Università degli Studi di Padova
Dipartimento di Scienze Statistiche
Corso di Laurea Triennale in Statistica per l'Economia e l'Impresa

RELAZIONE FINALE

Caplet and Floorlet Pricing under Backward-Looking Compounded Rates in a Discrete Hull-White Framework

Relatore
Prof. Massimiliano Caporin
Dipartimento di Scienze Statistiche
Correlatore
Prof. Claudio Fontana
Dipartimento di Matematica

Laureando
Sebastiano Pinotti
Matricola: 2101255

Anno Accademico 2025/2026

Contents

Introduction

The reform of interest-rate benchmarks has changed the object on which interest-rate derivatives are written. With the discontinuation of LIBOR, the reference rates in the major currencies are now the overnight risk-free rates (RFRs), SOFR in the United States, SONIA in the United Kingdom, €STR in the euro area, which are computed from actual overnight transactions rather than from panel submissions. Term rates are built from these overnight rates by compounding the daily fixings over the accrual period, and the resulting rate is *backward-looking*: it is realized only at the end of the period over which it applies, in contrast to the *forward-looking* LIBOR rate, which was fixed at the start of the accrual period.

A LIBOR caplet's payoff is determined at the beginning of its accrual period, so the option is effectively settled before the period begins. A backward-looking caplet's payoff, by contrast, keeps accumulating randomness throughout the accrual window and is known only at its end. The optionality therefore lives inside the accrual period, and any pricing model must account for the uncertainty accumulated during it. Backward-looking caps and floors are now the dominant non-linear RFR instruments, and the volume traded in them has grown rapidly since the transition.

The natural way to model an overnight rate is to specify the dynamics of the short rate, and within this approach the Gaussian Hull-White model has been the standard choice for RFRs since Mercurio [15] first adopted it for SOFR. Explicit pricing of compounded, in-arrears instruments in a Gaussian one-factor Hull-White model in fact can be found before the reform: Henrard [9] derived such formulas in a one-factor Gaussian HJM setting. After the transition, backward-looking caplet formulas under the Hull-White model were obtained by several authors [10, 20, 22], and the most general continuous-time treatment is due to Fontana [8], who prices backward- and forward-looking caplets and floorlets for general affine short-rate processes by Fourier methods. Like all Gaussian short-rate models, the framework considered here does not generate an implied-volatility smile and is calibrated only at the money [17]. What these contributions share is that they are set in continuous time and rely on the apparatus of stochastic calculus, Itô's formula, changes to a forward measure, partial differential equations, or Fourier inversion.

A separate part of the literature models the term structure directly in discrete time. Filipović and Zabczyk [7] characterize when a discrete-time Markovian short-rate model admits an affine term structure, and the multi-period discrete-time pricing framework is developed in textbook form by Pascucci and Runggaldier [16]. This line provides exactly the tools required to price an option whose payoff is a function of a sum of Gaussian short rates, yet it has not been used for backward-looking caplet pricing. The closest work to the present one is the paper of Iunevich and Kalikaeva [12], who also price discrete backward-looking caplets under the Hull-White model. Their model, however, retains the continuous-time Hull-White short rate and discretizes only the compounding convention, whereas the model adopted here is discrete at the level of the short rate itself.

This thesis prices backward-looking caplets and floorlets in a discrete Hull-White model in which the short rate evolves by a first-order autoregressive recursion, the Euler-Maruyama discretization of the Hull-White equation, which retains the properties of the continuous time model because the noise is additive. Under the single-period convention $p(k, k+1) = e^{-r_k}$, the compounded backward-looking rate telescopes exactly into the exponential of the sum of the short rates over the accrual period, so that the compounding is exact rather than an approximation. The entire pricing theory is then carried out within the discrete affine term structure of this model, using only elementary properties of the Gaussian distribution: the closed forms follow from truncated univariate and bivariate Gaussian moments, obtained by completing the square, and no step of any pricing argument uses stochastic calculus. The continuous-time Hull-White equation and its discretization, presented in Chapters 1 and 2, serve only to motivate and justify the discrete model and are not used in any subsequent derivation. To the author's knowledge, this is the first fully self-contained, discrete-time derivation of backward-looking caplet and floorlet prices that requires no stochastic calculus, and providing it is the principal contribution of the thesis.

This elementary, fully explicit form is also what makes the rest of the analysis possible. We obtain closed-form prices for the backward-looking caplet and floorlet both at the start of the accrual period and at any earlier valuation date, and the forward-looking prices derive from the same computation as a special case. Because the two conventions are produced by a single argument, they are seen to share an identical Black-type pricing formula and to differ only through the variance entering it: the backward-looking price uses the total variance ν_X, the forward-looking price the pre-accrual variance ν_W, and the two differ by exactly the in-accrual variance $\nu_{m,N}$, so that $\nu_X = \nu_W + \nu_{m,N}$. The entire difference between the two conventions thus reduces to a single, explicitly characterized term. This decomposition is the second contribution: it yields an exact integral representation of the price gap, a closed-form expression for it at the money, and a complete numerical study of its behavior in the model parameters, none of which requires simulation, since every quantity is available in closed form.

The thesis is organized as follows. Chapter 1 introduces the discrete-time market, the risk-neutral pricing framework, the one-step pricing recursion, and reviews the classical one-factor short-rate models in order to motivate the choice of the Hull-White model. Chapter 2 develops the discrete Hull-White model: its affine term structure, closed-form bond-pricing coefficients, and calibration to the initial discount curve. Chapter 3, the core of the thesis, derives the closed-form prices of the backward-looking caplet and floorlet at the start of the accrual period and at earlier times, together with their forward-looking counterparts and the variance decomposition relating them. Chapter 4 studies the resulting prices numerically, analyzing their sensitivity to volatility, mean reversion, strike, and the length of the pre-accrual and accrual windows, and the gap between the backward- and forward-looking prices.

Chapter 1

Stochastic Interest Rates in Discrete Time

This chapter develops the discrete-time framework on which the rest of the thesis is built: the bond market and the family of interest rates derived from it, the arbitrage-free pricing relations that connect them, and the class of short-rate models from which the Hull-White model will be selected. Working in discrete time is not merely a matter of convenience. It provides a rigorous and self-contained setting that avoids the technicalities of stochastic calculus while retaining full analytical tractability, and it is the natural formulation for backward-looking rates, which are themselves defined as discrete products of overnight fixings.

The fundamental traded assets are the *T-bonds*, contracts that guarantee the payment of one unit of currency at a fixed future date T and whose prices encode the market's valuation of money across time. Because the interest rate underlying a derivative is not itself a tradable asset, unlike the stock underlying an equity option, the pricing and hedging of interest-rate derivatives rest on the bond market rather than on the rate directly.

We begin in Section 1.1 by introducing the fundamental market quantities: the time grid, zero coupon bond prices, and the various notions of interest rates that will appear throughout the thesis. Section 1.2 develops the arbitrage-free pricing framework, introducing the risk-neutral measure and the one-step pricing recursions. Finally, Section 1.3.1 and Section 1.3.2 discuss the two main classes of interest rate models, short rate models and forward rate models, and motivate the choice of a short rate model as the basis for our analysis.

1.1 T-Bonds and Interest Rates Notation

We consider a discrete-time financial market defined over a finite time horizon $[0, \bar{T}]$. The interval $[0, \bar{T}]$ is divided into $\bar{N}$ subintervals of equal length, so that the time step is $\Delta = \bar{T}/\bar{N}$ and transactions occur only at the dates

$$t_n = n\Delta, \qquad n = 0, 1, \dots, \bar{N}.$$

For convenience, we refer to the n-th time instant indifferently by t_n or simply by n, and denote the n-th period $[t_{n-1}, t_n]$ simply by $[n-1, n]$.

The fundamental traded assets in the interest rate market are the *zero coupon bonds*, also called *T-bonds*.

Definition 1.1 (Zero Coupon Bond). *A T-bond with maturity $T = N\Delta$ is a contract that guarantees the payment of one unit of currency at time N and does not provide any intermediate cash flow. We denote by $p(n, N)$ the price at time n of the T-bond with maturity N, for $0 \le n \le N \le \bar{N}$.*

By definition, $p(N, N) = 1$ for all N. At time $n = 0$, the family of prices $\{p(0, N)\}_{N=1}^{\bar{N}}$ is observable on the market and constitutes the *term structure of zero coupon bonds*, also called the *initial discount curve*. We denote the market-observed term structure by $p^*(0, N)$ to distinguish it from the theoretical prices $p(0, N)$ generated by a given mathematical model.

Starting from the prices of T-bonds, it is natural to define interest rates as quantities that describe the return on investing over a given time interval. Consider three time instants $n \le N \le M$ and an investment at time n consisting of selling one unit of $p(n, N)$ in order to buy $p(n, N)/p(n, M)$ units of $p(n, M)$. This investment has zero cost at time n and consists of paying 1 at time N to receive $p(n, N)/p(n, M)$ at time M. The interest rates defined below are consistent with this investment scheme.

Definition 1.2 (Interest Rates). *Let $n \le N < M \le \bar{N}$. We define the following rates:*

1. *The **simple annualized forward rate** $L(n; N, M)$ for the period $[t_N, t_M]$, evaluated at time n, is defined by the capitalization formula*

$$\frac{p(n, N)}{p(n, M)} = 1 + L(n; N, M)(M - N)\Delta,$$

   or equivalently

$$L(n; N, M) = \frac{1}{(M - N)\Delta} \left(\frac{p(n, N)}{p(n, M)} - 1 \right). \tag{1.1}$$

   We also write $L(n, N) := L(n; N, N+1) = \frac{1}{\Delta}\left(\frac{p(n,N)}{p(n,N+1)} - 1\right)$ for the one-period simple forward rate.

2. *The* ***compounded forward rate*** $R(n;N,M)$ *for the period* $[t_N, t_M]$*, evaluated at time* n*, is defined by*

$$\frac{p(n,N)}{p(n,M)} = e^{R(n;N,M)}. \tag{1.2}$$

We also write $R(n,N) := R(n;N,N+1) = \log \frac{p(n,N)}{p(n,N+1)}$ *for the one-period compounded forward rate.*

3. *The* ***short rate*** $r_n := R(n,n)$ *is the compounded spot rate relative to the period* $[t_n, t_{n+1}]$*. By definition,*

$$p(n,n+1) = e^{-r_n}. \tag{1.3}$$

Remark 1.1. The short rate r_n is a *random* value that becomes known at time n. It is said to be *locally riskless* in the sense that an investment in the one-period bond at time n yields a deterministic return e^{r_n} over the period $[t_n, t_{n+1}]$, given that r_n is known at t_n. Over longer horizons, however, the return is stochastic since future short rates are unknown.

The short rate naturally gives rise to the *money market account*, which represents a rolling investment strategy that continually reinvests at the prevailing short rate.

Definition 1.3 (Money Market Account)**.** *The money market account* B_n *is defined by* $B_0 = 1$ *and the recursion* $B_{n+1} = B_n e^{r_n}$*, or more explicitly*

$$B_N = B_n \exp\left(\sum_{k=n}^{N-1} r_k\right), \qquad 0 \le n < N \le \bar{N}. \tag{1.4}$$

The money market account is a locally riskless asset: over each period $[n, n+1]$ it grows by the deterministic factor e^{r_n}. Over multiple periods, however, its value depends on the entire path of short rates and is therefore random at any future time. The associated *discount factor* is defined as follows.

Definition 1.4 (Discount Factor)**.** *The discount factor over the period* $[t_n, t_N]$ *is*

$$D(n,N) := \frac{B_n}{B_N} = \exp\left(-\sum_{k=n}^{N-1} r_k\right), \qquad 0 \le n < N \le \bar{N}. \tag{1.5}$$

It is important to distinguish between $p(n,N)$ and $D(n,N)$: both represent the value at time n of one unit of currency delivered at time N, but $p(n,N)$ is the *price of a traded contract* observable on the market at time n, whereas $D(n,N)$ is a *random quantity* that depends on the realized path of short rates from n to $N-1$ and is therefore not known at time n. As we shall see in the next section, the relationship between these two quantities is precisely established by the risk-neutral pricing formula.

1.2 Risk-Neutral Framework

In order to assign prices to interest rate derivatives in a way that is consistent with the absence of arbitrage, we need to introduce the probabilistic framework underlying the market model. We work on a filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_n))$, where P denotes the real-world probability measure and the filtration $(\mathcal{F}_n)_{n=0}^{\bar{N}}$ represents the information available to market participants at each discrete date. We assume that r and $p(\cdot, N)$, for $N = 1, \ldots, \bar{N}$, are adapted processes and that the bond price processes $p(\cdot, N)$ are strictly positive.

The central concept linking the absence of arbitrage to the existence of a pricing measure is that of equivalent probability measures, which we recall here for completeness.

Definition 1.5 (Equivalent Probability Measures). *Two probability measures P and Q on $(\Omega, \mathcal{F})$ are said to be equivalent, written $P \sim Q$, if they assign probability zero to the same events, i.e.*

$$P(A) = 0 \iff Q(A) = 0 \qquad \text{for all } A \in \mathcal{F}.$$

By the Radon-Nikodým theorem (for the proof we refer to theorem 28.2 in [13]), if $P \sim Q$ then there exists a strictly positive random variable $Z = \frac{dQ}{dP}$, called the Radon-Nikodym derivative of Q with respect to P, such that

$$Q(A) = E^P[Z \cdot \mathbf{1}_A] \qquad \text{for all } A \in \mathcal{F}.$$

To introduce the pricing measure, we define the *discounted price processes* of the T-bonds by

$$\widetilde{p}(n, N) := \frac{p(n, N)}{B_n}, \qquad 0 \leq n \leq N,$$

which represent bond prices expressed in units of the money market account, taken as numeraire.

Definition 1.6 (Martingale Measure). *A martingale measure with numeraire B is a probability measure Q equivalent to P such that, for every maturity $N = 1, \ldots, \bar{N}$, the discounted T-bond price process is a Q-martingale, i.e.*

$$\widetilde{p}(n, N) = E^Q[\widetilde{p}(n+1, N) \mid \mathcal{F}_n], \qquad 0 \leq n < N \leq \bar{N}. \tag{1.6}$$

The connection between absence of arbitrage and the existence of a martingale measure is provided by the First Fundamental Theorem of Asset Pricing. In the standard finite-state discrete-time setting, absence of arbitrage is equivalent to the existence of an equivalent martingale measure; see, for instance, [16, Theorem 1.14].

For the purposes of this thesis, the relevant framework is a finite-horizon discrete-time market with a finite number of traded maturities $N = 1, \ldots, \bar{N}$. Although the time horizon and the number of traded assets are finite, the state space of the Hull-White model introduced below is not finite, since the short rate is driven by Gaussian innovations. The corresponding finite-horizon discrete-time result on general probability spaces is due to Dalang, Morton and Willinger [5], who show that, under the appropriate no-arbitrage condition, an equivalent martingale measure exists for vector-valued securities price processes.

The following lemma provides an equivalent and more explicit characterization of the martingale condition in terms of bond prices and the discount factor.

Lemma 1.1. *The martingale condition* (1.6) *is equivalent to*

$$p(n, N) = E^Q[D(n, N) \mid \mathcal{F}_n], \qquad 0 \le n \le N \le \bar{N}, \tag{1.7}$$

where by convention $D(n, n) = 1$.

Proof. (⇒) Suppose (1.6) holds. By the martingale property applied iteratively from n to N:

$$\widetilde{p}(n, N) = E^Q[\widetilde{p}(N, N) \mid \mathcal{F}_n].$$

Since $p(N, N) = 1$, we have $\widetilde{p}(N, N) = 1/B_N$, so

$$\frac{p(n, N)}{B_n} = E^Q\left[\frac{1}{B_N} \mid \mathcal{F}_n\right],$$

and therefore, since B_n is $\mathcal{F}_n$ measurable

$$p(n, N) = E^Q\left[\frac{B_n}{B_N} \mid \mathcal{F}_n\right] = E^Q[D(n, N) \mid \mathcal{F}_n].$$

(⇐) Suppose (1.7) holds. Since r_n is $\mathcal{F}_n$-measurable, using (1.3) and the tower property of conditional expectation:

$$\begin{aligned} p(n, N) &= E^Q[D(n, N) \mid \mathcal{F}_n] \\ &= E^Q\left[e^{-r_n} D(n+1, N) \mid \mathcal{F}_n\right] \\ &= e^{-r_n} E^Q\left[E^Q[D(n+1, N) \mid \mathcal{F}_{n+1}] \mid \mathcal{F}_n\right] \\ &= e^{-r_n} E^Q[p(n+1, N) \mid \mathcal{F}_n]. \end{aligned}$$

Dividing both sides by B_n and using $B_{n+1} = B_n e^{r_n}$:

$$\frac{p(n, N)}{B_n} = E^Q\left[\frac{p(n+1, N)}{B_{n+1}} \mid \mathcal{F}_n\right],$$

which is precisely (1.6). □

The pricing formula (1.7) is the fundamental result that connects the market price of a T-bond to the expected discounted payoff under Q. The following proposition gives a further equivalent characterization in terms of a one-period forward relationship.

Proposition 1.1. *Let Q be a measure equivalent to P. Then Q is a martingale measure if and only if*

$$\frac{p(n,N)}{p(n,n+1)} = E^Q[p(n+1,N) \mid \mathcal{F}_n], \qquad 0 \le n < N \le \bar{N}. \tag{1.8}$$

Proof. Suppose first that Q is a martingale measure. Then, by Lemma 1.1,

$$p(n,N) = E^Q[D(n,N) \mid \mathcal{F}_n].$$

Since r_n is $\mathcal{F}_n$-measurable and $D(n,N) = e^{-r_n} D(n+1,N)$, we obtain

$$\begin{aligned} p(n,N) &= E^Q[D(n,N) \mid \mathcal{F}_n] \\ &= e^{-r_n} E^Q[D(n+1,N) \mid \mathcal{F}_n] \\ &= p(n,n+1)\, E^Q\big[E^Q[D(n+1,N) \mid \mathcal{F}_{n+1}] \mid \mathcal{F}_n\big] \\ &= p(n,n+1)\, E^Q[p(n+1,N) \mid \mathcal{F}_n], \end{aligned}$$

where in the last equality we used Lemma 1.1 again. Dividing both sides by $p(n,n+1)$ gives (1.8).

Conversely, suppose that (1.8) holds for every $0 \le n < N \le \bar{N}$. We prove that

$$p(n,N) = E^Q[D(n,N) \mid \mathcal{F}_n], \qquad 0 \le n \le N \le \bar{N},$$

and then the conclusion follows from Lemma 1.1.

The claim is trivial for $n = N$, since $p(N,N) = 1$ and, by convention, $D(N,N) = 1$. Now assume that, for some $n < N$,

$$p(n+1,N) = E^Q[D(n+1,N) \mid \mathcal{F}_{n+1}].$$

Using (1.8), the induction hypothesis, and the tower property of conditional expectation, we get

$$\begin{aligned} p(n,N) &= p(n,n+1)\, E^Q[p(n+1,N) \mid \mathcal{F}_n] \\ &= e^{-r_n} E^Q\big[E^Q[D(n+1,N) \mid \mathcal{F}_{n+1}] \mid \mathcal{F}_n\big] \\ &= E^Q\big[e^{-r_n} D(n+1,N) \mid \mathcal{F}_n\big] \\ &= E^Q[D(n,N) \mid \mathcal{F}_n], \end{aligned}$$

where we used that $e^{-r_n} = p(n,n+1)$ is $\mathcal{F}_n$-measurable. By backward induction, the pricing formula holds for all $0 \le n \le N \le \bar{N}$. Therefore, by Lemma 1.1, Q is a martingale measure. $\square$

The following corollary generalizes the one-step relationship to arbitrary adapted processes.

Corollary 1.1 (One-Step Pricing Recursion)**.** *Under a martingale measure Q, if a process $X = (X_n)_{n \le N}$ satisfies*

$$X_n = E^Q[D(n,N)X_N \mid \mathcal{F}_n], \qquad n \le N,$$

then $\widetilde{X} = (X_n/B_n)$ is a Q-martingale. In particular, for all $k \le n \le N$:

$$X_k = E^Q[D(k,n)X_n \mid \mathcal{F}_k], \tag{1.9}$$

and in the case $n = k+1$:

$$X_k = p(k,k+1)\, E^Q[X_{k+1} \mid \mathcal{F}_k], \qquad k < N. \tag{1.10}$$

Proof. Since r_n is $\mathcal{F}_n$-measurable, we have:

$$\begin{aligned}
X_n &= E^Q[D(n,N)X_N \mid \mathcal{F}_n] \\
&= E^Q\left[e^{-r_n} D(n+1,N)X_N \mid \mathcal{F}_n\right] \\
&= e^{-r_n} E^Q\left[E^Q[D(n+1,N)X_N \mid \mathcal{F}_{n+1}] \mid \mathcal{F}_n\right] \\
&= e^{-r_n} E^Q[X_{n+1} \mid \mathcal{F}_n].
\end{aligned}$$

Dividing both sides by B_n:

$$\frac{X_n}{B_n} = e^{-r_n} E^Q\left[\frac{X_{n+1}}{B_n} \mid \mathcal{F}_n\right] = E^Q\left[\frac{X_{n+1}}{B_{n+1}} \mid \mathcal{F}_n\right],$$

so $\widetilde{X}_n = E^Q[\widetilde{X}_{n+1} \mid \mathcal{F}_n]$, which shows that $\widetilde{X}$ is a Q-martingale. Equation (1.9) then follows from the martingale property applied iteratively. For (1.10), set $n = k+1$ in the first step above and use $e^{-r_k} = p(k,k+1)$:

$$X_k = e^{-r_k} E^Q[X_{k+1} \mid \mathcal{F}_k] = p(k,k+1)\, E^Q[X_{k+1} \mid \mathcal{F}_k]. \qquad \square$$

1.3 Stochastic Models for Interest Rates

In the previous sections we introduced the bond market, the family of interest rates derived from the term structure of zero-coupon bond prices, and the risk-neutral pricing framework. The pricing formula (1.7) shows that, once the dynamics of the short rate r

are specified under a martingale measure Q, bond prices are determined as conditional expectations of the discount factor. We now turn to the question of how to model these dynamics. Although this thesis works in discrete time, it is useful to recall the historical development of continuous-time short-rate models, since a direct Euler-Maruyama discretization may fail to preserve some features of the original model, such as positivity, support restrictions, or distributional properties. For this reason, the choice of the model to discretize is itself non-trivial.

1.3.1 Short-Rate Models versus Forward-Rate Models

The stochastic modeling of interest rates can be approached from two fundamentally different perspectives. One can either take the short rate as the primitive object of the model, or one can take the entire family of zero-coupon bond prices as primary assets and prescribe their joint dynamics directly. We discuss both approaches in turn.

Short-Rate Models

In a short-rate model one specifies the dynamics of the short rate r under a martingale measure Q. Once the dynamics of r are given, the prices of T-bonds are not assigned independently: they are obtained via the no-arbitrage pricing formula

$$p(n,N) = E^Q\left[\exp\left(-\sum_{k=n}^{N-1} r_k\right) \,\middle|\, \mathcal{F}_n\right], \qquad 0 \le n \le N \le \bar{N}, \tag{1.11}$$

where the T-bond is treated as a derivative of the underlying r with unit payoff at maturity. The entire term structure $\{p(n,N)\}_{N=n,\ldots,\bar{N}}$ is therefore determined by the short-rate process alone.

