CAPLET PRICING IN AFFINE MODELS FOR ALTERNATIVE RISK-FREE RATES

CLAUDIO FONTANA

Abstract. Alternative risk-free rates (RFRs) play a central role in the reform of interest rate benchmarks. We study a short-rate model for RFRs driven by a general affine process. In this context, under minimal assumptions, we derive explicit valuation formulas for forward-looking and backward-looking caplets/floorlets as well as term-basis caplets.

1. Introduction

The interest rate benchmarks reform is bringing a change of paradigm in fixed income markets. On 5 March 2021, the FCA announced that Libor rates will either cease to be provided or will no longer be representative benchmarks after 31 December 2021 (with the exception of US Libor rates for some tenors, that will be discontinued after June 2023)[1]. The new benchmark rates, as well as fallback rates for existing contracts, are provided by alternative nearly *risk-free rates* (RFRs), which are determined by overnight rates backed by actual transactions. Such overnight rates include SOFR in the US, SONIA in the UK, €STR in the Euro area.

The transition from Libor rates to RFRs has recently started to affect non-linear derivatives[2]. The trading volume in SOFR caps/floors reached 926.9 USD bn in the first 9 months of 2022, increasing from 85.6 USD bn in the whole year 2021 (source: ISDA). In the case of SONIA caps/floors, the trading volume in the first 9 months of 2022 amounts to 210.9 USD bn, against 72.5 USD bn in the whole year 2021 (source: ISDA).

In the post-Libor universe, one can consider *forward-looking* and *backward-looking* caps/floors, depending on the rate which defines their payoff. While forward-looking caps/floors are based on forward-looking term rates and are conceptually similar to classical Libor caps/floors, backward-looking caps/floors have a significantly different nature, since their payoff is determined by the compounded in-arrears RFR (see [LM19, Pit20]). This corresponds to a new type of payoff, which has Asian features and is fully determined only at the end of the accrual period. While both types of caps/floors coexist in the current post-Libor market, backward-looking caps/floors play a particularly important role. Indeed, forward-looking term rates (such as CME term SOFR and ICE term SOFR) have been introduced only recently and are not supported by the Alternative Reference Rates Committee (ARRC) for use in derivatives markets, being restricted to derivatives that hedge cash products referencing term SOFR. In addition, the Libor fallbacks protocol adopts backward-looking rates as fallback rates in existing contracts (see [ISD20]).

In this paper, we derive pricing formulas for forward-looking and backward-looking options in the context of a short-rate RFR model driven by a general affine process. While the valuation of forward-looking payoffs is relatively straightforward, the backward-looking case requires a more elaborate analysis and has never been considered in the literature on affine processes, except for the specific case of Gaussian Hull-White models. By relying on Fourier methods

Date: January 5, 2023.

Key words and phrases. Risk-free rate; Libor reform; backward-looking rate; affine process; Fourier pricing.

JEL classification: C02, C60, E43, G12, G13.

2020 Mathematics Subject Classification: 60J25, 91G15, 91G20, 91G30. The author is thankful to two anonymous reviewers for constructive comments and suggestions. Financial support from the Europlace Institute of Finance and the University of Padova (research programme BIRD190200/19) is gratefully acknowledged.

[1]See `https://www.fca.org.uk/news/press-releases/announcements-end-libor`.

[2]As part of the *SOFR First* initiative, the Market Risk Advisory Committee (MRAC) of the Commodity Futures Trading Commission recommended to switch from Libor to SOFR in non-linear derivatives starting from 8 November 2021 (see `https://www.cftc.gov/PressRoom/PressReleases/8449-21`). A similar recommendation has been issued by the working group on Sterling risk-free reference rates starting on 11 May 2021.

and a study of the integrability properties of certain functionals of the affine process, we obtain pricing formulas expressed as one-dimensional integrals, which can be efficiently implemented via fast Fourier transform (see [CM99]). We work under minimal technical assumptions, without imposing ad-hoc integrability requirements. We also study *term-basis* caplets, corresponding to exchange options between forward-looking and backward-looking rates (see [LM19]).

Short-rate modeling is probably the most natural approach for modeling RFRs. In [Mer18], one of the first papers on SOFR modeling, a Gaussian Hull-White short-rate model is adopted. In recent short-rate approaches to RFR modeling, the Hull-White model has remained dominant. This is for instance the case of [Hof20, Tur21, Xu22], where pricing formulae for backward-looking caplets are derived[3], and also of [RB21], where in addition collateralization and funding costs are taken into account. However, the Hull-White model does not support volatility smiles (see, e.g., [Pit20]). Moreover, as shown in [AB20], spikes and jumps are a prominent feature of RFRs and have a sizable effect on the pricing of backward-looking caplets. This motivates the modeling of RFRs by means of general affine processes and the study of their pricing aspects. Let us also mention that backward-looking caplets have been analyzed in the seminal work [LM19] in the context of an extended Libor market model, in [Wil20] for the SABR model and in [MS20] by adopting a rational model for the RFR savings account. A short-rate approach for modelling RFRs based on affine semimartingales has been developed in [FGS22b], also allowing for jumps at predetermined times, but without a specific focus on pricing applications.

The paper is structured as follows. In Section 2, we recall some essential notions on affine processes and prove two additional properties of the solutions to the associated Riccati ODEs. Section 3 contains the description of the modeling framework and the definition of forward/backward-looking rates, while all pricing results are derived in Section 4. In Section 5, we derive pricing formulas for forward-looking and backward-looking caplets in the context of a CIR++ model.

2. Preliminaries on affine processes

Let $(\Omega,\mathcal{F},P)$ be a probability space endowed with a filtration $\mathbb{F}=(\mathcal{F}_t)_{t\geq 0}$, satisfying the usual conditions of right-continuity and completeness, and $X=(X_t)_{t\geq 0}$ be a càdlàg adapted time-homogeneous conservative Markov process on $(\Omega,\mathcal{F},\mathbb{F},P)$, taking values in the state space $D:=\mathbb{R}^m_+\times\mathbb{R}^n$.[4] The family of the transition kernels of the Markov process X is given by $\{p_t : D\times\mathcal{B}_D\to[0,1]; t\geq 0\}$, where $\mathcal{B}_D$ denotes the Borel σ-algebra of D. We also introduce the set $\mathcal{U}:=\mathbb{C}^m_-\times \mathrm{i}\mathbb{R}^n$, where $\mathbb{C}_-:=\{u\in\mathbb{C} : \mathrm{Re}(u)\leq 0\}$. Setting $d:=m+n$, we recall the following definition (see [DFS03] and [KRM15, Definition 2.2]).

Definition 2.1. The process X is called *affine* with state space D if

(i) it is stochastically continuous (i.e., its transition kernels satisfy $\lim_{s\to t}p_s(x,\cdot)=p_t(x,\cdot)$ weakly for all $(t,x)\in\mathbb{R}_+\times D$) and

(ii) there exist functions $\phi:\mathbb{R}_+\times\mathcal{U}\to\mathbb{C}$ and $\psi:\mathbb{R}_+\times\mathcal{U}\to\mathbb{C}^d$ such that

$$\int_D e^{\langle u,\xi\rangle}p_t(x,\mathrm{d}\xi)=e^{\phi(t,u)+\langle\psi(t,u),x\rangle}, \tag{2.1}$$

for all $(t,u)\in\mathbb{R}_+\times\mathcal{U}$ and for every $x\in D$.

Requirement (i) in Definition 2.1 implies the regularity of the affine process X, meaning that the derivatives $\partial_t\phi(t,u)|_{t=0}$ and $\partial_t\psi(t,u)|_{t=0}$ exist for all $u\in\mathcal{U}$ and are continuous at $u=0$. As a consequence, in view of [DFS03, Theorem 2.7], the functions ϕ and ψ in (2.1) are determined

[3]We point out that, in the specific context of a Gaussian one-factor Hull-White model, a backward-looking caplet can be priced by relying on the valuation formulas obtained in the earlier work [Hen04].

