Interest Model Review
Under Single-Curve Framework

Jianing Yao
Department of MSIS-RUTCOR
Rutgers University, the State University of New Jersey
Piscataway, NJ 08854 USA

August 19, 2016

In this notes, we will review interest rate models that are widely used in practice, e.g., short rate model, forward rate model and market model. The standard setting for the discussion is – a filtered probability space $(\Omega, \mathcal{F}, \mathbb{P})$ *with* $\{\mathcal{F}_t\}_{t\geq 0}$ *generated by* d*-dimensional Brownian motion* $\{W_t\}_{t\geq 0}$.

1 Short Rate Model

The most natural approach to address the uncertainty in interest rate market is to directly model the short rate, namely, we assume the following SDE for short rate r_t:

$$dr_t = \alpha(t, r_t)dt + \sigma(t, r_t)dW_t^{\mathbb{P}}, \tag{1}$$

where α, σ are Borel measurable functions that are sufficiently regular. In this section, we give a comprehensive introduction of short rate model, including modeling philosophy and specific models.

1.1 General Framework of Short Rate Model

We denote by $P(t, T)$ a zero coupon bond maturing at time T and make the following assumption:

Assumption 1.1. *Zero coupon $P(t,T)$ is a smooth functional of t and r_t.*

The assumption above are two folded: it requires *Markovian structure* of the stochastic process driving short rate r_t; the smootheness of bond price with respect to time and current short rate. The first former part is true for almost all short rate model, as long as it's not path-dependent(equivalent, given by (1)), the second seems to be rather stronger, however, it turns out that those popular models all satisfy this condition. Apply Itô's formula:

$$\begin{aligned} dP(t,T) &= \left[\partial_t P(t,T) + \alpha(t,r_t)\partial_r P(t,T) + \frac{1}{2}\sigma(t,r_t)^2\partial_{rr}P(t,T)\right]dt \\ &\quad + \sigma(t,r_t)\partial_r P(t,T)dW_t^{\mathbb{P}} \\ &:= P(t,T)\big[m(t,r_t;T)dt + S(t,r_t;T)dW_t^{\mathbb{P}}\big] \end{aligned} \tag{2}$$

In equity model, we used to construct a portfolio to replicate contingent claim. Let's mimic the replication procedure here. For $t < T_1 < T_2$, we construct a self-financing portfolio involving the money bank account $B(t)$ and $P(t,T_2)$ to replicate $P(t,T_1)$. Specifically,

- *Portfolio A:* hold a_t units of $P(t,T_2)$ and b_t units of $B(t)$ at time t;
- *Portfolio B:* hold 1 unit of $P(t,T_1)$.

To determine a_t and b_t such that portfolio A is self-financing and the two portfolios have equal value, i.e., $V^A(t) = V^B(t)$, where $V^A(t) := a_t P(t,T^2) + b_t B(t)$ and $V^B(t) := P(t,T_2)$, we express, at an infinitesimal level,

$$dP(t,T^1) = a_t dP(t,T_2) + b_t dB(t). \tag{3}$$

Substituting (2) into (3) yields,

$$\begin{aligned} &a_t P(t,T_2)\big[m(t,r_t;T_2)dt + S(t,r_t;T_2)dW_t^{\mathbb{P}}\big] + b_t r_t B(t)dt \\ =&P(t,T_1)\big[m(t,r_t;T_1)dt + S(t,r_t;T_1)dW_t^{\mathbb{P}}\big]. \end{aligned}$$

which implies

$$a_t = \frac{S(t,r_t;T_1)P(t,T_1)}{S(t,r_t;T_2)P(t,T_2)}, \tag{4}$$

$$b_t = \frac{1}{r_t B(t)}\left[m(t,r_t;T_1)P(t,T_1) - \frac{m(t,r_t;T_2)P(t,T_1)S(t,r_t;T_1)}{S(t,r_t;T_2)}\right] \tag{5}$$

Since $V^A(t) = V^B(t)$, we then have

$$\begin{aligned} &\frac{S(t,r_t;T_1)P(t,T_1)}{S(t,r_t;T_2)P(t,T_2)}P(t,T_2) \\ +&\frac{1}{r_t B(t)}\left[m(t,r_t;T_1)P(t,T_1) - \frac{m(t,r_t;T_2)P(t,T_1)S(t,r_t;T_1)}{S(t,r_t;T_2)}\right]B(t) = P(t,T_1) \end{aligned}$$

After simplification, we end up with λ – *market-price-of-risk*:

$$\frac{m(t,r_t;T_1)-r_t}{S(t,r_t;T_1)}=\frac{m(t,r_t;T_2)-r_t}{S(t,r_t;T_2)}:=\lambda(t,r_t) \tag{6}$$

Observe $\lambda(t,r_t)$ is independent of the terminal time T.

Remark 1.2. *The above replication argument is valid for any non-dividend paying security. Therefore, for any derivative security with time t price X_t and follows:*

$$dX_t = X_t\big(\mu(t,r_t)dt+\sigma(t,r_t)dW_t^{\mathbb{P}}\big)$$

It must be true that

$$\lambda(t,r_t)=\frac{\mu(t,r_t)dt-r_t}{\sigma(t,r_t)}.$$

Using (6) and (2), we derive that $P(t,T)$ satisfies the following PDE:

$$\big[\alpha-\lambda\sigma\big](t,r_t)\partial_r P(t,T)+\partial_t P(t,T)+\frac{1}{2}\sigma^2(t,r_t)\partial_{tt}P(t,T)-r_tP(t,T)=0. \tag{7}$$

The above PDE[1] can be solved if we supply the boundary condition $P(T,T)=1$. To make clear, the replication argument leads to the equivalence between pricing and solving PDE below:

$$\begin{cases} \big[(\alpha-\lambda\sigma)\partial_x V\big](t,x)+\partial_t V(t,x)+\big[\frac{1}{2}\sigma^2\partial_{xx}V\big](t,x) \\ -xV(t,x)=0,\ (t,x)\in[0,T]\times\mathbb{R}^+, \\ V(T,x)=1, \quad x\in\mathbb{R}^+. \end{cases} \tag{8}$$

Here, $V(t,r_t):=P(t,T)$ is the value function corresponding time t value of the zero-coupon bond.

Remark 1.3. *Without change of measure (to risk-neutral framework), we use a purely replication argument to derive PDE whose value function corresponds to the value of the contract. This is a routine for no-arbitrage pricing.*

Alternatively, we can use risk-neutral pricing theory. Denote $\mathbb{Q}$ as an equivalent martingale measure, the dynamics of short rate is specified below:

$$dr_t=\alpha^*(t,r_t)dt+\sigma^*(t,r_t)dW_t^{\mathbb{Q}}. \tag{9}$$

[1](8) is called the *fundamental PDE for 1 factor model*.

Discounting the terminal value of zero-coupon bond back to time t under risk-neutral measure yiedls the present value, i.e., price of the bond $P(t, T)$,

$$P(t, T) = B(t)\mathbb{E}_t^{\mathbb{Q}}\left[\frac{P(T, T)}{B(T)}\right] \tag{10}$$

Since $P(t, T)/B(t)$ is a martingale, it leads to *Feyman Kac PDE* below:

$$\begin{cases} \left[\alpha^*\partial_x V\right](t, x) + \partial_t V(t, x) + \frac{1}{2}(\sigma^*)^2\partial_{xx}V(t, x) \\ -xV(t, x) = 0, \ (t, x) \in [0, T] \times \mathbb{R}^+, \\ V(T, x) = 1, \quad x \in \mathbb{R}^+. \end{cases} \tag{11}$$

Compare (11) and (8), it must be true that:

$$\sigma^*(t, x) = \sigma(t, x), \quad \lambda(t, x) = \frac{\alpha(t, x) - \alpha^*(t, x)}{\sigma(t, x)} \tag{12}$$

As a result, we find

$$dW_t^{\mathbb{P}} = dW_t^{\mathbb{Q}} - \lambda(t, r_t)dt, \tag{13}$$

Measure $\mathbb{Q}$ makes $dW_t^{\mathbb{P}} + \lambda(t, r_t)dt$ a Brownian motion suggesting it is defined by Radon Nikodym derivative in the following way:

$$\frac{d\mathbb{Q}}{d\mathbb{P}}(\omega) = \exp\left\{-\int_0^t \lambda(s, r_s)dW_s^{\mathbb{P}} - \frac{1}{2}\int_0^t \lambda(s, r_s)^2 ds\right\}. \tag{14}$$

Let's re-organize a little, we start with (1), short rate dynamics in real world. Given the market-price-of-risk $\lambda(t, r_t)$, we can define the risk neutral measure $\mathbb{Q}$ via (14), under which the dynamics of short rate becomes (9). By definition of $\mathbb{Q}$, we have a martingale (10) for pricing, from where Feyman Kac PDE (11) follows. The essential role here is λ, it connects the real world with risk-neutral world. The consistency between (1) and (9), (8) and (11) all follows from this fact.

1.2 Vasicek Model

For historical reason, we will start by the *Vasicek model(1977)*. It appears to be the first continuous-time model for short rates. It serves as a good example to understand the general framework in above section.

Remark 1.4. *For other short rates later on, we will not repeat all steps but just providing relevant results and specialties of the model. Reader can frequently refer to Vasicek model for clarification.*

1.2.1 Model Dynamics

Vasicek model assumes the instantaneous spot rate under the real world measure evolves as an *Orstein-Uhlenbeck process* with constant coefficients. That is, under physical measure $\mathbb{P}$,

$$dr_t = (\kappa\theta - (\kappa + \lambda\sigma)r_t)dt + \sigma dW_t^{\mathbb{P}}, \quad r_0 = r. \tag{15}$$

where λ is a parameter contributing to the market price of risk. If we assume that:

$$\lambda(t, r_t) = \lambda r_t, \tag{16}$$

then, by (12), the short rate dynamic under $\mathbb{Q}$ is:

$$dr_t = \kappa(\theta - r_t)dt + \sigma dW_t^{\mathbb{Q}}, \quad r_0 = r. \tag{17}$$

Remark 1.5. *In general, there is no reason why $\lambda(t, r_t)$ should be this form. However, under this choice we obtain a short rate process that is tractable under both measures.*

The real world data actually characterized the dynamics of *Vasicek model*. Thus, one way to calibrate the model is to use historical time-series. Moreover, if we want to measure the exposure, we would also like to simulate the dynamic under measure $\mathbb{P}$. For pricing, we focus on the risk-neutral dynamics (17). Notice simple integration yields, for $t \geq s$,

$$r_t = r_s e^{-\kappa(t-s)} + \theta\big(1 - e^{-\kappa(t-s)}\big) + \sigma \int_s^t e^{-\kappa(t-u)} dW_u. \tag{18}$$

This implies the following basic statistics of the stochastic process:

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}}\big[r_t \,\big|\, \mathcal{F}_s\big] &= r_s e^{-\kappa(t-s)} + \theta\big(1 - e^{-\kappa(t-s)}\big), \\ Var\big[r_t \,\big|\, \mathcal{F}_s\big] &= \frac{\sigma^2}{2\kappa}\big[1 - e^{-2k(t-s)}\big], \end{aligned} \tag{19}$$

which reveals the properties of Vasicek model:

(i) the model is a Gaussian model permits negative rates;

(ii) θ is interpreted the long run mean, as $t \to \infty$, $r_t \to \theta$;

(iii) $\kappa > 0$ controls the speed of mean reversion, and large κ implies strong intensity for mean reverting;

(iv) the process is mean reverting, if the interest rate is bigger than the long run mean θ, then the mean reversion term κ will pull down in the direction of θ. The same reasoning hold for when $r_t \leq \theta$.