This approach offers clear conceptual appeal: a single Markovian state variable drives the entire yield curve, leading to tractable and computationally efficient implementations. However, as noted in [16], three structural issues are intrinsic to this framework:

(i) The short rate is not a traded asset, so the market is in general incomplete and the martingale measure is not unique. The standard practice is to specify the dynamics of r directly under a chosen Q, thereby implicitly fixing the choice of measure.

(ii) The initial term structure $\{p^*(0,N)\}_{N=1,\ldots,\bar{N}}$ is an observable input that the model must reproduce. In time-homogeneous models with constant coefficients this calibration condition cannot in general be satisfied exactly.

(iii) A single state variable is unlikely to capture realistically the joint dynamics of all bond prices.

Forward-Rate Models

The forward-rate paradigm takes the opposite point of view: rather than modeling the short rate and deriving bond prices, one directly assigns the dynamics of the entire family of bond prices. The current term structure then becomes part of the initial condition of the model and is matched by construction.

The price to pay for this generality is dual. First, since the dynamics of bond prices are assigned exogenously, it is necessary to impose conditions ensuring absence of arbitrage. Second, the state variable is typically larger than in a one-factor short-rate model. In a continuous-maturity framework the state is, in general, the entire forward curve and is therefore infinite-dimensional. In the finite-maturity grid considered in this thesis, however, the state is finite-dimensional, although its dimension may still be large if many maturities are included. From a computational standpoint, this makes forward-rate models more demanding than the one-factor short-rate framework adopted below.

The Choice of Paradigm in This Thesis

For derivatives whose payoff depends on a single rate or on a limited section of the curve, as is the case for the backward-looking caplets and floorlets studied in this work, the short-rate approach is well suited and offers significant analytical advantages. We therefore adopt this framework throughout the thesis.

1.3.2 Classical One-Factor Short-Rate Models in Continuous Time

We now review the most influential one-factor short-rate models proposed in the literature (partially based on [3]). Each model is presented in continuous time under the risk-neutral measure Q. As anticipated, this exposition is not merely historical: of the four time-homogeneous models discussed below, the Dothan, Exponential-Vasicek and CIR models do not preserve the key properties of the continuous-time model if they are transformed through an Euler-Maruyama discretization. By contrast, the Vasicek model and its Hull-White extension have additive Gaussian noise, for which the Euler-Maruyama scheme preserves the properties of the continuous-time models. This is the decisive argument in favor of the discrete Hull-White model adopted in this thesis.

The Vasicek Model (1977)

Vasicek [21] proposed the first analytically tractable continuous-time short-rate model. Under the risk-neutral measure, the short rate follows an Ornstein-Uhlenbeck process:

$$dr(t) = k\big[\theta - r(t)\big]\,dt + \sigma\,dW(t), \qquad r_0 \in \mathbb{R}, \tag{1.12}$$

where $W = (W(t))_{t\geq 0}$ is a Q-Brownian motion.

Integrating equation (1.12), we obtain, for each $s \leq t$,

$$r(t) = r(s)e^{-k(t-s)} + \theta\left(1 - e^{-k(t-s)}\right) + \sigma \int_s^t e^{-k(t-u)}\, dW(u). \tag{1.13}$$

Therefore, $r(t)$ conditional on $\mathcal{F}_s$ is normally distributed with mean and variance given respectively by

$$\begin{aligned} \mathbb{E}\left[r(t) \mid \mathcal{F}_s\right] &= r(s)e^{-k(t-s)} + \theta\left(1 - e^{-k(t-s)}\right), \\ \operatorname{Var}\left[r(t) \mid \mathcal{F}_s\right] &= \frac{\sigma^2}{2k}\left[1 - e^{-2k(t-s)}\right]. \end{aligned} \tag{1.14}$$

where $k > 0$ is the speed of mean reversion, $\theta \in \mathbb{R}$ is the long-run mean and $\sigma > 0$ is a constant volatility. The solution shows that $r_t \,|\, \mathcal{F}_s$ is Gaussian with conditional mean converging to θ as $t \to \infty$ and conditional variance converging to $\sigma^2/(2k)$ as $t \to \infty$, a feature known as mean reversion.

The Gaussian distribution is the source of both the model's strength and its main weakness. On the positive side, it leads to closed-form expressions for zero-coupon bond prices and European options on bonds and it is an affine term structure model. On the negative side, the short rate can take negative values with positive probability. More importantly for practice, the three constant parameters k, θ, σ constrain the family of admissible yield curves and the model cannot be exactly calibrated to an arbitrary observed initial term structure. This endogenous nature of the term structure is the principal reason why the original Vasicek model is rarely used directly for calibration-sensitive derivative pricing.

The Dothan Model (1978)

Motivated by the desire for a strictly positive short rate, Dothan [6] proposed modeling r as a geometric Brownian motion:

$$dr(t) = a\, r(t)\, dt + \sigma\, r(t)\, dW(t), \tag{1.15}$$

with $a \in \mathbb{R}$ and $\sigma > 0$.

The dynamics are easily integrated as follows:

$$r(t) = r(s) \exp\left\{\left(a - \frac{1}{2}\sigma^2\right)(t-s) + \sigma\left(W(t) - W(s)\right)\right\}, \tag{1.16}$$

for $s \leq t$. Hence, $r(t)$ conditional on $\mathcal{F}_s$ is lognormally distributed with mean and variance given by

$$\begin{aligned} \mathbb{E}\left[r(t) \mid \mathcal{F}_s\right] &= r(s)e^{a(t-s)}, \\ \operatorname{Var}\left[r(t) \mid \mathcal{F}_s\right] &= r^2(s)e^{2a(t-s)}\left(e^{\sigma^2(t-s)} - 1\right). \end{aligned} \tag{1.17}$$

The rate is lognormally distributed and strictly positive. However, mean reversion towards a non-zero level is absent: the rate either grows or decays exponentially depending on the sign of a.

Analytically, the model is not affine and its bond pricing formula, while closed-form, is rather complex since it depends on two integrals of functions involving hyperbolic sines and cosines. No closed-form option pricing formula exists. Moreover, lognormally distributed short-rate models suffer from the *explosion problem*: the bank account $B_t = \exp(\int_0^t r_s\,ds)$ has infinite expectation for any $t > 0$, leading to infinite prices for certain instruments (see, for instance, Sandman and Sondermann 1997, [18]). From the standpoint of discretization, the multiplicative noise $\sigma r_t\,dW_t$ means that the standard Euler-Maruyama scheme does not preserve positivity, a sufficiently negative shock produces a negative r_{n+1}, contradicting the model's defining feature.

The Cox-Ingersoll-Ross Model (1985)

The Cox-Ingersoll-Ross model [4] combines mean reversion with strict positivity:

$$dr(t) = k\big[\theta - r(t)\big]\,dt + \sigma\sqrt{r(t)}\,dW(t), \qquad r_0 > 0, \tag{1.18}$$

with $k, \theta, \sigma > 0$. The square-root diffusion coefficient vanishes at zero and, under the condition $2k\theta \geq \sigma^2$ (which ensures the drift is strong enough to keep the process away from zero), $r(t)$ remains strictly positive almost surely.

Despite the non-Gaussian distribution, the CIR model is affine: bond prices take the form $p(t,T) = A(t,T)e^{-B(t,T)r_t}$ with explicit A and B, and European option prices are available in closed form in terms of the non-central chi-squared distribution. Like Vasicek, however, the time-homogeneous CIR model cannot be exactly calibrated to an arbitrary initial yield curve. From a discretization standpoint, the square-root noise introduces a fundamental difficulty: the Euler-Maruyama scheme may produce negative values of r_{n+1}, rendering $\sigma\sqrt{r_n}$ undefined at the next step.

The Exponential-Vasicek Model

The exponential-Vasicek model achieves positivity and mean reversion by requiring that $y(t) := \ln r(t)$ follows an Ornstein-Uhlenbeck process: $dy(t) = [\theta - ay(t)]\,dt + \sigma\,dW(t)$, so that by Itô's formula

$$dr(t) = r(t)\Big[\theta + \tfrac{\sigma^2}{2} - a\ln r(t)\Big]dt + \sigma\,r(t)\,dW(t). \tag{1.19}$$

Remembering the solution of the Vasicek model, the process r, for each $s \leq t$, is explicitly given by

$$r(t) = \exp\left\{\ln r(s) e^{-a(t-s)} + \frac{\theta}{a}\left(1 - e^{-a(t-s)}\right) + \sigma \int_s^t e^{-a(t-u)}\, dW(u)\right\}. \tag{1.20}$$

Therefore, $r(t)$ conditional on $\mathcal{F}_s$ is lognormally distributed with first and second moments given respectively by

$$\begin{aligned}
\mathbb{E}\left[r(t) \mid \mathcal{F}_s\right] &= \exp\left\{\ln r(s) e^{-a(t-s)} + \frac{\theta}{a}\left(1 - e^{-a(t-s)}\right) + \frac{\sigma^2}{4a}\left(1 - e^{-2a(t-s)}\right)\right\}, \\
\mathbb{E}\left[r^2(t) \mid \mathcal{F}_s\right] &= \exp\left\{2\ln r(s) e^{-a(t-s)} + \frac{2\theta}{a}\left(1 - e^{-a(t-s)}\right) + \frac{\sigma^2}{a}\left(1 - e^{-2a(t-s)}\right)\right\}.
\end{aligned} \tag{1.21}$$

Therefore, contrary to the Dothan process, this r is always mean-reverting since

$$\lim_{t\to\infty} \mathbb{E}\left[r(t) \mid \mathcal{F}_s\right] = \exp\left(\frac{\theta}{a} + \frac{\sigma^2}{4a}\right). \tag{1.22}$$

Notice that also the variance converges to a finite value since

$$\lim_{t\to\infty} \mathrm{Var}\left[r(t) \mid \mathcal{F}_s\right] = \exp\left(\frac{2\theta}{a} + \frac{\sigma^2}{2a}\right)\left[\exp\left(\frac{\sigma^2}{2a}\right) - 1\right]. \tag{1.23}$$

The short rate is lognormally distributed, strictly positive, and mean-reverting towards a finite level. However, bond prices and option prices cannot be expressed analytically. The model is not affine, inherits the explosion problem of all lognormal short-rate models, and the Euler-Maruyama scheme on r does not preserve positivity. The Black-Karasinski model [2] extends this by allowing time-dependent coefficients to fit the observed term structure, but retains all of these drawbacks.

Chapter 2

The Discrete Hull-White Model

In the previous chapter we reviewed the principal one-factor short-rate models in the literature. In this chapter we develop the discrete Hull-White model in full detail. We begin by recalling the continuous-time Hull-White model and its key properties, then motivate and carry out the Euler-Maruyama discretization. The core of the chapter is the derivation of the affine term structure possessed by the discrete model: we obtain explicit closed-form formulas for the bond pricing coefficients B_n^N and A_n^N, including the case of non-zero mean reversion $a \neq 0$, and derive an explicit calibration formula for the drift parameters Φ_n in terms of the observed initial term structure.

2.1 The Continuous-Time Hull-White Model

The Hull-White model [11] extends the Vasicek model (1.12) by replacing the constant long-run mean with a deterministic function of time. Under a risk-neutral measure Q, the short rate satisfies

$$dr(t) = \big[\vartheta(t) - a\, r(t)\big]\, dt + \sigma\, dW(t), \tag{2.1}$$

where $a > 0$ is the speed of mean reversion, $\sigma > 0$ is a constant volatility, and $\vartheta(t)$ is a deterministic function to be calibrated. The key innovation relative to Vasicek is precisely the time dependence of $\vartheta(t)$: by choosing it appropriately, the model can reproduce any observed initial term structure exactly, a property that was absent in all the time-homogeneous models reviewed in Section 1.3.2.

The solution of (2.1) is obtained by the same integrating-factor technique as for the Vasicek model. For $s \leq t$:

$$r(t) = r(s)\, e^{-a(t-s)} + \int_s^t e^{-a(t-u)} \vartheta(u)\, du + \sigma \int_s^t e^{-a(t-u)}\, dW(u). \tag{2.2}$$

Since ϑ is deterministic and the stochastic integral is Gaussian, $r_t \,|\, \mathcal{F}_s$ is normally distributed with

$$E^Q[r(t) \mid \mathcal{F}_s] = r_s\, e^{-a(t-s)} + \int_s^t e^{-a(t-u)} \vartheta(u)\, du,$$
$$\mathrm{Var}^Q[r(t) \mid \mathcal{F}_s] = \frac{\sigma^2}{2a}\big(1 - e^{-2a(t-s)}\big). \tag{2.3}$$

The conditional variance is identical to that of the Vasicek model and does not depend on ϑ: the time-varying drift affects only the conditional mean, not the volatility structure. This is a direct consequence of the additive noise in (2.1).

The Hull-White model inherits from Vasicek all the analytical properties that make Gaussian short-rate models attractive:

- It possesses an *affine term structure*: zero-coupon bond prices take the form $p(t,T) = A(t,T)\, e^{-B(t,T)\, r_t}$ with explicit functions A and B.
- Closed-form pricing formulas exist for European options on zero-coupon bonds, and consequently for caplets and floorlets.
- The model is analytically tractable for calibration: $\vartheta(t)$ can be expressed explicitly in terms of the observed market instantaneous forward rate curve.

The price of this tractability is the same as in Vasicek: the short rate is Gaussian and may take negative values with positive probability. For a suitable choice of the model parameters, this probability can be made very small over the relevant maturities, although it is not negligible for all possible parameter configurations.

The following table summarizes the properties of the models discussed in Section 1.3.2 and highlights why the Hull-White model is the natural choice for the purposes of this thesis.

Model	**Mean rev.**	$r \geq 0$	**Exact fit**	**Affine TS**
Vasicek	✓	–	–	✓
Dothan	–	✓	–	–
CIR	✓	✓	–	✓
Exp-Vasicek	✓	✓	–	–
Hull-White	✓	–	✓	✓

It should be noted that the table refers to the standard time-homogeneous versions of the models discussed above. In particular, the entry "Exact fit" for the CIR model refers to the classical CIR specification. Time-inhomogeneous extensions of the CIR model, such as the Hull-White extension, can also be introduced in order to fit a given initial term structure. In this thesis, however, we focus on the Gaussian Hull–White model because it combines exact fitting of the initial term structure with affine Gaussian dynamics and particularly tractable discrete-time pricing formulas.

2.2 The Euler-Maruyama Discretization

The derivatives studied in this thesis, SOFR caplets and floorlets, are written on backward-looking compounded rates defined by a discrete product of overnight fixings. The underlying rate is not a continuous-time process: it is observed and compounded at a finite set of discrete dates. A discrete-time model is therefore the natural and internally consistent framework for pricing these instruments.

The standard method for obtaining a discrete-time model from a continuous-time stochastic differential equation is the *Euler-Maruyama scheme*. The idea is to replace the infinitesimal increments dt and dW_t with finite differences over a time step Δ:

$$dt \longrightarrow \Delta, \qquad dW_t \longrightarrow W_{t+\Delta} - W_t \sim \mathcal{N}(0, \Delta).$$

Since a Gaussian random variable with variance Δ can be written as $\sqrt{\Delta}\,\omega$ with $\omega \sim \mathcal{N}(0,1)$, the scheme applied to the Hull-White SDE (2.1) gives:

$$r_{n+1} = r_n + \big[\Phi_n - a\,r_n\big]\Delta + \sigma\sqrt{\Delta}\,\omega_n, \qquad \omega_n \overset{\text{iid}}{\sim} \mathcal{N}(0,1), \tag{2.4}$$

where Φ_n denotes the discrete-time analogue of the continuous-time drift function $\vartheta(t)$ at the n-th grid point.

A useful feature of this discretization is that it preserves the Gaussian structure of the continuous-time model. In the continuous-time model, $r_t \,|\, \mathcal{F}_s$ is Gaussian for any $t > s$; the Euler-Maruyama scheme (2.4) produces a process that is also exactly Gaussian at every step, with one-step conditional mean and variance that are first-order approximations of the exact continuous-time conditional moments. This is a direct consequence of the additive noise structure: the diffusion coefficient σ does not depend on r_t, so no nonlinearity is introduced by the discretization. By contrast, for models with multiplicative noise, Dothan (σr_t), CIR ($\sigma\sqrt{r_t}$), Exponential-Vasicek (σr_t), the Euler-Maruyama scheme distorts the distribution and may produce values outside the support of the continuous-time process.

For notational convenience we define

$$c := 1 - a\Delta, \tag{2.5}$$

so that the recursion (2.4) can be written compactly as

$$r_{n+1} = c\,r_n + \Phi_n\Delta + \sigma\sqrt{\Delta}\,\omega_n. \tag{2.6}$$

In the remainder of this thesis we work exclusively with this discrete-time model. We assume throughout that a and σ are constant, while Φ_n is a deterministic sequence calibrated to the observed initial term structure, as derived in Section 2.5.

Remark 2.1 (Stability of the autoregressive recursion). The recursion (2.6) has the form of a first-order autoregressive process, $r_{n+1} = c\,r_n + \Phi_n\Delta + \sigma\sqrt{\Delta}\,\omega_n$, with autoregressive coefficient $c = 1 - a\Delta$ and a deterministic, time-varying intercept $\Phi_n\Delta$. Following the standard analysis of AR(1) models in econometrics, the recursion is stable, equivalently, the root $1/c$ of the lag polynomial $1 - cL$ lies outside the unit circle, if and only if

$$|c| = |1 - a\Delta| < 1 \quad \Longleftrightarrow \quad 0 < a\Delta < 2.$$

Under this condition the influence of the initial value r_0 and of past shocks decays geometrically, so the discrete process inherits the mean-reverting behavior of the continuous-time model. The time-varying intercept $\Phi_n\Delta$ affects only the conditional mean and not the stability of the recursion. The lower bound $a\Delta > 0$ holds automatically since $a > 0$ and $\Delta > 0$, so the binding constraint is the upper bound $a < 2/\Delta$. This bound is a discretization effect: the exact one-step coefficient of the continuous-time dynamics (2.2) is $e^{-a\Delta} \in (0,1)$ for every $a > 0$, whereas the Euler–Maruyama scheme replaces it by the first-order term $1 - a\Delta$, which can leave $(0,1)$ when $a\Delta$ is large. For $0 < a\Delta < 1$ one has $c \in (0,1)$ and the decay toward the mean is monotone, matching the sign of $e^{-a\Delta}$; for $1 < a\Delta < 2$ the coefficient is negative and the scheme exhibits oscillations with no continuous-time counterpart. We therefore assume throughout the natural regime $0 < a\Delta < 1$, which guarantees stability and reproduces the qualitative behavior of the continuous-time Hull–White model.

Stability should not be confused with stationarity. Even under $0 < a\Delta < 1$, the conditional mean $\mathbb{E}[r_n \mid \mathcal{F}_0]$ is a geometrically weighted average of the entire calibrated sequence $\Phi_0, \ldots, \Phi_{n-1}$ and does not converge, since Φ_q varies with q. The process is therefore non-stationary in mean, the non-stationarity originating entirely in the time-varying intercept fixed by the calibration to the initial term structure, and not in the instability of the recursion. The conditional variance, on the other hand, is unaffected by the drift because the noise is additive, and under $0 < a\Delta < 2$ it converges to the finite limit $\sigma^2/(a(2 - a\Delta))$, which recovers the continuous-time stationary variance $\sigma^2/(2a)$ as $\Delta \to 0$. Only the time-homogeneous case $\Phi_q \equiv \Phi$ would be asymptotically stationary, with long-run mean Φ/a.

2.3 The Affine Term Structure

We now show that the discrete Hull-White model possesses an affine term structure, following the general framework developed in [16], and specializing it to the Hull-White dynamics (2.4).

2.3.1 Short Models and the Transition Kernel

As discussed in Section 1.3.1, in a short model the dynamics of r are specified directly under a martingale measure Q, and bond prices are obtained through the pricing formula (1.11). We assume that r is a Markov process defined on a filtered probability space $(\Omega, \mathcal{F}, Q, (\mathcal{F}_n))$. We denote by Q_n the transition kernel of r_n, seen as a Markov chain under Q: for each integrable function φ on $\mathbb{R}$,

$$(Q_n\varphi)(r_n) := E^Q[\varphi(r_{n+1}) \mid \mathcal{F}_n] = \int_{\mathbb{R}} \varphi(\varrho)\, Q_n(r_n, d\varrho). \tag{2.7}$$

The Markov property ensures that the right-hand side depends on $\mathcal{F}_n$ only through r_n.

For the discrete Hull-White model (2.4), the transition $r_{n+1} \,|\, r_n$ is Gaussian:

$$r_{n+1} \sim \mathcal{N}\big(r_n + (\Phi_n - ar_n)\Delta,\ \sigma^2\Delta\big)\,, \tag{2.8}$$

so that $Q_n(r_n, \cdot) \sim \mathcal{N}_{r_n+(\Phi_n-ar_n)\Delta,\sigma^2\Delta}$.

2.3.2 Bond Prices as Iterated Expectations

To compute bond prices we set $\varphi_0(r) := e^{-r}$ and define the family of functions (φ_n^N), with $0 \le n < N \le \bar{N}$, by the backward recursion

$$\begin{cases} \varphi_{N-1}^N(r) = 1, \\ \varphi_{n-1}^N(r) = Q_{n-1}\big(\varphi_0 \cdot \varphi_n^N\big)(r), \qquad 1 \le n \le N-1. \end{cases} \tag{2.9}$$

Proposition 2.1. *Under the martingale measure Q,*

$$p(n, N) = e^{-r_n}\, \varphi_n^N(r_n), \qquad 0 \le n < N \le \bar{N}, \tag{2.10}$$

with φ_n^N defined in (2.9).

Proof. We proceed by backward induction on n, with N fixed. For $n = N-1$:

$$p(N-1, N) = e^{-r_{N-1}} = e^{-r_{N-1}} \cdot 1 = e^{-r_{N-1}}\, \varphi_{N-1}^N(r_{N-1}),$$

since $\varphi_{N-1}^N \equiv 1$.

Suppose the statement holds for n, we prove it for $n-1$. Since Q is a martingale measure, by Proposition 1.1:

$$\begin{aligned}
p(n-1,N) &= p(n-1,n)\, E^Q[p(n,N) \mid \mathcal{F}_{n-1}] \\
&= e^{-r_{n-1}} E^Q\big[e^{-r_n}\, \varphi_n^N(r_n) \mid \mathcal{F}_{n-1}\big] \\
&= e^{-r_{n-1}} Q_{n-1}\big(\varphi_0 \cdot \varphi_n^N\big)(r_{n-1}) = e^{-r_{n-1}}\, \varphi_{n-1}^N(r_{n-1}),
\end{aligned}$$

where we used the induction hypothesis, the definition of Q_{n-1}, and the recursion (2.9). $\square$

2.3.3 Affine Models and the Moment Generating Function

The representation (2.10) expresses bond prices in terms of the functions φ_n^N, but does not yet give a closed-form expression. To obtain one, we introduce the *moment generating function* of the transition kernel:

$$\widetilde{Q}_n(r,\lambda) := \int_{\mathbb{R}} e^{-\lambda \varrho}\, Q_n(r, d\varrho), \qquad r, \lambda \in \mathbb{R}. \tag{2.11}$$

This corresponds to the usual moment generating function evaluated at $-\lambda$, or equivalently to the Laplace transform. This choice is convenient in interest-rate pricing, where discount factors naturally involve negative exponentials. Since the exponential function is positive, the integral in (2.11) is well defined (possibly equal to $+\infty$) for each r, λ.

The following proposition shows that if $\widetilde{Q}_n$ has an exponential-affine form, then the functions φ_n^N, and consequently the bond prices, also have an exponential-affine form.