[4]We restrict our attention to affine processes on $\mathbb{R}^m_+\times\mathbb{R}^n$ for simplicity of presentation. The modeling framework developed in this paper can be readily extended to matrix-valued affine processes, as characterized in [CFMT11], and all main results remain valid in the matrix-valued case with identical statements. In particular, this is possible since the results of [KRM15] are also applicable to affine processes taking values in the cone of symmetric positive-semidefinite matrices, up to an adaptation of the notation and of the proof of Lemma 2.6.

by the following system of generalized Riccati ODEs:

$$\frac{\partial\phi(t,u)}{\partial t} = F\big(\psi(t,u)\big), \qquad \phi(0,u)=0,$$
$$\frac{\partial\psi(t,u)}{\partial t} = R\big(\psi(t,u)\big), \qquad \psi(0,u)=u\in\mathcal{U},$$

with the functions F and R admitting explicit representations of Lévy-Khintchine type as follows:

$$\begin{aligned} F(u) &= \langle\alpha_0 u,u\rangle + \langle\beta_0,u\rangle + \int_{D\setminus\{0\}}\big(e^{\langle u,\xi\rangle}-1-\langle u,h(\xi)\rangle\big)\mu_0(\mathrm{d}\xi), \\ R_i(u) &:= \langle\alpha_i u,u\rangle + \langle\beta_i,u\rangle + \int_{D\setminus\{0\}}\big(e^{\langle u,\xi\rangle}-1-\langle u,h(\xi)\rangle\big)\mu_i(\mathrm{d}\xi), \qquad \text{for all } i=1,\ldots,d, \end{aligned} \tag{2.2}$$

with respect to a set of parameters (α,β,μ) satisfying the admissibility requirements of [DFS03, Definition 2.6], where μ_i, $i=0,1,\ldots,d$, are Lévy measures on D and h is a truncation function.

In the following, we shall work with time integrals of the affine process X. More specifically, let $\Lambda\in\mathbb{R}^d$ and consider the process $Y=(Y_t)_{t\geq0}$ defined by $Y_t:=\int_0^t\langle\Lambda,X_s\rangle\mathrm{d}s$, for all $t\geq0$. Note that, since X is càdlàg, the integral is well-defined pathwise. The couple (X,Y) can be regarded as a process on the enlarged state space $D\times\mathbb{R}$. The following well-known result, which is a direct consequence of [DFS03, Proposition 11.2], asserts that (X,Y) is an affine process.

Proposition 2.2. *Let the process (X,Y) be defined as above. Then, (X,Y) is an affine process on the state space $D\times\mathbb{R}$ and, for all $(t,u,v)\in\mathbb{R}_+\times\mathcal{U}\times\mathrm{i}\mathbb{R}$, it holds that*

$$\mathbb{E}\big[e^{\langle u,X_T\rangle+vY_T}\big|\mathcal{F}_t\big] = e^{\Phi(T-t,u,v)+\langle\Psi(T-t,u,v),X_t\rangle+vY_t}, \tag{2.3}$$

where $\Phi:\mathbb{R}_+\times\mathcal{U}\times\mathrm{i}\mathbb{R}\to\mathbb{C}$ and $\Psi:\mathbb{R}_+\times\mathcal{U}\times\mathrm{i}\mathbb{R}\to\mathbb{C}^d$ are solutions to

$$\frac{\partial\Phi(t,u,v)}{\partial t} = F\big(\Psi(t,u,v)\big), \qquad \Phi(0,u,v)=0, \tag{2.4a}$$

$$\frac{\partial\Psi(t,u,v)}{\partial t} = R\big(\Psi(t,u,v)\big)+v\Lambda, \qquad \Psi(0,u,v)=u. \tag{2.4b}$$

Remark 2.3. For a generic vector $x\in\mathbb{C}^d$, let us introduce the notation $x=(x_I,x_J)\in\mathbb{C}^m\times\mathbb{C}^n$. Writing $\Psi(t,u,v)=(\Psi_I(t,u,v),\Psi_J(t,u,v))$, it can be seen that $\Psi_J(t,u,v)$ solves a linear ODE with initial value u_J and is therefore globally well-defined on $\mathbb{R}_+$ (see [DFS03, Section 11.2]). In particular, $\Psi_J(t,u,v)$ is linear in (u_J,v) and does not depend on u_I. Note also that the function $\Phi(\cdot,u,v)$ can be explicitly solved as $\Phi(t,u,v)=\int_0^t F(\Psi(s,u,v))\mathrm{d}s$, for all $(t,u,v)\in\mathbb{R}_+\times\mathcal{U}\times\mathrm{i}\mathbb{R}$.

For pricing applications, we need to extend the domain of the affine transform formula (2.3) of the joint process (X,Y) beyond the set $\mathcal{U}\times\mathrm{i}\mathbb{R}$. This issue has been studied in detail in [KRM15], whose approach is followed here. As a preliminary, noting that the admissibility requirements of [DFS03, Definition 2.6] imply that $\mu_j=0$, for all $j=m+1,\ldots,n$, let us define the set

$$\mathcal{Y} := \left\{y\in\mathbb{R}^d : \sum_{i=0}^m\int_{\{\xi\in D:|\xi|\geq1\}}e^{\langle y,\xi\rangle}\mu_i(\mathrm{d}\xi)<+\infty\right\}.$$

The set $\mathcal{Y}$ is non-empty and convex and constitutes the effective real domain of the functions F and R given in (2.2), which are therefore well-defined convex functions on $\mathcal{Y}$.

Assumption 2.4. It holds that $0\in\mathcal{Y}^\circ$, with $\mathcal{Y}^\circ$ denoting the interior of $\mathcal{Y}$.

Noting that $0\in\mathcal{Y}$ always holds, Assumption 2.4 represents a mild technical requirement that can be easily checked, since the measures μ_i, $i=0,1,\ldots,m$, are explicitly known in applications. In particular, Assumption 2.4 is trivially satisfied by every continuous affine process, since in that case $\mathcal{Y}=\mathbb{R}^d$. Under Assumption 2.4, the set $S(\mathcal{Y}^\circ):=\{u\in\mathbb{C}^d:\mathrm{Re}(u)\in\mathcal{Y}^\circ\}$ is non-empty and the functions F and R can be analytically extended to $S(\mathcal{Y}^\circ)$. This enables us to study the Riccati ODEs (2.4a)-(2.4b) for $(u,v)\in S(\mathcal{Y}^\circ)\times\mathbb{C}$, replacing F and R by their analytic extensions to $S(\mathcal{Y}^\circ)$. Since F and R are locally Lipschitz on $S(\mathcal{Y}^\circ)$, the (possibly local) solution (Φ,Ψ) to (2.4a)-(2.4b) is unique if constrained to stay in the open domain $S(\mathcal{Y}^\circ)$. For $(u,v)\in S(\mathcal{Y}^\circ)\times\mathbb{C}$, let us denote by $T_+(u,v)$ the maximal lifetime of the solution to the Riccati system (2.4a)-(2.4b) such that $\Psi(t,\mathrm{Re}(u),\mathrm{Re}(v))\in\mathcal{Y}^\circ$ for all $t\leq T_+(u,v)$.

The following proposition is a direct consequence of [KRM15, Theorems 2.14 and 2.26] and provides the desired extension of the affine transform formula (2.3).

Proposition 2.5. *Suppose that Assumption 2.4 holds. Then, the following hold:*
(i) for $(u,v)\in\mathcal{Y}^\circ\times\mathbb{R}$, it holds that $\mathbb{E}[e^{\langle u,X_T\rangle+vY_T}]<+\infty$ for all $T\leq T_+(u,v)$;
(ii) for $(u,v)\in S(\mathcal{Y}^\circ)\times\mathbb{C}$, it holds that

$$\mathbb{E}\big[e^{\langle u,X_T\rangle+vY_T}\big|\mathcal{F}_t\big]=e^{\Phi(T-t,u,v)+\langle\Psi(T-t,u,v),X_t\rangle+vY_t}, \tag{2.5}$$

for all $0\leq t\leq T<+\infty$ such that $T-t\leq T_+(u,v)$, where $\Phi(\cdot,u,v)$ and $\Psi(\cdot,u,v)$ solve (2.4a)-(2.4b) *with F and R analytically extended to $S(\mathcal{Y}^\circ)$.*

For convenience of notation, similarly as in [HKRS17], let us introduce the sets

$$\mathcal{Y}_t:=\big\{(u,v)\in\mathcal{Y}^\circ\times\mathbb{R}:T_+(u,v)>t\big\},\qquad\mathcal{D}_t:=S(\mathcal{Y}_t)=\big\{(u,v)\in\mathbb{C}^{d+1}:(\mathrm{Re}(u),\mathrm{Re}(v))\in\mathcal{Y}_t\big\},$$

for $t\geq0$. The sets $\mathcal{Y}_t$ and $\mathcal{D}_t$ are open and the results of [KRMS10] imply that $\mathcal{Y}_t$ is convex. Moreover, the affine transform formula (2.5) holds for all $(u,v)\in\mathcal{D}_{T-t}$.

In the following, we denote by $\preceq$ the partial order on $\mathbb{R}^m_+$ (i.e., for any $y,u\in\mathbb{R}^m_+$, the relation $y\preceq u$ means that $u-y\in\mathbb{R}^m_+$) and recall the notation $x=(x_I,x_J)$, for any $x\in D=\mathbb{R}^m_+\times\mathbb{R}^n$. It is easily seen that the set $\mathcal{Y}$ is order-preserving. In the following lemma, we show that this property extends to the set $\mathcal{Y}_t$, for all $t\geq0$.