Remark 1.6. *As a normal model Vasicek accepts negative rates, which is an advantage over log-normal model given prevailing negative rates in Europe countries and Japan. However, the normal distribution is symmetric, therefore, it allows the rates delving into very negative with relatively high probability. In reality, we'd rather want a model that have some stickiness around* 0 *so that it won't go very negative*[2].

The most attractive aspect of *Vasicek model* is its tractability when pricing a pure-discount bond. If we make an initial guess,

$$P(t,T) = A(t,T)e^{-B(t,T)r_t} := V(t,r_t). \tag{20}$$

It should satisfy (8)*(fundamental PDE for 1 factor model)*, namely,

$$\partial_t V(t,x) + \kappa\big(\theta - x\big)\partial_x V(t,x) + \frac{\sigma^2}{2}\partial_{xx}V(t,x) - rV(t,x) = 0. \tag{21}$$

By compare coefficients, it results in the following ODE system:

$$\begin{cases} A'(t,T) - \kappa\theta A(t,T)B(t,T) + \frac{\sigma^2}{2}A(t,T)B^2(t,T) = 0 \\ B'(t,T) - \kappa B(t,T) + 1 = 0 \\ A(T,T) = 1,\ \ B(T,T) = 0. \end{cases} \tag{22}$$

We can solve for,

$$A(t,T) = \exp\left\{\left(\theta - \frac{\sigma^2}{2\kappa^2}\right)\big[B(t,T) - T + t\big] - \frac{\sigma^2}{4\kappa}B^2(t,T)\right\} \tag{23}$$

$$B(t,T) = \frac{1}{\kappa}\left[1 - e^{-\kappa(T-t)}\right]. \tag{24}$$

Let's apply Itô's lemma to (20) to obtain(it's a messy calculation),

$$dP(t,T) = r_t P(t,T)dt - \sigma B(t,T)P(t,T)dW_t^{\mathbb{Q}}. \tag{25}$$

The explicit bond price also implies the instantaneous forward rate dynamics under Vasicek framework,

$$\begin{aligned} f(t,T) &= -\frac{\partial \ln P(t,T)}{\partial T} = -\partial_T \ln A(t,T) + \partial_T B(t,T)r_t \\ &= \big(\theta\kappa - \frac{\sigma^2}{2}B(t,T)\big)B(t,T) + r_t e^{-\kappa(T-t)} \end{aligned}$$

[2] The *free boundary SABR* owns such merits.

Again, by Itô's lemma, one can get instantaneous forward rate dynamics under risk-neutral measure $\mathbb{Q}$,

$$df(t,T) = \big(\cdots\big)dt + \sigma e^{-\kappa(T-t)} dW_t^{\mathbb{Q}} \tag{26}$$

Since the drift term is lengthy and not that important, we omitted here.

In interest rate modeling, very often, we would like to work undder T-forward measure. We now investigate the dynamic of Vasicek model under T-forward measure. Change of numéraire only modify the drift term, by the toolkit developed,

$$\begin{aligned}
&\textit{new drift}\\
=&\textit{old drfit} - \textit{old diffusion}\Big(\frac{\textit{old numéraire diffusion}}{\textit{old numéraire}} - \frac{\textit{new numéraire diffusion}}{\textit{new numéraire}}\Big)\\
=&\kappa(\theta - r_t) - \sigma^2 B(t,T)\\
=&\big(\kappa\theta - \sigma^2 B(t,T) - \kappa r_t\big)
\end{aligned}$$

Thus, the corresponding dynamics under T-forward measure is:

$$dr_t = \big(\kappa\theta - \sigma^2 B(t,T) - \kappa r_t\big)dt + \sigma dW_t^T, \tag{27}$$

where W^T is a $\mathbb{Q}^T$-Brownian motion. The mean and variance of r_t give r_s changes now to,

$$\mathbb{E}^T\big[r_t \,\big|\, \mathcal{F}_s\big] = r_s e^{-\kappa(t-s)} + M(s,t;T), \tag{28}$$

$$Var^T\big[r_t \,\big|\, \mathcal{F}_s\big] = \frac{\sigma^2}{2\kappa}\big[1 - e^{-2k(t-s)}\big], \tag{29}$$

where,

$$M(s,t;T) = \big(\theta - \frac{\sigma^2}{\kappa^2}\big)\big(1 - e^{-\kappa(t-s)}\big) + \frac{\sigma^2}{2\kappa^2}\big[e^{-\kappa(T-t)} - e^{-\kappa(T+t-2s)}\big].$$

As for the dynamics of $f(t,T)$ under T-forward measure, since forward rate is a martingale under $\mathbb{Q}^T$-forward measure and $f(t,T) = \lim_{T^+\to T} F(t;T,T^+)$, by invoking dominant convergence theorem, $f(t,T)$ can be proved to be a martingale as well under $\mathbb{Q}^T$-forward measure, thus drift-less, i.e.,

$$df(t,T) = \sigma e^{\kappa(T-t)} dW_t^T. \tag{30}$$

Despite the analytical tractability, the simple structure limits the possible shapes of the yield curve, which often leads into situation, where the theoretical yield curve does not correspond to the market yield curve. To improve this situation, one can make the parameter time dependent, i.e.,

$$dr_t = \kappa(\theta_t - r_t)dt + \sigma dW_t^{\mathbb{Q}} \tag{31}$$

The exogenous term gives more degree of freedoms to fit the yield curve, we will discuss more in Hull-White one factor model.

1.2.2 Derivatives with Zero-Coupon Bond as Underlying

We now use Vasicek model to price some basic interest rate derivatives. Among them, the simplest interest rate derivative is the *bond option*. Similar to equity option, bond option allows the owner to buy or sell a S-zero-coupon-bond(matures at time S) at option expiry, denoted by T, for a determined price, i.e., strike price K. For call option, we have the payoff $H\big(P(T,S),K\big) = \big(P(T,S)-K\big)^+$. Thus, the time t value of the bond is,

$$V(t,P(t,T)) = \mathbb{E}_t^{\mathbb{Q}}\big[D(t,T)\big(P(T,S)-K\big)^+\big] = P(t,T)\mathbb{E}_t^T\big[\big(\frac{P(T,S)}{P(T,T)}-K\big)^+\big]$$

Notice, the last identity follows because of $P(T,T)=1$. The benefit of using a quotient process is to motivating change of measure that will be paid out handsomely later on. Define $U_t := P(t,S)/P(t,T)$, which evolves under risk neutral measure $\mathbb{Q}$,

$$dU_t = U_t\sigma B(t,T)\big[\sigma B(t,T)-\sigma B(t,S)\big]dt + U_t\big[\sigma B(t,T)-\sigma B(t,S)\big]dW_t^{\mathbb{Q}} \quad (32)$$

By definition, U_t has to be a martingale under $\mathbb{Q}^T$-forward measure, one can validate by change of numéraire, thus,

$$dU_t = U_t\sigma\big[B(t,T)-B(t,S)\big]dW_t^{\mathbb{Q}^T} := U_t\Theta_t dW_t^T. \quad (33)$$

The log-normality gives an explicit solution for U_T given U_t,

$$U_T = U_t\exp\left\{-\int_t^T \Theta_s dW_s^T - \frac{1}{2}\int_t^T \Theta_s^2 ds\right\} \quad (34)$$

that is, $\ln(U_T/U_t)$ is normal distributed with mean and variance below:

$$\mathbb{E}_t^T\left[\ln\left(\frac{U_T}{U_t}\right)\right] = -\frac{1}{2}\int_t^T \Theta_s^2 ds = -\frac{\sigma^2}{2\kappa^3}\big(1-e^{-\kappa(S-t)}\big)^2\big(1-e^{-2\kappa(T-t)}\big), \quad (35)$$

$$Var_t^T\left[\ln\left(\frac{U_T}{U_t}\right)\right] = \int_t^T \Theta_s^2 ds = \frac{\sigma^2}{\kappa^3}\big(1-e^{-\kappa(S-t)}\big)^2\big(1-e^{-2\kappa(T-t)}\big). \quad (36)$$

Coming back to the pricing problem,

$$V(t,P(t,T)) = P(t,T)\frac{P(t,S)}{P(t,T)}\mathbb{E}_t^T\left[\left(e^X - \frac{K}{U_t}\right)^+\right]$$

By Black-Scholes formula, the zero-coupon-bond call price (ZBC) at time t is,

$$ZBC(t,P(t,T);T,S,K) = V(t,P(t,T)) = P(t,S)N(d_+) - KP(t,T)N(d_-) \quad (37)$$

where,

$$d_{\pm} = \frac{1}{\sqrt{\int_t^T \Theta_s^2 ds}} \ln\Big(\frac{P(t,S)}{KP(t,T)}\Big) \pm \frac{1}{2}\int_t^T \Theta_s^2 ds.$$

Similarly, the zero-coupon-bond-put price (ZBP) is

$$ZBP(t, P(t,T); T, S, K) = KP(t,T)N(-d_-) - P(t,S)N(-d_+). \tag{38}$$

1.2.3 Pricing Caps and Swaptions

In the previous notes, we have seen the BS formula pricing cap/floor, where we assumed forward rate and forward swap rate processes have log-normal dynamics. That's how market quotes those instruments, they can be treated as converters between price and implied volatility. In this section, we will derive pricing formulas based on short rate model, which has richer structure than drift-less log-normal model.