Proposition 2.2 (Affine Term Structure). *Suppose there exist functions f_n and g_n such that*

$$\widetilde{Q}_n(r,\lambda) = \exp(-f_n(\lambda) - g_n(\lambda)\, r)\,, \qquad r, \lambda \in \mathbb{R}. \tag{2.12}$$

Then for each n, $0 \le n < \bar{N}$, the functions φ_n^N defined in (2.9) *have the exponential-affine form*

$$\varphi_n^N(r) = \exp\big(-A_n^N - B_n^N\, r\big)\,, \qquad r \in \mathbb{R}, \tag{2.13}$$

where the constants A_n^N, B_n^N satisfy the recursion

$$\begin{cases} A_{N-1}^N = B_{N-1}^N = 0, \\ A_{n-1}^N = A_n^N + f_{n-1}(1 + B_n^N), \\ B_{n-1}^N = g_{n-1}(1 + B_n^N). \end{cases} \tag{2.14}$$

In particular, the bond prices are given by

$$p(n,N) = \exp\big(-A_n^N - (1 + B_n^N)\, r_n\big)\,, \qquad 0 \le n < N \le \bar{N}. \tag{2.15}$$

Proof. By backward induction on n, with N fixed. For $n = N-1$: $\varphi_{N-1}^N \equiv 1 = \exp(-0 - 0 \cdot r)$, which matches (2.13) with $A_{N-1}^N = B_{N-1}^N = 0$.

Suppose the statement holds for n; we prove it for $n-1$. By the recursion (2.9):

$$\begin{aligned}\varphi_{n-1}^N(r) &= Q_{n-1}(\varphi_0 \cdot \varphi_n^N)(r) = \int_{\mathbb{R}} e^{-\varrho}\, \varphi_n^N(\varrho)\, Q_{n-1}(r, d\varrho) \\ &= \int_{\mathbb{R}} e^{-\varrho} \exp\bigl(-A_n^N - B_n^N \varrho\bigr)\ Q_{n-1}(r, d\varrho) \\ &= e^{-A_n^N} \int_{\mathbb{R}} e^{-(1+B_n^N)\varrho}\, Q_{n-1}(r, d\varrho) \\ &= e^{-A_n^N}\, \widetilde{Q}_{n-1}\bigl(r,\, 1 + B_n^N\bigr)\,.\end{aligned}$$

Applying the assumption (2.12) with $\lambda = 1 + B_n^N$:

$$\begin{aligned}\varphi_{n-1}^N(r) &= \exp\bigl(-A_n^N - f_{n-1}(1 + B_n^N) - g_{n-1}(1 + B_n^N)\, r\bigr) \\ &= \exp\bigl(-A_{n-1}^N - B_{n-1}^N\, r\bigr)\,,\end{aligned}$$

provided we define A_{n-1}^N and B_{n-1}^N as in (2.14). The bond price formula (2.15) then follows from Proposition 2.1. □

Definition 2.1 (Affine Term Structure Model)**.** *A short rate model for which* (2.15) *holds is called an affine term structure model.*

2.3.4 Application to the Discrete Hull-White Model

It remains to verify that the discrete Hull-White model (2.4) satisfies the hypothesis of Proposition 2.2, i.e. that its moment generating function $\widetilde{Q}_n$ has the exponential-affine form (2.12). This is established by the following lemma.

Lemma 2.1. *For any $\mu \in \mathbb{R}$ and $\sigma^2 > 0$,*

$$\int_{\mathbb{R}} e^{-\lambda\rho}\, \mathcal{N}_{\mu,\sigma^2}(d\rho) = \exp\left(-\lambda\mu + \frac{\lambda^2\sigma^2}{2}\right). \tag{2.16}$$

Proof. This is the standard identity for the moment generating function of a Gaussian distribution $\mathcal{N}(\mu, \sigma^2)$, evaluated at $-\lambda$:

$$\begin{aligned}
&\int_{\mathbb{R}} e^{-\lambda\varrho}\, \frac{1}{\sigma\sqrt{2\pi}} \exp\Big(-\frac{(\varrho-\mu)^2}{2\sigma^2}\Big)\, d\varrho \\
&\quad= \frac{1}{\sigma\sqrt{2\pi}} \int_{\mathbb{R}} \exp\Big(-\frac{2\sigma^2\lambda\varrho + \varrho^2 - 2\varrho\mu + \mu^2}{2\sigma^2}\Big)\, d\varrho \\
&\quad= \exp\Big(\frac{-\mu^2 + (\lambda\sigma^2 - \mu)^2}{2\sigma^2}\Big)\, \frac{1}{\sigma\sqrt{2\pi}} \int_{\mathbb{R}} \exp\Big(-\frac{(\varrho - (\mu - \lambda\sigma^2))^2}{2\sigma^2}\Big)\, d\varrho \\
&\quad= \exp\Big(\frac{-\mu^2 + (\lambda\sigma^2 - \mu)^2}{2\sigma^2}\Big) \\
&\quad= \exp\Big(-\lambda\mu + \frac{\lambda^2\sigma^2}{2}\Big)\,. \qquad \square
\end{aligned}$$

Since for the discrete Hull-White model $Q_n(r_n, \cdot) = \mathcal{N}_{r_n+(\Phi_n - a r_n)\Delta,\, \sigma^2\Delta}$, applying Lemma 2.1 with $\mu = r_n + (\Phi_n - a r_n)\Delta$ and $\sigma^2\Delta$ in place of σ^2 gives:

$$\begin{aligned}
\widetilde{Q}_n(r_n, \lambda) &= \exp\Big(-\lambda\big(r_n + (\Phi_n - a r_n)\Delta\big) + \frac{\lambda^2\sigma^2\Delta}{2}\Big) \\
&= \exp\Big(-\Phi_n\Delta\lambda + \frac{\sigma^2\Delta}{2}\lambda^2 - \lambda(1 - a\Delta)\, r_n\Big) \\
&= \exp(-f_n(\lambda) - g_n(\lambda)\, r_n)\,, \qquad (2.17)
\end{aligned}$$

with

$$f_n(\lambda) = \Phi_n\Delta\lambda - \frac{\sigma^2\Delta}{2}\lambda^2, \qquad g_n(\lambda) = \lambda(1 - a\Delta). \qquad (2.18)$$

Since $\widetilde{Q}_n$ has the exponential-affine form (2.12), Proposition 2.2 applies and the discrete Hull-White model is an affine term structure model. Substituting (2.18) into the general recursion (2.14):

$$\boxed{\begin{cases}
A^N_{N-1} = B^N_{N-1} = 0, \\[2ex]
A^N_{n-1} = A^N_n + \Phi_{n-1}\Delta\,(1 + B^N_n) - \dfrac{\sigma^2\Delta}{2}\,(1 + B^N_n)^2, \\[2ex]
B^N_{n-1} = (1 + B^N_n)(1 - a\Delta).
\end{cases}} \qquad (2.19)$$

The bond prices are given by

$$p(n, N) = \exp\big(-A^N_n - (1 + B^N_n)\, r_n\big)\,, \qquad 0 \le n < N \le \bar{N}. \qquad (2.20)$$

Remark 2.2. The forward rates $L(n, N)$ and $R(n, N)$ can also be expressed in terms of

A_n^N and B_n^N. From (2.20) and the definitions in Section 1.1:

$$R(n,N) = (A_n^{N+1} - A_n^N) + (B_n^{N+1} - B_n^N)\, r_n, \tag{2.21}$$
$$L(n,N) = \frac{1}{\Delta}\big(\exp\big((A_n^{N+1} - A_n^N) + (B_n^{N+1} - B_n^N) r_n\big) - 1\big)\,. \tag{2.22}$$

2.4 Explicit Formulas for B_n^N and A_n^N

The recursion (2.19) can be solved explicitly. We treat the two coefficients separately, beginning with B_n^N which is the simpler of the two.

2.4.1 The Coefficient B_n^N

The recursion for B_n^N reads

$$B_{N-1}^N = 0, \qquad B_{n-1}^N = (1 + B_n^N)\, c, \tag{2.23}$$

where $c = 1 - a\Delta$ as defined in (2.5). To solve it, we introduce the auxiliary variable

$$C_n := 1 + B_n^N. \tag{2.24}$$

In terms of C_n, the recursion becomes $C_{n-1} = 1 + c\, C_n$ with terminal condition $C_{N-1} = 1$. We unroll this backward recursion step by step.

Step 1. $C_{N-1} = 1$.
Step 2. $C_{N-2} = 1 + c\, C_{N-1} = 1 + c$.
Step 3. $C_{N-3} = 1 + c\, C_{N-2} = 1 + c(1+c) = 1 + c + c^2$.
Step 4. $C_{N-4} = 1 + c\, C_{N-3} = 1 + c + c^2 + c^3$.

By induction, we have

$$C_n = \sum_{j=0}^{m-1} c^j = \frac{1-c^m}{1-c} = \frac{1-(1-a\Delta)^{N-n}}{a\Delta}, \qquad m = N - n \geq 1, \tag{2.25}$$

where the last equality uses the geometric sum formula and $1 - c = a\Delta$. Therefore:

$$\boxed{B_n^N = C_n - 1 = \sum_{j=1}^{N-n-1} c^j = \frac{(1-a\Delta)\big(1-(1-a\Delta)^{N-n-1}\big)}{a\Delta}, \qquad n \leq N-2.} \tag{2.26}$$

For $n = N - 1$ we have $B_{N-1}^N = 0$, consistent with (2.26) since the sum is empty.

Remark 2.3 (Correction of Pascucci-Runggaldier (4.29)). The explicit formula given in equation (4.29) of [16] reads $B_n^N = (1 - a\Delta)^{N-n-2}(N - n - 1 - a\Delta)$. This is correct for

$a = 0$, where both expressions reduce to $B_n^N = N - n - 1$, and also happens to agree with (2.26) for $N - n \leq 3$. However, for $a \neq 0$ and $N - n \geq 4$, the two formulas differ. For instance, when $N - n = 4$, the recursion gives

$$B_n^N = c + c^2 + c^3.$$

The formula in [16] gives

$$c^2(3 - a\Delta) = c^2(2 + c) = 2c^2 + c^3,$$

since $a\Delta = 1 - c$. The difference is therefore

$$(c + c^2 + c^3) - (2c^2 + c^3) = c - c^2 = c(1 - c) = a\Delta c,$$

which is non-zero for $a \neq 0$ and $c \neq 0$. Hence the formula in [16] cannot be correct in general.

2.4.2 The Coefficient A_n^N

The recursion for A_n^N reads

$$A_{N-1}^N = 0, \qquad A_{n-1}^N = A_n^N + \Phi_{n-1}\Delta\,(1 + B_n^N) - \frac{\sigma^2\Delta}{2}\,(1 + B_n^N)^2. \tag{2.27}$$

This is a telescoping recursion: each step adds a term that depends on the already-known coefficients B_k^N. Denoting

$$\tau_k := \Phi_k\Delta\,(1 + B_{k+1}^N) - \frac{\sigma^2\Delta}{2}\,(1 + B_{k+1}^N)^2, \tag{2.28}$$

the recursion is $A_{n-1}^N = A_n^N + \tau_{n-1}$, with $A_{N-1}^N = 0$. Telescoping from n to $N - 1$:

$$\boxed{A_n^N = \sum_{k=n}^{N-2} \left(\Phi_k\Delta\,(1 + B_{k+1}^N) - \frac{\sigma^2\Delta}{2}\,(1 + B_{k+1}^N)^2 \right), \qquad n \leq N - 2.} \tag{2.29}$$

2.5 Calibration of Φ_n to the Initial Term Structure

The drift parameters Φ_n are determined by the requirement that the model reproduces the observed initial term structure exactly:

$$p(0, N) = p^*(0, N), \qquad N = 1, \dots, \bar{N}. \tag{2.30}$$

We derive an explicit formula for Φ_n in terms of the observed compounded forward rates $R(0,n) = \log \frac{p^*(0,n)}{p^*(0,n+1)}$.

2.5.1 Preliminary Notation

For convenience we define the partial sums of powers of $c = 1 - a\Delta$:

$$S_m := \sum_{j=0}^{m-1} c^j = \frac{1-c^m}{1-c} = \frac{1-(1-a\Delta)^m}{a\Delta}, \qquad m \geq 1, \tag{2.31}$$

with the convention $S_0 = 0$. Note that $1 + B_n^N = S_{N-n}$ by (2.25) and (2.24).

2.5.2 Derivation of the Calibration Formula

From the affine bond price formula (2.15), evaluated at time 0, we have

$$-\log p(0,N) = A_0^N + (1 + B_0^N) r_0 = A_0^N + S_N r_0, \tag{2.32}$$

since $1 + B_0^N = S_N$.

The one-period compounded forward rate observed at time 0 is

$$R(0,n) = \log \frac{p(0,n)}{p(0,n+1)}. \tag{2.33}$$

Using (2.32), we obtain

$$\begin{aligned} R(0,n) &= \left(A_0^{n+1} - A_0^n\right) + \left(S_{n+1} - S_n\right) r_0 \\ &= \left(A_0^{n+1} - A_0^n\right) + c^n r_0, \end{aligned} \tag{2.34}$$

because $S_{n+1} - S_n = c^n$.

We now compute $A_0^{n+1} - A_0^n$. From (2.29) and the identity

$$1 + B_{k+1}^N = S_{N-k-1},$$

we have

$$A_0^N = \sum_{k=0}^{N-2} \left(\Phi_k \Delta S_{N-k-1} - \frac{\sigma^2 \Delta}{2} S_{N-k-1}^2 \right). \tag{2.35}$$

Therefore,

$$\begin{aligned}A_0^{n+1} - A_0^n &= \sum_{k=0}^{n-1} \left(\Phi_k \Delta S_{n-k} - \frac{\sigma^2 \Delta}{2} S_{n-k}^2 \right) - \sum_{k=0}^{n-2} \left(\Phi_k \Delta S_{n-k-1} - \frac{\sigma^2 \Delta}{2} S_{n-k-1}^2 \right) \\ &= \Delta \sum_{k=0}^{n-1} \Phi_k \big(S_{n-k} - S_{n-k-1} \big) - \frac{\sigma^2 \Delta}{2} \sum_{k=0}^{n-1} \big(S_{n-k}^2 - S_{n-k-1}^2 \big), \end{aligned} \tag{2.36}$$

where we use the convention $S_0 = 0$. Since

$$S_{n-k} - S_{n-k-1} = c^{n-k-1},$$

and the second sum telescopes, we obtain

$$A_0^{n+1} - A_0^n = \Delta \sum_{k=0}^{n-1} \Phi_k c^{n-1-k} - \frac{\sigma^2 \Delta}{2} S_n^2. \tag{2.37}$$

Substituting (2.37) into (2.34) gives

$$R(0, n) = c^n r_0 + \Delta \sum_{k=0}^{n-1} \Phi_k c^{n-1-k} - \frac{\sigma^2 \Delta}{2} S_n^2. \tag{2.38}$$

We now consider the difference $R(0, n+1) - cR(0, n)$. From (2.38),

$$\begin{aligned}R(0, n+1) - cR(0, n) = &\left[c^{n+1} r_0 + \Delta \sum_{k=0}^{n} \Phi_k c^{n-k} - \frac{\sigma^2 \Delta}{2} S_{n+1}^2 \right] \\ &- c \left[c^n r_0 + \Delta \sum_{k=0}^{n-1} \Phi_k c^{n-1-k} - \frac{\sigma^2 \Delta}{2} S_n^2 \right]. \end{aligned} \tag{2.39}$$

The terms involving r_0 cancel, and the sums involving $\Phi_0, \ldots, \Phi_{n-1}$ also cancel. Hence

$$R(0, n+1) - cR(0, n) = \Delta \Phi_n - \frac{\sigma^2 \Delta}{2} \big(S_{n+1}^2 - cS_n^2 \big). \tag{2.40}$$

Therefore,

$$\Delta \Phi_n = R(0, n+1) - cR(0, n) + \frac{\sigma^2 \Delta}{2} \big(S_{n+1}^2 - cS_n^2 \big). \tag{2.41}$$

It remains to simplify the last term. Since

$$S_m = \frac{1 - c^m}{1 - c},$$

we have

$$\begin{aligned} S_{n+1}^2 - cS_n^2 &= \frac{(1-c^{n+1})^2 - c(1-c^n)^2}{(1-c)^2} \\ &= \frac{(1-c)(1-c^{2n+1})}{(1-c)^2} \\ &= \frac{1-c^{2n+1}}{1-c} = S_{2n+1}. \end{aligned} \tag{2.42}$$

Thus

$$\Delta\Phi_n = R(0,n+1) - cR(0,n) + \frac{\sigma^2\Delta}{2} S_{2n+1}. \tag{2.43}$$

Dividing by Δ, and recalling that $c = 1 - a\Delta$ and $S_{2n+1} = (1-c^{2n+1})/(a\Delta)$, we obtain

$$\boxed{\Phi_n = \frac{R(0,n+1) - (1-a\Delta)R(0,n)}{\Delta} + \frac{\sigma^2}{2a\Delta}\left[1-(1-a\Delta)^{2n+1}\right].} \tag{2.44}$$

Using

$$R^*(0,n) = \log\frac{p^*(0,n)}{p^*(0,n+1)},$$

the calibration formula can be written directly in terms of the observed initial discount curve as

$$\boxed{\Phi_n = \frac{1}{\Delta}\left[\log\frac{p^*(0,n+1)}{p^*(0,n+2)} - c\log\frac{p^*(0,n)}{p^*(0,n+1)}\right] + \frac{\sigma^2}{2a\Delta}\left(1-c^{2n+1}\right),} \tag{2.45}$$

where $c = 1 - a\Delta$.

Chapter 3

Backward-Looking Rates and Derivative Pricing

The previous chapter developed the discrete Hull-White model and established its affine term structure. We now turn to the central problem of this thesis: the pricing of interest rate derivatives written on backward-looking compounded rates, of the kind that underlie SOFR-based caps and floors.

The chapter is organized as follows. We first recall the classical framework for caps and floors under forward-looking rates (Section 3.1) and introduce the backward-looking compounded rate together with its measurability properties (Section 3.2). The core of the chapter then derives closed-form prices for backward-looking caplets and floorlets under the discrete Hull-White model, first at the start of the accrual period (Sections 3.3–3.5) and then at any earlier time (Section 3.6).

Within the same section we also price the forward-looking caplet and floorlet at earlier times, obtained as a special case of the same bivariate Gaussian lemma. This makes the two conventions directly comparable: their prices turn out to share an identical Black-type structure and to differ only through the conditional variance. We exploit this to give a decomposition of the backward-looking variance into a pre-accrual and an in-accrual part, and to prove, via Jensen's inequality, that the backward-looking caplet always dominates its forward-looking counterpart with identical terms. The chapter concludes (Section 3.7) with an alternative simple-rate specification, for which we explain the structural obstructions that prevent a closed-form solution.

3.1 Caps, Floors and the Forward-Looking Rate

Consider a sequence of maturities $(N_j)_{j=1,\dots,J}$ with $N_j \in \mathbb{N}$ and

$$1 \leq N_0 < N_1 < \cdots < N_J \leq \bar{N}.$$

Suppose that an agent has to pay interest on a unitary notional capital in each of the intervals $[N_{j-1}, N_j]$. Payments are made at the end of each interval, while the interest rate is established at the beginning. The rate applicable to the j-th interval is the forward-looking simple rate

$$L_j^{\mathrm{fr}} := L(N_{j-1};\, N_{j-1},\, N_j) = \frac{1}{\alpha_j}\left(\frac{1}{p(N_{j-1}, N_j)} - 1\right), \tag{3.1}$$

where

$$\alpha_j := (N_j - N_{j-1})\,\Delta, \qquad j = 1, \dots, J, \tag{3.2}$$

is the accrual factor for the j-th interval. Note that L_j^{fr} is $\mathcal{F}_{N_{j-1}}$-measurable: it is known at the *beginning* of the accrual period. This is the defining feature of forward-looking rates; the superscript "fr" will distinguish it throughout from the backward-looking rate L_j^{br} introduced in Section 3.2.

A *caplet* with strike K on the j-th interval protects the agent against rises in the interest rate above K. Its payoff at time N_j is

$$\alpha_j(L_j^{\mathrm{fr}} - K)^+, \tag{3.3}$$

and its price at time $n \le N_{j-1}$ is

$$\mathrm{Caplet}_j^{\mathrm{fr}}(n) = E^Q\left[D(n, N_j)\,\alpha_j(L_j^{\mathrm{fr}} - K)^+ \,\middle|\, \mathcal{F}_n\right]. \tag{3.4}$$

A *floorlet* has payoff $\alpha_j(K - L_j^{\mathrm{fr}})^+$ and provides protection against falling rates.

A *cap* (respectively *floor*) is a portfolio of caplets (respectively floorlets) over all intervals:

$$\mathrm{Cap}(n) = \sum_{j=1}^{J} \mathrm{Caplet}_j^{\mathrm{fr}}(n), \qquad \mathrm{Floor}(n) = \sum_{j=1}^{J} \mathrm{Floorlet}_j^{\mathrm{fr}}(n). \tag{3.5}$$

Remark 3.1 (Forward-looking rates and measurability)*.* The rate L_j^{fr} is $\mathcal{F}_{N_{j-1}}$-measurable because it depends only on the bond price $p(N_{j-1}, N_j)$, which is known at time N_{j-1}. Consequently, in the pricing formula (3.4), the payoff $\alpha_j(L_j^{\mathrm{fr}} - K)^+$ is known at time N_{j-1} and Corollary 1.1 can be applied with $X_{N_{j-1}} = \alpha_j(L_j^{\mathrm{fr}} - K)^+$ to price the caplet by backward recursion from N_{j-1}, using only the one-step pricing formula.

3.2 The Backward-Looking Compounded Rate

The transition from LIBOR to overnight risk-free rates such as SOFR has introduced a fundamentally different rate structure. Unlike the forward-looking rate L_j^{fr}, which is determined at the beginning of the accrual period, the new rates are *backward-looking*: they are computed as compounded overnight rates and are only known at the *end* of the accrual period.

To lighten the notation for the derivation, we fix one accrual interval $[N_{j-1}, N_j]$ and define

$$m := N_{j-1}, \qquad N := N_j, \qquad \alpha_j = (N - m)\,\Delta. \tag{3.6}$$

The discount factor over the period $[m, N]$ is

$$D(m,N) = \exp\left(-\sum_{k=m}^{N-1} r_k\right). \tag{3.7}$$

We now define the backward-looking rate that the caplet is written on. Following the contractual construction of compounded SOFR (see, for instance, [19]), one starts from the overnight rate. Over the elementary period $[k, k+1]$, the overnight rate ρ_k is the rate that the one-period bond return implies, namely

$$\rho_k := \frac{1}{\Delta}\left(\frac{1}{p(k,k+1)} - 1\right), \qquad \text{equivalently} \qquad 1 + \Delta\rho_k = \frac{1}{p(k,k+1)}, \tag{3.8}$$

where Δ is the day-count fraction of the period. The contractual compounded SOFR rate over the accrual interval $[m, N]$ is then obtained by a backward-looking compounded average,

$$L_j^{\mathrm{br}} := \frac{1}{\alpha_j}\left(\prod_{k=m}^{N-1}(1+\Delta\rho_k) - 1\right), \qquad \alpha_j = (N-m)\Delta, \tag{3.9}$$

so that $\prod_{k=m}^{N-1}(1+\Delta\rho_k)$ is the realized gross return on a unit notional rolled overnight at the prevailing rate from m to N, and α_j is the accrual factor of the whole interval.

