2.22	2.68	7.61
256	1.98×10^{-8}	1.13×10^{-7}	1.80	1.42	31.73
512	4.13×10^{-9}	1.88×10^{-8}	2.26	2.59	138.48
1024	1.15×10^{-9}	6.79×10^{-9}	1.85	1.47	605.42

6.2. Higher-dimensional cases ($N = 3, 4, 5$)

Next, we focus on the pricing of higher dimensional RFR swaptions. More precisely, we consider the swaptions $T_1 \times (T_3 - T_1)$, $T_1 \times (T_4 - T_1)$ and $T_1 \times (T_5 - T_1)$ under the market data shown in Table 1. Given the conclusions from the previous example about the numerical solution of the PDE, we just present the results for non-uniform grids.

PDE prices and the corresponding reference Monte Carlo confidence intervals are shown in Table 5. For each swaption, the number of mesh points in all directions is the same and denoted by L. As in the 2-dimensional case, to keep temporal errors negligible compared to spatial errors, for the time integration, the method AMFR-W1 (30) has been applied with constant step size $\Delta t = \tau_1/(2L)$. As expected, all PDE prices lie inside the Monte Carlo confidence intervals. Note that as before, the 95% Monte Carlo confidence intervals were computed by using 10^7 simulations with 100 time steps in the Milstein time stepping.

Next, the numerical orders of convergence of the PDE method for the three and four-dimensional spatial examples are shown in Tables 6 and 7, respectively. In such cases, the time integration has been performed with constant time step size $\Delta t = \tau_1/2^r$ with $r = 9$ for $N = 3$ and $r = 7$ for $N = 4$. For the five-dimensional case, we did not measure the order of convergence since this test would require a huge amount of RAM, which is not available in our machine.

Then, in Fig. 4, we present a graph of some of the previously shown execution times per problem size for both Monte Carlo and PDEs. As expected, Monte Carlo method does not suffer from the curse of dimensionality.

Finally, we show some numerical results for pricing Bermudan swaptions. We consider the case of two callability dates, known as Canary option. More precisely, we focus on the Canary swaption $T_3 \times (T_4 - T_3)$, where the two early exercises opportunities are T_2 and T_1, allowing to enter in the swaps $T_2 \times (T_4 - T_2)$ and $T_1 \times (T_4 - T_1)$, respectively. The results are shown in Table 8, along with the corresponding European swaption prices, which are lower as expected.

Table 5
95% Monte Carlo confidence intervals for swaption prices, 10^7 simulations with a 100 time steps. PDE prices and corresponding implied volatilities.

Swaption $T_1 \times (T_3 - T_1)$			
K	Monte Carlo Confidence Interval	PDE	Impl vol
$1.2\,K_{ATM}$	$[5.007571 \times 10^{-6}, 5.070211 \times 10^{-6}]$	5.020028×10^{-6}	0.178879
$1.1\,K_{ATM}$	$[4.532638 \times 10^{-5}, 4.552660 \times 10^{-5}]$	4.538339×10^{-5}	0.177969
K_{ATM}	$[2.361209 \times 10^{-4}, 2.365753 \times 10^{-4}]$	2.364758×10^{-4}	0.177020
$0.9\,K_{ATM}$	$[7.014066 \times 10^{-4}, 7.020817 \times 10^{-4}]$	7.014788×10^{-4}	0.176040
$0.8\,K_{ATM}$	$[1.340121 \times 10^{-3}, 1.340854 \times 10^{-3}]$	1.340742×10^{-3}	0.175032
Time	112.94 s	4316.30 s, $L = 256$	

Swaption $T_1 \times (T_4 - T_1)$			
K	Monte Carlo Confidence Interval	PDE	Impl vol
$1.2\,K_{ATM}$	$[9.480228 \times 10^{-6}, 9.589930 \times 10^{-6}]$	9.523646×10^{-6}	0.184582
$1.1\,K_{ATM}$	$[7.775208 \times 10^{-5}, 7.808471 \times 10^{-5}]$	7.788910×10^{-5}	0.183922
K_{ATM}	$[3.794420 \times 10^{-4}, 3.801720 \times 10^{-4}]$	3.800981×10^{-4}	0.183272
$0.9\,K_{ATM}$	$[1.094727 \times 10^{-3}, 1.095804 \times 10^{-3}]$	1.095566×10^{-3}	0.182621
$0.8\,K_{ATM}$	$[2.081112 \times 10^{-3}, 2.082289 \times 10^{-3}]$	2.082134×10^{-3}	0.181977
Time	150.69 s	23410.36 s, $L = 128$	