Caps/floors can be viewed as portfolios of zero-bond options. Let's denote by, $\mathcal{T} = \{T_0, T_1, ..., T_n\}$, the payment schedule of cap/floor. For a cap with strike K and maturity T, the time t value of its caplet living in $[T_{i-1}, T_i]$ is, for $t \leq T_{i-1}$,

$$\begin{aligned}
Cpl(t, \mathcal{T}, N = 1, K) &= \mathbb{E}_t^{\mathbb{Q}}\Big[D(t, T_{i-1})\mathbb{E}_{T_{i-1}}^{\mathbb{Q}}\Big[D(T_{i-1}, T_i)\Big]\tau_i\big(L(T_{i-1}, T_i) - K\big)^+\Big] \\
&= \mathbb{E}_t^{\mathbb{Q}}\Big[D(t, T_{i-1})P(T_{i-1}, T_i)\Big(\frac{1}{P(T_{i-1}, T_i)} - 1 - K\tau_i\Big)^+\Big] \\
&= \mathbb{E}_t^{\mathbb{Q}}\Big[D(t, T_{i-1})\Big(1 - (1 + K\tau_i)P(T_{i-1}, T_i)\Big)^+\Big] \\
&= (1 + K\tau_i)\mathbb{E}_t^{\mathbb{Q}}\Big[D(t, T_{i-1})\Big(\frac{1}{1 + K\tau_i} - P(T_{i-1}, T_i)\Big)^+\Big] \\
&= (1 + K\tau_i)ZBP\Big(t, P(t, T_{i-1}); T_{i-1}, T_i, \frac{1}{1 + K\tau_i}\Big)
\end{aligned}$$

Therefore the cap price is:

$$Cap(t; \mathcal{T}, N = 1, K) = \sum_{i=1}^{n}(1 + K\tau_i)ZBP\Big(t, P(t, T_{i-1}); T_{i-1}, T_i, \frac{1}{1 + K\tau_i}\Big), \tag{39}$$

correspondingly,

$$Flr(t; \mathcal{T}, N = 1, K) = \sum_{i=1}^{n}(1 + K\tau_i)ZBC\Big(t, P(t, T_{i-1}); T_{i-1}, T_i, \frac{1}{1 + K\tau_i}\Big). \tag{40}$$

1.2.4 Pricing Coupon-Bearing Bond and Swaptions

To use short rate model pricing Swaptions, we have to first take a detour to pricing European options on a coupon-bearing bond. Such bond pays the cash flows $\mathcal{C} = \{c_1, c_2, ..., c_n\}$ at time instances $T_1, T_2, ..., T_n$. The price of coupon-bearing bond in T_0 is

$$CB(T_0; \mathcal{T}, \mathcal{C}) = \sum_{i=1}^{n} c_i P(T_0, T_i) := \sum_{i=1}^{n} c_i V(T_0, r_{T_0}; T_i) \tag{41}$$

where $V(T_0, r_{T_0}; T_i)$ is a function of T_0, T_i and short rate at time T_0. A time t, vanilla put option on the coupon-bearing bond with strike K and maturity T_0 has payoff,

$$\left[K - CB(T_0; \mathcal{T}, \mathcal{C})\right]^+ = \left[K - \sum_{i=1}^{n} c_i V(T_0, r_{T_0}; T_i)\right]^+ \tag{42}$$

As in the cap/floor case, we want to take advantage of zero-coupon-bond pricing formula. However, the payoff function is nonlinear that does not allow linear decomposition. To have a workaround, we have to introduce the brilliant idea of *Jamishidian*.

Remark 1.7. (Jamishidian trick) *Consider a monotone decreasing positive functions $f_i : \mathbb{R} \mapsto \mathbb{R}^+$, W is a random variable and $K \geq 0$ is a constant. The finite summation of f_i is also positive and monotone decreasing in the argument, thus, there exists a unique ω such that*

$$\sum_{i=1}^{n} f_i(\omega) = K. \tag{43}$$

The following chain of equalities is essential

$$\begin{aligned} &\left(K - \sum_{i=1}^{n} f_i(W)\right)^+ = \left(\sum_{i=1}^{n} \big(f_i(\omega) - f_i(W)\big)\right)^+ \\ =&\left(\sum_{i=1}^{n} \big(f_i(\omega) - f_i(W)\big)\right) \mathbf{1}_{\{\omega \leq W\}} = \sum_{i=1}^{n} \big(f_i(\omega) - f_i(W)\big)_+ \end{aligned}$$

In financial application, each of random variables $f_i(W)$ represents an asset value, the number K is the strike of the option on the portfolio of assets. We can therefore express the payoff of an option of assets in terms of a portfolio of options on the individual assets $f_i(W)$ with corresponding $f_i(\omega)$.

To apply *Jamshidian trick* on (42), we needs to verify that V is monotone decreasing in r. In Vasicek model, it is automatic by observing the format of solution. Thus, by any numerical method for root finding we can determine an r^* such that:

$$\sum_{i=1}^{n} c_i V(T_0, r^*; T_i) = K, \tag{44}$$

and the payoff can be re-written as:

$$\left[\sum_{i=1}^{n} c_i \bigg(V(T_0, r^*; T_i) - V(T_0, r_{T_0}; T_i) \bigg) \right]^+. \tag{45}$$

Now, pricing our coupon-bearing-bond option is equivalent to pricing a portfolio of put options on zero-coupon bonds. We have

$$\begin{aligned} CBP(t, T_0; \mathcal{T}, \mathcal{C}, K) &= \mathbb{E}_t\left[D(t, T_0) \left[\sum_{i=1}^{n} c_i \bigg(V(T_0, r^*; T_i) - V(T_0, r_{T_0}; T_i) \bigg) \right]^+ . \right] \\ &= \sum_{i=1}^{n} c_i \mathbb{E}_t\left[D(t, T_0) \bigg(V(T_0, r^*; T_i) - V(T_0, r_{T_0}; T_i) \bigg)^+ \right] \\ &= \sum_{i=1}^{n} c_i ZBP(t, T_0, T_i, V(T_0, r^*; T_i)) \end{aligned} \tag{46}$$

Let's adapt the above method to Swaption pricing. Without loss of generality, we choose a payer Swaption with notional $N = 1$, strike K and maturity T_0 to illustrate. At time T_0, the owner has the right to enter an interest rate swap with payment schedule $T_1, T_2, ..., T_n$, where fixed rate K is paid and LIBOR set "in arrears" is received. We denote τ_i the year fraction from T_{i-1} to T_i, then

$$\begin{aligned} \textit{Swaption Time } T_0 \textit{ Payoff} &= \left(\sum_{i=1}^{N} P(T_0, T_i)\tau_i \left(\frac{1}{\tau_i} \left(\frac{P(T_0, T_{i-1})}{P(T_0, T_i)} - 1 \right) - K \right) \right)^+ \\ &= \left(\sum_{i=1}^{n} P(T_0, T_{i-1}) - P(T_0, T_i) - K\tau_i P(T_0, T_i) \right)^+ \\ &= \left(1 - \big(P(T_0, T_n) + K \sum_{i=1}^{n} P(T_0, T_i) \big) \right)^+ \\ &= \big(1 - CB(T_0; \mathcal{T}, \mathcal{C}) \big)^+ \end{aligned} \tag{47}$$

with the coupon $c_i := \tau_i K$ for $i = 1, 2, ..., n-1$ and $c_n = (1 + K\tau_i)$. Denoting by r^* the value of the spot rate at time T_0 for which

$$\sum_{i=1}^{n} A(T_0, T_i) e^{-B(T_0, T_i) r^*} = 1, \tag{48}$$

and then define $K_i := A(T_0, T_i)\exp\{-B(T_0, T_i)r^*\}$, the swaption price at time t is given by:

$$PS(t, T_0; \mathcal{T}, K) = \sum_{i=1}^{n} c_i ZBP(t, T_0, T_i, K_i). \tag{49}$$

Remark 1.8. *Notice, in all those derivative pricing problems, we switched to a forward measure to over come the complexity of stochastic interest rate. As a result, in pricing formula we need to know the $T \mapsto P(t, T)$ curve which can be bootrstrapped as we discussed in the previous notes.*

1.3 Cox-Ingersoll-Ross Model

The *Cox-Ingersoll-Ross(CIR) model* is featured by the introduction of a "square root" term in front of diffusion, i.e.,

$$dr_t = \big(\kappa\theta - (\kappa + \lambda\sigma)r_t\big)dt + \sigma\sqrt{r_t}\,dW_t^{\mathbb{P}}, \quad r_0 = r. \tag{50}$$

Notice, the drift term is identical to Vasicek model. To connect to risk-neutral world, we set the market price of risk to be

$$\lambda(t, r_t) = \lambda\sqrt{r_t}, \tag{51}$$

Such choice preserves the natural of the model after change of measure. Indeed, the dynamic under $\mathbb{Q}$ is,

$$dr_t = \kappa\big(\theta - r_t\big)dt + \sigma\sqrt{r_t}\,dW_t^{\mathbb{Q}}, \quad r_0 = r. \tag{52}$$

In order to ensure the origin to be inaccessible to the process[3], the condition below has to be enforced:

$$2\kappa\theta > \sigma^2. \tag{53}$$

CIR model is neither normal nor log-normal model, the presence of the square root term makes the distribution of process to be a *non-centra Chi-squared distribution*(see *Figure* (1)). The distribution has support on $(0, +\infty]$ and has fat-tails. The mean and variance of r_t conditioning on r_s is straightforward,

$$\begin{aligned}
\mathbb{E}_s^{\mathbb{Q}}[r_t] &= r_s e^{-\kappa(t-s)} + \theta\big(1 - e^{-\kappa(t-s)}\big),\\
Var_s^{\mathbb{Q}}[r_t] &= r_s\frac{\sigma^2}{\kappa}\left(e^{-\kappa(t-s)} - e^{-2\kappa(t-s)}\right) + \theta\frac{\sigma^2}{\kappa}\big(1 - e^{-\kappa(t-s)}\big).
\end{aligned}$$

[3]It can be verified by a mimicking Bessel process.

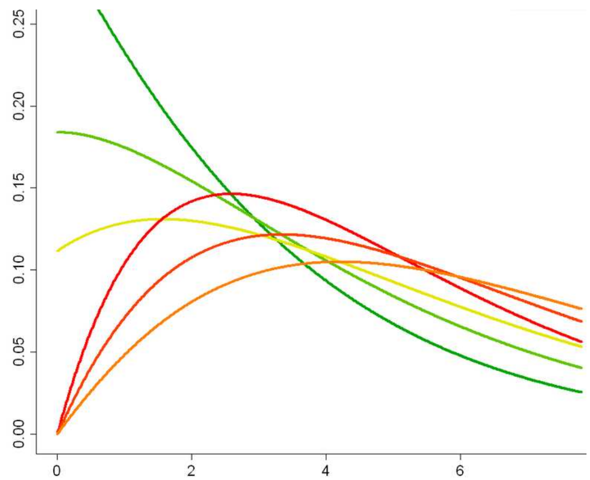

Figure 1: Noncentral Chi-square Distribution

By making an initial guess as in the Vasicek model for zero-coupon-price, i.e., $P(t,T) = A(t,T)e^{-B(t,T)r_t}$ and use fundamental PDE (8), we obtain:

$$\begin{aligned} A(t,T) &= \left[\frac{2h\exp\{(\kappa+h)(T-t)/2\}}{2h+(\kappa+h)(\exp\{(T-t)h\}-1)}\right]^{2\kappa\theta/\sigma^2}, \\ B(t,T) &= \frac{2(\exp\{(T-t)h\}-1)}{2h+(\kappa+h)(\exp\{(T-t)h\}-1)}, \\ h &= \sqrt{2\kappa^2+2\sigma^2}. \end{aligned}$$

By Itô's formula, the dynamic for $P(t,T)$, $f(t,T)$ can be obtained under risk-neutral measure:

$$dP(t,T) = r_t P(t,T)dt - B(t,T)P(t,T)\sigma\sqrt{r_t}dW_t^{\mathbb{Q}}, \tag{54}$$

$$df(t,T) = (\cdots)dt + \sigma\sqrt{r_t}B(t,T)dW_t^{\mathbb{Q}}. \tag{55}$$

The corresponding dynamics under T-forward measure are

$$\begin{aligned} dr_t &= \big(\kappa\theta - (\kappa+\sigma^2 B(t,T))r_t\big)dt + \sigma r_t dW_t^T \\ dP(t,T) &= \big(P(t,T) - B^2(t,T)P(t,T)\sigma^2\big)r_t dt + -B(t,T)P(t,T)\sigma\sqrt{r_t}dW_t^T \\ df(t,T) &= \sigma\sqrt{r_t}B(t,T)dW_t^T. \end{aligned}$$