This product is fully determined by the one-period bond prices, since each per-period factor is $1 + \Delta\rho_k = 1/p(k, k+1)$ by (3.8). Under the single-period convention $p(k, k+1) = e^{-r_k}$ adopted in (1.3), the per-period gross return is $1/p(k, k+1) = e^{r_k}$, and the compounding product telescopes into an exponential of the single-period rates:

$$\prod_{k=m}^{N-1}(1+\Delta\rho_k) = \prod_{k=m}^{N-1}\frac{1}{p(k,k+1)} = \prod_{k=m}^{N-1} e^{r_k} = \exp\Big(\sum_{k=m}^{N-1} r_k\Big) = \frac{1}{D(m,N)}. \tag{3.10}$$

The backward-looking rate (3.9) thus takes the equivalent form

$$L_j^{\mathrm{br}} = \frac{1}{\alpha_j}\left(\frac{1}{D(m,N)} - 1\right) = \frac{1}{\alpha_j}\left(\exp\Big(\sum_{k=m}^{N-1} r_k\Big) - 1\right). \tag{3.11}$$

Remark 3.2 (Measurability). The key structural difference between L_j^{br} and L_j^{fr} is their measurability. The forward-looking rate L_j^{fr} is $\mathcal{F}_m$-measurable: it is known at the start of the accrual period. The backward-looking rate L_j^{br} depends on the entire path of short rates $r_m, r_{m+1}, \ldots, r_{N-1}$ and is therefore only $\mathcal{F}_{N-1}$-measurable: it is known at the *end* of the accrual period. This distinction has important consequences for pricing, as the payoff is no longer a function of the bond price at the beginning of the period.

3.3 Pricing of the Backward-Looking Caplet

The backward-looking caplet with strike K on the j-th interval has payoff $\alpha_j(L_j^{\mathrm{br}} - K)^+$ at time N. Its price at time $n \le m$ is

$$\mathrm{Caplet}_j^{\mathrm{br}}(n) = E^Q\big[D(n,N)\,\alpha_j(L_j^{\mathrm{br}} - K)^+ \,\big|\, \mathcal{F}_n\big]\,. \tag{3.12}$$

In this section we compute the price at time $m = N_{j-1}$, the start of the accrual period. The derivation proceeds through a sequence of intermediate lemmas: a reduction of the payoff to a truncated lognormal form, the conditional Gaussian law of the accrual exponent, a bond-price identity for its conditional moment generating function, and a univariate truncated-Gaussian moment. These are then combined in Theorem 3.1.

3.3.1 Payoff Reduction and Conditional Law

Lemma 3.1 (Payoff reduction at the start of the accrual period)**.** *Let $k_j := 1 + \alpha_j K$ and $X_{m,N} := \sum_{k=m}^{N-1} r_k$. The price at time m of the backward-looking caplet admits the representation*

$$\mathrm{Caplet}_j^{\mathrm{br}}(m) = E^Q\Big[\big(1 - k_j\, e^{-X_{m,N}}\big)^+ \,\big|\, \mathcal{F}_m\Big]\,. \tag{3.13}$$

Proof. We first simplify the payoff. Using the definition of L_j^{br}:

$$\begin{aligned}\alpha_j(L_j^{\mathrm{br}} - K)^+ &= \left(\frac{1}{D(m,N)} - 1 - \alpha_j K\right)^+ \\ &= \big(D(m,N)^{-1} - k_j\big)^+\,,\end{aligned} \tag{3.14}$$

where we define

$$k_j := 1 + \alpha_j K. \tag{3.15}$$

Evaluating (3.12) at $n = m$,

$$\mathrm{Caplet}_j^{\mathrm{br}}(m) = E^Q\Big[D(m,N)\,\big(D(m,N)^{-1} - k_j\big)^+ \,\big|\, \mathcal{F}_m\Big]\,. \tag{3.16}$$

Since $D(m,N) > 0$ almost surely, we can multiply through:

$$D(m,N)\,\big(D(m,N)^{-1} - k_j\big)^+ = (1 - k_j\, D(m,N))^+\,. \tag{3.17}$$

Therefore

$$\mathrm{Caplet}_j^{\mathrm{br}}(m) = E^Q\Big[\big(1 - k_j\, e^{-X_{m,N}}\big)^+ \,\big|\, \mathcal{F}_m\Big]\,,$$

where we define

$$X_{m,N} := \sum_{k=m}^{N-1} r_k. \tag{3.18}$$

□

Lemma 3.2 (Conditional Gaussian law of $X_{m,N}$)**.** *Conditionally on $\mathcal{F}_m$, the accrual exponent $X_{m,N}$ is Gaussian,*

$$X_{m,N} \,\big|\, \mathcal{F}_m \;\sim\; \mathcal{N}(\mu_{m,N},\, \nu_{m,N}), \tag{3.19}$$

with

$$\mu_{m,N} = (1 + B_m^N)\, r_m + \Delta \sum_{q=m}^{N-2} (1 + B_{q+1}^N)\, \Phi_q, \tag{3.20}$$

$$\nu_{m,N} = \sigma^2 \Delta \sum_{q=m}^{N-2} (1 + B_{q+1}^N)^2. \tag{3.21}$$

Proof. From the forward iteration of the Hull-White recursion (2.6), for $h = 0, 1, \dots, N - m - 1$:

$$r_{m+h} = c^h r_m + \Delta \sum_{q=m}^{m+h-1} c^{m+h-1-q} \Phi_q + \sigma\sqrt{\Delta} \sum_{q=m}^{m+h-1} c^{m+h-1-q} \omega_q. \tag{3.22}$$

Therefore

$$X_{m,N} = \sum_{h=0}^{N-m-1} r_{m+h}. \tag{3.23}$$

We collect the coefficients. The total coefficient of r_m is $\sum_{h=0}^{N-m-1} c^h = 1 + B_m^N$, using (2.25). Each Φ_q, with $q \in \{m, \dots, N-2\}$, appears in $r_{q+1}, r_{q+2}, \dots, r_{N-1}$ with weights $1, c, c^2, \dots, c^{N-q-2}$, so its total coefficient is $\Delta \sum_{j=0}^{N-q-2} c^j = \Delta(1 + B_{q+1}^N)$. Similarly, each ω_q appears with total coefficient $\sigma\sqrt{\Delta} \sum_{j=0}^{N-q-2} c^j = \sigma\sqrt{\Delta}(1 + B_{q+1}^N)$. Thus:

$$X_{m,N} = (1 + B_m^N)\, r_m + \Delta \sum_{q=m}^{N-2} (1 + B_{q+1}^N)\, \Phi_q + \sigma\sqrt{\Delta} \sum_{q=m}^{N-2} (1 + B_{q+1}^N)\, \omega_q. \tag{3.24}$$

Conditionally on $\mathcal{F}_m$, the only random part is the last sum. Since the ω_q are iid $\mathcal{N}(0,1)$, we conclude that $X_{m,N} \mid \mathcal{F}_m \sim \mathcal{N}(\mu_{m,N}, \nu_{m,N})$ with $\mu_{m,N}$ and $\nu_{m,N}$ given by (3.20)–(3.21). □

Lemma 3.3 (Bond-price identity for the conditional moment generating function)**.** *The conditional moment generating function of $X_{m,N}$ evaluated at -1 recovers the bond price,*

$$E^Q\big[e^{-X_{m,N}} \,\big|\, \mathcal{F}_m\big] = e^{-\mu_{m,N} + \frac{1}{2}\nu_{m,N}} = e^{-A_m^N - (1+B_m^N) r_m} = p(m, N). \tag{3.25}$$

Equivalently,

$$-\mu_{m,N} + \tfrac{1}{2}\,\nu_{m,N} \;=\; \log p(m,N). \tag{3.26}$$

Proof. From the explicit formula for A_n^N in (2.29), we can verify that

$$\mu_{m,N} = A_m^N + (1+B_m^N)\,r_m + \tfrac{1}{2}\nu_{m,N}.$$

Therefore, using the identity $E^Q[e^{-X}] = e^{-\mu+\nu/2}$ for $X \sim \mathcal{N}(\mu,\nu)$, we obtain (3.25); this is the conditional moment generating function of $X_{m,N}$ evaluated at -1, and it recovers the bond price directly. Taking logarithms in (3.25) gives (3.26). □

3.3.2 A Univariate Truncated-Gaussian Moment

The closed-form price requires one truncated exponential moment of a univariate Gaussian, which we isolate as a lemma. It is the one-dimensional version of the bivariate computation used later in Section 3.6.

Lemma 3.4 (Truncated exponential moment of a univariate Gaussian)**.** *Let $X \sim \mathcal{N}(\mu,\nu)$ with $\nu > 0$. Then for every $\kappa \in \mathbb{R}$,*

$$E\big[e^{-X}\mathbf{1}_{\{X>\kappa\}}\big] = e^{\frac{\nu}{2}-\mu}\,\Phi\!\left(\frac{\mu-\nu-\kappa}{\sqrt{\nu}}\right), \tag{3.27}$$

where Φ denotes the standard normal cumulative distribution function.

Proof. Writing out the integral:

$$E\big[e^{-X}\mathbf{1}_{\{X>\kappa\}}\big] = \frac{1}{\sqrt{2\pi\nu}}\int_\kappa^\infty e^{-x}\exp\!\left(-\frac{(x-\mu)^2}{2\nu}\right)dx. \tag{3.28}$$

We complete the square in the exponent:

$$\begin{aligned} -x - \frac{(x-\mu)^2}{2\nu} &= -\frac{2\nu x + x^2 - 2\mu x + \mu^2}{2\nu} \\ &= -\frac{x^2 - 2x(\mu-\nu) + \mu^2}{2\nu}. \end{aligned}$$

Thus,

$$-x - \frac{(x-\mu)^2}{2\nu} = -\frac{(x-(\mu-\nu))^2}{2\nu} + \frac{\nu}{2} - \mu. \tag{3.29}$$

Substituting (3.29) into (3.28):

$$\begin{aligned} E\big[e^{-X}\mathbf{1}_{\{X>\kappa\}}\big] &= e^{\frac{\nu}{2}-\mu}\,\frac{1}{\sqrt{2\pi\nu}}\int_\kappa^\infty \exp\!\left(-\frac{(x-(\mu-\nu))^2}{2\nu}\right)dx \\ &= e^{\frac{\nu}{2}-\mu}\,\Phi\!\left(\frac{\mu-\nu-\kappa}{\sqrt{\nu}}\right). \end{aligned}$$ □

3.3.3 Closed-Form Caplet Price

Theorem 3.1 (Backward-Looking Caplet Pricing). *Under the discrete Hull-White model (2.4), the price at time $m = N_{j-1}$ of the backward-looking caplet with strike K on the j-th accrual interval $[N_{j-1}, N_j]$ is*

$$\boxed{\mathrm{Caplet}_j^{\mathrm{br}}(m) \;=\; \Phi\big(d_1^{(j)}\big) \;-\; k_j\, p(m,N)\, \Phi\big(d_2^{(j)}\big),} \tag{3.30}$$

where

$$d_1^{(j)} = \frac{\log\left(\dfrac{1}{k_j\, p(m,N)}\right) + \frac{1}{2}\,\nu_{m,N}}{\sqrt{\nu_{m,N}}}, \tag{3.31}$$

$$d_2^{(j)} = d_1^{(j)} - \sqrt{\nu_{m,N}}, \tag{3.32}$$

and

$$k_j = 1 + \alpha_j K, \qquad \alpha_j = (N_j - N_{j-1})\Delta, \tag{3.33}$$

$$\nu_{m,N} = \sigma^2 \Delta \sum_{q=m}^{N-2} \big(1 + B_{q+1}^N\big)^2. \tag{3.34}$$

Proof. By Lemma 3.1, the price at time m is (3.13), and by Lemma 3.2 $X_{m,N} \,|\, \mathcal{F}_m \sim \mathcal{N}(\mu_{m,N}, \nu_{m,N})$. The payoff is positive when

$$1 - k_j\, e^{-X_{m,N}} > 0 \quad\iff\quad e^{-X_{m,N}} < \frac{1}{k_j} \quad\iff\quad X_{m,N} > \log k_j.$$

Therefore, writing the positive part explicitly:

$$\begin{aligned}\mathrm{Caplet}_j^{\mathrm{br}}(m) &= E^Q\big[\big(1 - k_j\, e^{-X_{m,N}}\big)\, \mathbf{1}_{\{X_{m,N} > \log k_j\}} \,\big|\, \mathcal{F}_m\big] \\ &= Q\big(X_{m,N} > \log k_j \,\big|\, \mathcal{F}_m\big) - k_j\, E^Q\big[e^{-X_{m,N}} \mathbf{1}_{\{X_{m,N} > \log k_j\}} \,\big|\, \mathcal{F}_m\big]\,.\end{aligned} \tag{3.35}$$

First term. Since $X_{m,N} \,|\, \mathcal{F}_m \sim \mathcal{N}(\mu_{m,N}, \nu_{m,N})$:

$$Q\big(X_{m,N} > \log k_j \,\big|\, \mathcal{F}_m\big) = \Phi\left(\frac{\mu_{m,N} - \log k_j}{\sqrt{\nu_{m,N}}}\right). \tag{3.36}$$

Second term. Applying Lemma 3.4 with $X = X_{m,N}$, $\mu = \mu_{m,N}$, $\nu = \nu_{m,N}$, and $\kappa = \log k_j$:

$$E^Q\big[e^{-X_{m,N}} \mathbf{1}_{\{X_{m,N} > \log k_j\}} \,\big|\, \mathcal{F}_m\big] = e^{\frac{\nu_{m,N}}{2} - \mu_{m,N}}\, \Phi\left(\frac{\mu_{m,N} - \nu_{m,N} - \log k_j}{\sqrt{\nu_{m,N}}}\right). \tag{3.37}$$

By Lemma 3.3,

$$e^{\frac{\nu_{m,N}}{2}-\mu_{m,N}} = p(m,N). \tag{3.38}$$

Combining the two terms. Using (3.38),

$$\mathrm{Caplet}_j^{\mathrm{br}}(m) = \Phi\left(\frac{\mu_{m,N} - \log k_j}{\sqrt{\nu_{m,N}}}\right) - k_j\, p(m,N)\, \Phi\left(\frac{\mu_{m,N} - \nu_{m,N} - \log k_j}{\sqrt{\nu_{m,N}}}\right). \tag{3.39}$$

The arguments of the two Gaussian cumulative distribution functions still involve the conditional moment $\mu_{m,N}$, which is not directly observable on the market. By Lemma 3.3, however, this quantity is tied to the bond price through the identity (3.26). Rearranging (3.26) as $\mu_{m,N} = -\log p(m,N) + \frac{1}{2}\nu_{m,N}$ and substituting into the numerators of (3.39) yields

$$\mu_{m,N} - \log k_j \;=\; \log\left(\frac{1}{k_j\, p(m,N)}\right) + \tfrac{1}{2}\,\nu_{m,N}, \tag{3.40}$$

so that (3.39) reduces to (3.30) with $d_1^{(j)}$ and $d_2^{(j)}$ as in (3.31)–(3.32). □

3.4 Floorlet Pricing

The backward-looking floorlet with strike K on the j-th interval has payoff $\alpha_j(K - L_j^{\mathrm{br}})^+$ at time N. Using the same notation as before:

$$\begin{aligned}\mathrm{Floorlet}_j^{\mathrm{br}}(m) &= E^Q\left[D(m,N)\left(k_j - D(m,N)^{-1}\right)^+ \,\middle|\, \mathcal{F}_m\right] \\ &= E^Q\left[\left(k_j\, e^{-X_{m,N}} - 1\right)^+ \,\middle|\, \mathcal{F}_m\right].\end{aligned} \tag{3.41}$$

The payoff is positive when $X_{m,N} < \log k_j$. An analogous computation gives:

Theorem 3.2 (Backward-Looking Floorlet Pricing)**.** *Under the discrete Hull-White model, the price at time $m = N_{j-1}$ of the backward-looking floorlet with strike K is*

$$\boxed{\mathrm{Floorlet}_j^{\mathrm{br}}(m) = k_j\, p(m,N)\, \Phi(-d_2^{(j)}) - \Phi(-d_1^{(j)}),} \tag{3.42}$$

where $d_1^{(j)}$ and $d_2^{(j)}$ are defined in (3.31)–(3.32).

Proof. The payoff is positive when $k_j e^{-X_{m,N}} - 1 > 0$, i.e. $X_{m,N} < \log k_j$. Splitting:

$$\mathrm{Floorlet}_j^{\mathrm{br}}(m) = k_j\, E^Q\left[e^{-X_{m,N}}\mathbf{1}_{\{X_{m,N}<\log k_j\}} \,\middle|\, \mathcal{F}_m\right] - Q\left(X_{m,N} < \log k_j \,\middle|\, \mathcal{F}_m\right).$$

For the first term, the same completing-the-square argument as before gives $k_j\, p(m,N)\, \Phi(-d_2^{(j)})$. For the second term, $Q(X_{m,N} < \log k_j \,|\, \mathcal{F}_m) = \Phi(-d_1^{(j)})$. □

3.5 Caplet-Floorlet Parity

Proposition 3.1 (Caplet-Floorlet Parity)**.** *Under the discrete Hull-White model, for any strike K:*

$$\boxed{\mathrm{Caplet}_j^{\mathrm{br}}(m) - \mathrm{Floorlet}_j^{\mathrm{br}}(m) = 1 - k_j\, p(m, N).} \tag{3.43}$$

Proof. At time N, the caplet and floorlet payoffs (before discounting by $D(m, N)$) are

$$\begin{aligned}
\pi_{\mathrm{cap}} &= \left(D(m, N)^{-1} - k_j\right)^+, \\
\pi_{\mathrm{floor}} &= \left(k_j - D(m, N)^{-1}\right)^+.
\end{aligned}$$

Using the identity $x^+ - (-x)^+ = x$ with $x = D(m, N)^{-1} - k_j$:

$$\pi_{\mathrm{cap}} - \pi_{\mathrm{floor}} = D(m, N)^{-1} - k_j.$$

Multiplying both sides by $D(m, N)$ and taking the conditional expectation:

$$\begin{aligned}
\mathrm{Caplet}_j^{\mathrm{br}}(m) - \mathrm{Floorlet}_j^{\mathrm{br}}(m) &= E^Q\left[D(m, N)(D(m, N)^{-1} - k_j) \,\middle|\, \mathcal{F}_m\right] \\
&= E^Q\left[1 - k_j\, D(m, N) \,\middle|\, \mathcal{F}_m\right] \\
&= 1 - k_j\, E^Q\left[D(m, N) \,\middle|\, \mathcal{F}_m\right] \\
&= 1 - k_j\, p(m, N),
\end{aligned}$$

where the last step uses the pricing formula (1.7). □

Remark 3.3. The parity relation (3.43) can also be verified directly from the closed-form formulas (3.30) and (3.42), using $\Phi(x) + \Phi(-x) = 1$.

3.6 Pricing at Earlier Times

The formulas derived in Sections 3.3 and 3.4 give the price of the caplet and floorlet at time $m = N_{j-1}$, the start of the accrual period. We now show that the price can be obtained in closed form at any earlier time $n < m$ as well. The result has exactly the same structure as Theorem 3.1, with the constant 1 replaced by the bond price $p(n, m)$ and the conditional variance enlarged by a pre-accrual contribution.

Throughout this section we abbreviate

$$E_n^Q[\,\cdot\,] := E^Q[\,\cdot \mid \mathcal{F}_n].$$

3.6.1 Setup and Payoff Reduction

Fix $n < m < N$. Recall from Section 3.3 that

$$X_{m,N} := \sum_{k=m}^{N-1} r_k, \qquad k_j := 1 + \alpha_j K,$$

and that the caplet payoff at time N, multiplied by $D(m, N)$, equals $(1 - k_j e^{-X_{m,N}})^+$. To handle the additional discounting between n and m, we introduce the pre-accrual sum

$$Y_{n,m} := \sum_{k=n}^{m-1} r_k, \tag{3.44}$$

so that

$$D(n, N) = \exp\left(-\sum_{k=n}^{N-1} r_k \right) = e^{-Y_{n,m} - X_{m,N}}.$$

The caplet price at time n is

$$\begin{aligned} \mathrm{Caplet}_j^{\mathrm{br}}(n) &= E_n^Q\left[e^{-Y_{n,m} - X_{m,N}} \left(e^{X_{m,N}} - k_j \right)^+ \right] \\ &= E_n^Q\left[e^{-Y_{n,m}} \mathbf{1}_{\{X_{m,N} > \log k_j\}} \right] - k_j\, E_n^Q\left[e^{-Y_{n,m} - X_{m,N}} \mathbf{1}_{\{X_{m,N} > \log k_j\}} \right]. \end{aligned} \tag{3.45}$$

The problem reduces to the computation of two truncated exponential moments of the bivariate Gaussian vector $(Y_{n,m}, X_{m,N})$ conditional on $\mathcal{F}_n$.

Remark 3.4 (Bivariate Gaussian structure). By iterating the Hull-White recursion (2.6) forward from time n, both $Y_{n,m}$ and $X_{m,N}$ are functions of r_n, of the deterministic sequence (Φ_q), and of the innovations $\omega_n, \omega_{n+1}, \ldots, \omega_{N-2}$, which are i.i.d. standard Gaussian under Q. Conditionally on $\mathcal{F}_n$, the variable r_n is known, and the only remaining randomness is Gaussian. Hence $(Y_{n,m}, X_{m,N}) \mid \mathcal{F}_n$ is bivariate Gaussian.

3.6.2 A Bivariate Gaussian Lemma

The two expectations in (3.45) are particular cases of a single Gaussian computation, which we isolate in the following lemma. It generalizes the completing-the-square argument used in the proof of Lemma 3.4 to the bivariate case.