Swaption $T_1 \times (T_5 - T_1)$			
K	Monte Carlo Confidence Interval	PDE	Impl vol
$1.2\,K_{ATM}$	$[1.485427 \times 10^{-5}, 1.501782 \times 10^{-5}]$	1.500055×10^{-5}	0.188628
$1.1\,K_{ATM}$	$[1.139641 \times 10^{-4}, 1.144421 \times 10^{-4}]$	1.143997×10^{-4}	0.187909
K_{ATM}	$[5.350862 \times 10^{-4}, 5.361152 \times 10^{-4}]$	5.357548×10^{-4}	0.187452
$0.9\,K_{ATM}$	$[1.515406 \times 10^{-3}, 1.516917 \times 10^{-3}]$	1.516010×10^{-3}	0.187002
$0.8\,K_{ATM}$	$[2.869551 \times 10^{-3}, 2.871208 \times 10^{-3}]$	2.870076×10^{-3}	0.186816
Time	196.45 s	77738.56 s, $L = 64$	

Table 6
ATM Swaption $T_1 \times (T_3 - T_1)$: spatial error and estimated orders on the non-uniform grids (25) with $L = M_1 = M_2 = M_3$.

L	l_2-error	l_∞-error	l_2-order	l_∞-order	Time (s)
8	1.30×10^{-5}	9.00×10^{-5}	-	-	0.19
16	3.97×10^{-6}	3.76×10^{-5}	1.72	1.26	1.05
32	1.62×10^{-6}	1.88×10^{-5}	1.29	1.00	7.26
64	3.15×10^{-7}	4.25×10^{-6}	2.37	2.15	58.23
128	6.31×10^{-8}	7.62×10^{-7}	2.32	2.48	482.55
256	1.24×10^{-8}	1.46×10^{-7}	2.34	2.39	4316.30

Table 7
ATM Swaption $T_1 \times (T_4 - T_1)$: spatial error and estimated orders on the non-uniform grids (25) with $L = M_1 = M_2 = M_3 = M_4$.

L	l_2-error	l_∞-error	l_2-order	l_∞-order	Time (s)
8	6.96×10^{-4}	1.83×10^{-2}	-	-	0.59
16	1.45×10^{-5}	3.33×10^{-4}	5.58	5.78	6.62
32	2.35×10^{-6}	6.27×10^{-5}	2.63	2.41	97.18
64	3.91×10^{-7}	1.53×10^{-5}	2.59	2.03	1473.89

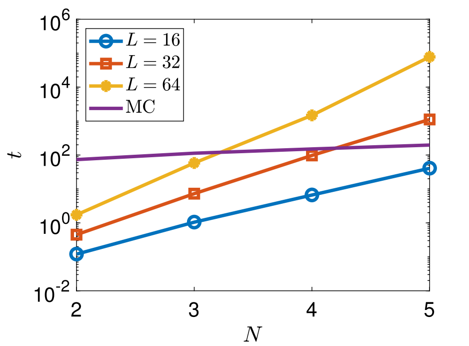

Fig. 4. Plot of execution times per problem size ($\log_{10}$ scale in the y axis). For the PDEs, the cases with $L = 16, 32, 64$ are shown. For Monte Carlo method, 10^7 paths were considered.

Table 8
PDE prices for the Swaption $T_3 \times (T_4 - T_1)$ considering $L = 128$. Execution time is 73522.17 seconds.