Lastly, the forward rate $F(t;T,S)$ dynamics under S-forward measure is, for $t \le t \le S$,

$$
\begin{aligned}
dF(t;T,S) = \sigma\bigg(F(t;T,S) + \frac{1}{\tau(T,S)} + \frac{1}{\tau(T,S)}\bigg)\times \\
\sqrt{\big(B(t,S) - B(t,T)\big)\ln\bigg(\big(\tau(T,S)F(t;T,S)+1\big)\frac{A(t,S)}{A(t,T)}\bigg)}dW_t^S
\end{aligned}
\tag{56}
$$

As observed in Vasicek case, for pricing caps/floors, coupon-bearing-bond-options and Swaptions, the building block is the zero-coupon-bond option price. In CIR model, by solving corresponding PDE or integration on the distribution of r_T, it can be achieved that, for option maturing at T with strike K whose underlying is S-zero-coupon-bond, the time t value of the option is:

$$
\begin{aligned}
ZBC(t,T,S,K) = \; & P(t,S)\mathcal{X}^2\bigg(2\bar{r}(\rho+\psi+B(T,S)); \frac{4\kappa\theta}{\sigma^2}, \frac{2\rho^2 r_t \exp\{h(T-t)\}}{\rho+\psi+B(T,S)}\bigg) \\
& - KP(t,T)\mathcal{X}^2\bigg(2\bar{r}(\rho+\psi); \frac{4\kappa\theta}{\sigma^2}, \frac{2\rho^2 r_t \exp\{h(T-t)\}}{\rho+\psi}\bigg)
\end{aligned}
\tag{57}
$$

where

$$
\begin{aligned}
\rho &= \rho(T-t) := \frac{2h}{\sigma^2(\exp(h(T-t))-1)} \\
\psi &= \frac{\kappa+h}{\sigma^2}, \quad \bar{r} = \bar{r}(S-T) := \frac{A(T,S)/K}{B(T,S)}.
\end{aligned}
$$

1.4 Affine Term-Structure Model

Although differs in dynamics, most of the derivations are similar for Vasicek model and CIR model. This is because they all fall into the category of *affine term-structure model*.

Definition 1.1. Affine term-structure models *are interest rate models where the continuously compounded spot rate $R(t,T)$ is an affine function in the short rate r_t, i.e.,*

$$
R(t,T) = -\frac{\ln P(t,T)}{\tau(t,T)} = \alpha(t,T) + \beta(t,T)r_t \tag{58}
$$

where α and β are deterministic functions of time t. Equivalently, $P(t,T)$ is represented as

$$
P(t,T) = A(t,T)e^{-B(t,T)r_t} \tag{59}
$$

where

$$
\alpha(t,T) = \frac{-\ln A(t,T)}{\tau(t,T)}, \quad \beta(t,T) = \frac{B(t,T)}{\tau(t,T)}. \tag{60}
$$

This explicit form of $P(t, T)$, (59), makes derivations of bond dynamic, forward rate dynamic and instantaneous forward rate dynamic convenient under different measures.

We can identify affine term-structure model by only examining the dynamic of short rates under risk-neutral measure. Let's suppose a general time-homogeneous risk-neutral dynamics for short rate as below:

$$dr_t = b(t, r_t)dt + \sigma(t, r_t)dW_t^{\mathbb{Q}}. \tag{61}$$

If $b : [0, T] \times \mathbb{R} \mapsto \mathbb{R}$ and $\sigma : [0, T] \times \mathbb{R} \mapsto \mathbb{R}$ are of the following form:

$$b(t, x) = \lambda(t)x + \eta(t), \quad \sigma^2(t, x) = \gamma(t)x + \delta(t), \tag{62}$$

for suitable chosen deterministic time functions $\lambda, \eta, \gamma, \delta$, then the model has affine term structure. Then, as seen in both CIR and Vasicek case, we can solve the following *Riccati differential equation* to obtain the concrete expressions for both A and B,

$$\begin{cases} \partial_t B(t, T) + \lambda(t)B(t, T) - \frac{\gamma}{2}B^2(t, T) + 1 = 0, \quad B(T, T) = 0, \\ \partial_t \ln A(t, T) - \eta(t)B(t, T) + \frac{1}{2}\delta(t)B^2(t, T) = 0, \quad A(T, T) = 1. \end{cases} \tag{63}$$

The converse is true conditioning on the coefficients b and σ^2 independent of time. That is, if

$$b(x) = \lambda x = \eta, \quad \sigma^2(x) = \gamma x + \delta, \tag{64}$$

then the model is of affine term-structure.

1.5 Log-normal Models

1.5.1 Dothan Model

Dothan model is a driftless log-normal model under original measure, namely, they dynamic is defined by:

$$dr_t = \sigma r_t dW_t^{\mathbb{P}}. \tag{65}$$

Moving to risk-neutral measure, a constant market price of risk is specified, i.e., $\lambda(t, r_t) = \alpha$, which results in:

$$dr_t = ar_t dt + \sigma r_t dW_t^{\mathbb{Q}}. \tag{66}$$

(66) can be solved directly,

$$r_t = r_s \exp\left\{(\alpha - \frac{\sigma^2}{2})(t - s) + \sigma\big(W_t - W_s\big)\right\}, \tag{67}$$

which gives conditional mean and variance,

$$\mathbb{E}_t^{\mathbb{Q}}[r_t] = r_s e^{-\alpha(t-s)}, \tag{68}$$

$$Var_t^{\mathbb{Q}}[r_t] = r_s^2 e^{2\alpha(t-s)}\big(e^{\sigma^2(t-s)} - 1\big). \tag{69}$$

The nature of log-normal distribution prohibits the rates in Dothan model to be negative, and the drift term only allows mean reversion around 0 level. The zero-coupon-bond still has an analytical form, which is a surprising result, since the bond price requires exponentiating the exponential function that is highly non-linear. It turns out that *Dothan model* is the only log-normal model has closed form of zero-coupon-bond, i.e.,

$$\begin{aligned} P(t,T) = &\frac{\bar{r}^p}{\pi^2}\int_0^\infty \sin(2\sqrt{\bar{r}}\sinh y)\int_0^\infty f(z)\sin(yz)dzdy \\ &+ \frac{2}{\sqrt{\Gamma(2p)}}\bar{r}^p K_{2p}(2\sqrt{\bar{r}}), \end{aligned}$$

where

$$\begin{aligned} f(z) &= \exp\left\{\frac{-\sigma^2(4p^2+z^2)(T-t)}{8}\right\} z\left|\Gamma(-p+i\frac{z}{2}\right|^2\cosh\frac{\pi z}{2} \\ \bar{r} &= 2r_t/\sigma^2, \quad p = \frac{1}{2} - \alpha. \end{aligned}$$

and K_q denotes the modified Bessel function of the second kind of order q. As one can see the explicit formula is not that nice, especially the evaluation of double integral makes the computational efficiency dramatically reduced.

1.5.2 EV Model

Since vasicek model is a normal model, the exponentiation is a log-normal model, we call it *Exponential-Vasicek model(EV)*. In other words, the logarithm of short rate r_t follows *Ornstein-Uhlenbeck process* under risk neutral measure $\mathbb{Q}$,

$$dy_t = \big(\theta - \alpha y_t\big)dt + \sigma dW_t^{\mathbb{Q}}, \quad y_0 = y, \tag{70}$$

with $y \in \mathbb{R}$, θ, α and σ positive constant. Then, $r_t = \exp\{y_t\}$ follows

$$dr_t = r_t\left[\theta + \frac{\sigma^2}{2} - \alpha\ln r_t\right]dt + \sigma r_t dW_t^{\mathbb{Q}} \tag{71}$$

The conditional mean and variance can be obtained easily, which, as in Vasicek model, exhibits mean-reverting. However, even zero-coupon-bond can not be priced analytically, let alone other derivatives. Numerical procedures have to be implemented for derivative pricing.

1.5.3 Issue with Log-normal Model

Log-normal model has fat-tail distribution of the underlying process, which is more realistic in the market. It also prohibits the possibility of negative rates, which is used to be a good property. However, since the emergence of negative rates on market, log-normal model has to be shifted in the first place. Another concern for log-normal model is the explosion of bank account. We know that the expected value of our position at time Δt will be

$$\mathbb{E}[B(\Delta t)] = \mathbb{E}\big[e^{\int_0^{\Delta t} r_s ds}\big] \tag{72}$$

If Δt is small, we can approximate the integral as follows:

$$\mathbb{E}[B(\Delta t)] \approx \mathbb{E}\big[e^{\Delta t(r_0+r(\Delta t))/2}\big] \tag{73}$$

Given that short rate $r(\Delta t)$ is log-normally distributed, we face an expectation of the type

$$\mathbb{E}\big[\exp\big(\exp(Y)\big)\big] \tag{74}$$

where Y is normally distributed. Such expectation can be proved to be infinite. This means that in arbitrarily small time the bank account growths to infinity on average. This drawback can be partially overcame when using numerical recipes, because the discretization can only be finite.

1.6 Time-varying Short Rate Models

The models introduced up to now are all endogenous term-structure model, that is, the current term structure of rates is an output rather than an input of the model. For example, in vasicek model, the bond price $P(t,T)$ is a function of κ, θ, σ and r_t, once they are determined by calibration of derivatives, the yield curve $T \mapsto P(t,T;\kappa,\theta,\sigma,r_t)$ is known. On the other hand, we have $P^{Market}(t,T)$ observed from market. Are they compatible with each other, at least close? It is very unlikely to happen, the number of parameters we can control is too less to reproduce the curve. Let's put in another way, if we try to calibrate the model by zero-coupon price, the optimization should render us the best it can do to fit the initial term-structure. Three parameters, however, can not output a satisfactory yield curve.

To improve this situation, *exogenous term* has to be considered. We want to both preserve the fundamental structure of endogenous models and better fit initial term structure. The basic strategy is to include "time-varying" parameters. In the following, we will discuss couple of models that are of this kind.

