Derivations for Short-Rate Models

John A. Dodson∗

April 4, 2009

Abstract

Notes for Parts I-III, “No Arbitrage, Short Rate Models, & Market Models” of Brigo-Mercurio.

Contents

1 Black-Scholes  2

2 Affine Coefficients  2

2.1 Riccati equation  3
2.2 Time-homogeneous version  3
2.3 Deterministic shift extension  4
2.4 Conditional risk-neutral density  5

3 Gaussian Models  6

3.1 Conditional risk-neutral density  7
3.2 Multi-factor gaussian model  7
3.3 Deterministic shift extension  8
3.4 Multi-factor conditional risk-neutral density  9

4 Forward Rate Analogs  9

4.1 Affine forward rates  9
4.2 Gaussian forward rates  10
4.3 Drift condition  11

5 Pure-Discount Calls  11

5.1 Forward Measure  11
5.2 Single-Factor Affine Models  13
5.3 Single-Factor Gaussian Model  15

∗University of Minnesota Financial Math. program & The Options Clearing Corp. quantitative risk management
1 Black-Scholes

Assume that all bonds are perfectly correlated, so that any bond can be used to hedge any other bond. Consider \( P(t, T_1) \) and \( P(t, T_2) \) for \( T_1 \neq T_2 \). The value of the hedged portfolio is

\[
\Pi(t) = P(t, T_1) - \Delta \cdot P(t, T_2)
\]

The stochastic process for this value is

\[
d\Pi(t) = \left( P_t(t, T_1) + \frac{1}{2} \cdot \sigma^2(t, r) \cdot P_{rr}(t, T_1) \right) dt + P_r(t, T_1) \cdot dr - \Delta \cdot \left( P_t(t, T_2) + \frac{1}{2} \cdot \sigma^2(t, r) \cdot P_{rr}(t, T_2) \right) dt + P_r(t, T_2) \cdot dr
\]

where \( dr \, dr = \sigma^2(t, r) \cdot dt \).

No-arbitrage requires \( d\Pi = r \cdot \Pi dt \), so we have

\[
\Delta = \frac{P_r(t, T_1)}{P_r(t, T_2)}
\]

in order to eliminate the \( dr \) term, and

\[
\frac{P_t(t, T_1) + \frac{1}{2} \cdot \sigma^2(t, r) \cdot P_{rr}(t, T_1) - r \cdot P(t, T_1)}{P_t(t, T_2) + \frac{1}{2} \cdot \sigma^2(t, r) \cdot P_{rr}(t, T_2) - r \cdot P(t, T_2)} = \frac{P(t, T_1)}{P(t, T_2)}
\]

Since the LHS above depends only on the dynamics of bond 1, and the RHS above depends only on the dynamics of bond 2, both must equal some quantity \(-b(t, r)\) which depends on neither. Thus we have in general the Black-Scholes PDE for the single-factor risk-free discount factor,

\[
P_t(t, T) + b(t, r) \cdot P_r(t, T) + \frac{1}{2} \sigma^2(t, r) \cdot P_{rr}(t, T) = r \cdot P(t, T)
\]

(1)

which, along with the terminal condition \( P(T, T) = 1 \), determines \( P(t, T) \) for \( t \leq T \).

2 Affine Coefficients

A useful special case for the coefficients \( b \) and \( \sigma^2 \) are affine functions of the short-rate.

\[
b(t, r) \overset{\text{def}}{=} \eta(t) + \lambda(t) \cdot r
\]

\[
\sigma^2(t, r) \overset{\text{def}}{=} \delta(t) + \gamma(t) \cdot r
\]

In this case, we can simplify the Black-Scholes PDE using separation. In particular, let

\[
P(t, T) \overset{\text{def}}{=} A(t, T) \cdot e^{-r \cdot B(t, T)}
\]

with \( A(T, T) = 1 \) and \( B(T, T) = 0 \). Substituting these definitions into the Black-Scholes PDE (and dividing through by \( P \)), we get

\[
\frac{A_t(t, T)}{A(t, T)} - \eta(t) \cdot B(t, T) + \frac{1}{2} \cdot \delta(t) \cdot B^2(t, T) = r \cdot (1 + B_t(t, T) + \lambda(t) \cdot B(t, T) - \frac{1}{2} \cdot \gamma(t) \cdot B^2(t, T))
\]
Since this must hold for all \( r \), we must have that \( B \) satisfies the Riccati equation

\[
B_t(t, T) = -1 - \lambda(t) \cdot B(t, T) + \frac{1}{2} \cdot \gamma(t) \cdot B^2(t, T) \tag{2}
\]

and

\[
\frac{A_t(t, T)}{A(t, T)} = \eta(t) \cdot B(t, T) - \frac{1}{2} \cdot \delta(t) \cdot B^2(t, T) \tag{3}
\]