K	European	Canary
$1.2K_{ATM}$	1.075029×10^{-4}	1.205543×10^{-4}
$1.1K_{ATM}$	1.938612×10^{-4}	2.355759×10^{-4}
K_{ATM}	3.317602×10^{-4}	4.483170×10^{-4}
$0.9K_{ATM}$	5.339683×10^{-4}	8.331122×10^{-4}
$0.8K_{ATM}$	8.032271×10^{-4}	1.553333×10^{-3}

7. Conclusions and future work

After the recent scandal of manipulation of IBORs, the worldwide financial authorities, and regulators started the replacement of IBORs by the so-called RFRs, that rely on real transactions. In this so-called IBOR transition banks started to offer interest rate derivatives based on RFRs, thus motivating the need for an appropriate modeling RFRs dynamics to price these derivative products. The recent seminal rigorous article [1] introduced the generalized FMM. Pricing with Monte Carlo under FMM is very natural.

Having in view the limitations of Monte Carlo pricing techniques, in the present article we have rigorously stated for the first time in the literature a PDE formulation for pricing RFRs derivatives. Note that the spatial dimension is equal to the number of RFRs involved in the derivative payoff and can become large. In order to solve efficiently the PDE formulation, we propose the use of an AMFR-W1 finite differences method, which is specially appropriate and efficient to cope with the presence of mixed derivatives in the spatial variables. The numerical results illustrate the correctness of the method, by comparing them with a reference solution obtained with converged Monte Carlo simulations. Moreover, order two in space is verified for the different examples.

In future work, we aim to adapt the previous FMM PDE of Proposition 4.1 to allow the pricing of derivatives with payoffs including past fixings of RFRs. Many interest-rate contracts have payoffs depending not only on $R_k(T_k) = R(T_{k-1}, T_k)$, but also on the backward-looking rates

$$R(t,T) = \frac{1}{T-t}\left(e^{\int_t^T r(u)\mathrm{d}u} - 1\right), \tag{31}$$

for general $t < T$. Such derivatives with payment times and settings outside the FMM tenor structure $\{T_0, T_1, \ldots, T_N\}$ can not be priced directly with the FMM or its corresponding PDE. In [21], Lyashenko and Mercurio completed the FMM by embedding it into a Markovian HJM model with separable Cheyette volatility structure by aligning the HJM and FMM dynamics of the forward rates modeled by the FMM. Under this aligned FMM-Cheyette Markovian HJM model, it is possible to derive the dynamics of the short rate r. Besides, this short rate can be simulated by leveraging the realized paths of the simulated generalized FMM rates $R_k(t)$ and their volatilities $v_k(t)$. In order to adapt the FMM PDE to be able to price such path-dependent products, one possible approach is to introduce additional path-dependent variables. For example, to price a derivative whose payoff depends on the evolution of $R_k(\tau)$ for all times $\tau \in [T_{k-1}, t]$, we can define a path-dependent variable of the general form

$$I_k(t) = \int_{T_{k-1}}^{t} \zeta_k(\tau, R_k(\tau))\mathrm{d}\tau, \quad T_{k-1} < t < T_k. \tag{32}$$

Here, the given function ζ_k vanishes for the time values outside the specified interval $[T_{k-1}, t]$. The specification of the function ζ_k is part of future work and needs to be consistent with the FMM-HJM model in [21]. From now on, one must compute the SDE for I_k, and apply Itô's lemma to the derivative price function Π depending on time, forward rates, and the additional state variable I_k. Finally, the PDE to price path-dependent RFR derivatives will be obtained by imposing no-arbitrage conditions. This strategy is classical and is explained for Asian options in equity markets in [29], for example. Discrete sampling is also possible within this path-dependent framework and can be implemented by forcing Π to satisfy the so-called jump conditions.

Another possible extension is the consideration of sparse grid combination techniques for solving PDEs with higher spatial dimensions. The authors have already developed this strategy for pricing interest rate derivatives under the SABR/LIBOR model in [2]. This approach will also open the door to parallel computing.

Data availability

No data was used for the research described in the article.