Lemma 3.5 (Truncated exponential moment of a bivariate Gaussian)**.** *Let (Y, X) be a Gaussian vector with mean (μ_Y, μ_X) and covariance matrix*

$$\Sigma = \begin{pmatrix} \nu_Y & \gamma \\ \gamma & \nu_X \end{pmatrix},$$

with Σ positive definite. Then for every $a, b, \kappa \in \mathbb{R}$,

$$E\big[e^{aY+bX}\mathbf{1}_{\{X>\kappa\}}\big] = \exp\!\big(a\mu_Y + b\mu_X + \tfrac{1}{2}\big(a^2\nu_Y + b^2\nu_X + 2ab\gamma\big)\big)\,\Phi\!\left(\frac{\mu_X + a\gamma + b\nu_X - \kappa}{\sqrt{\nu_X}}\right). \tag{3.46}$$

Proof. Set $Z = (Y, X)^\top$, $\mu = (\mu_Y, \mu_X)^\top$, and $u = (a, b)^\top$, so that $aY + bX = u^\top Z$. Since Σ is positive definite, Z admits the density

$$f_Z(z) = \frac{1}{2\pi\sqrt{\det\Sigma}} \exp\!\big(-\tfrac{1}{2}(z-\mu)^\top\Sigma^{-1}(z-\mu)\big)\,,$$

so the quantity to compute is

$$E\Big[e^{u^\top Z}\mathbf{1}_{\{X>\kappa\}}\Big] = \int_{\{x>\kappa\}} e^{u^\top z}\, f_Z(z)\, dz. \tag{3.47}$$

Step 1: Completing the square. We expand the quadratic form

$$\begin{aligned}(z-\mu-\Sigma u)^\top\Sigma^{-1}(z-\mu-\Sigma u) &= (z-\mu)^\top\Sigma^{-1}(z-\mu) - (z-\mu)^\top\Sigma^{-1}\Sigma u \\ &\quad - u^\top\Sigma\Sigma^{-1}(z-\mu) + u^\top\Sigma\Sigma^{-1}\Sigma u \\ &= (z-\mu)^\top\Sigma^{-1}(z-\mu) - 2u^\top(z-\mu) + u^\top\Sigma u,\end{aligned}$$

where we used $\Sigma^{-1}\Sigma = I$ together with the fact that the scalar $(z-\mu)^\top u$ equals its transpose $u^\top(z-\mu)$. Rearranging,

$$u^\top(z-\mu) - \tfrac{1}{2}(z-\mu)^\top\Sigma^{-1}(z-\mu) = -\tfrac{1}{2}(z-\mu-\Sigma u)^\top\Sigma^{-1}(z-\mu-\Sigma u) + \tfrac{1}{2}u^\top\Sigma u.$$

Adding $u^\top\mu$ to both sides and using $u^\top z = u^\top(z-\mu) + u^\top\mu$ gives

$$u^\top z - \tfrac{1}{2}(z-\mu)^\top\Sigma^{-1}(z-\mu) = -\tfrac{1}{2}(z-\mu-\Sigma u)^\top\Sigma^{-1}(z-\mu-\Sigma u) + u^\top\mu + \tfrac{1}{2}u^\top\Sigma u. \tag{3.48}$$

Step 2: Tilting. Exponentiating (3.48) and dividing by $2\pi\sqrt{\det\Sigma}$, the integrand in (3.47) factorizes as

$$e^{u^\top z}\, f_Z(z) = \exp\!\big(u^\top\mu + \tfrac{1}{2}u^\top\Sigma u\big)\, f_u(z),$$

where

$$f_u(z) := \frac{1}{2\pi\sqrt{\det\Sigma}} \exp\!\big(-\tfrac{1}{2}(z-\mu-\Sigma u)^\top\Sigma^{-1}(z-\mu-\Sigma u)\big)$$

is the density of a Gaussian vector with mean $\mu + \Sigma u$ and covariance Σ. Since the exponential prefactor is constant in z, we may take it outside the integral (3.47):

$$E\left[e^{u^\top Z}\mathbf{1}_{\{X>\kappa\}}\right] = \exp\bigl(u^\top\mu + \tfrac{1}{2}u^\top\Sigma u\bigr)\int_{\{x>\kappa\}} f_u(z)\,dz = \exp\bigl(u^\top\mu + \tfrac{1}{2}u^\top\Sigma u\bigr)\,P^u(X>\kappa), \tag{3.49}$$

where P^u denotes the probability measure under which Z has density f_u, that is $Z \sim \mathcal{N}(\mu + \Sigma u,\, \Sigma)$.

Step 3: Marginal of X under P^u. The shifted mean is

$$\mu + \Sigma u = \begin{pmatrix}\mu_Y\\ \mu_X\end{pmatrix} + \begin{pmatrix}\nu_Y & \gamma\\ \gamma & \nu_X\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix} = \begin{pmatrix}\mu_Y + a\nu_Y + b\gamma\\ \mu_X + a\gamma + b\nu_X\end{pmatrix}.$$

The covariance is unchanged, so the marginal distribution of the second component X under P^u is

$$X \sim \mathcal{N}\bigl(\mu_X + a\gamma + b\nu_X,\ \nu_X\bigr).$$

Standardizing,

$$P^u(X>\kappa) = \Phi\left(\frac{\mu_X + a\gamma + b\nu_X - \kappa}{\sqrt{\nu_X}}\right). \tag{3.50}$$

Step 4: Conclusion. The two scalar terms in the exponent of (3.49) are

$$u^\top\mu = a\mu_Y + b\mu_X, \qquad u^\top\Sigma u = \begin{pmatrix}a & b\end{pmatrix}\begin{pmatrix}\nu_Y & \gamma\\ \gamma & \nu_X\end{pmatrix}\begin{pmatrix}a\\ b\end{pmatrix} = a^2\nu_Y + b^2\nu_X + 2ab\gamma.$$

Substituting these into (3.49) together with (3.50) yields (3.46). □

Remark 3.5 (Alternative approach via the N-forward measure). An alternative to Lemma 3.5 is to change the probability measure toward the N-forward measure Q^N, defined by taking the N-maturity bond as numéraire. Under Q^N, the discounting $D(n,N)$ is absorbed into the change of measure and the caplet price reduces to $p(n,N)\,E_n^{Q^N}[(e^{X_{m,N}} - k_j)^+]$. The Hull-White model preserves the Gaussian distribution of the short rate under any forward measure change [3]: only the conditional mean is shifted by a deterministic function of time, while the conditional variance is unchanged. Since $X_{m,N}$ is a linear functional of the short rate path, it remains Gaussian under Q^N with the same conditional variance ν_X as under Q, making $e^{X_{m,N}}$ lognormal and yielding a formula of the same Black-76 [1] structure as Theorem 3.3. This forward-measure approach is standard in continuous-time Gaussian interest rate models [3] and is carried out explicitly for backward-looking caplets in the Hull-White model by Hofmann [10].

3.6.3 The Conditional Variance

Lemma 3.6 (Conditional variance of $X_{m,N}$ given $\mathcal{F}_n$). *For $n < m < N$, the conditional variance $\nu_X := \mathrm{Var}^Q(X_{m,N} \mid \mathcal{F}_n)$ is*

$$\nu_X = \sigma^2 \Delta \left[(1 + B_m^N)^2 \sum_{q=n}^{m-1} c^{2(m-1-q)} + \sum_{q=m}^{N-2} (1 + B_{q+1}^N)^2 \right]. \tag{3.51}$$

Proof. Iterating the Hull-White recursion (2.6) forward from time n, for $h = 0, \dots, N - m - 1$:

$$r_{m+h} = c^{m+h-n}\, r_n + \Delta \sum_{q=n}^{m+h-1} c^{m+h-1-q}\, \Phi_q + \sigma\sqrt{\Delta} \sum_{q=n}^{m+h-1} c^{m+h-1-q}\, \omega_q. \tag{3.52}$$

Summing over h,

$$X_{m,N} = \sum_{h=0}^{N-m-1} r_{m+h},$$

the r_n and Φ_q contributions are deterministic conditionally on $\mathcal{F}_n$ and do not affect the variance. The random part is

$$\sigma\sqrt{\Delta} \sum_{h=0}^{N-m-1} \sum_{q=n}^{m+h-1} c^{m+h-1-q}\, \omega_q.$$

We collect the coefficient of each innovation ω_q by distinguishing shocks before and inside the accrual period.

Pre-accrual shocks ($q = n, \dots, m-1$). The shock ω_q appears in every r_{m+h}, $h = 0, \dots, N - m - 1$. Its total coefficient in $X_{m,N}$ is

$$\sum_{h=0}^{N-m-1} c^{m+h-1-q} = c^{m-1-q} \sum_{h=0}^{N-m-1} c^h = c^{m-1-q}\, S_{N-m},$$

where $S_\ell := \sum_{i=0}^{\ell-1} c^i$ as in (2.31).

In-accrual shocks ($q = m, \dots, N-2$). The shock ω_q first enters r_{q+1}, so it appears in r_{m+h} only for $h \geq q - m + 1$. Its total coefficient in $X_{m,N}$ is

$$\sum_{h=q-m+1}^{N-m-1} c^{m+h-1-q} = \sum_{i=0}^{N-q-2} c^i = S_{N-q-1}.$$

Since the innovations are i.i.d. standard Gaussian, the conditional variance is the sum of squared coefficients:

$$\nu_X = \sigma^2 \Delta \left[S_{N-m}^2 \sum_{q=n}^{m-1} c^{2(m-1-q)} + \sum_{q=m}^{N-2} S_{N-q-1}^2 \right]. \tag{3.53}$$

Recalling from (2.26) and (2.31) that $S_{N-m} = 1 + B_m^N$ and $S_{N-q-1} = 1 + B_{q+1}^N$, we obtain (3.51). □

Remark 3.6 (Structure of the variance). The variance (3.51) is the sum of two non-negative contributions:

- The *pre-accrual* term $(1 + B_m^N)^2 \sum_{q=n}^{m-1} c^{2(m-1-q)}$ accounts for the uncertainty accumulated from n to m, which propagates into the future short rates that determine $X_{m,N}$.
- The *in-accrual* term $\sum_{q=m}^{N-2}(1+B_{q+1}^N)^2$ is the variance already obtained in Section 3.3 for the price at $n = m$.

For $n = m$ the pre-accrual sum is empty and (3.51) reduces to $\nu_{m,N}$.

3.6.4 The Caplet at Earlier Times

We now combine the bivariate lemma and the conditional variance into the main result of this section.

Theorem 3.3 (Backward-looking caplet price at earlier times). *Under the discrete Hull-White model (2.4), for every $n \leq m = N_{j-1} < N = N_j$, the price at time n of the backward-looking caplet with strike K on the j-th accrual interval is*

$$\boxed{\mathrm{Caplet}_j^{\mathrm{br}}(n) = p(n,m)\, \Phi(d_{1,n}^{(j)}) - k_j\, p(n,N)\, \Phi(d_{2,n}^{(j)}),} \tag{3.54}$$

where

$$d_{1,n}^{(j)} = \frac{\log\left(\dfrac{p(n,m)}{k_j\, p(n,N)}\right) + \frac{1}{2}\nu_X}{\sqrt{\nu_X}}, \qquad d_{2,n}^{(j)} = d_{1,n}^{(j)} - \sqrt{\nu_X}, \tag{3.55}$$

and the conditional variance is

$$\nu_X = \sigma^2 \Delta \left[(1 + B_m^N)^2 \sum_{q=n}^{m-1} c^{2(m-1-q)} + \sum_{q=m}^{N-2} (1 + B_{q+1}^N)^2 \right]. \tag{3.56}$$

Proof. By Remark 3.4, $(Y_{n,m}, X_{m,N}) \mid \mathcal{F}_n$ is bivariate Gaussian. Denote

$$\mu_Y := E_n^Q[Y_{n,m}], \qquad \mu_X := E_n^Q[X_{m,N}], \tag{3.57}$$

$$\nu_Y := \mathrm{Var}^Q(Y_{n,m} \mid \mathcal{F}_n), \quad \nu_X := \mathrm{Var}^Q(X_{m,N} \mid \mathcal{F}_n), \quad \gamma := \mathrm{Cov}^Q(Y_{n,m}, X_{m,N} \mid \mathcal{F}_n), \tag{3.58}$$

and apply Lemma 3.5 conditionally on $\mathcal{F}_n$.

First term. Setting $a = -1$, $b = 0$, and $\kappa = \log k_j$ in (3.46),

$$E_n^Q\left[e^{-Y_{n,m}} \mathbf{1}_{\{X_{m,N} > \log k_j\}}\right] = \exp\bigl(-\mu_Y + \tfrac{1}{2}\nu_Y\bigr)\, \Phi\left(\frac{\mu_X - \gamma - \log k_j}{\sqrt{\nu_X}}\right). \tag{3.59}$$

The prefactor admits a direct probabilistic interpretation. Since $-Y_{n,m} \mid \mathcal{F}_n$ is Gaussian with mean $-\mu_Y$ and variance ν_Y, its moment generating function evaluated at -1 is $\exp\bigl(-\mu_Y + \frac{1}{2}\nu_Y\bigr)$, which by Lemma 1.1 equals the bond price $p(n,m)$:

$$\exp\bigl(-\mu_Y + \tfrac{1}{2}\nu_Y\bigr) = E_n^Q\left[e^{-Y_{n,m}}\right] = p(n,m). \tag{3.60}$$

Second term. Setting $a = -1$, $b = -1$, and $\kappa = \log k_j$,

$$E_n^Q\left[e^{-Y_{n,m} - X_{m,N}} \mathbf{1}_{\{X_{m,N} > \log k_j\}}\right] = \exp\bigl(-\mu_Y - \mu_X + \tfrac{1}{2}(\nu_Y + \nu_X + 2\gamma)\bigr)\, \Phi\left(\frac{\mu_X - \gamma - \nu_X - \log k_j}{\sqrt{\nu_X}}\right). \tag{3.61}$$

By the same argument applied to the random variable $Y_{n,m} + X_{m,N}$, the prefactor equals

$$\exp\bigl(-\mu_Y - \mu_X + \tfrac{1}{2}(\nu_Y + \nu_X + 2\gamma)\bigr) = E_n^Q\left[e^{-Y_{n,m} - X_{m,N}}\right] = p(n,N). \tag{3.62}$$

Combining (3.45) with (3.59), (3.60), (3.61) and (3.62) yields

$$\mathrm{Caplet}_j^{\mathrm{br}}(n) = p(n,m)\, \Phi(d_{1,n}^{(j)}) - k_j\, p(n,N)\, \Phi(d_{2,n}^{(j)}), \tag{3.63}$$

with

$$d_{1,n}^{(j)} = \frac{\mu_X - \gamma - \log k_j}{\sqrt{\nu_X}}, \qquad d_{2,n}^{(j)} = d_{1,n}^{(j)} - \sqrt{\nu_X}. \tag{3.64}$$

The arguments $d_{1,n}^{(j)}$ and $d_{2,n}^{(j)}$ in (3.64) involve the conditional moments μ_X and γ, which are not yet visible market quantities and have long analytical forms. We now show that the combination $\mu_X - \gamma$ can be expressed entirely in terms of the bond prices $p(n,m)$ and $p(n,N)$, which yields a cleaner formulation. From (3.60) and (3.62),

$$\log \frac{p(n,m)}{p(n,N)} = \bigl(-\mu_Y + \tfrac{1}{2}\nu_Y\bigr) - \bigl(-\mu_Y - \mu_X + \tfrac{1}{2}(\nu_Y + \nu_X + 2\gamma)\bigr) = \mu_X - \gamma - \tfrac{1}{2}\nu_X,$$

so that

$$\mu_X - \gamma = \log \frac{p(n,m)}{p(n,N)} + \tfrac{1}{2}\nu_X. \tag{3.65}$$

Substituting (3.65) into (3.64) gives (3.55). Finally, the conditional variance ν_X is computed in Lemma 3.6, which is (3.56). $\square$

Remark 3.7 (Comparison with Theorem 3.1). As n approaches m, the pre-accrual sum $Y_{n,m}$ vanishes, $p(n,m) \to 1$, and the conditional moments reduce to those of Section 3.3. The formula (3.63) therefore reproduces Theorem 3.1 exactly.

Remark 3.8 (Consistency with the one-step recursion). The price (3.54) is consistent with the one-step recursion

$$\mathrm{Caplet}_j^{\mathrm{br}}(n) = p(n,n+1)\, E_n^Q\left[\mathrm{Caplet}_j^{\mathrm{br}}(n+1)\right], \qquad n < m,$$

which is itself a consequence of Corollary 1.1. The closed-form expression makes the recursion redundant for analytical purposes, since the price at every intermediate time is a deterministic function of the bond prices $p(n,m)$ and $p(n,N)$ and of the model parameters through ν_X.

3.6.5 The Floorlet at Earlier Times

Theorem 3.4 (Backward-looking floorlet price at earlier times). *Under the discrete Hull-White model, for every $n \leq m < N$,*

$$\boxed{\mathrm{Floorlet}_j^{\mathrm{br}}(n) = k_j\, p(n,N)\, \Phi(-d_{2,n}^{(j)}) - p(n,m)\, \Phi(-d_{1,n}^{(j)}),} \tag{3.66}$$

where $d_{1,n}^{(j)}$ and $d_{2,n}^{(j)}$ are as in Theorem 3.3.

Proof. The floorlet payoff at time N, multiplied by $D(m,N)$, equals $(k_j e^{-X_{m,N}} - 1)^+$. The price at time $n \leq m$ is therefore

$$\mathrm{Floorlet}_j^{\mathrm{br}}(n) = k_j\, E_n^Q\left[e^{-Y_{n,m}-X_{m,N}}\mathbf{1}_{\{X_{m,N}<\log k_j\}}\right] - E_n^Q\left[e^{-Y_{n,m}}\mathbf{1}_{\{X_{m,N}<\log k_j\}}\right].$$

Applying Lemma 3.5 (with the indicator $\mathbf{1}_{\{X_{m,N}<\log k_j\}}$ producing $\Phi(-\,\cdot\,)$ in place of $\Phi(\,\cdot\,)$) and the prefactor identifications (3.60) and (3.62) yields (3.66). ☐

Remark 3.9 (Caplet-floorlet parity at earlier times). The caplet-floorlet parity of Proposition 3.1 extends to any $n \leq m$: subtracting (3.66) from (3.54) and using $\Phi(x)+\Phi(-x) = 1$ gives

$$\mathrm{Caplet}_j^{\mathrm{br}}(n) - \mathrm{Floorlet}_j^{\mathrm{br}}(n) = p(n,m) - k_j\, p(n,N),$$

which is the analogue at time n of the parity relation (3.43). The same identity can be obtained directly from the payoff decomposition $x^+ - (-x)^+ = x$ applied to $D(m,N)^{-1} - k_j$ at time N, followed by discounting under the pricing measure.

3.6.6 The Forward-Looking Caplet at Earlier Times

The same bivariate Gaussian technique prices the *forward-looking* caplet of Section 3.1 at any time $n \le m = N_{j-1}$. We develop it here as a special case of Lemma 3.5, both to complete the pricing of the contracts introduced in Section 3.1 and, more importantly, because the resulting formula has the *identical* structure of Theorem 3.3 and differs from it *only* through the conditional variance. This isolates the effect of the backward-looking convention on the price and will be the basis for comparing the sensitivities of the two prices.

Recall from Section 3.1 that the forward-looking rate applicable to the interval $[m, N]$ is

$$L_j^{\mathrm{fr}} = \frac{1}{\alpha_j}\left(\frac{1}{p(m,N)} - 1\right),$$

which is $\mathcal{F}_m$-measurable, and that the caplet pays $\alpha_j(L_j^{\mathrm{fr}} - K)^+$ at time N. As in (3.14), the payoff reduces to

$$\alpha_j(L_j^{\mathrm{fr}} - K)^+ = \left(\frac{1}{p(m,N)} - k_j\right)^+, \qquad k_j = 1 + \alpha_j K, \tag{3.67}$$

and its price at time $n \le m$ is, by (3.4),

$$\mathrm{Caplet}_j^{\mathrm{fr}}(n) = E_n^Q\left[D(n,N)\left(\frac{1}{p(m,N)} - k_j\right)^+\right]. \tag{3.68}$$

The feature that distinguishes the forward-looking case from the backward-looking one is that the payoff (3.67) is already $\mathcal{F}_m$-measurable, so the discounting from m to N can be carried out first. Writing $D(n,N) = D(n,m)\,D(m,N)$ and conditioning on $\mathcal{F}_m$ (the tower property applies since $\mathcal{F}_n \subseteq \mathcal{F}_m$), the factors $D(n,m)$ and the payoff are $\mathcal{F}_m$-measurable while $E^Q[D(m,N) \mid \mathcal{F}_m] = p(m,N)$ by (1.7), hence

$$\begin{aligned}
\mathrm{Caplet}_j^{\mathrm{fr}}(n) &= E_n^Q\left[D(n,m)\left(\frac{1}{p(m,N)} - k_j\right)^+ E^Q[D(m,N) \mid \mathcal{F}_m]\right] \\
&= E_n^Q\left[D(n,m)\,p(m,N)\left(\frac{1}{p(m,N)} - k_j\right)^+\right] \\
&= E_n^Q\left[D(n,m)\,(1 - k_j\,p(m,N))^+\right],
\end{aligned} \tag{3.69}$$

where the last step uses $p(m,N) > 0$ to move $p(m,N)$ inside the positive part. The problem has thus been reduced to the conditional law of the single bond price $p(m,N)$, which by the affine representation (2.20) is an exponential of r_m alone.

Define

$$W := -\log p(m,N) = A_m^N + (1 + B_m^N)\,r_m, \tag{3.70}$$

so that $1 - k_j\, p(m,N) > 0$ if and only if $W > \log k_j$. Using $D(n,m) = e^{-Y_{n,m}}$ with $Y_{n,m} = \sum_{k=n}^{m-1} r_k$ as in (3.44), and $p(m,N) = e^{-W}$, the price (3.69) splits into two truncated exponential moments,

$$\mathrm{Caplet}_j^{\mathrm{fr}}(n) = E_n^Q\big[e^{-Y_{n,m}}\mathbf{1}_{\{W>\log k_j\}}\big] - k_j\, E_n^Q\big[e^{-Y_{n,m}-W}\mathbf{1}_{\{W>\log k_j\}}\big]\,. \tag{3.71}$$

This is *formally identical* to the backward-looking decomposition (3.45), with the path-sum $X_{m,N}$ replaced by the single-rate functional W. The reduction to the bivariate Gaussian lemma is therefore immediate.

Lemma 3.7 (Forward-looking conditional variance and pre-accrual decomposition). *For $n \le m < N$, the variable W in* (3.70) *satisfies $\nu_W := \mathrm{Var}^Q(W \mid \mathcal{F}_n) = (1+B_m^N)^2\,\mathrm{Var}^Q(r_m \mid \mathcal{F}_n)$, and explicitly*

$$\boxed{\nu_W = \sigma^2\Delta\,(1+B_m^N)^2 \sum_{q=n}^{m-1} c^{2(m-1-q)}.} \tag{3.72}$$

Proof. By (3.70), W is an affine function of r_m with deterministic coefficients, hence

$$\nu_W = (1+B_m^N)^2\,\mathrm{Var}^Q\big(r_m \mid \mathcal{F}_n\big). \tag{3.73}$$

Iterating the recursion (2.6) forward from time n,

$$r_m = c^{m-n} r_n + \Delta \sum_{q=n}^{m-1} c^{m-1-q}\Phi_q + \sigma\sqrt{\Delta}\sum_{q=n}^{m-1} c^{m-1-q}\omega_q, \tag{3.74}$$

the first two terms are deterministic conditionally on $\mathcal{F}_n$ and the innovations ω_q are i.i.d. standard Gaussian, so

$$\mathrm{Var}^Q\big(r_m \mid \mathcal{F}_n\big) = \sigma^2\Delta \sum_{q=n}^{m-1} c^{2(m-1-q)}, \tag{3.75}$$

and consequently (3.72) follows from (3.73). □

Theorem 3.5 (Forward-looking caplet price at earlier times). *Under the discrete Hull-White model* (2.4)*, for every $n \le m = N_{j-1} < N = N_j$, the price at time n of the forward-looking caplet with strike K on the j-th accrual interval is*

$$\boxed{\mathrm{Caplet}_j^{\mathrm{fr}}(n) = p(n,m)\,\Phi(\hat{d}_{1,n}^{(j)}) - k_j\, p(n,N)\,\Phi(\hat{d}_{2,n}^{(j)}),} \tag{3.76}$$

where

$$\hat{d}_{1,n}^{(j)} = \frac{\log\left(\dfrac{p(n,m)}{k_j\, p(n,N)}\right) + \frac{1}{2}\nu_W}{\sqrt{\nu_W}}, \qquad \hat{d}_{2,n}^{(j)} = \hat{d}_{1,n}^{(j)} - \sqrt{\nu_W}, \tag{3.77}$$

and the conditional variance ν_W is given explicitly in Lemma 3.7.