1.6.1 Hull-White One-Factor Model

The *Hull-White model* is a direct generalization of vasicek model, the dynamic of which under risk-neutral measure is:

$$dr_t = \big(\theta(t) - \alpha r_t\big)dt + \sigma dW_t^{\mathbb{Q}}, \tag{75}$$

where α and σ are positive constant, θ is a deterministic function of time used for fitting the term structure. With same initial guess for affine model, *Raccati equation* gives:

$$\begin{aligned} A(t,T) &= -\frac{\sigma^2}{2}\int_t^T B^2(s,T)ds + \int_t^T \theta_s B(s,T)ds, \\ B(t,T) &= \frac{1}{\alpha}\big(1 - e^{-\alpha(T-t)}\big). \end{aligned}$$

Let's take a look at the instantaneous forward rate,

$$\begin{aligned} f(0,t) &= -\frac{\partial \ln P(0,t)}{\partial T} = -\partial_T \ln A(0,t) + \partial_T B(0,t) r_0 \\ &= -\frac{\sigma^2}{2\alpha^2}\big(e^{-\beta t} - 1\big)^2 + \int_0^t \theta_s e^{-\alpha(t-s)} ds + e^{-\alpha t} r_0 \end{aligned}$$

By inverting $f(0,t)$,

$$\theta_t = \frac{\partial f(0,t)}{\partial T} + \alpha f(0,t) + \frac{\sigma}{2\alpha}(1 - e^{-2\alpha t}). \tag{76}$$

To be consistent with market data and the Hull-White model, θ_t has to be chosen as in (76). Let's denote the observed curve to be $f^M(0,T)$, substitute (76) into (75) and integrate from s to t,

$$r_t = r_s e^{-\alpha(t-s)} + \Lambda_t - \Lambda_s e^{-\alpha(t-s)} + \sigma \int_s^t e^{-\alpha(t-u)} dW_u^{\mathbb{Q}}, \tag{77}$$

where

$$\Lambda_t = f^{Market}(0,t) + \frac{\sigma^2}{2\alpha^2}\big(1 - e^{-\alpha t}\big)^2,$$

Conditioning on $\mathcal{F}_s$, r_t is normally distributed with mean and variance below:

$$\mathbb{E}_s[r_t] = r_s e^{-\alpha(t-s)} + \Lambda_t - \Lambda_s e^{-\alpha(t-s)}, \tag{78}$$

$$Var_s[r_t] = \frac{\sigma^2}{2\alpha}\big[1 - e^{-2\alpha(t-s)}\big]. \tag{79}$$

Interestingly, r_t can be decomposed linearly into x_t and Λ_t, where x_t is defined by certain SDE. From (77), let's write

$$
\begin{aligned}
& x_t + \Lambda_t = (x_s + \Lambda_s)e^{-\alpha(t-s)} + \Lambda_t - \Lambda_s e^{-\alpha(t-s)} + \sigma \in_s^t e^{-\alpha(t-u)} dW_u. \\
\rightarrow \quad & x_t = x_s e^{-\alpha(t-s)} + \sigma \int_s^t e^{-\alpha(t-u)} dW_u^{\mathbb{Q}}. \\
\rightarrow \quad & dx_t = -\alpha x_t dt + \sigma dW_t^{\mathbb{Q}}.
\end{aligned}
$$

Therefore, at any time t, r_t can be recovered by:

$$
r_t = \Lambda_t + x_t \tag{80}
$$

Notice Λ_t is determined by market data thus deterministic.

Next, we want to price zero-coupon-bond options in Hull-White setting. The procedure is almost the same as in Vasicek case. First of all, we want to obtain the bond dynamic under risk neutral measure. To do so, let's first plug θ_t to $A(t,T)$, which gives us(the derivation is a little bit involved):

$$
A(t,T) = \frac{P^{Market}(0,T)}{P^{Market}(0,t)} \exp\left\{ B(t,T) f^{Market}(0,t) - \frac{\sigma^2}{4\alpha}(1-e^{-2\alpha t})B^2(t,T) \right\} \tag{81}
$$

Apply Itô's lemma on $P(t,T) = A(t,T)e^{-B(t,T)r_t}$,

$$
\begin{aligned}
dP(t,T) = & \left[P(t,T) - B(t,T)P(t,T)(\theta_t - \alpha r_t) + \frac{1}{2} B^2(t,T)P(t,T)\sigma^2 \right] dt \\
& - P(t,T)B(t,T)\sigma dW_t^{\mathbb{Q}}
\end{aligned} \tag{82}
$$

Then, we define $U_t = \frac{P(t,S)}{P(t,T)}$ and get dynamic of U_t under T-forward measure. Lastly, the option price can be computed by a BS like formula:

$$
ZBC(t,T,S,K) = P(t,S)N(h) - KP(t,T)N(h - \sigma_p), \tag{83}
$$

$$
ZBP(t,T,S,K) = KP(t,T)N(\sigma_p - h) - P(t,S)N(h). \tag{84}
$$

where

$$
\begin{aligned}
\sigma_p &= \sigma \sqrt{\frac{1 - e^{-2\alpha(T-t)}}{2\alpha}} B(T,S), \\
h &= \frac{1}{\sigma_p} \ln \frac{P(t,S)}{P(t,T)K} + \frac{\sigma_p}{2}.
\end{aligned}
$$

The derivation of cap/floor, swaption and option on coupon-bearing-bond follows immediately.

Remark 1.9. *The derivation of option price is almost the same as in the Vasicek case, however it is much more sophisticated and a little bit messy. To not obscure the main information we want to deliver, the derivation is not present (reader interested can try to work it out).*

1.6.2 Other Affine Term Structure Model – Extended CIR Model & Ho-Lee Model

Ho-Lee model is the simplest normal model that is in the category of affine term-structure:

$$dr_t = \theta_t dt + \sigma dW_t^{\mathbb{Q}}, \quad r_t = r. \tag{85}$$

where θ_t is a deterministic function of time. To fit initial yield curved structure, θ_t has to be the form below:

$$\theta_t = f^{Market}(0,t) + \sigma^2 t. \tag{86}$$

Because of the simple structure, the building block of derivative pricing, option on zero-coupon-bond, has closed form under Ho-Lee model. Nevertheless, as a normal model, the symmetric normal distribution is controversial. In addition, the mean-reversion level is fixed to be 0.

The time-dependent version of CIR model was proposed by Hull and White in the following version under risk-neutral measure

$$dr_t = \left[\theta_t - \alpha_t r_t\right]dt + \sigma_t \sqrt{r_t} dW_t^{\mathbb{Q}}, \quad r_0 = r. \tag{87}$$

where α, θ, σ are deterministic functions of time. However, such extension is not analytically tractable. Although it has affine term structure, the resulting *Riccati equation* can not be explicitly solved. As a result, one has to resort to numerical solution. Even we reduce to the case only θ is time dependent, it still does no have closed form solution for zero-coupon bond.

1.6.3 Log-normal Model – BK Model

The *Black Karasinski(BK)* model is actually a time-dependent version of *EV* model, it assumes the logarithm of short rate follows a time-dependent Ornstein-Uhlenbeck process,

$$d \ln r_t = \left[\theta_t - \alpha \ln r_t\right]dt + \sigma dW_t^{\mathbb{Q}}, \quad r_0 = r. \tag{88}$$

As in the previous mean reversion models, the coefficients α and σ can be interpreted as the speed of mean reversion and standard deviation per time unit of the instantaneous return of r_t, respectively. Itô's formula can transform to the original dynamic for r_t under risk-neutrality,

$$dr_t = r_t\big[\theta_t + \frac{\sigma^2}{2} - \alpha \ln r_t\big]dt + \sigma r_t dW_t^{\mathbb{Q}}, \quad r_0 = r. \tag{89}$$

whose explicit solution satisfies

$$r_t = \exp\left\{\ln r_s e^{-\alpha(t-s)} + \int_s^t e^{-\alpha(t-u)}\theta_u du + \sigma \int_s^t e^{-\alpha(t-u)} dW_u^{\mathbb{Q}}\right\} \tag{90}$$

As expected, the conditional distribution of r_t is log-normal distributed. As EV model, it is not analytically tractable, one has to use numerical procedure for zero-coupon-bond pricing and also other derivatives.

1.7 Calibration Remark

Let's define calibration operator $\mathcal{L}$, which takes market price and model parameters as input. We try to find the parameters, Θ, that minimize the difference, in some sense(for example, L^2-distance),

$$\min_{\Theta} \mathcal{L}(Market\ Price, \Theta) \tag{91}$$

The candidate instruments can be bond options, caps/floors and swaptions. For endogenous model, Vasicek and CIR model, has nothing special, all those derivatives have tractable expressions(see (37), (38), (39), (40), (46) and (49)). Thus, the $\mathcal{L}$ operator can compare market price and theoretical price very efficiently. While for other log-normal model, Monte Carlo simulation or numerical solution to PDE has to be adopted for pricing which complicates the analysis of calibration operator $\mathcal{L}$.

For exogenous model, one has to first determine the time-varying function. As illustrated previously, those "function parameter" can be obtained by looking up the forward instantaneous curve interpolated, namely, $f^{Market}(t, T)$. Notice, f should be stripped from the bootstrapped discounting curve. Then, the rest of the parameters are constant, it can be taken care by calibration operator, in Hull-White and Ho-Lee model, the calibration is efficient for existence of explicit formula for derivative. For other model, numerical methods are needed, especially, the tree lattice model is of great interests.

1.8 Short Rate Model Summary & Mulit-Factor Models

Let's summarize the strengths and weakness of short rate model.

Strengths:

- the model structure is generally simple, some of affine-term-structure models are analytical tractable, others are very amenable to numerical and Monte-Carlo simulation methods;
- derivatives pricing, i.e., cap/floor, Swaptions, is very efficient for some short rate models, Vasicek, CIR, Hull White, Ho-Lee, because the existence of explicit expression similar to BS formula;
- they can provide *"sanity checks"* on more sophisticated models that can be calibrated to "fit everything". Models that are calibrated to "fit everything" can be unreliable due to over-fitting. Of course, exogenous short rate models are also susceptible to over-fitting.

Weakness:

- the one-factor short rate models imply that movements in the entire term-structure can be hedged with only two securities. Equivalently, instantaneous returns on zero-coupon bonds of different maturities are perfectly correlated. Neither of these features is realistic but these problems can be overcome by using multi-factor models. In fact, models with just 2 or 3 factors can afford considerably more modeling flexibility. Moreover, they generally retain their numerical tractability. Models with 3 or more factors, however, tend to suffer from the curse-of-dimensionality in which case Monte Carlo simulation becomes the only practical pricing technique;
- they are not as "close to reality" as LIBOR market models(we will discuss in a moment), which directly model observable market quantities, i.e., LIBOR rates, and this feature makes these models relatively straightforward to calibrate, because most of derivatives are directly related to LIBOR, modeling short rate needs one extra step to recover LIBOR.

Let's finish by comments on multi-factor model. As mentioned above, the main drawback of single factor short rate model is the perfect correlation among bonds of different maturities. To be specific, let's illustrate with asicek model. The continuously compounded spot rates are given by:

$$R(t,T) = -\frac{\ln A(t,T)}{T-t} + \frac{B(t,T)}{T-t} r_t := a(t,T) + b(t,T) r_t. \tag{92}$$

Consider a derivative whose payoff depends on two such rates at time t. We can set $T_1 = t + 1yr$ and $T_2 = t + 10yrs$. Obviously, the payoff is determined by the joint distribution of the one-year and ten-year continuously-compounded spot interest rate at expiry t. As long as joint distribution is involved, the correlation plays a significant role, in Vasicek model, the correlation can be computed

$$Corr\big(R(t,T_1), R(t,T_2)\big) = Corr\big(a(t,T_1) + b(t,T_1) r_t, a(t,T_2) + b(t,T_2) r_t\big) = 1 \tag{93}$$

Obviously, this is really not what is happening in the market. The real world interests are known to exhibit some de-correlation. Therefore, we need a model exhibits more realistic correlation patterns. This motivates the multi-factor short rate model. Again, let's use a *two-factor Vasicek* model to demonstrate. The dynamics of r_t is decomposed to two stochastic process, each of which follows single factor vasicek with correlated Brownian motion, i.e.,

$$\begin{aligned}
r_t &= x_t + y_t, \\
dx_t &= \kappa_x(\theta_x - x_t)dt + \sigma_x dW_1^{\mathbb{Q}}(t), \\
dy_t &= \kappa_y(\theta_y - y_t)dt + \sigma_y dW_2^{\mathbb{Q}}(t), \\
\mathbb{E}^{\mathbb{Q}}\big[dW_1^{\mathbb{Q}}(t)dW_2^{\mathbb{Q}}(t)\big] &= \rho dt.
\end{aligned} \tag{94}$$

The bond price is of the form:

$$P(t,T) = A(t,T)\exp\big\{-B^x(t,T)x_t - B^y(t,T)y_t\big\} \tag{95}$$

Now, if we compute the continuously compounded spot rate,

$$\begin{aligned}
&Corr\big(R(t,T_1), R(t,T_2)\big) \\
=&Corr\big(b^x(t,T_1)x_t + b^y(t,T)y_t, b^x(t,T_2)x_t + b^y(t,T_2)y_t\big)
\end{aligned}$$

This quantity is in general not equal to 1, depending on the correlation ρ. Thus, we gain more flexibilities to match the market correlation pattern. The number of factors to be used really depends. According to a *PCA(principle component analysis)* of the yield curve, it is found that two components can already explain 85% to 90% of the variation, and three factors can go up to 94%. Thus, the typical choice will be two or three factors. Although increasing number of factors can boost the model, it is, on other hand, negatively affect the analytical tractability of the model.