Lemma 2.6. *Let $(u,v)\in\mathcal{Y}_t$, for some $t\geq0$. If $(y,z)\in D\times\mathbb{R}$ is such that $y_I\preceq u_I$, $y_J=u_J$ and $z=v$, then $(y,z)\in\mathcal{Y}_t$.*

Proof. By [KRM15, Lemma 5.8], it holds that $y\in\mathcal{Y}^\circ$. Since X_t takes values in D, it holds that

$$\mathbb{E}\big[e^{\langle y,X_t\rangle+zY_t}\big]=\mathbb{E}\big[e^{\langle y_I,X_{I,t}\rangle+\langle y_J,X_{J,t}\rangle+zY_t}\big]\leq\mathbb{E}\big[e^{\langle u_I,X_{I,t}\rangle+\langle u_J,X_{J,t}\rangle+vY_t}\big]=\mathbb{E}\big[e^{\langle u,X_t\rangle+vY_t}\big]<+\infty,$$

due to the assumption that $(u,v)\in\mathcal{Y}_t$ together with part (i) of Proposition 2.5. By [KRM15, Theorem 2.14-(a)], this implies the existence of a solution $(\Phi(\cdot,y,z),\Psi(\cdot,y,z))$ to (2.4a)-(2.4b) up to time t. It remains to prove that $\Psi(s,y,z)\in\mathcal{Y}^\circ$ for all $s\leq t$. To this effect, note first that $\Psi_J(s,u,v)=\Psi_J(s,y,z)$, for all $s\geq0$, since $\Psi_J(\cdot,u,v)$ does not depend on u_I (see Remark 2.3) and $(u_J,v)=(y_J,z)$. Hence, considering the first m components of (2.4b), we have that

$$\frac{\partial}{\partial s}\Psi_I(s,y,z)=R_I\big(\Psi_I(s,y,z),\Psi_J(s,u,v)\big)+v\Lambda_I,\qquad\Psi_I(0,y,z)=y_I.$$

By [KRM15, Lemma 5.7], the map $x_I\to R_I(x_I,\Psi_J(s,u,v))$ is quasi-monotone increasing with respect to the natural cone $\mathbb{R}^m_+$, for all $s\geq0$. By classical comparison results for ODEs (see, e.g., [KRMS10, Lemma 3]), $\Psi_I(s,y,z)\preceq\Psi_I(s,u,v)$, for all $s<T_+(y,z)\wedge T_+(u,v)$. Arguing by contradiction, suppose that $T_+(y,z)<T_+(u,v)$. Since we already know that $\Psi(\cdot,y,z)$ cannot explode before time t, this means that $\Psi(\cdot,y,z)$ reaches the boundary of $\mathcal{Y}$ before time $T_+(u,v)$. However, continuity implies that $\Psi_I(T_+(y,z),y,z)\preceq\Psi_I(T_+(y,z),u,v)$ and, since $T_+(y,z)<T_+(u,v)$, we have that $\Psi(T_+(y,z),u,v)\in\mathcal{Y}^\circ$. Again by [KRM15, Lemma 5.8], this implies that $\Psi(T_+(y,z),y,z)\in\mathcal{Y}^\circ$, thus obtaining a contradiction. Therefore, $T_+(y,z)\geq T_+(u,v)$ must necessarily hold. Since $T_+(u,v)>t$ by assumption, this shows that $(y,z)\in\mathcal{Y}_t$. □

We close this section with the following result on the convexity of the function Ψ_I. The following lemma can be easily proved by relying on the techniques of [KRMS10].

Lemma 2.7. *Let $(u_1,v_1),(u_2,v_2)\in\mathcal{Y}_t$, for some $t\geq0$, and define $(u_\lambda,v_\lambda):=\lambda(u_1,v_1)+(1-\lambda)(u_2,v_2)$, for $\lambda\in[0,1]$. Then it holds that $\Psi_I(s,u_\lambda,v_\lambda)\preceq\lambda\Psi_I(s,u_1,v_1)+(1-\lambda)\Psi_I(s,u_2,v_2)$, for all $s\in[0,t]$.*

Remark 2.8. For later use, we recall that the functions $\Phi(\cdot,u,v)$ and $\Psi(\cdot,u,v)$ appearing in Proposition 2.5 satisfy the following semiflow relations, for all $(u,v)\in\mathcal{Y}^\circ\times\mathbb{R}$:

$$\begin{aligned}\Phi(s+t,u,v)&=\Phi(t,u,v)+\Phi(s,\Psi(t,u,v),v),\\ \Psi(s+t,u,v)&=\Psi(s,\Psi(t,u,v),v),\end{aligned}\tag{2.6}$$

for all $0\leq s\leq t<+\infty$ such that $s+t\leq T_+(u,v)$ (see [KRM15, Lemma 4.3]).

3. Model setup

In this section, we present the general affine modeling framework for a risk-free overnight rate. Let $X = (X_t)_{t\geq 0}$ be an affine process on $(\Omega, \mathcal{F}, \mathbb{F}, P)$ taking values in the space $D = \mathbb{R}^m_+ \times \mathbb{R}^n$. Adopting a short rate approach, we model the instantaneous RFR process $r = (r_t)_{t\geq 0}$ as follows:

$$r_t := \ell(t) + \langle \Lambda, X_t \rangle, \qquad \text{for all } t \geq 0, \tag{3.1}$$

where $\Lambda \in \mathbb{R}^d$, with $d = m+n$, and $\ell : \mathbb{R}_+ \to \mathbb{R}$ is a function satisfying $\int_0^T |\ell(t)|\mathrm{d}t < +\infty$, for all $T > 0$. In line with [BM01], the function ℓ in (3.1) serves to fit the RFR term structure at $t = 0$. We assume that r also represents the collateral rate for collateralized OTC transactions, as well as the price alignment interest (PAI) for cleared derivatives. As pointed out by [RB21], this assumption is consistent with the collateralization schemes currently prevailing in the market. Adopting a martingale approach, we can therefore assume that P is a pricing measure with respect to the savings account $B = (B_t)_{t\geq 0}$ given by $B_t := \exp(\int_0^t r_s \mathrm{d}s)$, $t \geq 0$. Denoting by $P_t(T)$ the price at time t of a zero-coupon bond (ZCB) with maturity T, it holds that

$$P_t(T) = \mathbb{E}[B_t/B_T|\mathcal{F}_t], \qquad \text{for all } 0 \leq t \leq T < +\infty. \tag{3.2}$$

We denote $Y := \int_0^{\cdot} \langle \Lambda, X_s \rangle \mathrm{d}s$, similarly as in Section 2, and introduce the following assumption.

Assumption 3.1. For every $T > 0$, it holds that $(0, -1) \in \mathcal{Y}_T$.

In view of Proposition 2.5, Assumption 3.1 ensures that ZCB prices are well-defined by (3.2) for all maturities $T > 0$. Note also that Assumption 3.1 implies the validity of Assumption 2.4. In the following, we always suppose that Assumption 3.1 is satisfied without further mention.

In line with the terminology of [LM19], we give the following definition, denoting by $\mathbb{E}^T$ the expectation under the T-forward measure P^T defined by $\mathrm{d}P^T/\mathrm{d}P := 1/(B_T P_0(T))$, for $T > 0$.

Definition 3.2. For all $0 \leq S < T < +\infty$, we define:

(i) the *backward-looking rate* $R(S,T) := (B_T/B_S - 1)/(T - S)$;

(ii) the *forward-looking rate* $F(S,T) := \mathbb{E}^T[R(S,T)|\mathcal{F}_S]$.

We point out that part (ii) of Definition 3.2 is well-posed since ZCB prices are well-defined for all maturities. More specifically, for all $0 \leq S < T < +\infty$, it holds that

$$1 + (T - S)F(S,T) = \mathbb{E}^T[B_T/B_S|\mathcal{F}_S] = 1/P_S(T). \tag{3.3}$$

The fundamental difference between forward-looking and backward-looking rates consists in the fact that the forward-looking rate $F(S,T)$ is observable at time S, while the backward-looking rate $R(S,T)$ is only known at the end of the accrual period $[S,T]$. According to Definition 3.2, the forward-looking rate $F(S,T)$ represents the fixed rate K that makes equal to zero the value at time S of a swaplet (overnight indexed swap) delivering payoff $R(S,T) - K$ at maturity T.