Proof. Conditionally on $\mathcal{F}_n$, both $Y_{n,m}$ and W are affine functions of r_n and of the innovations $\omega_n, \ldots, \omega_{m-1}$ (for W, through r_m); hence $(Y_{n,m}, W) \mid \mathcal{F}_n$ is bivariate Gaussian. Denote

$$\mu_Y := E_n^Q[Y_{n,m}], \quad \mu_W := E_n^Q[W], \quad \nu_Y := \mathrm{Var}^Q(Y_{n,m} \mid \mathcal{F}_n), \quad \nu_W := \mathrm{Var}^Q(W \mid \mathcal{F}_n), \tag{3.78}$$

and $\gamma_W := \mathrm{Cov}^Q(Y_{n,m}, W \mid \mathcal{F}_n)$. Applying Lemma 3.5 with $(Y, X) = (Y_{n,m}, W)$ to the two terms of (3.71) gives, with $\kappa = \log k_j$,

$$E_n^Q\left[e^{-Y_{n,m}}\mathbf{1}_{\{W>\log k_j\}}\right] = \exp\!\left(-\mu_Y + \tfrac{1}{2}\nu_Y\right)\Phi\!\left(\frac{\mu_W - \gamma_W - \log k_j}{\sqrt{\nu_W}}\right), \qquad (a=-1,\ b=0), \tag{3.79}$$

$$E_n^Q\left[e^{-Y_{n,m}-W}\mathbf{1}_{\{W>\log k_j\}}\right] = \exp\!\left(-\mu_Y - \mu_W + \tfrac{1}{2}(\nu_Y + \nu_W + 2\gamma_W)\right)\Phi\!\left(\frac{\mu_W - \gamma_W - \nu_W - \log k_j}{\sqrt{\nu_W}}\right), \quad (a=-1,\ b=-1). \tag{3.80}$$

The two prefactors are again bond prices. The first is the moment generating function of $-Y_{n,m}$ evaluated at -1, which by Lemma 1.1 is

$$\exp\!\left(-\mu_Y + \tfrac{1}{2}\nu_Y\right) = E_n^Q\left[e^{-Y_{n,m}}\right] = p(n,m), \tag{3.81}$$

exactly as in (3.60). For the second, note that $e^{-Y_{n,m}-W} = D(n,m)\,p(m,N) = E^Q[D(n,N) \mid \mathcal{F}_m]$, so

$$\exp\!\left(-\mu_Y - \mu_W + \tfrac{1}{2}(\nu_Y + \nu_W + 2\gamma_W)\right) = E_n^Q\left[e^{-Y_{n,m}-W}\right] = E_n^Q[D(n,N)] = p(n,N). \tag{3.82}$$

Substituting (3.81) and (3.82) into (3.71) yields

$$\mathrm{Caplet}_j^{\mathrm{fr}}(n) = p(n,m)\,\Phi(\hat{d}_{1,n}^{(j)}) - k_j\,p(n,N)\,\Phi(\hat{d}_{2,n}^{(j)}), \tag{3.83}$$

with

$$\hat{d}_{1,n}^{(j)} = \frac{\mu_W - \gamma_W - \log k_j}{\sqrt{\nu_W}}, \qquad \hat{d}_{2,n}^{(j)} = \hat{d}_{1,n}^{(j)} - \sqrt{\nu_W}.$$

Exactly as in the proof of Theorem 3.3, the conditional moments μ_W and γ_W need not be computed separately: subtracting the logarithms of (3.81) and (3.82) gives

$$\log\frac{p(n,m)}{p(n,N)} = \mu_W - \gamma_W - \tfrac{1}{2}\nu_W,$$

so that $\mu_W - \gamma_W = \log\frac{p(n,m)}{p(n,N)} + \frac{1}{2}\nu_W$ and the arguments collapse to bond prices and the conditional variance alone, yielding (3.77). The explicit form of ν_W is given in Lemma 3.7. $\square$

3.6.7 Forward versus Backward: the Variance Decomposition and Price Dominance

Proposition 3.2 (Variance decomposition). *For every $n \leq m < N$, the backward-looking variance ν_X of Lemma 3.6 and the forward-looking variance ν_W of Lemma 3.7 satisfy*

$$\boxed{\nu_X = \nu_W + \nu_{m,N}, \qquad \nu_{m,N} = \sigma^2 \Delta \sum_{q=m}^{N-2} (1 + B_{q+1}^N)^2,} \tag{3.84}$$

where $\nu_{m,N}$ is the in-accrual variance (3.21) *obtained for the backward-looking caplet at time m. In particular ν_W is exactly the pre-accrual term of ν_X.*

Proof. Comparing (3.72) with the backward-looking variance (3.56),

$$\nu_X = \sigma^2 \Delta \left[\underbrace{(1 + B_m^N)^2 \sum_{q=n}^{m-1} c^{2(m-1-q)}}_{\text{pre-accrual}} + \underbrace{\sum_{q=m}^{N-2} (1 + B_{q+1}^N)^2}_{\text{in-accrual}} \right],$$

shows that the forward-looking variance is *exactly* the pre-accrual term of the backward-looking variance, which is (3.84). $\square$

Proposition 3.3 (Backward-looking caplets dominate forward-looking ones). *Fix an accrual interval $[m, N] = [N_{j-1}, N_j]$ and a strike K, and let $k_j = 1 + \alpha_j K$. Then for every $n \leq m$,*

$$\text{Caplet}_j^{\text{fr}}(n) \;\leq\; \text{Caplet}_j^{\text{br}}(n),$$

with equality if and only if $D(m, N)$ is $\mathcal{F}_m$-measurable (equivalently, $\nu_{m,N} = 0$).

Proof. Consider first $n = m$. The function $\phi(x) := (1 - k_j x)^+$ is convex on $(0, \infty)$, being the maximum of the affine functions 0 and $1 - k_j x$. The forward-looking caplet at m is the discounted intrinsic value $\text{Caplet}_j^{\text{fr}}(m) = (1 - k_j \, p(m, N))^+ = \phi\big(p(m, N)\big)$, and by the pricing formula (1.7) $p(m, N) = E^Q[D(m, N) \mid \mathcal{F}_m]$. The backward-looking caplet at m is $\text{Caplet}_j^{\text{br}}(m) = E^Q[\phi(D(m, N)) \mid \mathcal{F}_m]$ by (3.16)–(3.17). Conditional Jensen's inequality applied to the convex ϕ gives

$$\text{Caplet}_j^{\text{fr}}(m) = \phi\big(E^Q[D(m, N) \mid \mathcal{F}_m]\big) \leq E^Q\big[\phi(D(m, N)) \mid \mathcal{F}_m\big] = \text{Caplet}_j^{\text{br}}(m).$$

For general $n \leq m$, both prices are obtained by applying $E^Q[D(n, m) \cdot \mid \mathcal{F}_n]$ to the respective time-m prices (Corollary 1.1). Since $D(n, m) \geq 0$ and conditional expectation is monotone, the inequality is preserved. Equality in conditional Jensen holds precisely when $D(m,N)$ is F_m-measurable, which in the Gaussian model means $\nu_{m,N}=0$. $\square$

This is a standard result known in the continuous time literature, for instance see [14].

Remark 3.10 (Forward versus backward: a pure variance effect). Theorems 3.3 and 3.5 have the *identical* Black-type structure: the same prefactors $p(n,m)$ and $p(n,N)$, the same constant k_j, and $\hat{d}_{1,n}^{(j)}$, $\hat{d}_{2,n}^{(j)}$ are obtained from $d_{1,n}^{(j)}$, $d_{2,n}^{(j)}$ by the single substitution $\nu_X \mapsto \nu_W$. By Proposition 3.2 the two prices therefore differ *only* through their total variance, the gap being governed entirely by the in-accrual variance $\nu_{m,N}\geq 0$.

Seen from time $n<m$, the forward-looking rate L_j^{fr} carries only the uncertainty accumulated over the pre-accrual window $[n,m]$, through which the fixing $p(m,N)$, and hence L_j^{fr}, is determined; once m is reached the rate is locked and no further randomness enters its payoff. The backward-looking rate L_j^{br}, by contrast, keeps compounding the overnight fixings $r_m,\dots,r_{N-1}$ throughout the accrual window $[m,N]$, so it carries the *same* pre-accrual uncertainty *plus* the in-accrual uncertainty $\nu_{m,N}$. Since the Black-type price is increasing in the total variance and the two contracts share the same forward and strike, the backward-looking caplet is always at least as valuable as the forward-looking one with identical terms.

This is sharpest at $n=m$, where $\nu_W=0$: (3.76) then collapses to the discounted intrinsic value $\mathrm{Caplet}_j^{\mathrm{fr}}(m)=\big(1-k_j\,p(m,N)\big)^+$, reflecting that L_j^{fr} is already fixed, whereas the backward-looking caplet retains at $n=m$ the strictly positive variance $\nu_{m,N}$ and hence optionality (Theorem 3.1).

3.6.8 The Forward-Looking Floorlet at Earlier Times

Theorem 3.6 (Forward-looking floorlet price at earlier times). *Under the discrete Hull-White model, for every $n\leq m<N$,*

$$\boxed{\mathrm{Floorlet}_j^{\mathrm{fr}}(n)=k_j\,p(n,N)\,\Phi(-\hat{d}_{2,n}^{(j)})-p(n,m)\,\Phi(-\hat{d}_{1,n}^{(j)}),}\tag{3.85}$$

where $\hat{d}_{1,n}^{(j)}$ and $\hat{d}_{2,n}^{(j)}$ are as in Theorem 3.5.

Proof. The forward-looking floorlet pays $\alpha_j(K-L_j^{\mathrm{fr}})^+$ at time N. The same payoff reduction and collapse of the m-to-N discounting give

$$\mathrm{Floorlet}_j^{\mathrm{fr}}(n)=E_n^Q\big[D(n,m)\,(k_j\,p(m,N)-1)^+\big]=k_j\,E_n^Q\big[e^{-Y_{n,m}-W}\mathbf{1}_{\{W<\log k_j\}}\big]-E_n^Q\big[e^{-Y_{n,m}}\mathbf{1}_{\{W<\log k_j\}}\big],$$

the exercise region now being $\{W<\log k_j\}$. Applying Lemma 3.5 (the indicator $\mathbf{1}_{\{W<\log k_j\}}$ producing $\Phi(-\,\cdot\,)$) together with the prefactor identifications (3.81) and (3.82) yields (3.85). $\square$

Remark 3.11 (Shared caplet-floorlet parity). Subtracting (3.85) from (3.76) and using $\Phi(x) + \Phi(-x) = 1$ gives

$$\text{Caplet}_j^{\text{fr}}(n) - \text{Floorlet}_j^{\text{fr}}(n) = p(n,m) - k_j\, p(n,N),$$

which coincides with the backward-looking spread obtained from Proposition 3.1.

3.7 Absence of a Closed Form under Simple-Rate Compounding

The pricing formulas derived in the previous sections were obtained under the single-period convention $p(k,k+1) = e^{-r_k}$, for which the compounded backward-looking rate equals $1/D(m,N) - 1$ exactly and the accrual factor reduces to $\exp(\sum_{k=m}^{N-1} r_k)$. We now consider the alternative specification in which the single-period rate is treated as a simple rate, so that the accrual factor retains the product form $\prod_{k=m}^{N-1}(1 + r_k)$. We discuss whether a closed-form price is available under this specification, and show that the product structure obstructs it.

3.7.1 Fully Simple-Rate Specification

Define

$$A_{m,N} := \prod_{k=m}^{N-1}(1 + r_k), \qquad D(m,N) = A_{m,N}^{-1}. \tag{3.86}$$

The caplet price under this specification is

$$\text{Caplet}_j^{\text{br}}(m) = E^Q\Big[\big(1 - k_j\, A_{m,N}^{-1}\big)^+ \;\big|\, \mathcal{F}_m\Big]. \tag{3.87}$$

From the forward iteration (3.22) of the Hull-White recursion, each r_{m+h} is an affine function of r_m and the innovations $\omega_m, \ldots, \omega_{m+h-1}$. Therefore

$$\prod_{k=m}^{N-1}(1 + r_k) = P(z), \tag{3.88}$$

where P is a polynomial of degree $N - m - 1$ in the $(N - m - 1)$-dimensional Gaussian vector $z = (\omega_m, \ldots, \omega_{N-2})$. The caplet price can then be written as

$$\int_{\{z:\, P(z) > k_j\}} \left(1 - \frac{k_j}{P(z)}\right) \varphi_{N-m-1}(z)\, dz, \tag{3.89}$$

where $\varphi_{N-m-1}(z) = \prod_{i=1}^{N-m-1} \varphi(z_i)$ is the density of an $(N - m - 1)$-dimensional vector of independent standard Gaussians.

The integral (3.89) does not lead to a tractable closed-form expression analogous to the one obtained under the compounded convention. The main obstruction is that the exercise region $\{P(z) > k_j\}$ is no longer described by a linear inequality in a Gaussian random variable, but by a polynomial inequality in several Gaussian variables. Consequently, the problem cannot be reduced to a one-dimensional Gaussian tail probability.

For long accrual periods, the boundary of the region $\{P(z) > k_j\}$ is determined by a high-degree polynomial equation. In general, such boundaries cannot be represented in a simple analytical form, and even in low-degree cases the resulting expressions would be too complicated to yield a useful pricing formula. Moreover, the integrand contains the reciprocal term $1/P(z)$, which further prevents the integral from being expressed in terms of elementary functions or standard Gaussian distribution functions.

3.7.2 Mixed Specification

Consider instead a mixed specification in which the compounding factor is kept in its simple-rate product form, while the discount factor is modeled through exponential compounding:

$$A_{m,N} = \prod_{k=m}^{N-1}(1 + r_k), \qquad D(m, N) = \exp\left(-\sum_{k=m}^{N-1} r_k\right). \tag{3.90}$$

The corresponding caplet price is

$$\text{Caplet}_j^{\text{br}}(m) = E^Q\left[e^{-\sum_{k=m}^{N-1} r_k}\,(A_{m,N} - k_j)^+ \;\middle|\; \mathcal{F}_m\right]. \tag{3.91}$$

Conditionally on $\mathcal{F}_m$, the variable r_m is known, while $r_{m+1}, \ldots, r_{N-1}$ depend on the Gaussian innovations

$$z = (\omega_m, \ldots, \omega_{N-2}) \in \mathbb{R}^{N-m-1}.$$

Therefore, the sum $\sum_{k=m}^{N-1} r_k$ can be written as a linear function $Q(z)$, while the compounding factor $A_{m,N}$ can be written as a polynomial $P(z)$ of degree $N - m - 1$. The price then takes the form

$$\int_{\{z:\, P(z)>k_j\}} e^{-Q(z)}\,(P(z) - k_j)\,\varphi_{N-m-1}(z)\,dz, \tag{3.92}$$

where φ_{N-m-1} denotes the density of an $(N - m - 1)$-dimensional standard Gaussian vector.

The exponential factor $e^{-Q(z)}$ is not the main difficulty: since $Q(z)$ is linear, it can be absorbed into the Gaussian density by completing the square, as in the proof of the closed-form caplet formula. The obstruction comes instead from the exercise region $\{P(z) > k_j\}$. Unlike the compounded convention, this region is not described by a linear inequality in a Gaussian random variable, but by a polynomial inequality in several Gaussian variables.

Consequently, the pricing problem cannot be reduced to a one-dimensional Gaussian tail probability, and no tractable closed-form expression is available in general.

Remark 3.12. The simple-rate product destroys the linear-Gaussian form that made the closed-form formula possible. Under the compounded convention, the exercise condition is

$$\sum_{k=m}^{N-1} r_k > \log k_j,$$

which is a linear inequality in a Gaussian random variable. The pricing problem therefore reduces to the computation of truncated Gaussian expectations. By contrast, under the simple-rate specification, the exercise condition becomes

$$P(z) > k_j,$$

where P is a polynomial in several Gaussian shocks. The boundary of the exercise region is then a nonlinear hypersurface, and the resulting integral does not reduce to standard Gaussian distribution functions.

In practice, the price under this specification can still be computed numerically, for example by Monte Carlo simulation or numerical integration. This is feasible because the number of compounding dates $N - m$ is finite, but the analytical closed form obtained under the compounded convention is lost.

Chapter 4

Numerical Analysis of Backward-Looking Caplet Prices

The previous chapter derived closed-form prices for backward-looking caplets and floorlets under the discrete Hull-White model, both at the start of the accrual period (Theorem 3.1) and at any earlier valuation date (Theorem 3.3), together with the variance decomposition $\nu_X = \nu_W + \nu_{m,N}$ that separates the pre-accrual and in-accrual sources of uncertainty (Proposition 3.2). All of these results are analytic: the price at any date is a deterministic function of the observed bond prices and of the model parameters a and σ. This chapter exploits that fact to study, purely numerically, how the price responds to the volatility, the mean reversion, the strike, and the length of the pre-accrual and accrual windows. Because every quantity is available in closed form, no Monte Carlo simulation is required and, as explained in Section 4.1, the drift sequence Φ_n never enters: the calibration to the initial term structure has already been absorbed into the bond prices that appear in the pricing formulas.

The chapter is organized as follows. Section 4.1 fixes the time grid, the base-case parameters, a stylized discount curve, and the normalization used throughout, and records the closed forms in the notation used by the implementation. Section 4.2 then studies the backward-looking caplet one parameter at a time, beginning with a curve-independent illustration of the variance decomposition and ending with the joint effect of time and curve shape. Section 4.3 builds on that to describe the gap between the backward- and forward-looking caplets analytically, reducing it to a single integral and, at the money, to a closed form in the variance ratio $\nu_{m,N}/\nu_W$. Section 4.4 closes the chapter by tracing that gap across the model inputs, so that its analytical description and its numerical illustration stand side by side.

4.1 Numerical Setup

4.1.1 Time Grid, Base-Case Parameters, and the Volatility Convention

We work on the daily grid $\Delta = 1/360$, consistent with the ACT/360 day-count convention used for SOFR, so that one step represents one calendar day and a three-month accrual corresponds to $L = N - m = 90$ steps. Unless stated otherwise, the base case uses a pre-accrual window of the same length, $m = 90$, and a mean reversion $a = 0.75$. Since this chapter carries out a sensitivity analysis over a reasonable range of parameters rather than a calibration to market data, $a = 0.75$ is taken as a representative value. It is also the choice made in the numerical study of discrete backward-looking caplets by Iunevich and Kalikaeva [12]. It sits well inside the stability regime $0 < a\Delta < 1$ discussed in Section 2.2: here $a\Delta \approx 2.1 \times 10^{-3}$, so the autoregressive coefficient $c = 1 - a\Delta \approx 0.9979$ lies strictly in $(0, 1)$, close to its upper boundary.

A word is needed on the volatility parameter, because the short rate in this model is a per-period quantity. By definition $r_n = -\log p(n, n+1)$ is the rate compounded over a single step of length Δ, so its natural annualized counterpart is $\hat{r}_n := r_n/\Delta$. The one-step conditional standard deviation of r_n under the recursion (2.6) is $\sigma\sqrt{\Delta}$; that of $\hat{r}_n$ is therefore $\sigma/\sqrt{\Delta}$, and accumulating the innovation variance over the $1/\Delta$ steps of one year gives an annualized short-rate volatility

$$\sigma_{\text{ann}} = \frac{\sigma}{\Delta}, \qquad \text{equivalently} \qquad \sigma = \sigma_{\text{ann}}\,\Delta. \tag{4.1}$$

We fix the base case by the interpretable quantity: an annualized short-rate volatility of $\sigma_{\text{ann}} = 100$ basis points, i.e. $\sigma = \sigma_{\text{ann}}\Delta \approx 2.78 \times 10^{-5}$. This is the parameter that enters the variance formulas of Chapter 3; quoting it through (4.1) avoids the otherwise misleading appearance of an extremely small number and ties the volatility sweep below to a recognizable scale.

4.1.2 Stylized Discount Curve and Price Normalization

Two distinct curves are built from the bonds and must not be confused. The *price curve* $N \mapsto p(0, N) = e^{-y(N)N\Delta}$ is decreasing in N for any non-negative yield function, simply because a more distant unit payment is discounted over a longer horizon; longer-dated zero-coupon bonds are always cheaper. The *yield curve* $N \mapsto y(N)$, by contrast, may slope either way, and it is to this object that the terms "upward" (normal) and "inverted" refer. The distinction matters because the yield-curve shape affects the experiments below only through the moneyness of the contract, and, as we now explain, in most of them it does not enter at all.

The price at $n = 0$ of the backward-looking caplet (Theorem 3.3 with $n = 0$) is

$$\text{Caplet}_j^{\text{br}}(0) = p(0,m)\,\Phi(d_1) - k_j\,p(0,N)\,\Phi(d_2), \qquad d_1 = \frac{\log \frac{p(0,m)}{k_j\,p(0,N)} + \frac{1}{2}\nu_X}{\sqrt{\nu_X}}, \quad d_2 = d_1 - \sqrt{\nu_X}. \tag{4.2}$$

The bonds enter only through the prefactors $p(0,m)$, $k_j p(0,N)$ and through the moneyness term $\log \frac{p(0,m)}{k_j p(0,N)}$. Writing $p(0,N) = p(0,m)/(1+\alpha_j K^\star)$, where

$$K^\star := L(0;m,N) = \frac{1}{\alpha_j}\left(\frac{p(0,m)}{p(0,N)} - 1\right) \tag{4.3}$$

is the time-0 forward rate of (1.1) and hence the at-the-money (ATM) strike, the price factorizes as

$$\text{Caplet}_j^{\text{br}}(0) = p(0,m)\,\widehat{C}(K;\nu_X), \qquad \widehat{C}(K;\nu_X) = \Phi(d_1) - \frac{1+\alpha_j K}{1+\alpha_j K^\star}\,\Phi(d_2), \tag{4.4}$$

with $d_1 = \left[\log \frac{1+\alpha_j K^\star}{1+\alpha_j K} + \frac{1}{2}\nu_X\right]/\sqrt{\nu_X}$. The *normalized price* $\widehat{C}$ depends on the curve only through the single scalar $K^\star$, not through the level of the bonds; the prefactor $p(0,m)$ is a positive constant whenever the maturities are held fixed, and rescales the price without altering the shape of any sensitivity. We therefore report $\widehat{C}$ (equivalently, the price per unit $p(0,m)$, in basis points) in the volatility, mean-reversion and strike experiments, where the maturities are fixed. In particular, at the money the moneyness term vanishes and

$$\widehat{C}^{\text{ATM}}(\nu_X) = \Phi\big(\tfrac{1}{2}\sqrt{\nu_X}\big) - \Phi\big(-\tfrac{1}{2}\sqrt{\nu_X}\big) = 2\Phi\big(\tfrac{1}{2}\sqrt{\nu_X}\big) - 1 \tag{4.5}$$

depends on ν_X alone. The full curve is reinstated only in the time experiment of Section 4.2.5, where the maturities move and the yield-curve shape genuinely affects the price; there we take the linear yield $y(N) = y_0 + s\,N\Delta$ with base level $y_0 = 3\%$ and slope s set to 0 (flat), +0.4% per year of maturity (upward) and −0.4% (inverted).

4.1.3 Closed Forms in Implementation Notation

For the implementation we collect the variance components from Lemmas 3.6–3.7 and Proposition 3.2 in terms of the partial sums S_ℓ of (2.31), recalling $1 + B_m^N = S_L$ and $1 + B_{q+1}^N = S_{N-q-1}$ with $L = N - m$:

$$\nu_{m,N} = \sigma^2\Delta\sum_{\ell=1}^{L-1} S_\ell^2, \qquad \nu_W = \sigma^2\Delta\,S_L^2\,\frac{1-c^{2m}}{1-c^2}, \qquad \nu_X = \nu_W + \nu_{m,N}. \tag{4.6}$$

Two features of (4.6) drive the entire analysis and are worth stating before any plot. First, the variance is *curve-independent*: no bond price appears in ν_W, $\nu_{m,N}$ or ν_X. Second, the volatility factorizes as a pure scale,

$$\nu_X = \sigma^2 \cdot \Delta\, H(a; m, L), \tag{4.7}$$

with H independent of σ, so that σ sets the level of the variance and a (through c and the sums S_ℓ) sets its shape.