Setting aside the former for now, we notice that

\[
A(t, T) = e^{-\int_T^t \eta(t') \cdot B(t', T) - \frac{1}{2} \cdot \delta(t') \cdot B^2(t', T) \, dt'} \tag{4}
\]

The integral above can be simplified by the substitution \( u = B(t', T) \), since we know that \( \frac{du}{dt'} = B_t(t', T) = -1 - \lambda(t') \cdot u + \frac{1}{2} \cdot \gamma(t') \cdot u^2 \)

so

\[
\log A(t, T) = -\int_0^{B(t, T)} \frac{\eta \cdot u - \frac{1}{2} \cdot \delta \cdot u^2}{1 + \lambda \cdot u - \frac{1}{2} \cdot \gamma \cdot u^2} \, du
\]

and finally

\[
P(t, T) = \exp \left( -\int_0^{B(t, T)} r + \frac{\eta \cdot u - \frac{1}{2} \cdot \delta \cdot u^2}{1 + \lambda \cdot u - \frac{1}{2} \cdot \gamma \cdot u^2} \, du \right) \tag{5}
\]

### 2.1 Riccati equation

The non-linear Riccati ODE (2) can be reduced to a linear second-order ODE by the substitution

\[
B(t, T) \overset{\text{def.}}{=} -\frac{2}{\gamma(t)} \cdot \frac{\dot{v}(t)}{v(t)}
\]

where \( v \) solves

\[
\ddot{v} + \left( \lambda - \frac{\dot{\gamma}}{\gamma} \right) \cdot \dot{v} - \frac{\gamma}{2} \cdot v = 0 \tag{6}
\]

with \( v(T) \) arbitrary (non-zero) and \( \dot{v}(T) = 0 \).

### 2.2 Time-homogeneous version

For constant \( \lambda \) and \( \gamma \), (6) can be solved (e.g. using the Fourier technique), leading to

\[
B(t, T) = \frac{1}{\lambda} \cdot \frac{1}{\frac{1}{2} \cdot \left( \sqrt{1 + \frac{2}{\lambda^2}} - 1 \right) + \frac{1}{e^\lambda \cdot \sqrt{1 + \frac{2}{\lambda^2}} \cdot (T-t) - 1}} \tag{7}
\]

Furthermore, for constant \( \eta \) and \( \delta \), the integral (5) can be evaluated, leading to

\[
P(t, T) = \exp \left( \left( \frac{\delta}{\gamma} + \frac{\lambda^2}{\gamma} \cdot \left( \frac{\delta - \eta}{\lambda} \right) \right) \cdot (T-t) - \left( \frac{\delta}{\gamma} + r \right) \cdot B(t, T) - \left( \frac{\delta}{\gamma} - \frac{\eta}{\lambda} \right) \cdot \frac{\lambda}{\gamma} \cdot \log \left( 1 + \lambda \cdot B(t, T) - \frac{\gamma}{2} \cdot B^2(t, T) \right) \right) \tag{8}
\]
2.3 Deterministic shift extension

The time-homogeneous version of the affine model can accommodate a wide range of shapes for the current discount term structure, but this may not be sufficient in practice. There is a special case of the general affine model that can accommodate calibration of the initial term structure with arbitrary precision. Let

\[ \eta(t) = \eta + \dot{\varphi}(t) - \lambda \cdot \varphi(t) \]
\[ \delta(t) = \delta - \gamma \cdot \varphi(t) \]

for some arbitrary function of time \( \varphi(\cdot) \).