For brevity of notation, we introduce the following shorthand notation, for $(u,v) \in S(\mathcal{Y}^\circ) \times \mathbb{C}$:

$$A^0(t,t+\tau,v) := \Phi(\tau,0,-v) - vL(t,t+\tau), \quad B^0(\tau,v) := \Psi(\tau,0,-v), \qquad \text{for all } \tau \leq T_+(0,-v),$$

$$A^1(t,t+\tau,u) := \Phi(\tau,u,-1) - L(t,t+\tau), \quad B^1(\tau,u) := \Psi(\tau,u,-1), \qquad \text{for all } \tau \leq T_+(u,-1),$$

where $L(t,t+\tau) := \int_t^{t+\tau} \ell(z)\mathrm{d}z$ and (Φ, Ψ) solves (2.4a)-(2.4b). Making use of this notation and in view of Proposition 2.5 and Assumption 3.1, ZCB prices can be expressed as follows:

$$P_t(T) = e^{A^0(t,T,1) + \langle B^0(T-t,1), X_t \rangle}, \qquad \text{for all } 0 \leq t \leq T < +\infty. \tag{3.4}$$

Remark 3.3. In reality, the backward-looking rate is computed by compounding at a daily frequency the overnight rate over the period $[S,T]$, corresponding to the following quantity:

$$R'(S,T) := \frac{1}{T-S}\left(\prod_{i=1}^n \frac{1}{P_{t_i}(t_i + \delta_i)} - 1\right), \tag{3.5}$$

where the product is taken over the business days $(t_1, \ldots, t_n)$ comprised between S and T, with δ_i denoting the day-count fraction, for $i = 1, \ldots, n$. In our model setup, it holds that

$$R'(S,T) = \frac{1}{T-S}\Big(e^{-\sum_{i=1}^n A^0(t_i,t_i+\delta_i,1) - \sum_{i=1}^n \langle B^0(\delta_i,1), X_{t_i} \rangle} - 1\Big). \tag{3.6}$$

The specification of the backward-looking rate $R(S,T)$ adopted in Definition 3.2 corresponds to the usual continuous-time approximation of (3.5) (see, e.g., [LM19]) and is adopted for simplicity of computation only. The approximation error can be quantified exactly by relying on formula (3.6) and can be shown to be practically negligible (see, e.g., [SS21, Appendix D]).

Remark 3.4. The modeling framework can be extended to time-inhomogeneous affine processes, with time-dependent parameters (α, β, μ) in (2.2). All pricing results derived in next section can be easily generalized to this case, provided that suitable exponential moments exists. For time-inhomogeneous affine processes, under an additional integrability condition on the jump measure, it has been shown in [KMK10, Theorem 5.1] that formula (2.5) holds for real (u,v) if there exists a solution to the Riccati equations (2.4a)-(2.4b) up to time T. However, while in the case of time-homogeneous affine processes we are able to characterize and provide conditions for the existence of solutions to the Riccati equations up to a certain time, for time-inhomogeneous affine processes one has to assume it a priori, similarly as in [KMK10, Section 5].

4. Pricing of options on forward and backward-looking rates

As explained in Section 1, one can consider two distinct types of caplets/floorlets, depending on whether the payoff is determined by the forward-looking or the backward-looking rate.

Definition 4.1. For $0 \leq S < T < +\infty$ and $K > 0$,

(i) a *forward-looking caplet* is defined by the payoff $(T-S)(F(S,T)-K)^+$ at maturity T;
(ii) a *backward-looking caplet* is defined by the payoff $(T-S)(R(S,T)-K)^+$ at maturity T.

Forward-looking and backward-looking floorlets are defined in an analogous way. Since the reference probability P is assumed to be a pricing measure with respect to the numéraire B, arbitrage-free prices of forward-looking and backward-looking caplets can be expressed as follows:

$$
\begin{aligned}
\pi_t^{c,\mathrm{F}}(S,T,K) &:= (T-S)\,\mathbb{E}\left[\frac{B_t}{B_T}(F(S,T)-K)^+ \Big| \mathcal{F}_t\right], \\
\pi_t^{c,\mathrm{B}}(S,T,K) &:= (T-S)\,\mathbb{E}\left[\frac{B_t}{B_T}(R(S,T)-K)^+ \Big| \mathcal{F}_t\right].
\end{aligned}
\tag{4.1}
$$

The prices of forward-looking and backward-looking floorlets can be expressed in a similar way and shall be denoted by $\pi_t^{f,\mathrm{F}}(S,T,K)$ and $\pi_t^{f,\mathrm{B}}(S,T,K)$, respectively.

Remark 4.2. (1) As pointed out by several authors (see [LM19, MS20, Pit20]), Definition 3.2 together with Jensen's inequality implies the following relation between the two types of caplets:

$$
\begin{aligned}
\pi_t^{c,\mathrm{F}}(S,T,K) &= (T-S)P_t(T)\mathbb{E}^T\left[(F(S,T)-K)^+ \big| \mathcal{F}_t\right] \\
&\leq (T-S)P_t(T)\mathbb{E}^T\left[(R(S,T)-K)^+ \big| \mathcal{F}_t\right] = \pi_t^{c,\mathrm{B}}(S,T,K),
\end{aligned}
$$

with the same inequality holding in the case of floorlets.

(2) Always as a consequence of Definition 3.2, it holds that

$$
\pi_t^{c,\mathrm{B}}(S,T,K) - \pi_t^{f,\mathrm{B}}(S,T,K) = \pi_t^{c,\mathrm{F}}(S,T,K) - \pi_t^{f,\mathrm{F}}(S,T,K), \qquad \text{for all } t \in [0,S]. \tag{4.2}
$$

This relation implies that, before the beginning of the accrual period, forward-looking and backward-looking caplets/floorlets satisfy the same put-call parity relation.

Remark 4.3. According to the prevailing ISDA protocol (see [ISD20]), the primary fallback for derivatives based on Libor rates is the compounded overnight risk-free rate with the addition of a *credit adjustment spread*. In the case of Libor caplets, this means that a tenor-dependent spread $c(T-S)$ has to be added to the rate in (4.1). For all combinations of tenors and currencies, the values of the credit adjustment spread have been fixed on 5 March 2021 on the basis of historical data[5]. Therefore, since the spread $c(T-S)$ is predetermined, the valuation of caplets/floorlets under the ISDA fallback provisions reduces to formulae (4.1), replacing K with $K - c(T-S)$.

[5] See `https://www.isda.org/2021/03/05/isda-statement-on-uk-fca-libor-announcement`.

Until the end of this section, let us consider fixed but arbitrary $0 \leq S < T < +\infty$, $K > 0$. For brevity of notation, we denote

$$K' := 1 + (T-S)K.$$

Moreover, we define

$$h^K(w,\lambda) := \frac{1}{2\pi}\frac{(1+(T-S)K)^{w+\mathrm{i}\lambda}}{(w+\mathrm{i}\lambda)(w-1+\mathrm{i}\lambda)}, \qquad \text{for all } w \in \mathbb{R}\setminus\{0,1\} \text{ and } \lambda \in \mathbb{R}.$$

Note that the function $\lambda \mapsto h^K(w,\lambda)$ is integrable on $\mathbb{R}$, for every $w \in \mathbb{R}\setminus\{0,1\}$.

We start our analysis from forward-looking caplets/floorlets, by relying on a Fourier decomposition of their payoff. Since forward-looking caplets/floorlets can be reduced to put/call options on ZCBs, the proof follows the structure of [Fil09, Corollary 10.4] and is therefore omitted.

Theorem 4.4. *There exist constants $w_- < 0$ and $w_+ > 1$ such that the prices of forward-looking caplets and floorlets can be expressed respectively as follows, for all $w \in (w_-, w_+)\setminus\{0,1\}$:*

$$\pi_t^{c,\mathrm{F}}(S,T,K) = \begin{cases} \Pi_t^{\mathrm{F}}(w), & \text{for } w \in (w_-,0), \\ \Pi_t^{\mathrm{F}}(w) + P_t(S), & \text{for } w \in (0,1), \end{cases} \tag{4.3}$$

$$\pi_t^{f,\mathrm{F}}(S,T,K) = \begin{cases} \Pi_t^{\mathrm{F}}(w), & \text{for } w \in (1,w_+), \\ \Pi_t^{\mathrm{F}}(w) + K'P_t(T), & \text{for } w \in (0,1), \end{cases} \tag{4.4}$$

for all $t \in [0,S]$, where

$$\Pi_t^{\mathrm{F}}(w) := \int_{\mathbb{R}} e^{(w+\mathrm{i}\lambda)A^0(S,T,1)+A^1(t,S,(w+\mathrm{i}\lambda)B^0(T-S,1))+\langle B^1(S-t,(w+\mathrm{i}\lambda)B^0(T-S,1)),X_t\rangle} h^K(w,\lambda)\mathrm{d}\lambda.$$

The valuation of backward-looking caplets/floorlets requires a more delicate analysis, mainly as a consequence of the different measurability properties of their payoffs. In particular, we shall make use of the following integrability property of a functional of the affine process.