Acknowledgements

S. Pérez-Rodríguez has been partially supported by the Spanish Ministry of Science and Innovation through project PID2022-141385NB-I00. J.G. López-Salas and C. Vázquez acknowledge the funding by the Spanish Ministry of Science and Innovation through the projects PID2019-10858RB-I00 and PID2022-141058OB-I00, as well as from the Galician Government through grant ED431C 2022/47, including FEDER financial support. Also J.G. López-Salas and C. Vázquez acknowledge the support received from the Centro de Investigación en Tecnologías de la Información y las Comunicaciones de Galicia, CITIC, as a center accredited for excellence within the Galician University System and a member of the CIGUS Network, receives subsidies from the Department of Education, Science, Universities, and Vocational Training of the Xunta de Galicia. Additionally, it is co-financed by the EU through the FEDER Galicia 2021-27 operational program ED431G 2023/01.

Also, the authors would like to thank the anonymous reviewers whose very appropriate remarks improved the article.

Appendix A. Matrices in problem (29)

In this Appendix, we detail the coefficients of the matrices of the semi-discretized initial value problem (29) and how the linear systems involved in the one-stage AMFR-W1-method (30) are solved.

In order to simplify this explanation, we denote $I^{(k)}$ as the identity matrix of dimension $M_k + 1$, $k = 1, \ldots, N$. By using the tensor product notation ($A \otimes B = (a_{ij}B)$), we define the matrices

$$\tilde{A}_1 = I^{(N)} \otimes I^{(N-1)} \otimes \cdots \otimes I^{(2)} \otimes A_1,$$
$$\tilde{A}_k^{(l)} = I^{(N)} \otimes \cdots \otimes I^{(k+1)} \otimes A_k^{(l)} \otimes I^{(k-1)} \otimes \cdots \otimes I^{(1)},$$
$$l = 1, 2, \; k = 2, \ldots, N.$$

The matrix $A_1 = ((A_1)_{ij})_{i,j=0}^{M}$ is a tridiagonal matrix of dimension $M_1 + 1$. The entries of its diagonals are

$$(A_1)_{j_1,j_1-1} = \begin{cases} \left(-\dfrac{1}{h_1}\dfrac{\tau_1}{1+\tau_1 x_{1,j_1}} + \dfrac{1}{h_1^2}\right)\dfrac{x_{1,j_1}^2}{2}, & \text{if } 1 \le j_1 \le M_1 - 1, \\ -\dfrac{1}{h_1}\dfrac{\tau_1}{1+\tau_1 x_{1,j_1}} x_{1,j_1}^2, & \text{if } j_1 = M_1, \end{cases}$$

$$(A_1)_{j_1,j_1} = \begin{cases} -\dfrac{1}{h_1^2} x_{1,j_1}^2, & \text{if } 0 \le j_1 \le M_1 - 1, \\ \dfrac{1}{h_1}\dfrac{\tau_1}{1+\tau_1 x_{1,j_1}} x_{1,j_1}^2, & \text{if } j_1 = M_1, \end{cases}$$

$$(A_1)_{j_1,j_1+1} = \left(\frac{1}{h_1}\frac{\tau_1}{1+\tau_1 x_{1,j_1}} + \frac{1}{h_1^2}\right)\frac{x_{1,j_1}^2}{2}, \quad \text{if } 0 \le j_1 \le M_1 - 1.$$

For each $k = 2, \ldots, N$, both matrices $A_k^{(1)}$ and $A_k^{(2)}$ are tridiagonal matrices of dimension $M_k + 1$. Their respective diagonals' elements are

$$(A_k^{(1)})_{j_k,j_k-1} = \left(\frac{\tau_k}{1+\tau_k x_{k,j_k}}\beta_{j_k,-1}^{(k)} + \frac{1}{2}\eta_{j_k,-1}^{(k)}\right)x_{k,j_k}^2, \quad \text{if } 1 \le j_k \le M_k,$$

$$(A_k^{(1)})_{j_k,j_k} = \left(\frac{\tau_k}{1+\tau_k x_{k,j_k}}\beta_{j_k,0}^{(k)} + \frac{1}{2}\eta_{j_k,0}^{(k)}\right)x_{k,j_k}^2, \quad \text{if } 0 \le j_k \le M_k,$$