2 Heath-Jarrow-Morton Philosophy

Heath-Jarrow-Morton(HJM) is a quite general framework for the modeling of interest rate dynamics. The idea behind the HJM approach to model the bond market is to follow the procedure we took for risky asset: first propose a model for how the assets behave stochastically in the real world; Secondly, derive an equivalent risk-neutral model. HJM models apply this strategy to the bond market. Fix a $\widetilde{T}$, for each $0 \le T \le \widetilde{T}$, we can think of the zero-coupon bond maturing at T, with price $\{P(t,T); 0 \le t \le T\}$ as a risky asset or investment opportunity. Thus, instead of a market consisting of finite number of assets, we have the bond market presents a whole continuum of financial instruments, indexed by maturity date T, with prices $P(t,T)$.

Certainly, one can directly model each T-maturity bond price $P(t,T)$ by SDE. However, such approach has several complications: the bond price satisfies $P(T,T) = 1$ that is hard to enforce; when modeling stock prices, the short rate r_t was specified exogenously, but the short rate cannot be specified independently of the zero-coupon bond price; the bond prices has to stay non-negative with probability 1 which also narrows the models that can be chosen.

The HJM approach, however, directly models the forward rate $f(t,T)$, which defines the zero-coupon price as,

$$P(t,T) = e^{-\int_t^T f(t,u)du}, \quad f(0,T) = -\frac{\partial \ln P(0,T)}{\partial T}. \tag{96}$$

Obviously, now, there is no concern about the terminal value, since $P(T,T) = 1$ by definition, also the bond price won't delve into negative terrain. To ease the notation, let's only consider one-dimensional Brownian motion, the HJM follows:

$$df(t,T) = \alpha(t,T)dt + \sigma(t,T)dW_t, \quad 0 \le t \le T \le \widetilde{T}. \tag{97}$$

At time $t = 0$, we know the forward rate curve, $T \mapsto f(0,T)$(numerical apporximation of (96)), which provides the initial condition for (97), i.e.,

$$f(t,T) = f(0,T) + \int_0^t \alpha(u,T)du + \int_0^t \sigma(u,T)dW_u \tag{98}$$

In particular, since $f(t,t) = R(t)$, the short rate at time t,

$$R(t) = f(t,t) = f(0,t) + \int_0^t \alpha(u,t)du + \int_0^t \sigma(u,t)dW_u. \tag{99}$$

The dynamic of the bond price can be derived by applying Itô's formula to $e^{-\int_t^T f(t,v)dv}$, by noticing that,

$$\begin{aligned}
d\left(-\int_t^T f(t,v)dv\right) &= f(t,t)dt - \int_t^T df(t,v)dv \quad \textit{(Lebniz Rule)} \\
&= R(t)dt - \int_t^T \big[\alpha(t,v) + \sigma(t,v)\big]dv \\
&= R(t)dt - \alpha^*(t,T)dt - \sigma^*(t,T)dW_t
\end{aligned}$$

with

$$\begin{aligned}
\alpha^*(t,T) &= \int_t^T \alpha(t,v)dtdv = \int_t^T \alpha(t,v)dvdt \\
\sigma^*(t,T) &= \int_t^T \sigma(t,v)dW_t dv = \int_t^T \sigma(t,v)dvdW_t
\end{aligned}$$

Therefore, the zero-coupon price $P(t, T)$ follows:

$$\begin{aligned} dP(t,T) = & P(t,T)\big[R(t) - \alpha^*(t,T) + \frac{1}{2}\big(\sigma^*(t,T)\big)^2\big]dt \\ & - P(t,T)\sigma^*(t,T)dW_t \end{aligned} \tag{100}$$

An interesting observation is that $\sigma^*(T, T) = 0$, which makes sense because the volatility of $P(T, T)$ is zero leading $P(T, T) = 1$.

In order that there be no arbitrage in HJM model, it is necessary that there exists an equivalent probability measure $\mathbb{Q}$ with respect to $\{D(t)P(t, T)\}_{\{t \leq T\}}$ is a martingale for $0 \leq t \leq T$. A necessary condition for this to be true is the market risk equation. In the case of stock, we have only a finite number, i.e., indexed by i, but now there is a continuum, each $T \leq \widetilde{T}$. Thus, $\forall T \in [t, \widetilde{T}]$, there exists one adatped process $\{\theta_t\}_{0 \leq t \leq \widetilde{T}}$ such that

$$\begin{aligned} -\sigma^*(t,T)\theta_t = & R(t) - \alpha^2(t,T) + \frac{1}{2}\big(\sigma^*(t,T)\big)^2 - R(t) \\ = & -\alpha^*(t,T) + \frac{1}{2}\big(\sigma^*(t,T)\big)^2 \end{aligned}$$

By differentiating both sides with respect to T one derives the equivalent

$$\sigma(t,T)\theta_t = \alpha(t,T) - \sigma^*(t,T)\sigma(t,T), \quad for\ 0 \leq t \leq T \leq \widetilde{T}. \tag{101}$$

Conversely, if this equation has a solution and if

$$\mathbb{E}\Big[\exp\Big\{-\int_0^{\tilde{T}} \theta_u dW_u - \int_0^{\tilde{T}} \theta_u^2 du\Big\}\Big] = 1,$$

then there exists a risk-neutral measure.

Suppose the the existence of risk-neutral measure is justified, then $W_t^{\mathbb{Q}} = W_t + \int_0^t \theta_u du$ is a Brownian motion under $\mathbb{Q}$, and

$$df(t,T) = \sigma^*(t,T)\sigma(t,T)dt + \sigma(t,T)dW_t^{\mathbb{Q}}. \tag{102}$$

This is an interesting and important equation. There is no freedom in choosing the drift of the forward rate under risk-neutral model. It must be $[\sigma^*\sigma](t, T)$, which is completely determined by the volatility $\sigma(t, T)$. Although it is not immediately obvious, this fact limits the nature of the models that can be proposed for the forward rate. For example, suppose we want to have a log-normal model for $f(t, T)$ in the risk-neutral world. Then we would like to have $\sigma(t, T) = \sigma(t) f(t, T)$ which is obviously not possible in the HJM framework. Forgetting a bout the dirft α in the original model, we can use (102) to pose a risk neutral model, where

$$dP(t,T) = R(t)P(t,T)dt - \sigma^*(t,T)P(t,T)dW_t^{\mathbb{Q}}.$$

The *Vasicek*, *Hull-White* are both *HJM* models. Actually, any exogenous term=structure interest rate model can be derived within such a framework. However, only a restricted class of volatilities is known to imply a Markovian short-rate process. This means, in general, burdensome procedure are needed to price interest rate derivatives. Substantially, the problem remains of defining a suitable volatility function for practical purposes.

3 LIBOR Market Model

In this section, we will focus on an interest rate model of different nature – *LIBOR Market Model(LMM)*, which is a popular model for practitioners. Most of interest rate derivatives pricing are associated with LIBOR rate, when modeling short rate, however, the LIBOR turns out to be a very non-linear functional of short rate, especially when the underlying short rate model is sophisticated. LIBOR market model, instead, directly models the dynamics of forward rate(which at maturity equals LIBOR). The model is intrinsically multi-factor, which enables capturing various aspects of the forward curve; parallel shifts, steepenings/flaettenings, butterflies, e.t.c..

3.1 Mechanism of General LIBOR Market Model

Let's start by creating a schedule, $\mathcal{T} = \{T_0, T_1, ..., T_N\}$ with observation time $0 < T_0$. By convention, we usually set $N = 120$ and the $\delta_i = T_i - T_{i-1} = 3M$. We denote by $F_j(t) := F(t; T_{j-1}, T_j)$ the forward rate, for $j = 1, ..., N$. The idea of LMM model is to use a SDE to characterize the evolution of a family $F_j(t)$ (parameterized by tenor). To do so, let's observe that any $F_j(t)$ process gets killed before time T_{j-1}, which is the LIBOR fixing date[4]. To correlate those forward rate process, we use N-dimensional Brownian motion $W_1(t), W_2(t), ..., W_N(t)$ that is correlated by ρ_{jk}, for $1 \le j, k \le N$, i.e.,

$$\mathbb{E}\big[dW_j(t)dW_k(t)\big] = \rho_{jk}dt, \quad 0 \le T_j, T_k \le T_N.$$

We use a general SDE for $F_j(t)$ under real world mesure $\mathbb{P}$,

$$dF_j(t) = \Delta_j\big(t, F_j(t)\big)dt + C_j\big(t, F_j(t)\big)dW_j(t), \quad 0 \le t \le T_{j-1}. \tag{103}$$

Let's fix a $k \in \{1, ..., N\}$, the forward rate $F_k(t)$ is a $\mathbb{Q}^{T_k}$-martingale. As a result, under T_k-forward measure,

$$dF_k(t) = C_k\big(t, F_k(t)\big)dW_k^{T_k}(t), \quad 0 \le t \le T_{k-1}. \tag{104}$$

[4]In practice, the fixing date of LIBOR is usually two days preceding corresponding starting tenor, but we consider them as the same day for simplicity.