Lemma 4.5. *For every $0 \leq S \leq T < +\infty$, there exist constants $\eta_1, \eta_2 \in (0,1)$ such that*

$$\mathbb{E}\big[e^{-(1+\eta_1)Y_S+\eta_1(Y_T-Y_S)}\big] < +\infty \qquad \text{and} \qquad \mathbb{E}\big[e^{-(1-\eta_2)Y_S-(1+\eta_2)(Y_T-Y_S)}\big] < +\infty. \tag{4.5}$$

Moreover, the constants η_1 and η_2 do not depend on S.

Proof. Since $\mathcal{Y}_T$ is open and $(0,-1) \in \mathcal{Y}_T$, due to Assumption 3.1, there exists $\varepsilon > 0$ such that $(0,-1-\varepsilon) \in \mathcal{Y}_T$. Moreover, since $(0,0) \in \mathcal{Y}_T$ always holds, we can assume that $(0,\varepsilon) \in \mathcal{Y}_T$, for ε small enough. Taking $\eta_1 = \varepsilon^2/(1+3\varepsilon)$, Hölder's inequality with $p = (1+3\varepsilon)/(1+2\varepsilon)$ and $q = p/(p-1)$ implies that

$$\begin{aligned}\mathbb{E}\big[e^{-(1+\eta_1)Y_S+\eta_1(Y_T-Y_S)}\big] &= \mathbb{E}\big[e^{-(1+2\eta_1)Y_S+\eta_1 Y_T}\big] \\ &\leq \mathbb{E}\big[e^{-p(1+2\eta_1)Y_S}\big]^{\frac{1}{p}}\mathbb{E}\big[e^{q\eta_1 Y_T}\big]^{\frac{1}{q}} = \mathbb{E}\big[e^{-(1+\varepsilon)Y_S}\big]^{\frac{1}{p}}\mathbb{E}\big[e^{\varepsilon Y_T}\big]^{\frac{1}{p}} < +\infty,\end{aligned}$$

as a consequence of part (i) of Proposition 2.5, together with the fact that $\mathcal{Y}_S \subseteq \mathcal{Y}_T$, for every $S \in [0,T]$. Similarly, taking $\eta_2 = \varepsilon^2/(2+3\varepsilon)$, $p = (2+3\varepsilon)/(2\varepsilon)$ and $q = p/(p-1)$, we have that

$$\begin{aligned}\mathbb{E}\big[e^{-(1-\eta_2)Y_S-(1+\eta_2)(Y_T-Y_S)}\big] &= \mathbb{E}\big[e^{2\eta_2 Y_S-(1+\eta_2)Y_T}\big] \\ &\leq \mathbb{E}\big[e^{2p\eta_2 Y_S}\big]^{\frac{1}{p}}\mathbb{E}\big[e^{-q(1+\eta_2)Y_T}\big]^{\frac{1}{q}} = \mathbb{E}\big[e^{\varepsilon Y_S}\big]^{\frac{1}{p}}\mathbb{E}\big[e^{-(1+\varepsilon)Y_T}\big]^{\frac{1}{p}} < +\infty,\end{aligned}$$

again by part (i) of Proposition 2.5. □

Remark 4.6. In view of Lemma 4.5, the moment generating function $\mathbb{E}[\exp(uY_T)]$ is well-defined for all $u \in [-(1+\eta_2),\eta_1]$, for suitable constants η_1 and η_2 (depending on T). This avoids the potential explosion of the moment generating function, which is a well-known phenomenon for some non-Gaussian affine processes (see, e.g., [AP07, Corollary 3.3]). The constants η_1 and η_2 can be explicitly determined by relying on [FGS22a, Theorem 3.4] when X is a CBI process.

By relying on Lemma 4.5, we are now in a position to state the following theorem, which provides a general Fourier-based pricing formula for backward-looking caplets/floorlets. In the next theorem, we consider the case of valuation before the accrual period (i.e., for $t \leq S$), referring to Remark 4.8 for the valuation inside the accrual period (i.e., for $t \in [S,T]$).

Theorem 4.7. *There exist constants $w_- < 0$ and $w_+ > 1$ such that the prices of backward-looking caplets and floorlets can be expressed respectively as follows, for all $w \in (w_-, w_+) \setminus \{0,1\}$:*

$$\pi_t^{c,\mathrm{B}}(S,T,K) = \begin{cases} \Pi_t^{\mathrm{B}}(w), & \text{for } w \in (w_-, 0), \\ \Pi_t^{\mathrm{B}}(w) + P_t(S), & \text{for } w \in (0,1), \end{cases} \tag{4.6}$$

$$\pi_t^{f,\mathrm{B}}(S,T,K) = \begin{cases} \Pi_t^{\mathrm{B}}(w), & \text{for } w \in (1, w_+), \\ \Pi_t^{\mathrm{B}}(w) + K' P_t(T), & \text{for } w \in (0,1), \end{cases} \tag{4.7}$$

for all $t \in [0,S]$, where

$$\Pi_t^{\mathrm{B}}(w) := \int_{\mathbb{R}} e^{A^0(S,T,w+\mathrm{i}\lambda)+A^1(t,S,B^0(T-S,w+\mathrm{i}\lambda))+\langle B^1(S-t,B^0(T-S,w+\mathrm{i}\lambda)),X_t\rangle} h^K(w,\lambda)\mathrm{d}\lambda.$$

Proof. We first show that there exist constants $w_- < 0$ and $w_+ > 1$ such that

$$(B^0(T-S,w_-),-1) \in \mathcal{Y}_S \qquad \text{and} \qquad (B^0(T-S,w_+),-1) \in \mathcal{Y}_S. \tag{4.8}$$

By (4.5) together with [KRM15, Theorem 2.14-(a)], the ODEs (2.4a)-(2.4b) admit a solution up to time S starting from $(u,v) = (\Psi(T-S,0,\eta_1), -(1+\eta_1))$ (however, note that it may happen that $(\Psi(T-S,0,\eta_1), -(1+\eta_1)) \in \partial\mathcal{Y}_S$). In addition, since $(0,\eta_1) \in \mathcal{Y}_T$ by Lemma 4.5, the semiflow property (2.6) implies that $(\Psi(T-S,0,\eta_1),\eta_1) \in \mathcal{Y}_S$. Let $\gamma := \eta_1/(1+2\eta_1) \in (0,1)$. Noting that $\gamma\eta_1 - (1-\gamma)(1+\eta_1) = -1$, the convexity of $\mathcal{Y}_S$ implies that

$$(\Psi(T-S,0,\eta_1),-1) = \gamma(\Psi(T-S,0,\eta_1),\eta_1) + (1-\gamma)(\Psi(T-S,0,\eta_1),-(1+\eta_1)) \in \mathcal{Y}_S.$$

Letting $w_- := -\eta_1$ proves the first property in (4.8). Arguing in a similar way, (4.5) implies the existence of a solution up to time S of the ODEs (2.4a)-(2.4b) starting from $(u,v) = (\Psi(T-S,0,-(1+\eta_2)),-(1-\eta_2))$. Moreover, it holds that $(\Psi(T-S,0,-(1+\eta_2)),-(1+\eta_2)) \in \mathcal{Y}_S$ and convexity of $\mathcal{Y}_S$ then implies that $(\Psi(T-S,0,-(1+\eta_2)),-1) \in \mathcal{Y}_S$, thus proving the second property in (4.8) with $w_+ := 1+\eta_2$. Moreover, Lemma 2.6 and Lemma 2.7 together imply that $(B^0(T-S,w),-1) \in \mathcal{Y}_S$ for all $w \in (w_-,w_+)$. By Proposition 2.5, it follows that

$$\begin{aligned} \mathbb{E}\left[\frac{1}{B_S}\int_{\mathbb{R}}\left|\left(\frac{B_S}{B_T}\right)^{w+\mathrm{i}\lambda} h^K(w,\lambda)\right|\mathrm{d}\lambda\right] &\leq \mathbb{E}\left[\frac{1}{B_S}\left(\frac{B_S}{B_T}\right)^{w}\right]\int_{\mathbb{R}}|h^K(w,\lambda)|\mathrm{d}\lambda \\ &= \mathbb{E}\left[e^{-Y_S-L(0,S)+A^0(S,T,w)+\langle B^0(T-S,w),X_S\rangle}\right]\int_{\mathbb{R}}|h^K(w,\lambda)|\mathrm{d}\lambda < +\infty, \end{aligned} \tag{4.9}$$

for all $w \in (w_-,w_+) \setminus \{0,1\}$. We can therefore apply Fubini's theorem and obtain