All variance components are evaluated directly from the closed forms (4.6), with the partial sums S_ℓ taken from their definition (2.31). Every result below is produced by the accompanying R script.[1]

Quantity	**Base-case value**
Time step Δ	1/360
Pre-accrual / accrual length m, L	90, 90 (3M / 3M, daily)
Mean reversion a	0.75
Annualized volatility σ_{ann}	100 bp ($\sigma = \sigma_{\text{ann}}\Delta \approx 2.78 \times 10^{-5}$)
$c = 1 - a\Delta$, c^L, S_L	0.99792, 0.829, 82.1 (vs. $L = 90$)
Pre-accrual variance ν_W	1.09×10^{-6} ($\sqrt{\nu_W} = 1.04 \times 10^{-3}$)
In-accrual variance $\nu_{m,N}$	4.47×10^{-7} ($\sqrt{\nu_{m,N}} = 6.69 \times 10^{-4}$)
Total variance ν_X	1.53×10^{-6} ($\sqrt{\nu_X} = 1.24 \times 10^{-3}$)
ATM normalized caplet $\widehat{C}^{\text{ATM}}$	4.94×10^{-4} (= 4.94 bp of $p(0, m)$)

Table 4.1: Base-case parameters and the resulting variance components and ATM caplet price. At the base maturities the pre-accrual component accounts for 70.9% of the total variance and the in-accrual component for 29.1%.

4.2 Sensitivity of the Backward-Looking Caplet

We now study the backward-looking caplet price as a function of one input at a time, holding all others at the base case of Table 4.1. Moving a single parameter at a time is what lets each figure be read directly against the closed-form structure of Chapter 3. We begin with the variance itself, which is the channel through which σ, a, and the time windows act, and which is moreover entirely free of the curve.

4.2.1 The Variance Decomposition

Figure 4.1 plots the three variance components of (4.6) directly, with no price and no curve. Panel (i) fixes the accrual length at $L = 90$ and varies the pre-accrual horizon m. As m grows, the pre-accrual variance ν_W rises and saturates at the limit $\sigma^2\Delta\, S_L^2/(1 - c^2)$

[1]The R code reproducing all figures and numerical results of this chapter is available at `https://github.com/Seba-Pinotti/discrete-hull-white-caplets-floorlets`.

obtained by letting $c^{2m} \to 0$ in (4.6): uncertainty accumulated further and further in the past is discounted by the mean reversion and contributes a bounded amount. The in-accrual variance $\nu_{m,N}$, in contrast, is exactly flat in m: for fixed L it is translation-invariant, since the summands S_ℓ^2 depend only on the accrual length and not on its position on the grid. The total ν_X is their sum, and the figure is thus a direct rendering of the decomposition $\nu_X = \nu_W + \nu_{m,N}$ of Proposition 3.2: one growing and saturating component, one constant component. Panel (ii) varies the accrual length L and shows $\nu_{m,N}$ increasing in L, the convex growth reflecting that each additional accrual date adds a squared partial sum S_ℓ^2 to the sum.

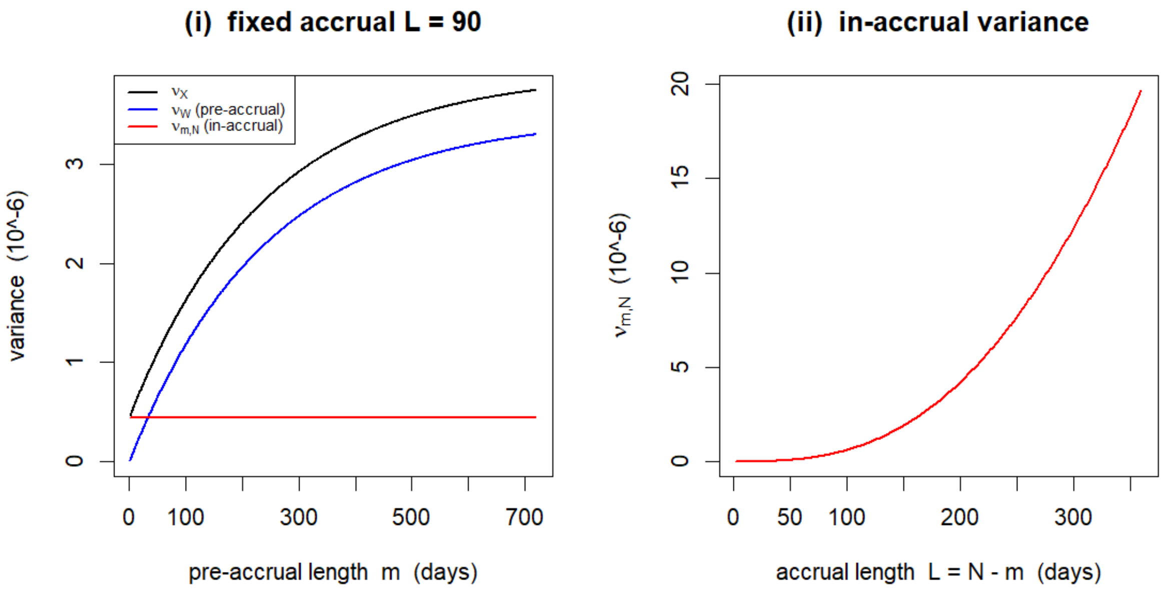

Figure 4.1: Variance decomposition $\nu_X = \nu_W + \nu_{m,N}$, curve-independent. *(i)* Fixed accrual $L = 90$: the pre-accrual variance ν_W grows and saturates while the in-accrual variance $\nu_{m,N}$ is flat in m; their sum is ν_X. *(ii)* In-accrual variance $\nu_{m,N}$ as a function of the accrual length L.

At the base maturities the pre-accrual component already dominates, accounting for about 71% of ν_X (Table 4.1); this is a consequence of the base pre-accrual window being as long as the accrual window, so that ν_W has had time to build up most of its saturation value. The relative weight of the two components shifts as the windows are varied, and it is precisely the in-accrual component $\nu_{m,N}$ that will distinguish the backward-looking caplet from its forward-looking counterpart in the following sections.

4.2.2 Sensitivity to Volatility

Figure 4.2 shows the ATM normalized caplet as a function of the annualized volatility σ_{ann}, with maturities fixed at the base case. By (4.7) the variance is exactly quadratic in σ, so $\sqrt{\nu_X}$ is linear in σ and, by (4.5), the ATM price is the standard Black response $2\Phi(\frac{1}{2}\sqrt{\nu_X}) - 1$ to a volatility input. The curve is increasing and, over the realistic range shown, very nearly linear: for the small total volatilities involved ($\sqrt{\nu_X} \approx 1.2 \times 10^{-3}$ at the base case), Φ is close to its linear regime around the origin, so $\widehat{C}^{\mathrm{ATM}} \approx \sqrt{\nu_X/(2\pi)}$ to leading order.

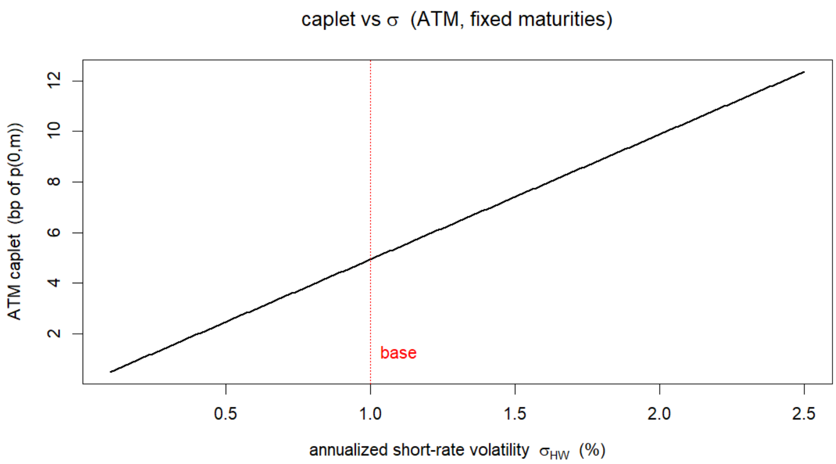

Figure 4.2: ATM backward-looking caplet (normalized, in basis points of $p(0,m)$) as a function of the annualized short-rate volatility $\sigma_{\mathrm{ann}} = \sigma/\Delta$, at fixed base-case maturities. The dotted line marks the base case $\sigma_{\mathrm{ann}} = 100$ bp.

Remark 4.1 (Small-volatility ATM limit). The approximation $\widehat{C}^{\mathrm{ATM}} \approx \sqrt{\nu_X/(2\pi)}$ is the first-order Taylor expansion of the exact ATM price (4.5) around zero volatility. Writing $x := \frac{1}{2}\sqrt{\nu_X}$ and expanding the standard normal distribution function around the origin,

$$\Phi(x) = \Phi(0) + \varphi(0)\,x + \tfrac{1}{2}\varphi'(0)\,x^2 + O(x^3) = \tfrac{1}{2} + \frac{1}{\sqrt{2\pi}}\,x + o(x^2),$$

where the constant term is $\Phi(0) = \frac{1}{2}$, the linear coefficient is the density at the origin $\varphi(0) = 1/\sqrt{2\pi}$, and the quadratic term vanishes because $\varphi'(0) = -0 \cdot \varphi(0) = 0$, so the leading correction is of order x^2. Hence $2\Phi(x) - 1 = 2x/\sqrt{2\pi} + o(x^2)$, and substituting $x = \frac{1}{2}\sqrt{\nu_X}$ the factor 2 and the inner $\frac{1}{2}$ cancel:

$$\widehat{C}^{\mathrm{ATM}}(\nu_X) = 2\Phi\big(\tfrac{1}{2}\sqrt{\nu_X}\big) - 1 = \frac{2}{\sqrt{2\pi}} \cdot \tfrac{1}{2}\sqrt{\nu_X} + o\big(\nu_X\big) = \sqrt{\frac{\nu_X}{2\pi}} + o\big(\nu_X\big).$$

4.2.3 Sensitivity to Mean Reversion

The mean reversion acts differently. Whereas σ sets the level of the variance, a sets its shape: through $c = 1 - a\Delta$ and $S_\ell = (1 - c^\ell)/(a\Delta)$, a larger a produces a smaller c, compresses the partial sums S_ℓ, and hence reduces the variance, with the compression biting harder over longer windows. Figure 4.3(i) shows the ATM caplet decreasing smoothly in

a, and panel (ii) decomposes the effect: both ν_W and $\nu_{m,N}$ fall with a, but the pre-accrual component falls considerably faster. This asymmetry is structural: ν_W carries the factor S_L^2 *and* the geometric pre-accrual sum $(1-c^{2m})/(1-c^2)$, both of which contract as a grows, whereas $\nu_{m,N}$ contracts only through the S_ℓ^2 in its sum. Mean reversion thus reshapes the split between the two variance sources.

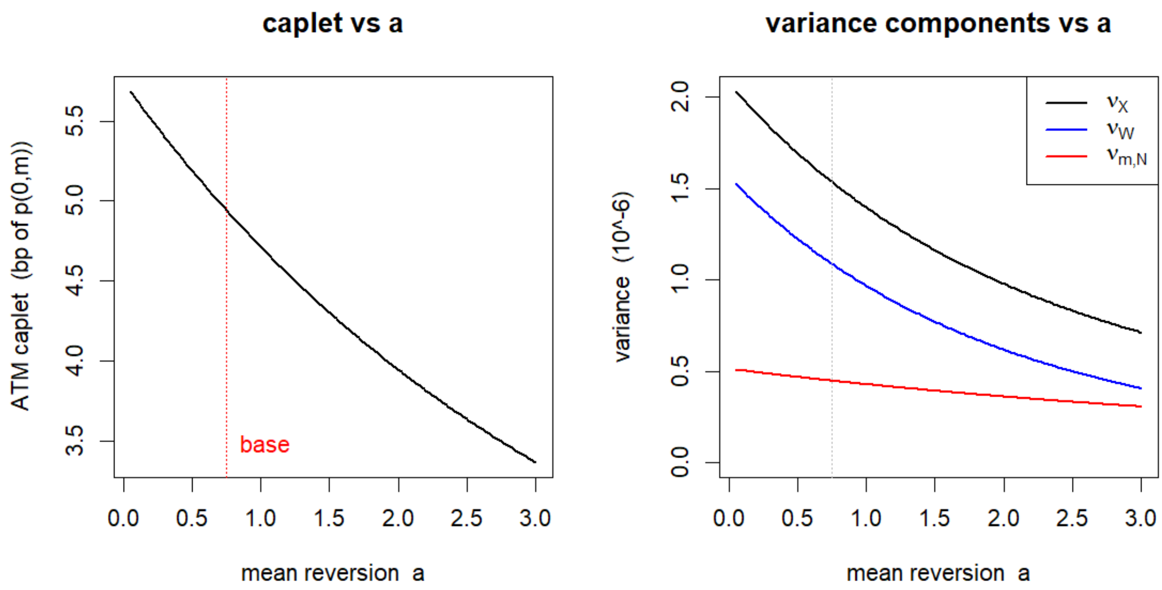

Figure 4.3: Sensitivity to mean reversion. *(i)* ATM normalized caplet (bp of $p(0,m)$) versus a. *(ii)* Variance components ν_W, $\nu_{m,N}$ and their sum ν_X versus a. The dotted line marks the base case $a = 0.75$.

4.2.4 Sensitivity to the Strike

Figure 4.4 sweeps the strike K around the ATM level $K^\star$ of (4.3), taken here at a representative $K^\star = 3.5\%$, and plots both the caplet and the floorlet. The caplet is decreasing and convex in K, the floorlet increasing and convex, as expected from their payoffs. The two cross exactly at the money: at $K = K^\star$ one has $k_j = p(0,m)/p(0,N)$, so the parity relation of Proposition 3.1,

$$\mathrm{Caplet}_j^{\mathrm{br}}(0) - \mathrm{Floorlet}_j^{\mathrm{br}}(0) = p(0,m) - k_j\, p(0,N),$$

vanishes and the two prices coincide. The figure is thus also a numerical check of parity: the crossing point is pinned to $K^\star$ independently of the volatility and mean reversion, which enter only through ν_X and shift the common height at the crossing but not its location. Because the floorlet is tied to the caplet by parity, we do not study it separately in the remaining experiments, its sensitivities are those of the caplet plus a deterministic term linear in the bond prices.

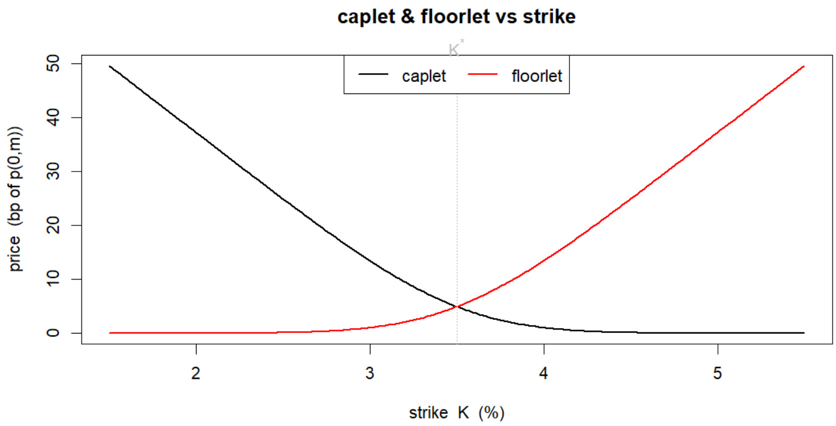

Figure 4.4: Backward-looking caplet and floorlet (normalized, bp of $p(0, m)$) versus strike K, at fixed base-case maturities and variance. The dotted line marks the ATM strike $K^\star$, where the two prices cross by caplet-floorlet parity (Proposition 3.1).

4.2.5 Sensitivity to Pre-Accrual and In-Accrual Time

The final experiment combines the curve-free variance growth of Section 4.2.1 with the moneyness drift that arises when the maturities are moved along a sloped yield curve. Here we report the actual price (4.2) rather than the normalized one, since the curve is the object of interest, and we hold the strike fixed at an external value $K = 3.01\%$, chosen as the ATM forward at the base maturities on the *flat* curve, so that on the flat curve the contract remains close to the money as the windows are varied. Figure 4.5 shows the result on the upward, flat, and inverted yield curves of Section 4.1.2.

On the flat curve (dashed) the ratio $p(0, m)/p(0, N)$ is constant as the pre-accrual window is lengthened, so the moneyness term in (4.2) is frozen and the price moves only through the variance. Panel (i) shows the flat-curve price tracing the same growing and saturating profile as ν_W in Figure 4.1(i): with moneyness held fixed, the caplet is a monotone image of the total variance. The sloped curves add a tilt on top of this common behavior. On the upward curve the forward rate over the accrual window rises as the window is pushed out, the contract drifts in the money, and the price tilts upward. On the inverted curve the forward falls, the contract drifts out of the money, and the price is pulled down and eventually decays. Panel (ii), which lengthens the accrual window, shows the same separation: the flat curve grows through the rising in-accrual variance, while the upward and inverted curves diverge from it according to the direction of the moneyness drift. The broad time sensitivity is common to all three regimes in that the variance-driven growth of Figure 4.1 is always present, the curve shape then governs the magnitude and the direction of the tilt, and on a sufficiently inverted curve the out-of-the-

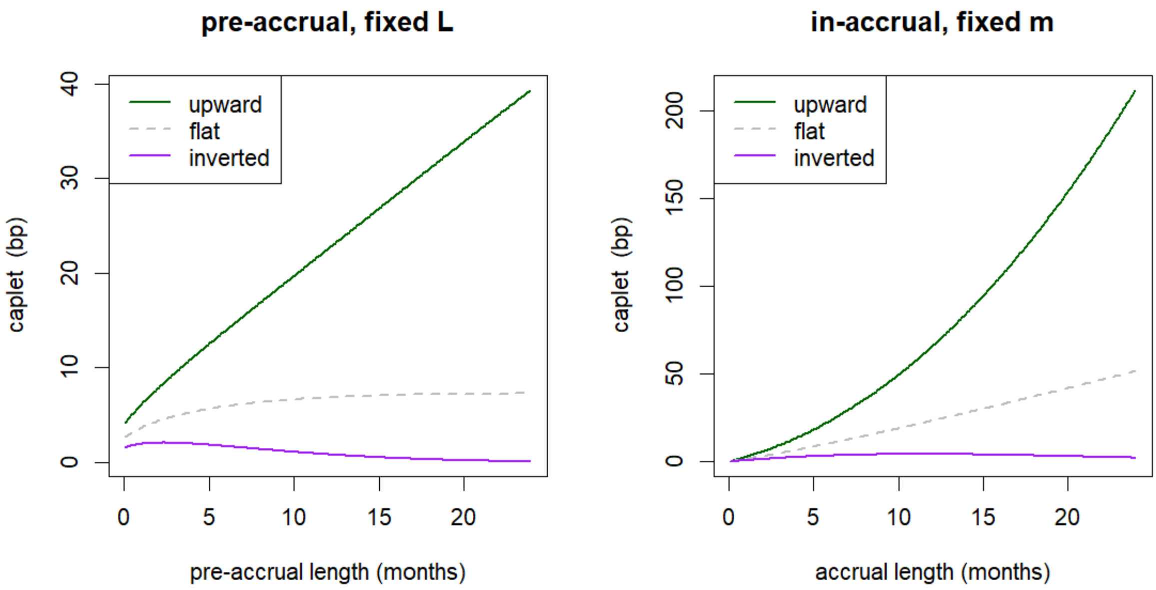

Figure 4.5: Backward-looking caplet price (bp) under upward, flat and inverted yield curves, at a fixed external strike. *(i)* Pre-accrual sensitivity: fixed accrual $L = 90$, varying pre-accrual length m. *(ii)* In-accrual sensitivity: fixed pre-accrual $m = 90$, varying accrual length L. The flat curve isolates the pure variance effect. The sloped curves add a moneyness tilt.

money moneyness drift can eventually outweigh that growth and pull the price back down. The underlying mechanism of Chapter 3 is thus robust to the term-structure environment.

4.3 Analytical Description of the Backward–Forward Gap

We now turn to the comparison that motivates the whole construction: the backward-looking caplet, written on the rate compounded *over* the accrual window, against the forward-looking caplet, written on the rate fixed at the *start* of that window. Chapter 3 established that the former is never cheaper than the latter (Proposition 3.3), the economic content being that the backward rate carries the additional in-accrual uncertainty that the forward rate locks out. This section makes that statement quantitative: it derives an exact expression for the price difference, identifies the single quantity responsible for it, and extracts a closed form.

The starting point is a structural observation already implicit in Chapter 3. Valued at a date $n < m$ before the accrual begins, both caplets carry the *same* forward bond prices and the same strike, and differ only in the variance fed into the pricing formula: the forward-looking caplet (Theorem 3.5) uses the pre-accrual variance ν_W, while the backward-looking caplet (Theorem 3.3) uses the total variance $\nu_X = \nu_W + \nu_{m,N}$ of Proposition 3.2. It is therefore natural to regard the price as a function of the variance alone. Fix n, m, N and the strike, write the log-moneyness

$$M := \log \frac{p(n,m)}{k_j\, p(n,N)}, \tag{4.8}$$

and define, for $v > 0$,

$$C(v) := p(n,m)\,\Phi\big(d_1(v)\big) - k_j\, p(n,N)\,\Phi\big(d_2(v)\big), \qquad d_{1,2}(v) = \frac{M \pm \frac{1}{2}v}{\sqrt{v}}. \tag{4.9}$$

The forward-looking caplet is then $C(\nu_W)$, the backward-looking caplet is $C(\nu_X)$, and the object of interest is the gap

$$G := C(\nu_X) - C(\nu_W) = C(\nu_W + \nu_{m,N}) - C(\nu_W). \tag{4.10}$$

Proposition 4.1 (Integral representation of the gap). *For any strike and any valuation date $n < m$,*

$$G \;=\; \int_{\nu_W}^{\nu_X} \frac{p(n,m)\,\varphi\big(d_1(v)\big)}{2\sqrt{v}}\,\mathrm{d}v \;=\; \int_{\nu_W}^{\nu_W+\nu_{m,N}} \frac{p(n,m)\,\varphi\big(d_1(v)\big)}{2\sqrt{v}}\,\mathrm{d}v, \tag{4.11}$$

where φ is the standard normal density. The integrand is strictly positive, so $G > 0$ whenever $\nu_{m,N} > 0$.