Since in general we have

\[ P(t, T) = e^{-r(t) \cdot B(t, T) - \int_t^T \eta(t') \cdot B(t', T) - \frac{1}{2} \delta(t') \cdot B^2(t', T) \, dt'} \]

and

\[ -\lambda \cdot B(t, T) + \frac{1}{2} \cdot \gamma \cdot B^2(t, T) = 1 + \dot{B}(t, T) \]

In the deterministic shift version, we get

\[ P(t, T) \overset{\text{def}}{=} P(t, T; r, \eta(t), \lambda, \delta(t), \gamma) \]
\[ = e^{-\int_t^T \varphi(t') \, dt'} \cdot P(t, T; r - \phi(t), \eta, \lambda, \delta, \gamma) \]  

(9)

This suggests that we can calibrate to an arbitrary \( P^M(0, T) \) by fitting a time-homogeneous affine model to a scaled version of the observed term structure.

\[ P(0, T; r(0) - \varphi(0), \eta, \lambda, \delta, \gamma) = e^{\int_0^T \varphi(t') \, dt'} \cdot P^M(0, T) \]

Subsequently, we get

\[ P(t, T) = \frac{P^M(0, T)}{P^M(0, t)} \cdot \frac{P(0, t; r(0) - \varphi(0), \eta, \lambda, \delta, \gamma)}{P(0, T; r(0) - \varphi(0), \eta, \lambda, \delta, \gamma)} \cdot P(t, T; r(t) - \varphi(t), \eta, \lambda, \delta, \gamma) \]  

(10)
2.4 Conditonal risk-neutral density

The stochastic process for the short rate under the risk-neutral measure is

\[ dr = b(t, r) \, dt + \sigma(t, r) \, dW \]

For the time-homogeneous affine model\(^1\), this is

\[ dr = (\eta + \lambda \cdot r) \, dt + \sqrt{\delta + \gamma \cdot r} \, dW \]  \hspace{1cm} (11)

which is clearly bounded below by \(-\frac{\delta}{\lambda}\) and mean-reverts (for \(\lambda < 0\)) to \(-\frac{\eta}{\lambda}\).

Consider the transformation

\[ x = \frac{4}{\gamma} \cdot \left( \frac{\delta}{\gamma} + r \right) \]

with the corresponding process

\[ dx = \left( 4 \cdot \frac{\lambda}{\gamma} \cdot \left( \frac{\eta}{\lambda} - \frac{\delta}{\gamma} \right) + \lambda \cdot x \right) \, dt + 2 \cdot \sqrt{x} \, dW \]

Assume for a moment that

\[ \nu \overset{\text{def.}}{=} 4 \cdot \frac{\lambda}{\gamma} \cdot \left( \frac{\eta}{\lambda} - \frac{\delta}{\gamma} \right) \]  \hspace{1cm} (12)

is a positive integer, and consider the \(\nu\)-dimensional process

\[ dy = \frac{\lambda}{2} \cdot y \, dt + dW \]  \hspace{1cm} (13)

with initial condition

\[ y|_{\mathcal{F}_t} = \sqrt{\frac{\delta + \gamma \cdot r}{\eta \cdot \gamma - \delta \cdot \lambda} \cdot 1} \]

It turns out that the process for \(\|y\|^2\) coincides with the process for \(x\). Furthermore, it is clear that increments of the components of \(y\) are i.i.d. normal, so it would seem that \(x\) is some version of chi-squared and \(r\) is in the same location-scale family.

In particular, we can integrate (13) to determine that

\[ y(T)|_{\mathcal{F}_t} \sim N \left( \frac{1}{e^{2 \cdot \lambda \cdot (T-t)}} \cdot \sqrt{\frac{\delta + \gamma \cdot r}{\eta \cdot \gamma - \delta \cdot \lambda}} \cdot 1, \frac{e^{\lambda \cdot (T-t)} - 1}{\lambda} \cdot I \right) \]

This gives a value

\[ c \overset{\text{def.}}{=} \mathbb{E}_t \left[ y(T) \right]^T \cdot \text{cov}_t \left[ y(T) \right]^{-1} \cdot \mathbb{E}_t \left[ y(T) \right] = \frac{4}{\gamma} \cdot \frac{\lambda}{1 - e^{-\lambda \cdot (T-t)}} \cdot \left( \frac{\gamma \cdot r}{\gamma} \right) \]  \hspace{1cm} (14)

for the non-centrality parameter. Finally, if \(\chi^2_\nu(c)\) has the distribution of a non-central chi-squared random variable, then

\[ r(T)|_{\mathcal{F}_t} \sim \frac{\gamma}{4} \cdot \frac{e^{\lambda \cdot (T-t)} - 1}{\lambda} \cdot \chi^2_\nu(c) - \frac{\delta}{\gamma} \]  \hspace{1cm} (15)

\(^1\)\(r \mapsto r - \varphi(t)\) for the deterministic shift version.
3 Gaussian Models