This is only one of the family member has been addressed. For $1 \leq j < k$, we need to find the evolution of $F_j(t)$ under measure $\mathbb{Q}^{T_k}$. Notice $F_j(t)$ under T_j-forward measure is also driftless. Since the associated numérairés are $P(t, T_j)$ and $P(t, T_k)$, we have to change from T_j-forward measure to T_k-forward measure. By change of numérairé toolkit we developed, the drift term of $F_j(t)$ under T_k is

$$\Delta_j(t, F_j(t)) = -\left\langle F_j(t), \ln \frac{P(t, T_k)}{P(t, T_j)} \right\rangle. \tag{105}$$

To calculate the quadratic variation above, we have to express the numérairé in terms of forward rate, that is,

$$P(t, T_j) = \prod_{t \leq T_i \leq T_j} \frac{1}{1 + \delta_i F_i(t)}, \tag{106}$$

then, (105) can be re-written as:

$$\begin{aligned}
\Delta_j\big(t, F_j(t)\big) &= -\left\langle F_j(t), \ln \sum_{j+1 \leq i \leq k} \big(1 + \delta_i F_i(t)\big) \right\rangle &(107)\\
&= - \sum_{j+1 \leq i \leq k} \left\langle F_j(t), \ln \big(1 + \delta_i F_i(t)\big) \right\rangle &(108)\\
&= -C_j\big(t, F_j(t)\big) \sum_{j+1 \leq i \leq k} \frac{\rho_{ji} \delta_i C_i\big(t, F_i(t)\big)}{1 + \delta_i F_i(t)}. &(109)
\end{aligned}$$

By the same token, one can easily check, for $j > k$,

$$\Delta_j\big(t, F_j(t)\big) = C_j\big(t, F_j(t)\big) \sum_{k+1 \leq i \leq j} \frac{\rho_{ji} C_i\big(t, F_i(t)\big)}{1 + \delta_i F_i(t)}. \tag{110}$$

To summary, the dynamics of LMM is given by the following systeem of SDEs:

$$\begin{aligned}
&dF_j(t) = C_j\big(t, F_j(t)\big)\\
&\times \begin{cases} -C_j\big(t, F_j(t)\big) \sum_{j+1 \leq i \leq k} \frac{\rho_{ji} \delta_i C_i(t, F_i(t))}{1+\delta_i F_i(t)} dt + dW_j^{T_k}(t), & if\ j < k, \\ dW_j^{T_k}(t), & if\ j = k, \\ C_j\big(t, F_j(t)\big) \sum_{k+1 \leq i \leq j} \frac{\rho_{ji} C_i(t, F_i(t))}{1+\delta_i F_i(t)} dt + dW_j^{T_k}(t), & if\ j > k, \end{cases}
\end{aligned} \tag{111}$$

with initial condition:

$$F_j(0) = F_0^j, \quad \forall j \in \{1, ..., N\}, \tag{112}$$

that are quoted on the market, i.e., FRA rate.

There are considerable number of cases when we want to obtain LMM dynamics under spot measure. For example, in the convexity adjustment for Euro-dollar future, which has to be marked to market. Recall the numéraire of spot measure:

$$B^d(t) = \prod_{1\le i\le \gamma(t)} \big(1+\delta_i F_i(T_{i-1})\big) P(t, T_{\gamma(t)}) \tag{113}$$

where $\gamma(t) = i$ when $t \in [T_{i-1}, T_i]$. Fix a $j \in \{1, ..., N\}$, under T_j-forward measure,

$$dF_j(t) = C_j\big(t, F_j(t)\big) dW_j^{T_j}(t), \quad t \le T_{j-1}. \tag{114}$$

The new drift term is:

$$\begin{aligned}\Delta_j\big(t, F_j(t)\big) &= -\left\langle F_j(t), \ln \frac{\prod_{1\le i\le\gamma(t)} \big(1+\delta_i F(T_{i-1})\big) P(t, T_{\gamma(t)})}{P(t, T_{\gamma(t)})} \right\rangle \\ &= C_j\big(t, F_j(t)\big) \sum_{\gamma(t)\le i\le j} \frac{\rho_{ji}\delta_i C_i\big(t, F_i(t)\big)}{1+\delta_i F_i(t)}.\end{aligned}$$

Since j is arbitrary, we obtain the dynamics of LMM under spot measure,

$$dF_j(t) = C_j\big(t, F_j(t)\big) \sum_{\gamma(t)\le i\le j} \frac{\rho_{ji}\delta_i C_i\big(t, F_i(t)\big)}{1+\delta_i F_i(t)} dt + C_j\big(t, F_j(t)\big) dW_j^{spot}(t) \tag{115}$$

3.2 Structure of Volatility Functional & Correlation Matrix

The discussion so far assumes very general functional form of volatility, i.e., $C_j\big(t, F_j(t)\big)$ for $F_j(t)$. When implementing LMM, the explicit form of $C_j(\cdot,\cdot)$ has to be specified. Several common choices are:

- *Normal model:* $C_j\big(t, F_j(t)\big) = \sigma_j(t)$;
- *CEV model:* $C_j\big(t, F_j(t)\big) = \sigma_j(t) F_j(t)^{\beta_j}$;
- *Log-normal model:* $C_j\big(t, F_j(t)\big) = \sigma_j(t) F_j(t)$;
- *Shifted log-normal model:* $C_j\big(t, F_j(t)\big) = \big(\sigma_j(t) F_j(t) + \eta_j\big)$.

where $\sigma_j(t)$ are deterministic[5], $\beta_j \in [0, 1]$ and $\eta_j \ge 0$. As one may expect, the $\sigma_j(t)$ is usually assumed to be piecewise constant functions. In other words, before the current $F_j(t)$ processes dies, $\sigma(t)$ keeps constant in each interval $[T_i, T_{i+1}]$ where $i \le j$. The table below makes a clear presentation(see *Table 1*). If $N = 120$, the table contains 7140

[5]Notice, there exists stochastic LMM model, we here restrict our attention to the classcial LMM where the diffusion term is deterministic.

Table 1: First Parameterization of $\sigma_j(t)$

$\sigma_j(t)$ / $t \in$	$[0, T_0)$	$[T_0, T_1)$	$[T_1, T_2)$	$\cdots$	$[T_{N-2}, T_{N-1})$
$\sigma_1(t)$	$\sigma_{1,0}$	0	0	$\cdots$	0
$\sigma_2(t)$	$\sigma_{2,0}$	$\sigma_{2,1}$	0	$\cdots$	0
$\sigma_3(t)$	$\sigma_{3,0}$	$\sigma_{3,1}$	$\sigma_{3,2}$	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	0
$\sigma_N(t)$	$\sigma_{N,0}$	$\sigma_{N,1}$	$\sigma_{N,2}$	$\cdots$	$\sigma_{N,N-1}$

parameters, obviously, it is overparameterized.

A more natural choice of parameterization is assuming that the instantaneous volatility is *stationary*, i.e., $\sigma_{j,i} = \sigma_{j-i}$, for all $i < j$. That is the volatility of the forward rate depends on the difference between observation time and maturity(the time when the forward rate being said dies). Table 2 summarizes the staionarity assumption. Such parame-

Table 2: Second Parameterization of $\sigma_j(t)$

$\sigma_j(t)$ / $t \in$	$[0, T_0)$	$[T_0, T_1)$	$[T_1, T_2)$	$\cdots$	$[T_{N-2}, T_{N-1})$
$\sigma_1(t)$	σ_1	0	0	$\cdots$	0
$\sigma_2(t)$	σ_2	σ_1	0	$\cdots$	0
$\sigma_3(t)$	σ_3	σ_2	σ_1	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	0
$\sigma_N(t)$	σ_N	σ_{N-1}	σ_{N-2}	$\cdots$	σ_1

terization reduces the number of parameters, nevertheless, this assumption is not suitable for accurate calibration of the model. The financial reason behind this fact appears to be the phenomenon of *mean reveresion of long term rates*. Unlike in Vasicek model, it is impossible to take this phnomenon into account by adding an *Ornstein-Uhlenbeck style model* since this would violate the arbitrage free condition of the model.

To overcome the hurdle, *volatility kernel* function $K(\tau, \lambda)$ can be introduced, which takes the form[6]:

$$K(\tau, \lambda) = \exp(-\lambda\tau). \tag{116}$$

For each T_j we choose a parameter λ_j and set

$$K_{j,i} = K(T_j - T_i, \lambda_j). \tag{117}$$

[6]Notice now the volatility function is not piecewise constant.

The structure assumed in this case in organized in table 3.

Table 3: Third Parameterization of $\sigma_j(t)$

$\sigma_j(t)$ / $t \in$	$[0, T_0)$	$[T_0, T_1)$	$[T_1, T_2)$	$\cdots$	$[T_{N-2}, T_{N-1})$
$\sigma_1(t)$	$\sigma_1 K_{1,0}$	0	0	$\cdots$	0
$\sigma_2(t)$	$\sigma_2 K_{2,0}$	$\sigma_1 K_{2,1}$	0	$\cdots$	0
$\sigma_3(t)$	$\sigma_3 K_{3,0}$	$\sigma_2 K_{3,1}$	$\sigma_1 K_{3,2}$	$\cdots$	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	0
$\sigma_N(t)$	$\sigma_N K_{N,0}$	$\sigma_{N-1} K_{N,1}$	$\sigma_{N-2} K_{N,2}$	$\cdots$	$\sigma_N K_{N,N-1}$

The structure of correlation matrix is another important input of LMM. We need to determine correlation, $\{\rho_{jk}\}_{0 \leq j,k \leq N-1}$, in a reasonable way. The dimensionality of ρ is $N(N-1)/2$ that is too high as calibration is considered. A convenient approach is to use a parameterized form of ρ_{jk}. An intuitive and flexible parameterization is given by the formula:

$$\rho_{jk} = \rho_\infty + (1 - \rho_\infty) \exp\left\{ -\frac{\lambda |T_i - T_j|}{1 + \kappa \min\{T_i, T_j\}} \right\} \tag{118}$$

where ρ_∞ is asymptotic level of correlations, λ is a decay rate of correlations, and κ is an asymmetry parameter. The parameters in this formula can be calibrated by using, for example, history data.

3.3 Calibration of Caps/Floors

For simplicity, let's choose log-normal LMM, which is actually a common choice. We want to fit our model into market quotes of caps/floors. A quick review of cap/pricing, the discounted payoff at time $t = 0$ of a cap with first reset date T_0 and payment dates $T = \{T_1, ..., T_N\}$ is given by:

$$\sum_{i=1}^{N} \delta_i D(0, T_i) \big(F_i(T_{i-1}) - K\big)^+ \tag{119}$$

Notice we consider unit notional. Thus, the fair price at $t = 0$ is:

$$\begin{aligned} Cap\big(0, T_0, \mathcal{T}, K\big) &= \sum_{i=1}^{N} \delta_i P(0, T_i) \mathbb{E}^{T_i} \big[\big(F_{i-1}(T_{i-1}) - K\big)^+ \big] \\ &= \sum_{i=1}^{N} \delta_i Cplt\big(0, F_i(0); T_{i-1}, T_i\big) \end{aligned} \tag{120}$$

Remark 3.1. *The correlation will have no effect on the expectation in* (120)*, indeed, the integral only implements on the marginal distribution.*

Let's pick one caplet that matures at T_i, $i = 1, ..., N$, the key is to compute

$$\mathbb{E}^{T_i}\big[\big(F_i(T_{i-1}) - K\big)^+\big] \tag{121}$$

with underlying process $F_i(t)$ under T_i-forward measure obeys,

$$dF_i(t) = \sigma_i(t)F_i(t)dW_i^{T_i}, \quad t \leq T_{i-1}. \tag{122}$$

If you have read the *equity model notes*, you probably immediately recognized it is the place to use BS time-dependent volatility formula. Therefore, the caplet price is can be represented as:

$$\begin{aligned} &Cplt^{LMM}(0, F_i(0); T_i - 1, T_i) = P(0, T_i)Bl(0, F_i(0), \sqrt{T_{i-1}}\sigma, K) \\ =&P(0, T_{i-1})\big(F_i(0)N(d_+(F_i(0), \nu_i, L)) - KN(d_-(F_i(0), \nu, K)\big) \end{aligned} \tag{123}$$

where

$$d_\pm(F, \nu, K) = \frac{\ln(F/K) \pm \nu^2/2}{\nu} \tag{124}$$

with

$$\nu_i^2 = T_{i-1}\nu_{T_{i-1}-cplt}^2 := T_{i-1} \times \big(\frac{1}{T_{i-1}}\int_0^{T_{i-1}} \sigma_i^2(t)dt\big). \tag{125}$$

The quantity $\nu_{T_{i-1}-cplt}$ is called T_{i-1}-*caplet volatility*, which is defined as the square root of the average percentage variance of the forward rate $F_i(t)$ for $t \in [0, T_{i-1})$.