$$(A_k^{(1)})_{j_k,j_k+1} = \left(\frac{\tau_k}{1+\tau_k x_{k,j_k}}\beta_{j_k,1}^{(k)} + \frac{1}{2}\eta_{j_k,1}^{(k)}\right)x_{k,j_k}^2, \quad \text{if } 0 \le j_k \le M_k - 1,$$

$$(A_k^{(2)})_{j_k,j_k-1} = \beta_{j_k,-1}^{(k)} x_{k,j_k}, \quad \text{if } 1 \le j_k \le M_k,$$

$$(A_k^{(2)})_{j_k,j_k} = \beta_{j_k,0}^{(k)} x_{k,j_k}, \quad \text{if } 0 \le j_k \le M_k,$$

$$(A_k^{(2)})_{j_k,j_k+1} = \beta_{j_k,1}^{(k)} x_{k,j_k}, \quad \text{if } 0 \le j_k \le M_k - 1,$$

where the finite difference coefficients $\beta_{j_k,\cdot}^{(k)}$ and $\eta_{j_k,\cdot}^{(k)}$ are given in (28). Moreover, in order to compute the Jacobian $\mathcal{A}_k$ for $k = 2, \ldots, N$, the diagonal matrix $\mathcal{D}_k(t) = \mathrm{diag}((d_k(t))_J)_{J=0}^{M-1}$ of dimension M is needed. Its entries are

$$(d_k(t))_J = \sum_{l=1}^{k-1} \lambda_l(t)\,\rho_{kl}\,\frac{\tau_l}{1+\tau_l x_{l,j_l}}\,x_{l,j_l}, \quad \text{with } (j_1, \ldots, j_N) = \vartheta^{-1}(J). \tag{A.1}$$

Note that the first row of all of these matrices is null, since $x_{k,0} = 0$, for all $k = 1, \ldots, N$.

As a consequence, the linear systems of the form

$$(I - \nu\tau\mathcal{A}_1(t_n))K = R$$

of dimension M in the AMFR-W1-method (30) are decoupled into $\prod_{k=2}^{N}(M_k + 1)$ systems of dimension $M_1 + 1$ with coefficient matrix $(I - \nu\tau\lambda_1^2(t_n)A_1)$.
Moreover, since the coefficients of the diagonal matrix $\mathcal{D}_k(t)$ defined in (A.1) only depend on the indices $j_1, \ldots, j_{k-1}$, the linear systems in (30) of the form

$$(I - \nu\tau\mathcal{A}_k(t_n))K = R, \quad \text{for } k = 2, \ldots, N,$$

can be decoupled into $\prod_{l \ne k}(M_l + 1)$ linear systems of dimension $M_k + 1$. More precisely, for each $k = 2, \ldots, N$, for each multi-index $(\ldots, j_{k-1}, j_{k+1}, \ldots)$ of $(N-1)$ integers with $0 \le j_r \le M_r$, $r \ne k$, the code computes

$$d_k(t_n) = \sum_{l=1}^{k-1} \lambda_l(t_n)\,\rho_{kl}\,\frac{\tau_l}{1+\tau_l x_{l,j_l}}\,x_{l,j_l},$$