$$\begin{aligned} \mathbb{E}\left[\frac{B_t}{B_S}\int_{\mathbb{R}}\left(\frac{B_S}{B_T}\right)^{w+\mathrm{i}\lambda} h^K(w,\lambda)\mathrm{d}\lambda\middle|\mathcal{F}_t\right] &= \int_{\mathbb{R}}\mathbb{E}\left[\frac{B_t}{B_S}\left(\frac{B_S}{B_T}\right)^{w+\mathrm{i}\lambda}\middle|\mathcal{F}_t\right] h^K(w,\lambda)\mathrm{d}\lambda \\ &= \int_{\mathbb{R}}\mathbb{E}\left[e^{-(Y_S-Y_t)-L(t,S)+A^0(S,T,w+\mathrm{i}\lambda)+\langle B^0(T-S,w+\mathrm{i}\lambda),X_S\rangle}\middle|\mathcal{F}_t\right] h^K(w,\lambda)\mathrm{d}\lambda = \Pi_t^{\mathrm{B}}(w), \end{aligned} \tag{4.10}$$

where the second equality follows from formula (2.5) using that $(0,-(w+\mathrm{i}\lambda)) \in \mathcal{D}_{T-S} \subseteq \mathcal{D}_T$, for every $w \in (w_-,w_+)$, which is in turn a consequence of $(0,-w_-) \in \mathcal{Y}_T$ and $(0,-w_+) \in \mathcal{Y}_T$ together with the convexity of $\mathcal{Y}_T$. To justify the last equality in (4.10), note that, in view of Remark 2.3 and arguing as in the proof of [KRM15, Proposition 5.1], it holds that

$$\mathrm{Re}(B_I^0(T-S,w+\mathrm{i}\lambda)) \preceq B_I^0(T-S,w) \qquad \text{and} \qquad \mathrm{Re}(B_J^0(T-S,w+\mathrm{i}\lambda)) = B_J^0(T-S,w).$$

Recalling that $(B^0(T-S,w),-1) \in \mathcal{Y}_S$ for all $w \in (w_-,w_+)$, Lemma 2.6 therefore implies that $(\mathrm{Re}(B^0(T-S,w+\mathrm{i}\lambda)),-1) \in \mathcal{Y}_S$. The last equality in (4.10) then follows by applying formula (2.5) with $(u,v) = (B^0(T-S,w+\mathrm{i}\lambda),-1)$. Formula (4.6) follows by noting that

$$\pi_t^{c,\mathrm{B}}(S,T,K) = \mathbb{E}\left[\frac{B_t}{B_T}\left(\frac{B_T}{B_S}-K'\right)^+\middle|\mathcal{F}_t\right] = \mathbb{E}\left[\frac{B_t}{B_S}\left(1-K'\frac{B_S}{B_T}\right)^+\middle|\mathcal{F}_t\right] \tag{4.11}$$

and applying [Fil09, Lemma 10.2], making use of equation (4.10). In the case of a backward-looking floorlet, formula (4.7) can be proved in an analogous way. □

Remark 4.8. Backward-looking caplets/floorlets can also be evaluated inside the accrual period $[S,T]$. In this case, the price of a backward-looking caplet for $t\in[S,T]$ can be expressed as

$$\pi_t^{c,\mathrm{B}}(S,T,K)=\mathbb{E}\bigg[\bigg(\frac{B_t}{B_S}-K'\frac{B_t}{B_T}\bigg)^+\bigg|\mathcal{F}_t\bigg]=\int_{\mathbb{R}}\mathbb{E}\bigg[\bigg(\frac{B_t}{B_T}\bigg)^{w+\mathrm{i}\lambda}\bigg|\mathcal{F}_t\bigg]\bigg(\frac{B_t}{B_S}\bigg)^{-(w-1+\mathrm{i}\lambda)}h^K(w,\lambda)\mathrm{d}\lambda$$
$$=\int_{\mathbb{R}}e^{A^0(t,T,w+\mathrm{i}\lambda)+\langle B^0(T-t,w+\mathrm{i}\lambda),X_t\rangle}\bigg(\frac{B_t}{B_S}\bigg)^{-(w-1+\mathrm{i}\lambda)}h^K(w,\lambda)\mathrm{d}\lambda,$$

for any $w\in(w_-,0)$. The application of Fubini's theorem is justified by (4.9), while the last equality follows from formula (2.5) using the fact that $(0,-w-\mathrm{i}\lambda)\in\mathcal{D}_T$ for all $w\in(w_-,w_+)$.

As pointed out in [LM19], in a market where forward-looking and backward-looking rates are present, one may also consider a *term-basis caplet*, corresponding to an exchange option between forward-looking and backward looking rates. The corresponding arbitrage-free price is given by

$$\pi_t^{c,\mathrm{TB}}(S,T)=(T-S)\mathbb{E}\left[\frac{B_t}{B_T}\big(R(S,T)-F(S,T)\big)^+\bigg|\mathcal{F}_t\right],\qquad\text{for } t\in[0,T].$$

For $t\in[S,T]$, term-basis caplets/floorlets reduce to backward-looking caplets/floorlets that can be priced as in Remark 4.8. In the following, we therefore restrict our attention to the valuation of term-basis caplets on $[0,S]$. As a preliminary, let us define

$$k(w,\lambda):=\frac{1}{2\pi}\frac{1}{(w+\mathrm{i}\lambda)(w-1+\mathrm{i}\lambda)},\qquad\text{for all } w\in\mathbb{R}\setminus\{0,1\}\text{ and }\lambda\in\mathbb{R}.$$

Similarly as above, the next theorem relies on a Fourier representation of the payoff of a term-basis caplet. The proof requires a specific analysis of the integrability properties of the terms appearing in the representation and of the applicability of the affine transform formula (2.5).

Theorem 4.9. *There exists a constant $w_-<0$ such that the price of a term-basis caplet can be expressed as follows, for all $w\in(w_-,1)\setminus\{0\}$:*

$$\pi_t^{c,\mathrm{TB}}(S,T)=\begin{cases}\Pi_t^{\mathrm{TB}}(w), & \text{for } w\in(w_-,0),\\ \Pi_t^{\mathrm{TB}}(w)+P_t(S), & \text{for } w\in(0,1),\end{cases}\tag{4.12}$$

for all $t\in[0,S]$, where

$$\Pi_t^{\mathrm{TB}}(w):=\int_{\mathbb{R}}e^{A^1(t,S,B^0(T-S,w+\mathrm{i}\lambda)-(w+\mathrm{i}\lambda)B^0(T-S,1))+A^0(S,T,w+\mathrm{i}\lambda)-(w+\mathrm{i}\lambda)A^0(S,T,1)}$$
$$\times e^{\langle B^1(S-t,B^0(T-S,w+\mathrm{i}\lambda)-(w+\mathrm{i}\lambda)B^0(T-S,1)),X_t\rangle}k(w,\lambda)\mathrm{d}\lambda.$$

Proof. Observe that the price $p_t^{c,\mathrm{TB}}(S,T)$ of a term-basis caplet can be expressed as follows:

$$\pi_t^{c,\mathrm{TB}}(S,T)=\mathbb{E}\bigg[\frac{B_t}{B_S}\bigg(1-\frac{B_S}{B_TP_S(T)}\bigg)^+\bigg|\mathcal{F}_t\bigg].\tag{4.13}$$

In order to apply [Fil09, Lemma 10.2], we first note that

$$\mathbb{E}\bigg[\frac{1}{B_S}\int_{\mathbb{R}}\bigg|\bigg(\frac{B_S}{B_TP_S(T)}\bigg)^{w+\mathrm{i}\lambda}k(w,\lambda)\bigg|\mathrm{d}\lambda\bigg]\le\mathbb{E}\left[\frac{1}{B_S}\bigg(\frac{B_S}{B_TP_S(T)}\bigg)^{w}\right]\int_{\mathbb{R}}|k(w,\lambda)|\mathrm{d}\lambda.\tag{4.14}$$

The expected value on the right-hand side of the above inequality is obviously finite for $w=0$ and $w=1$ and, therefore, for all $w\in(0,1)$ by convexity. Moreover, Hölder's inequality implies the existence of $w_->0$ such that (4.14) is finite-valued for all $w\in(-w_-,1)\setminus\{0\}$. We can therefore apply Fubini's theorem and obtain

$$\mathbb{E}\bigg[\frac{B_t}{B_S}\int_{\mathbb{R}}\bigg(\frac{B_S}{B_TP_S(T)}\bigg)^{w+\mathrm{i}\lambda}k(w,\lambda)\mathrm{d}\lambda\bigg|\mathcal{F}_t\bigg]=\Pi_t^{\mathrm{TB}}(w).\tag{4.15}$$