Take one of those parameterization of volatility from above section, we have in general, we will have an explicit formula for calculating the caplet price. For instance, if we take table 2,

$$\nu_i^2 = \sum_{j=1}^{i} \tau_{j-1}\sigma_{i-j+1}^2$$

where $\tau_j = T_j - T_{j-1}$. Notice, in this case, σ should exactly fit the market Caplet volatilities quotes.

To match our forward rate dynamics to market cap quotes, let's notice the following relationship:

$$\begin{aligned} &\sum_{i=1}^{j} \tau_i P(0, T_i)Bl(0, F_i(0), \sqrt{T_{i-1}}\nu_{T_j-cap}) \\ =&\sum_{i=1}^{j} \tau_i P(0, T_i)Bl(0, F_i(0), \sqrt{T_{i-1}}\nu_{T_{i-1}-cplt}) \end{aligned}$$

The quantity $\nu_{T_{j-1}-cap}$ is called *forward volatilities* and $\nu_{T_{i-1}-cplt}$ is called *forward forward volatilities*. We discussed stripping method before, it allows us to get $\nu_{T_{i-1}-cplt}$. Consequently, we can find constraints for parameters of volatility function by matching caplet volatilities. If we assume table 2, all parameter can be determined and the calibration is accomplished. In general, however, we can only reduce the number of unknowns. To fully calibrate the parameters, we have to take Swaption into consideration as well.

3.4 Swaption Calibration

The calibration of Swaption under LIBOR market model is challenging because the pricing requires the joint distribution of a spanning of forward rates:

$$F_1(t), F_2(t), ..., F_N(t). \tag{126}$$

Because of this, the correlation structure ignored by caps/floors can be taken into consideration. In the following, we will introduce some simple approaches for Swaption pricing and comment on the advanced approaches in the end.

3.5 Modeling Forward Swap Rate(LSM)

To recall, the price of payer Swaption at time 0 is:

$$PS(0, \mathcal{T}, N = 1, K) = \mathbb{E}^{swap}\left[\left(S^{1,N}(T_0) - K\right)^+\right]\sum_{i=1}^{N}\delta_i P(T_0, T_i) \tag{127}$$

The underlying process here is the forward swap rate, which is a martingale under swap measure. If we assume log-normal dynamics of $S^{1,N}(t)$, i.e.,

$$dS^{1,N}(t) = \sigma^{1,N}(t)S^{1,N}(t)dW_t^{Swap}, \tag{128}$$

with deterministic function $\sigma^{1,N}(\cdot)$ represents the instantaneous percentage volatility.

We denote $\nu_{1,N}^2(T)$ the average percentage variance of the forward rate in the interval $[0, T_0]$ times the interval length, i.e.,

$$\nu_{1,N}^2(T_0) = \int_0^{T_0}\left(\sigma^{1,N}(t)\right)2dt. \tag{129}$$

As in the case of caps/floors, we are in a position to apply time-dependent-BS formula. In the language of Swaption, the model is known as *log-normal forward-swap model(LSM)*. Associating with the Black formula:

$$PS(0, T_0, \mathcal{T}, N = 1, K) = Bl(S^{1,N}(0), \nu_{1,N}(T_0), N, K)\sum_{i=1}^{N}\delta_i P(T_0, T_i) \tag{130}$$

Thus, the parameterization of $\sigma_{1,N}(\cdot)$ must match the market log-normal quotes of Swaption.

It may seem to be confusing because we migrate to a completely new model LSM while the calibration is for LIBOR forward model. We will see later on the connection between this two model. For now, we want to point out, however, the incompatibility. The following calculation is important:

$$\begin{aligned} S^{1,N}(t) &= \frac{P(t,T_0) - P(t,T_N)}{\sum_{i=1}^{n} P(t,T_i)} \\ &= \frac{1 - \prod_{j=1}^{N} \frac{1}{1+\delta_j F_j(t)}}{\sum_{i=1}^{N} \delta_i \prod_{j=1}^{i} \frac{1}{1+\tau_j F_j(t)}} \\ &:= \sum_{i=1}^{N} \omega_i F_i(t) \end{aligned} \tag{131}$$

where

$$\omega_i(t) = \frac{\frac{\delta_i}{1+\delta_i F_i(t)}}{\sum_{k=1^N} \frac{\delta_k}{1+\delta_k F_k(t)t}} = \frac{\tau_i P(t,T_i)}{\sum_{k=1}^{N} \delta_k P(t,T_k)} \tag{132}$$

The take-away from above derivation is that the forward swap rate can be viewed as a weighted average of forward rates. However, those weights are actually also function of forward rates. To make it tractable, we write the following approximation:

$$S^{1,N}(t) = \sum_{i=1}^{N} \omega_i(0) F_i(t). \tag{133}$$

The empricial studies showed that the variability of the ω's are smaller compared to the variability of F's, thus we can approximate ω's by their initial values $\omega(0)$.

If we directly models forward rate, we shall be able to use (133)(or, exact form) to find the dynamic of forward swap rate $S^{1,N}(t)$. Notice, we shall first switch to the swap measure to have a comparison. The LSM under swap measure has log-normal engine, while, the one we derive from forward rates are almost impossible to get log-normal distribution of the process. In practice, the distribution of swap rates under these two models are not that different, thus, LSM is still a practical model to use.

3.6 Monte Carlo Simulation

Let's write the pricing formula of payer Swaption in terms of T_0-forward measure:

$$PS(0,T_0,\mathcal{T},N=1,K) = P(0,T_0)\mathbb{E}^{T_0}\left[\left(S^{1,N}(T_0) - K\right)^+ \sum_{i=1}^{N} \delta_i P(T_0,T_i)\right] \tag{134}$$

Obviously, the expectation appears above depends on the *joint distribution* of all forward rates. For Monte-Carlo pricing, recall the forward rates under T_0-forward measure,

$$dF_k(t) = \sigma_k(t)F_k(t)\sum_{i=1}^{k}\frac{\rho_{k,j}\delta_j\sigma_j(t)F_j(t)}{1+\tau_j F_j(t)}dt + \sigma_k(t)F_k(t)dW_k^{T_0}(t) \tag{135}$$

where $k = \alpha + 1, ..., \beta$. The dynamics above allows us to generate sample paths for $F_1(T_0), F_2(T_0), ..., F_N(T_0)$. Therefore, the payoff can be evaluate for each element in the outcome space:

$$\left(S^{1,N}(T_0) - K\right)^+ \sum_{i=1}^{N}\delta_i P(T_0, T_i) \tag{136}$$

A simple averaging can lead to the present value. When calibrating, the volatility function in (135) has to be parameterized, the optimization operator is highly nonlinear because of the appearance of the Monte Carlo simulation. Nevertheless, a suitable choice of optimization method should yield decent optimal parameters.

3.7 LFM Formula for Swaption Volatilities

This is an approximation method without resorting to simulation. Let's recall the dynamic of forward swap rate under swap measure:

$$dS^{1,N}(t) = \sigma^{1,N}(t)S^{1,N}(t)dW_t^{Swap}, \quad t \in [0, T_0]. \tag{137}$$

Also, as we discussed previously, the Black equivalent volatility quoted on market is(multiplied by T_0):

$$\left(\nu_{1,N}(T_0)\right)^2 := \int_0^{T_0}\sigma_{1,N}^2(t)dt = \int_0^{T_0}(d\ln S^{1,N}(t))^2. \tag{138}$$

Our objective is represents $\nu_{1,N}(T_0)$ in terms of the volatilities appearing in LIBOR forward rate model.

The forward swap rates can be expressed as a weighted average of forward rates as in (131) with its approximation (133). Let's differentiate both sides of the approximation formula:

$$dS^{1,N}(t) \approx \sum_{i=1}^{N}\omega_i(0)dF_i(t) = (\cdots)dt + \sum_{i=1}^{N}\omega_i(0)\sigma_i(t)F_i(t)dW_i^{adj}(t) \tag{139}$$

Notice, the measure can be any forward measure, i.e., T_i-forward measure, for $i = 1, 2, 3, ..., N$. The drift term is omitted, because change of numérairé can results very messy drift term, but diffusion coefficients will not be affected. The approximated quadratic variation is:

$$dS^{1,N}(t)dS^{1,N}(t) \approx \sum_{i,j=1}^{N} \omega_i(0)\omega_j(0)F_i(t)F_j(t)\sigma_i(t)\sigma_j(t)dt. \tag{140}$$

The percentage quadratic variation is thus:

$$\left(\frac{dS^{1,N}(t)}{S^{1,N}(t)}\right)^2 = \left(d\ln S^{1,N}(t)\right)^2 = \frac{\sum_{i,j=1}^{N} \omega_i(0)\omega_j(0)F_i(t)F_j(t)\rho_{i,j}\sigma_i(t)\sigma_j(t)}{S^{1,N}(t)}dt \tag{141}$$

A further approximation is to freeze all F to their time-zero value:

$$\left(d\ln S^{1,N}(t)\right)^2 \approx \sum_{i,j=1}^{N} \frac{\omega_i(0)\omega_j(0)F_i(0)F_j(0)\rho_{i,j}}{S^{1,N}(0)^2}\sigma_i(t)\sigma_j(t)dt. \tag{142}$$

Using (142), we can now compute an approximation of the integrated percentage variance of S as:

$$\int_0^{T_0} \left(d\ln S^{1,N}(t)\right)^2 dt \approx \sum_{i,j=1}^{N} \frac{\omega_i(0)\omega_j(0)F_i(0)F_j(0)\rho_{i,j}}{S^{1,N}(0)^2} \int_0^{T_0} \sigma_i(t)\sigma_j(t)dt \tag{143}$$

This approximation formula is called *Rebonato's formula*, which uses volatility appearing in LIBOR forward rate model to match the Swaption Black volatility quotes. One may wonder, this is a coarse approximation because of over-simplification. But the truth is the quality is not poor at all, based on researchers' empirical test results.

3.8 Other Approximation Method

Because of the joint distribution of forward rates under LIBOR market model, exact solution of Swaption pricing is unavailable. Thus, main efforts were made on approximated analytical solution, as *Rebonato's formula* discussed above. The *low noise expansion* studied by *Andrew Lesniewski* provides an efficient approximation. The derivation is quite lengthy because of the perturbation technique adopted there. Apart from that, rank-r analytical methods proposed by *Brace* is also widely used among practitioners. Such methods require a root searching and eigen-value/vector analysis, we will not discuss in details but refer readers to *Brigo, D., Mercurio, F., "INterest Rate Models – Theory and Practice", Chapter 6.*