Proof. Since C in (4.9) is differentiable in v on $(0,\infty)$, the fundamental theorem of calculus gives $G = \int_{\nu_W}^{\nu_X} C'(v)\,\mathrm{d}v$. Differentiating,

$$C'(v) = p(n,m)\,\varphi(d_1)\,\frac{\partial d_1}{\partial v} - k_j\, p(n,N)\,\varphi(d_2)\,\frac{\partial d_2}{\partial v}.$$

The two density terms are equal. Indeed, from (4.9), $d_1 + d_2 = 2M/\sqrt{v}$ and $d_1 - d_2 = \sqrt{v}$, so $d_1^2 - d_2^2 = (d_1+d_2)(d_1-d_2) = 2M$, whence

$$\frac{\varphi(d_1)}{\varphi(d_2)} = e^{-\frac{1}{2}(d_1^2 - d_2^2)} = e^{-M} = \frac{k_j\, p(n,N)}{p(n,m)}, \qquad \text{i.e.} \qquad p(n,m)\,\varphi(d_1) = k_j\, p(n,N)\,\varphi(d_2).$$

Substituting this identity collapses C' to $C'(v) = p(n,m)\,\varphi(d_1)\,\frac{\partial(d_1-d_2)}{\partial v}$, and since $d_1 - d_2 = \sqrt{v}$ we have $\partial_v(d_1 - d_2) = 1/(2\sqrt{v})$. Hence $C'(v) = p(n,m)\,\varphi(d_1(v))/(2\sqrt{v}) > 0$, which is the integrand in (4.11); positivity of the integrand over the non-degenerate interval $[\nu_W, \nu_X]$ gives $G > 0$. $\square$

Remark 4.2 (The gap as priced in-accrual variance). The quantity

$$\mathcal{V}(v) := C'(v) = \frac{p(n,m)\,\varphi\big(d_1(v)\big)}{2\sqrt{v}}$$

is the sensitivity of the price to the total variance, the option's *variance vega.* Proposition 4.1 therefore reads: the backward-looking premium is the extra variance $\nu_{m,N}$ integrated against the variance vega as the variance is raised from ν_W to ν_X. This is a differential counterpart to the Jensen argument of Proposition 3.3: rather than comparing two convex payoffs, one observes that the price climbs along a strictly positive vega over exactly the interval of length $\nu_{m,N}$. The premium vanishes if and only if $\nu_{m,N} = 0$, that is, if and only if there is no in-accrual uncertainty and the two rates coincide. The component $\nu_{m,N}$, shown in Section 4.2.1 to be the non-saturating, in-accrual part of the variance, is thus the exact and *sole* source of the backward-looking premium. Everything that makes $\nu_{m,N}$ large increases the gap.

Proposition 4.2 (At-the-money gap). *At the money, $k_j = p(n,m)/p(n,N)$, so $M = 0$ and $d_1(v) = \frac{1}{2}\sqrt{v}$. The integrand of* (4.11) *is then $\dfrac{p(n,m)}{2\sqrt{2\pi}}\,\dfrac{e^{-v/8}}{\sqrt{v}}$, and for the small variances relevant here $\big(v \le \nu_X \sim 10^{-6}$, so $e^{-v/8} = 1 - O(10^{-7})\big)$,*

$$G^{\mathrm{ATM}} \;\approx\; \frac{p(n,m)}{\sqrt{2\pi}}\Big(\sqrt{\nu_X} - \sqrt{\nu_W}\Big). \tag{4.12}$$

The gap relative to the forward-looking caplet is then independent of the volatility, of the bond level, and of $p(n,m)$:

$$\frac{G^{\mathrm{ATM}}}{C^{\mathrm{ATM}}(\nu_W)} \;\approx\; \sqrt{\frac{\nu_X}{\nu_W}} - 1 \;=\; \sqrt{1 + \frac{\nu_{m,N}}{\nu_W}} - 1, \tag{4.13}$$

a pure function of the variance ratio $\nu_{m,N}/\nu_W$.

Expression (4.12) is the difference of the two single-caplet at-the-money prices, each in the small-volatility form $C^{\mathrm{ATM}}(v) \approx p(n,m)\sqrt{v/(2\pi)}$ of Remark 4.1. It agrees with the exact gap to a relative error of 1.6×10^{-7} at the base case. Numerically, the forward-looking caplet is worth 4.16 bp of $p(n,m)$ there and the backward-looking caplet 4.94 bp, a gap of 0.78 bp, i.e. a premium of (4.13) $\approx$ 18.8%.

The closed forms (4.12)-(4.13) and the integral (4.11) already determine how the gap responds to each input. Section 4.4 then displays each dependence in turn. Since both variances scale as σ^2 (equation (4.7)), the at-the-money gap (4.12) is exactly linear in σ, while the relative gap (4.13) is independent of it. The gap grows with the accrual length L without bound: $\nu_{m,N}$ increases in L while ν_W, as a function of L at fixed n and m, saturates with S_L, so the ratio $\nu_{m,N}/\nu_W$ and hence the relative premium (4.13) diverge, from about 6% at a one-month accrual to roughly 134% at two years. The backward-versus-forward distinction is therefore small for very short accruals and dominant for long ones. The gap also grows, both absolutely and, more sharply, in relative terms, as the mean-reversion parameter a rises: although a compresses $\nu_{m,N}$, it compresses ν_W faster (Section 4.2.3), so the ratio $\nu_{m,N}/\nu_W$ governing (4.13) increases. Away from the money, finally, the gap is governed by the vega weight $\varphi(d_1)$ in (4.11), which is largest near the money and decays for deep in- and out-of-the-money strikes, so the gap is a single-peaked, non-monotone function of the strike, maximized around $K^\star$. The leading-order formula (4.12) tracks the exact gap throughout.

4.4 Numerical Tests of the Gap

This section displays numerically the dependences derived in Section 4.3, reusing the same closed-form engine and again moving one input at a time. All gaps are reported in the normalized units of Section 4.1.2 (per unit $p(0, m)$, in basis points), the valuation date is $n = 0$, and the forward- and backward-looking caplets are evaluated at the pre-accrual and total variances ν_W and $\nu_X = \nu_W + \nu_{m,N}$ respectively, exactly as in (4.10). Where a closed form is available we overlay it on the exact gap, so that formula and computation can be read off the same axes.

4.4.1 Sensitivity to Volatility

Figure 4.6 sweeps the annualized volatility with all maturities held at the base case. The left panel plots the exact at-the-money gap together with the leading-order formula (4.12). The two are visually indistinguishable, the relative discrepancy being of order 10^{-7}, which is the promised validation of (4.12). The gap is exactly linear in σ: both ν_W and ν_X scale as σ^2 (equation (4.7)), so $\sqrt{\nu_X} - \sqrt{\nu_W}$ is proportional to σ and so is the gap. The right panel shows the complementary fact predicted by (4.13): the gap measured relative to the forward-looking caplet is a flat line, independent of σ, pinned at 18.8% for the base maturities. Volatility sets the absolute size of the premium but not its proportion.

4.4.2 Sensitivity to Accrual Length

Figure 4.7 varies the accrual length L at fixed pre-accrual window $m = 90$. The absolute gap (left) grows without bound and convexly, while the relative gap (right) rises from about 6% at a one-month accrual through 18.8% at the base three-month case, crossing 100% near a seventeen-month accrual and reaching roughly 134% at two years. The mechanism is the one isolated in Section 4.2.1: the in-accrual variance $\nu_{m,N}$ grows without saturating, whereas the pre-accrual variance ν_W saturates with S_L as a function of L, so the ratio $\nu_{m,N}/\nu_W$ governing (4.13) diverges. The backward-looking premium is therefore

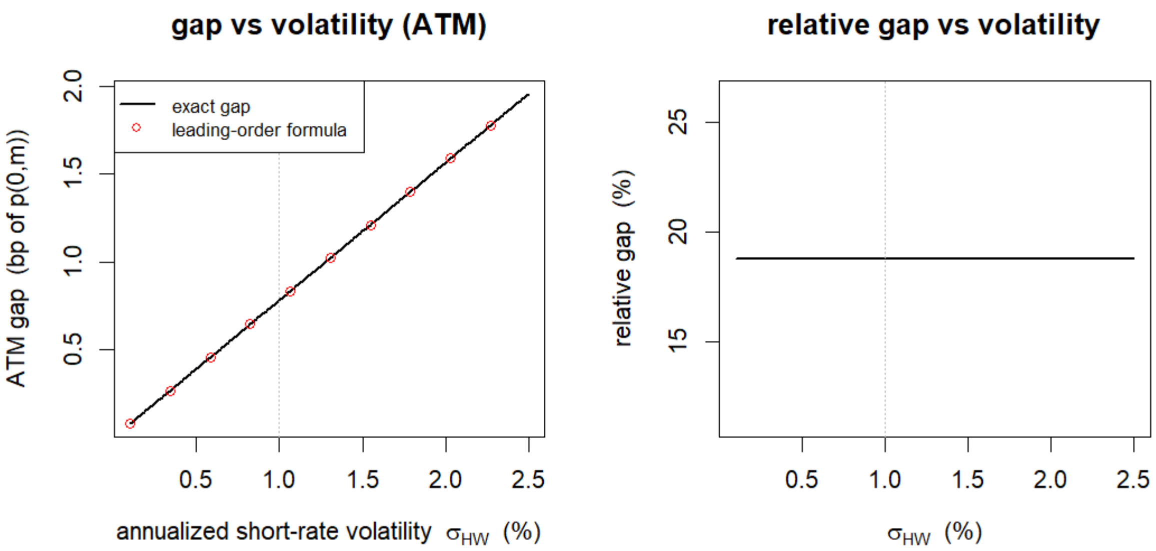

Figure 4.6: Backward-forward gap versus volatility (ATM, base maturities). *Left:* the exact gap (line) and the leading-order formula (4.12) (circles) coincide, and the gap is linear in σ_{HW}. *Right:* the relative gap (4.13) is independent of σ.

a second-order effect for the short accruals typical of standard SOFR caplets, but becomes first-order, and eventually dominant, for long-dated contracts.

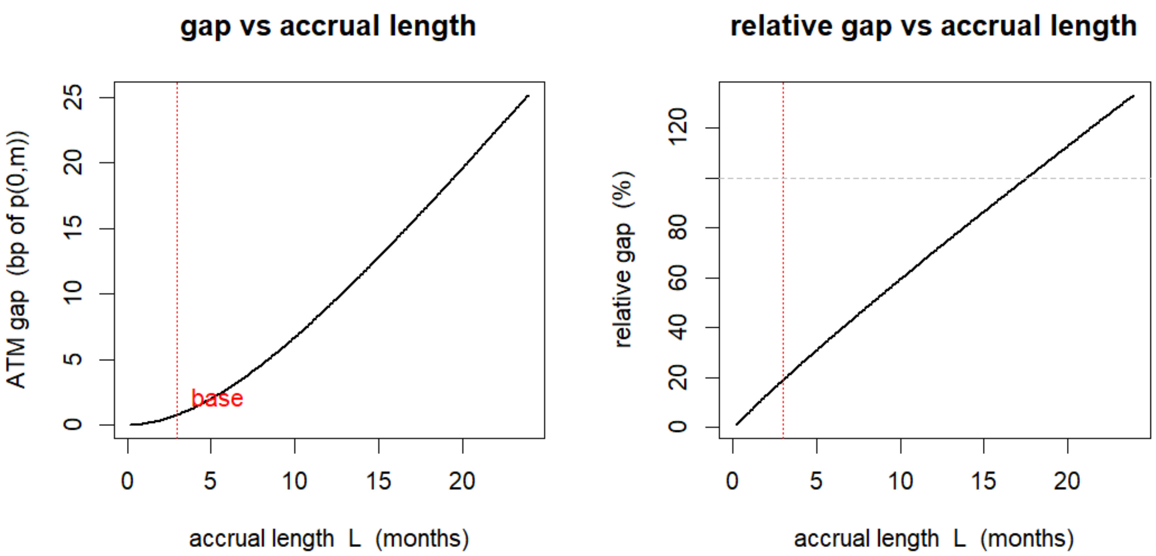

Figure 4.7: Backward-forward gap versus accrual length L (ATM, fixed $m = 90$). *Left:* absolute gap. *Right:* relative gap, which crosses 100% (dashed) near a seventeen-month accrual. The growth reflects the non-saturation of $\nu_{m,N}$ against the saturation of ν_W.

4.4.3 Sensitivity to Mean Reversion

The dependence on the mean reversion is more subtle than a one-line argument suggests, and the two panels of Figure 4.8 are best read together. A larger a compresses *both* variance components, so one might expect the premium to shrink. The absolute gap, however, rises mildly with a (from 0.76 to 0.82 bp across $a \in [0.05, 3]$), and the relative gap rises markedly (from 15% to 33%). The reason is the asymmetry established in Section 4.2.3: mean reversion compresses the pre-accrual variance ν_W faster than the in-accrual variance $\nu_{m,N}$, because ν_W carries both the factor S_L^2 and the geometric pre-accrual sum while $\nu_{m,N}$ contracts only through its S_ℓ^2. Consequently the ratio $\nu_{m,N}/\nu_W$, which alone determines the relative gap (4.13), *increases* with a, and the absolute gap $\propto \sqrt{\nu_W}\big(\sqrt{1 + \nu_{m,N}/\nu_W} - 1\big)$ inherits a mild net increase as the rising ratio outweighs the falling $\sqrt{\nu_W}$. Stronger mean reversion thus makes the backward-looking feature *relatively* more important, not less.

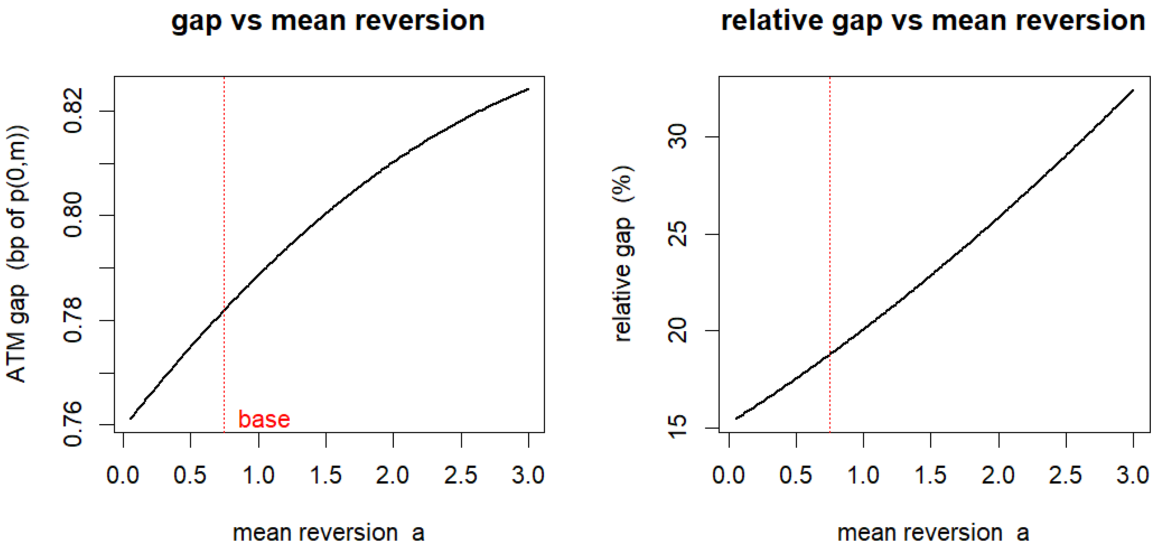

Figure 4.8: Backward-forward gap versus mean reversion a (ATM, base maturities). *Left:* the absolute gap rises mildly. *Right:* the relative gap rises clearly, since a compresses ν_W faster than $\nu_{m,N}$ and so raises the ratio $\nu_{m,N}/\nu_W$.

4.4.4 Sensitivity to the Strike

Figure 4.9 sweeps the strike around the ATM level $K^\star$ at fixed base-case variances. The gap is a single-peaked, bell-shaped function of the strike, maximized at $K^\star$ and decaying to zero for deep in- and out-of-the-money strikes. This is the vega weighting $\varphi(d_1)$ of the integral representation (4.11): the extra variance $\nu_{m,N}$ is most valuable where the option's variance vega is largest, that is, at the money, and is worth almost nothing once the strike is far from the forward and the payoff is locally linear. The premium is therefore an at-the-money phenomenon, concentrated precisely where caplets are most actively traded.

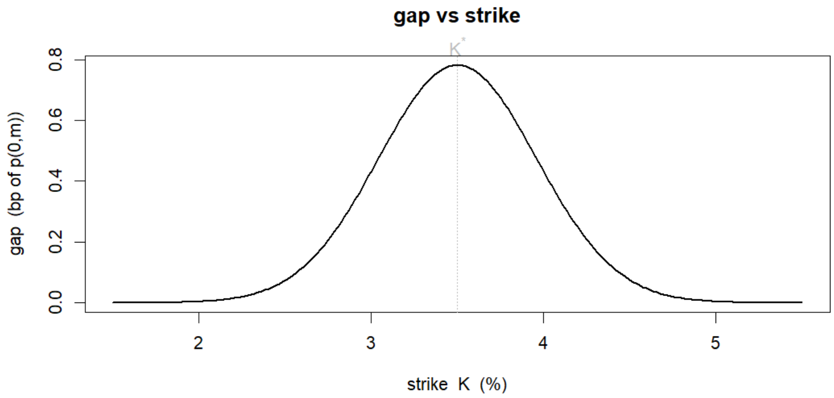

Figure 4.9: Backward-forward gap versus strike (base-case variances). The gap is single-peaked at the ATM strike $K^\star$ (dotted) and decays for deep in- and out-of-the-money strikes, tracing the variance-vega weight of (4.11).

Taken together, the four experiments display exactly the behavior described analytically in Section 4.3. The premium scales linearly with volatility but is proportionally volatility-invariant, it grows without bound with the accrual length, it is concentrated at the money and, contrary to the naive expectation that compressing $\nu_{m,N}$ shrinks it, it becomes proportionally larger as the mean reversion increases, because the pre-accrual variance is compressed faster still. In every case the single governing quantity is the ratio $\nu_{m,N}/\nu_W$ of in-accrual to pre-accrual variance: the backward-looking premium is large exactly when the in-accrual uncertainty is large relative to the uncertainty already resolved before the accrual begins.

Conclusion

This thesis has priced backward-looking caplets and floorlets in a discrete Hull-White model in which the short rate itself, and not only the compounding convention, is discrete. The short rate evolves by a first-order autoregressive recursion, the Euler-Maruyama discretization of the Hull-White equation, and the entire pricing theory is carried out inside the discrete affine term structure of this model, using only elementary properties of the Gaussian distribution. The continuous-time material of Chapters 2 and earlier serves only to motivate the discrete dynamics and to justify, through the additive-noise, why the Hull-White specification is the one that survives discretization with its distributional properties intact.

The contribution is twofold. The first is methodological. Under the single-period convention $p(k, k+1) = e^{-r_k}$ the compounded backward-looking rate telescopes exactly into the exponential of the sum of the short rates over the accrual window. This reduces the caplet payoff to a function of a single Gaussian sum, and the closed-form prices, at the start of the accrual period (Theorem 3.1) and at any earlier valuation date (Theorem 3.3), follow from truncated univariate and bivariate Gaussian moments obtained by completing the square. To the author's knowledge this is the first fully self-contained, discrete-time derivation of backward-looking caplet and floorlet prices that uses no stochastic calculus, and it occupies a gap in the pedagogical literature alongside the continuous-time works.

The second contribution is the variance decomposition $\nu_X = \nu_W + \nu_{m,N}$ (Proposition 3.2). Because the backward- and forward-looking caplets are produced by a single argument, they are seen to share an identical Black-type formula and to differ *only* through the variance fed into it: the forward-looking price uses the pre-accrual variance ν_W, the backward-looking price the total variance ν_X, and the two differ by exactly the in-accrual variance $\nu_{m,N}$. The entire distinction between the two conventions is thereby reduced to a single, explicitly characterized term. This yields an exact integral representation of the price gap (Proposition 4.1), a closed form for it at the money, and a complete numerical study of its behavior in the model parameters (Chapter 4), none of which requires simulation. The numerical analysis confirms that the ratio $\nu_{m,N}/\nu_W$ governs every comparative-static property of the gap: the backward-looking premium is small for the shortest accruals, is already material at standard three-month maturities, grows without bound with the accrual length, is concentrated at the money, and, against naive intuition, becomes proportionally larger as mean reversion increases, because ν_W is compressed faster than $\nu_{m,N}$.

The framework has clear limitations, all inherited from its modeling choices. It is Gaussian, hence generates no implied-volatility smile and is calibrated only at the money, with a short rate that may turn negative. It is one-factor, and the numerical study runs on a stylized curve rather than on quoted prices, so the formulas are never confronted with market data. The exact telescoping is moreover tied to the exponential single-period convention: under the simple-rate specification of Section 3.7 the exercise region ceases to be linear in the Gaussian shocks and the closed form is lost, leaving only numerical evaluation.

These limitations can be explored in further work. The most immediate direction is empirical: calibration to quoted SOFR caps and floors, with the predicted backward-looking premium tested against the market. On the modeling side, the additive-noise criterion that singles out Hull-White among the classical short-rate models also suggests which extensions could remain tractable in discrete time. In particular, introducing non-Gaussian innovations, jumps, or stochastic volatility would move the terminal distribution away from the Gaussian setting considered here and would therefore provide a natural route to non-flat Black-implied volatility smiles.

Bibliography

[1] F. Black, *The pricing of commodity contracts*, Journal of Financial Economics, 3:167–179, 1976.

[2] F. Black and P. Karasinski, *Bond and option pricing when short rates are lognormal*, Financial Analysts Journal, 47(4):52–59, 1991.

[3] D. Brigo and F. Mercurio, *Interest Rate Models – Theory and Practice: With Smile, Inflation and Credit*, 2nd edition, Springer Finance, Springer-Verlag, Berlin, 2006.

[4] J. C. Cox, J. E. Ingersoll, and S. A. Ross, *A theory of the term structure of interest rates*, Econometrica, 53(2):385–407, 1985.

[5] R. C. Dalang, A. Morton, and W. Willinger, *Equivalent martingale measures and no-arbitrage in stochastic securities market models*, Stochastics and Stochastic Reports, 29:185–201, 1990.

[6] L. U. Dothan, *On the term structure of interest rates*, Journal of Financial Economics, 6(1):59–69, 1978.

[7] D. Filipović and J. Zabczyk, *Markovian term structure models in discrete time*, Annals of Applied Probability, 12(2):710–729, 2002.

[8] C. Fontana, *Caplet pricing in affine models for alternative risk-free rates*, SIAM Journal on Financial Mathematics, 14(1):SC1–SC16, 2023.

[9] M. Henrard, *Overnight indexed swaps and floored compounded instruments in HJM one-factor models*, Working paper, 2004.

[10] K. Hofmann, *Implied volatilities for options on backward-looking term rates*, Working paper, 2020.

[11] J. Hull and A. White, *Pricing interest-rate-derivative securities*, The Review of Financial Studies, 3(4):573–592, 1990.

[12] D. Iunevich and D. Kalikaeva, *Caplet pricing for discrete backward-looking RFR term rates under the Hull-White model*, SSRN, 2025.

[13] J. Jacod and P. Protter, *Probability Essentials*, 2nd edition, Universitext, Springer-Verlag, Berlin, 2004.

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

[15] F. Mercurio, *A simple multi-curve model for pricing SOFR futures and other derivatives*, Working paper, SSRN 3225872, 2018.

[16] A. Pascucci and W. J. Runggaldier, *Financial Mathematics: Theory and Problems for Multi-period Models*, Unitext, Springer-Verlag Italia, Milan, 2012.

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

[18] K. Sandmann and D. Sondermann, *A note on the stability of lognormal interest rate models and the pricing of Eurodollar futures*, Mathematical Finance, 7:119–125, 1997.

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

[20] C. Turfus, *Caplet pricing with backward-looking rates*, Working paper, 2020.

[21] O. Vasicek, *An equilibrium characterization of the term structure*, Journal of Financial Economics, 5(2):177–188, 1977.

[22] M. Xu, *SOFR derivative pricing using a short-rate model*, Working paper, 2022.