To obtain (4.15), we have applied formula (2.5), using the fact that $(0,-(w+\mathrm{i}\lambda))\in\mathcal{D}_{T-S}\subseteq\mathcal{D}_T$ for all $w\in(w_-,1)\setminus\{0\}$. To justify (4.15), note first that $B^0_J(T-S,w)-wB^0_J(T-S,1)=0$. For all $w\in(0,1)$, Lemma 2.7 implies that $B^0_I(T-S,w)-wB^0_I(T-S,1)\preceq0$. Lemma 2.6 and Assumption 3.1 then imply $(B^0(T-S,w)-wB^0(T-S,1),-1)\in\mathcal{Y}_S$, for all $w\in(0,1)$. Similarly as in the proof of [KRM15, Proposition 5.1], we have that $\mathrm{Re}(B^0_I(T-S,w+\mathrm{i}\lambda))\preceq B^0_I(T-S,w)$

and a further application of Lemma 2.6 yields $(B^0(T-S,w+\mathrm{i}\lambda)-(w+\mathrm{i}\lambda)B^0(T-S,1),-1)\in\mathcal{D}_S$. This justifies (4.15) for all $w\in(0,1)$. Considering now the case $w\in(w_-,0)$, note that there exists a solution up to time S to (2.4a)-(2.4b) starting from $(B^0(T-S,w_-)-w_-B^0(T-S,1),-1)$, but this point may lie on the boundary $\partial\mathcal{Y}_S$. However, setting $\gamma:=w/w_-\in(0,1)$, we have that

$$B_I^0(T-S,w)-wB_I^0(T-S,1)\preceq\gamma\left(B_I^0(T-S,w_-)-w_-B_I^0(T-S,1)\right).$$

In particular, since $(0,-1)\in\mathcal{Y}_S$ by Assumption 3.1 and the set $\mathcal{Y}_S$ is convex, this implies that $(B^0(T-S,w)-wB^0(T-S,1),-1)\in\mathcal{Y}_S$, for all $w\in(w_-,0)$. Similarly as above, this yields $(B^0(T-S,w+\mathrm{i}\lambda)-(w+\mathrm{i}\lambda)B^0(T-S,1),-1)\in\mathcal{D}_S$, thus completing the proof of (4.15). Formula (4.12) follows by applying [Fil09, Lemma 10.2] to equation (4.13), making use of (4.15). □

Remark 4.10. In this remark, we discuss briefly some general aspects that are relevant for the numerical implementation of the pricing formulas derived above:

(1) When computing the quantities $\Pi_t^{\mathrm{F}}(w)$ and $\Pi_t^{\mathrm{B}}(w)$ that appear in Theorems 4.4 and 4.7, respectively, the parts that depend on the characteristic function have to be evaluated only once for all different strikes. This fact is important for model calibration, especially in models for which the Riccati ODEs (2.4a)-(2.4b) have to be solved numerically.
(2) The computation of the quantities $\Pi_t^{\mathrm{F}}(w)$, $\Pi_t^{\mathrm{B}}(w)$ and $\Pi_t^{\mathrm{TB}}(w)$ is subject to a truncation error (due to the truncation of the integral at some suitably chosen upper/lower bounds) and a discretization error (due to approximating the integral with a finite sum). Explicit error bounds that are applicable to our setting have been derived in [Lee04].
(3) The choice of the parameter w depends on the specific properties of the affine process under consideration and has to be suitably chosen in order to exclude large oscillations of the integrand. The choice of w can also depend on the maturity and the strike of the products to be priced (see, e.g., [Lee04]). As can be seen from the proofs above, the range of possible values of w is crucially related to the integrability properties of the model. Precise recommendations on the choice of w, as well as of the other numerical parameters, are given in [Lev16], in an affine setup that also covers our setting.

Alternative pricing formulae can be obtained by passing to forward measures, exploiting the fact that the characteristic function of the affine process X under any forward measure can be explicitly determined. The proof of the following proposition is similar to [Fil09, Corollary 10.4] and is based on (4.11) together with (3.4) and the fact that $B_T/B_S=\exp(L(S,T)+(Y_T-Y_S))$.

Proposition 4.11. *It holds that*

$$\pi_t^{c,\mathrm{F}}(S,T,K)=P_t(S)p_t^S(\mathcal{I}^{\mathrm{F}})-K'P_t(T)p_t^T(\mathcal{I}^{\mathrm{F}}),\tag{4.16}$$

$$\pi_t^{c,\mathrm{B}}(S,T,K)=P_t(S)q_t^S(\mathcal{I}^{\mathrm{B}})-K'P_t(T)q_t^T(\mathcal{I}^{\mathrm{B}}),\tag{4.17}$$

where $\mathcal{I}^{\mathrm{F}}:=(-\infty,-A^0(S,T,1)-\log(K'))$, $\mathcal{I}^{\mathrm{B}}:=(\log(K')-L(S,T),+\infty)$ *and*

- $p_t^S(\mathrm{d}y)$ *and* $p_t^T(\mathrm{d}y)$ *denote respectively the* $\mathcal{F}_t$*-conditional distributions of the random variable* $\langle B^0(T-S,1),X_S\rangle$ *under the* S*-forward and* T*-forward measures;*
- $q_t^S(\mathrm{d}y)$ *and* $q_t^T(\mathrm{d}y)$ *denote respectively the* $\mathcal{F}_t$*-conditional distributions of the random variable* Y_T-Y_S *under the* S*-forward and* T*-forward measures.*

The distributions appearing in Proposition 4.11 can be recovered from the $\mathcal{F}_t$-conditional characteristic function of $(\langle B^0(T-S,1),X_S\rangle,Y_T-Y_S)$ under the S-forward and T-forward measures. In view of Proposition 2.5, the latter can be computed as follows, for all $(\zeta_1,\zeta_2)\in\mathbb{R}^2$:

$$\mathbb{E}^S\left[e^{\mathrm{i}\zeta_1\langle B^0(T-S,1),X_S\rangle+\mathrm{i}\zeta_2(Y_T-Y_S)}\middle|\mathcal{F}_t\right]$$
$$=\frac{1}{P_t(S)}e^{-\mathrm{i}\zeta_2L(S,T)+A^0(S,T,-\mathrm{i}\zeta_2)+A^1(t,S,\mathrm{i}\zeta_1B^0(T-S,1)+B^0(T-S,-\mathrm{i}\zeta_2))+\langle B^1(S-t,\mathrm{i}\zeta_1B^0(T-S,1)+B^0(T-S,-\mathrm{i}\zeta_2)),X_t\rangle}$$

and

$$\mathbb{E}^T\left[e^{\mathrm{i}\zeta_1\langle B^0(T-S,1),X_S\rangle+\mathrm{i}\zeta_2(Y_T-Y_S)}\middle|\mathcal{F}_t\right]$$
$$=\frac{1}{P_t(T)}e^{-\mathrm{i}\zeta_2L(S,T)+A^0(S,T,1-\mathrm{i}\zeta_2)+A^1(t,S,\mathrm{i}\zeta_1B^0(T-S,1)+B^0(T-S,1-\mathrm{i}\zeta_2))+\langle B^1(S-t,\mathrm{i}\zeta_1B^0(T-S,1)+B^0(T-S,1-\mathrm{i}\zeta_2)),X_t\rangle}.$$

Remark 4.12. For some specific models, the conditional distributions appearing in Proposition 4.11 can be explicitly computed. In particular, this is the case for (multi-factor) Hull-White models, which have the property of preserving the Gaussian distribution of X_S and $Y_T - Y_S$ under any forward measure. In this setting, one can deduce from Proposition 4.11 the caplet pricing formulas recently derived in [Hof20, RB21, Tur21, Xu22].

5. An example: CIR++ model for an RFR process

In this section, we illustrate the applicability of Proposition 4.11 in the context of the CIR++ model introduced in [BM01]. We assume that the RFR process is given by $r := \ell(\cdot) + X$, where $X = (X_t)_{t\geq 0}$ is a square-root process:

$$\mathrm{d}X_t = (b - \beta X_t)\mathrm{d}t + \sigma\sqrt{X_t}\mathrm{d}W_t, \qquad X_0 > 0, \tag{5.1}$$

where $b, \beta, \sigma > 0$. The explicit expression of the unique function ℓ that fits the term structure at $t = 0$ is given in [BM01, Section 6]. For the CIR++ model, the explicit form of the functions $A^0(t, T, v)$ and $B^0(T - t, v)$ can be deduced from [JYC09, Corollary 6.3.4.2]:

$$\begin{aligned}
A^0(t,T,v) &= \frac{2b}{\sigma^2}\log\left(\frac{2\theta_v e^{\frac{(\theta_v+\beta)(T-t)}{2}}}{2\theta_v + (\beta+\theta_v)(e^{\theta_v(T-t)} - 1)}\right) - v\int_t^T \ell(u)\mathrm{d}u,\\
B^0(T-t,v) &= \frac{-2v}{\beta + \theta_v\coth(\frac{\theta_v(T-t)}{2})},
\end{aligned}$$

with $\theta_v := \sqrt{\beta^2 + 2v\sigma^2}$. Setting $v = 1$ enables us to explicitly compute ZCB prices by (3.4). [JYC09, Proposition 6.3.4.1] implies that the functions $A^1(t, T, u)$ and $B^1(T - t, u)$ are given by

$$\begin{aligned}
A^1(t,T,u) &= \frac{2b}{\sigma^2}\log\left(\frac{2\theta e^{\frac{(\theta+\beta)(T-t)}{2}}}{\theta(e^{\theta(T-t)}+1) + \beta(e^{\theta(T-t)}-1) - u\sigma^2(e^{\theta(T-t)}-1)}\right) - \int_t^T \ell(u)\mathrm{d}u,\\
B^1(T-t,u) &= \frac{u(\theta+\beta+e^{\theta(T-t)}(\theta-\beta)) - 2(e^{\theta(T-t)}-1)}{\theta(e^{\theta(T-t)}+1) + \beta(e^{\theta(T-t)}-1) - u\sigma^2(e^{\theta(T-t)}-1)},
\end{aligned}$$

where we denote $\theta := \theta_1$ for brevity of notation.