and solves a linear system of dimension $M_k + 1$ with coefficient matrix

$$(I - \nu\tau(\lambda_k^2(t_n)A_k^{(1)} + \lambda_k(t_n)d_k(t_n)A_k^{(2)})).$$

References

[1] A. Lyashenko, F. Mercurio, LIBOR replacement: a modelling framework for in-arrears term rates, Risk (June 2019) 57–62.
[2] J.G. López-Salas, S. Pérez-Rodríguez, C. Vázquez, AMFR-W numerical methods for solving high dimensional SABR/LIBOR PDE models, SIAM J. Sci. Comput. 43 (2021) B30–B54.
[3] M. Fabrizi, X. Huan, A. Parbonetti, When LIBOR becomes LIEBOR: reputational penalties and bank contagion, Financ. Rev. 56 (1) (2021) 157–178.
[4] X. Huan, G.J. Previts, A. Parbonetti, Understanding the LIBOR scandal: the historical, the ethical, and the technological, J. Bank. Regul. 24 (2023) 403–419.
[5] P. McConnell, Systemic operational risk: the LIBOR manipulation scandal, J. Oper. Risk 8 (3) (2013) 59–99.
[6] P. Gandhi, B. Golez, J.C. Jackwerth, A. Plazzi, Financial market misconduct and public enforcement: the case of libor manipulation, Manag. Sci. 65 (11) (2013) 5268–5289.
[7] D. Duffie, Notes on LIBOR conversion, Available at https://www.darrellduffie.com/uploads/policy/Duffie-Conversion-Auction-Notes-2017.pdf.
[8] W. Huan, K. Todorov, The post-Libor world: a global view from the BIS derivatives statistics, BIS Q. Rev. (2022).
[9] A. Brace, D. Gatarek, M. Musiela, The market model of interest rate dynamics, Math. Finance 7 (2) (1997) 127–155.
[10] F. Jamshidian, LIBOR and swap market models and measures, Finance Stoch. 1 (1997) 293–330.
[11] D. Brigo, F. Mercurio, Interest Rate Models - Theory and Practice. With Smile, Inflation and Credit, 2nd edition, Springer, 2007.
[12] R. Pietersz, PDE pricing for BGM, Available at SSRN: https://ssrn.com/abstract=302266, February 28, 2002.
[13] A. Pascucci, M. Suárez, C. Vázquez, Mathematical analysis and numerical methods for a partial differential equation model governing a ratchet cap pricing problem, Math. Models Methods Appl. Sci. 21 (2011) 1479–1498.
[14] M. Suárez, C. Vázquez, A numerical method for pricing spread options on LIBOR rates with a PDE model, Math. Comput. Model. 52 (2010) 1074–1080.
[15] M. Suárez, C. Vázquez, Numerical solution of a PDE model for a rachet-cap pricing with BGM interest rate dynamic, Appl. Math. Comput. 218 (2012) 5217–5230.
[16] F. Mercurio, M. Morini, No-arbitrage dynamics for a tractable SABR term structure libor model, in: Modeling Interest Rates: Advances in Derivatives Pricing, Risk Books, 2009.
[17] R. Rebonato, R. White, Linking caplets and swaptions prices in the LMM-SABR model, J. Comput. Finance 13 (2) (2009) 19–45.
[18] R. Rebonato, The Q-measure dynamics of forward rates, Annu. Rev. Financ. Econ. 15 (2023) 493–522.
[19] A. Brace, Engineering BGM, 1st edition, Chapman & Hall/CRC Financial Mathematics Series, 2008.
[20] J.G. López-Salas, C. Vázquez, PDE formulation of some SABR/LIBOR market models and its numerical solution with a sparse grid combination technique, Comput. Math. Appl. 75 (2018) 1616–1634.
[21] A. Lyashenko, F. Mercurio, LIBOR replacement II: completing the generalized Forward Market Model, Risk (July 2020) 61–66.
[22] S. González-Pinto, E. Hairer, D. Hernández-Abreu, S. Pérez-Rodríguez, AMF-type W-methods for parabolic problems with mixed derivatives, SIAM J. Sci. Comput. 40 (5) (2018) A2905–A2929.
[23] S. González-Pinto, D. Hernández-Abreu, S. Pérez-Rodríguez, AMFR-W-methods for parabolic problems with mixed derivates. Applications to the Heston model, J. Comput. Appl. Math. 387 (2021) 1–20.
[24] T. Mikosh, Elementary Stochastic Calculus with Finance in View, first edition, World Scientific, 1998.
[25] L. Andresen, V. Piterbarg, Interest Rate Modeling, Atlantic Financial Press, 2010.
[26] P. Glasserman, X. Zhao, Arbitrage-free interpolation of lognormal forward LIBOR and swap rate models, Finance Stoch. 4 (2000) 35–68.
[27] B. Oksendal, Stochastic Differential Equations. An Introduction with Applications, 5th edition, Universitext, Springer, 1998.
[28] K. in 't Hout, Numerical Partial Differential Equations in Finance Explained, first edition, Palgrave Macmillan, London, 2017.
[29] P. Wilmott, J. Dewynne, S. Howison, Option Pricing. Mathematical Models and Computation, Oxford Financial Press, 1993.