In the CIR++ model, the price of a forward-looking caplet can be computed in closed form, since the $\mathcal{F}_t$-conditional distributions $p_t^S(\mathrm{d}y)$ and $p_t^T(\mathrm{d}y)$ appearing in formula (4.16) can be explicitly determined. Indeed, for all $0 \leq t \leq S \leq T < +\infty$, the $\mathcal{F}_t$-conditional density $p^T_{(t,S)}(x)$ of X_S under the T-forward measure is given by (see, e.g., [BM01, Section 6])

$$p^T_{(t,S)}(x) = f_{\chi^2(\nu,\delta(t,S)X_t)}(x),$$

where $f_{\chi^2(\nu,\delta(t,S)X_t)}$ denotes the density function of a non-central χ^2 distribution with ν degrees of freedom and non-centrality parameter $\delta(t, S)X_t$, with $\nu := 4b/\sigma^2$ and

$$\delta(t,S) := \frac{4\rho^2(S-t)e^{\theta(S-t)}}{4(\rho(S-t) + (\beta+\theta)/\sigma^2 - B^0(T-S,1))^2}, \quad \text{where} \quad \rho(S-t) := \frac{2\theta}{\sigma^2(e^{\theta(S-t)}-1)}.$$

The price of a backward-looking caplet can be computed by relying on formula (4.17). For a square-root process X, the $\mathcal{F}_t$-conditional distribution of $\int_S^T X_u\mathrm{d}u$ under a forward measure is not known. However, it can be retrieved by the Gil-Pelaez inversion formula, making use of the explicit knowledge of the $\mathcal{F}_t$-conditional characteristic function of $\int_S^T X_u\mathrm{d}u$ under any forward measure. This leads to the following semi-closed pricing formula:

$$\pi_t^{c,\mathrm{B}}(S,T,K) = \frac{P_t(S) - K'P_t(T)}{2} + \frac{1}{\pi}\int_0^{+\infty}\frac{\mathrm{Im}\left((K')^{-\mathrm{i}x}(e^{g_1(x)} - K'e^{g_2(x)})\right)}{x}\mathrm{d}x,$$

where

$$\begin{aligned}
g_1(x) &:= A^0(S,T,-\mathrm{i}x) + A^1(t,S,B^0(T-S,-\mathrm{i}x)) + \langle B^1(S-t,B^0(T-S,-\mathrm{i}x)), X_t\rangle,\\
g_2(x) &:= A^0(S,T,1-\mathrm{i}x) + A^1(t,S,B^0(T-S,1-\mathrm{i}x)) + \langle B^1(S-t,B^0(T-S,1-\mathrm{i}x)), X_t\rangle.
\end{aligned}$$

Remark 5.1. The CIR++ model of this section can be generalized to a Wishart driving process, as considered in [Gno12]. By [CFG19, Lemma 5.3], a Wishart process has a non-central Wishart distribution with known parameters under any forward measure. Similarly as above, this allows for the explicit computation of the conditional probabilities appearing in formula (4.16), while the conditional probabilities appearing in formula (4.17) can be recovered by Fourier inversion.

References

[AB20] L. Andersen and D. Bang. Spike modeling for interest rate derivatives with an application to SOFR caplets. Working paper, 2020.

[AP07] L. Andersen and V. Piterbarg. Moment explosions in stochastic volatility models. *Finance and Stochastics*, 11:29–50, 2007.

[BM01] D. Brigo and F. Mercurio. A deterministic-shift extension of analytically-tractable and time-homogeneous short-rate models. *Finance and Stochastics*, 5(3):369–387, 2001.

[CFG19] C. Cuchiero, C. Fontana, and A. Gnoatto. Affine multiple yield curve models. *Mathematical Finance*, 29(2):568–611, 2019.

[CFMT11] C. Cuchiero, D. Filipović, E. Mayerhofer, and J. Teichmann. Affine processes on positive semidefinite matrices. *Annals of Applied Probability*, 21(2):397–463, 2011.

[CM99] P. Carr and D. B. Madan. Option valuation using the fast Fourier transform. *Journal of Computational Finance*, 2(4):61–73, 1999.

[DFS03] D. Duffie, D. Filipović, and W. Schachermayer. Affine processes and applications in finance. *Annals of Applied Probability*, 13(3):984–1053, 2003.

[FGS22a] C. Fontana, A. Gnoatto, and G. Szulda. CBI-time-changed Lévy processes. Working paper (available at `https://arxiv.org/abs/2205.12355`), 2022.

[FGS22b] C. Fontana, Z. Grbac, and T. Schmidt. Term structure modelling with overnight rates beyond stochastic continuity. Working paper (available at `https://arxiv.org/abs/2202.00929`), 2022.

[Fil09] D. Filipovic. *Term-structure Models. A Graduate Course.* Springer, Berlin - Heidelberg, 2009.

[Gno12] A. Gnoatto. The Wishart short rate model. *International Journal of Theoretical and Applied Finance*, 15(8):1250056, 2012.

[Hen04] M.P. Henrard. Overnight indexed swaps and floored compounded instruments in HJM one-factor model. Working paper, 2004.

[HKRS17] F. Hubalek, M. Keller-Ressel, and C. Sgarra. Geometric asian option pricing in general affine stochastic volatility models with jumps. *Quantitative Finance*, 17(6):873–888, 2017.

[Hof20] K. Hofman. Implied volatilities for options on backward-looking term rates. Working paper, 2020.

[ISD20] ISDA. 2020 IBOR fallbacks protocol. Protocol (available at `https://www.isda.org`), 2020.

[JYC09] M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical Methods for Financial Markets.* Springer, London, 2009.

[KMK10] J. Kallsen and J. Muhle-Karbe. Exponentially affine martingales, affine measure changes and exponential moments of affine processes. *Stochastic Processes and their Applications*, 120(2):163–181, 2010.

[KRM15] M. Keller-Ressel and E. Mayerhofer. Exponential moments of affine processes. *Annals of Applied Probability*, 25(2):714–752, 2015.

[KRMS10] M. Keller-Ressel, E. Mayerhofer, and A.G. Smirnov. On convexity of solutions of ordinary differential equations. *Journal of Mathematical Analysis and Applications*, 368(1):247–253, 2010.

[Lee04] R. Lee. Option pricing by transform methods: extensions, unification and error control. *Journal of Computational Finance*, 7(3):51–86, 2004.

[Lev16] S. Levendorskii. Pitfalls of the Fourier transform method in affine models, and remedies. *Applied Mathematical Finance*, 23(2):81–134, 2016.

[LM19] A. Lyashenko and F. Mercurio. Looking forward to backward-looking rates: a modeling framework for term rates replacing LIBOR. Working paper, 2019.

[Mer18] F. Mercurio. A simple multi-curve model for pricing SOFR futures and other derivatives. Working paper, 2018.

[MS20] A. Macrina and D. Skovmand. Rational savings account models for backward-looking interest rate benchmarks. *Risks*, 8(1):23, 2020.

[Pit20] V. Piterbarg. Interest rates benchmark reform and options markets. Working paper, 2020.

[RB21] M. Rutkowski and M. Bickersteth. Pricing and hedging of SOFR derivatives under differential funding costs and collateralization. Working paper, 2021.

[SS21] J.B. Skov and D. Skovmand. Dynamic term structure models for SOFR futures. *Journal of Futures Markets*, 41(10):1520–1544, 2021.

[Tur21] C. Turfus. Caplet pricing with backward-looking rates. Working paper, 2021.

[Wil20] S. Willems. SABR smiles for RFR caplets. Working paper, 2020.

[Xu22] M. Xu. SOFR derivative pricing using a short rate model. Working paper (available at `https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4007604`), 2022.