The special case of the affine model with $\gamma = 0$ can be generalized in a natural way to multiple factors, allowing us to model imperfect correlations between bonds. First, we need to re-cast the results from above (which depend on $\gamma \neq 0$) for the single-factor case. Consider first the simplified Riccati equation.

$$B_t(t, T) = -1 - \lambda(t) \cdot B(t, T)$$

The general solution to this is

$$B(t, T) = \int_t^T e^{\int_t^{t'} \lambda(t') \, dt'} \, dT'$$

and the time-homogeneous version is

$$B(t, T) = e^{\lambda(T-t)} - \frac{1}{\lambda}$$

We can evaluate the integral in (5) for $\gamma = 0$ to get the discount factor for the time-homogeneous case.

$$P(t, T) = \exp\left(\frac{\eta}{\lambda} + \frac{\delta}{2 \cdot \lambda^2} \cdot (T - t) - \left( r + \frac{\eta}{\lambda} + \frac{\delta}{2 \cdot \lambda^2} \right) \cdot B(t, T) + \frac{\delta}{2 \cdot \lambda^2} \cdot \frac{\lambda}{2} \cdot B^2(t, T) \right)$$ (16)

Similarly with the time-homogenous affine version, the discount rate

$$R(t, T) \triangleq -\frac{\log P(t, T)}{T - t}$$

is a linear combination of three functions of term, $T - t$. The scales of these functions are determined by $\lambda$; and $r, \eta, \text{ and } \delta$, uniquely determine the coefficients. In most cases, the term structure can evolve between upward-sloping, humped, and downward-sloping as the short-rate evolves. Notice also that the limiting long-term discount rate is fixed.
3.1 Conditional risk-neutral density

The risk-free short-rate under the risk-neutral measure for the time-homogeneous gaussian model follows an Ornstein-Uhlenbeck process,

\[ dr = (\eta + \lambda \cdot r) \, dt + \sqrt{\delta} \, dW \]

This is easily integrated².

\[
    r(T) | \mathcal{F}_t \sim N \left( r + (\eta + \lambda \cdot r) \cdot e^{\lambda (T-t)} - \frac{1}{\lambda}, \delta \cdot \frac{e^{2 \cdot \lambda (T-t)} - 1}{2 \cdot \lambda} \right) \tag{17}
\]

The general time-homogeneous affine model has shifted and scaled non-central chi-squared finite increments. With \( \gamma = 0 \), the finite increments are normal. The more general model is clearly superior in the single-factor setting. The advantage of the gaussian model emerges when we consider multiple factors.

3.2 Multi-factor gaussian model

In the single-factor setting, all bonds are perfectly correlated. This is unrealistic; and since a linear combination of normal random variables is normal, it makes sense to consider a multi-factor extension in the gaussian setting.

Let us assume that there are \( \nu \) factors driving the short rate

\[ r \overset{\text{def.}}{=} r_1 + \cdots + r_\nu \]

and furthermore that

\[
    dr_i \overset{\text{def.}}{=} b_i(t, r_i) \, dt + \sum_{j=1}^{\nu} l_{ij}(t) \, dW_j
\]

where \( W \) is a standard \( \nu \)-dimensional brownian motion under the risk-neutral measure. Note that the form of the brownian term allows for correlation amongst the factors; but the absence of any level dependence guarantees that increments are normal.

The corresponding drift of the discount factor is

\[
    P_i(t, T) + \sum_{i=1}^{\nu} b_i(t, r_i) \cdot P_{r_i}(t, T) + \frac{1}{2} \cdot \sum_{i,j,k=1}^{\nu} P_{r_i r_j}(t, T) \cdot l_{ij}(t) \cdot l_{ki}(t)
\]

This must equal \( r \cdot P(t, T) \). Letting \( \Sigma_{j,k}(t) = \sum_{i=1}^{\nu} l_{ji}(t) \cdot l_{ki}(t) \), we get a multi-factor version of the Black-Scholes PDE³.

\[
    P_i(t, T) + \sum_{i=1}^{\nu} b_i(t, r_i) \cdot P_{r_i}(t, T) + \frac{1}{2} \cdot \sum_{j,k=1}^{\nu} \Sigma_{j,k}(t) \cdot P_{r_j r_k}(t, T) = \sum_{i=1}^{\nu} r_i \cdot P(t, T) \tag{18}
\]

²Hint: \( d \left( e^{\lambda (T-t)} \cdot r \right) = \eta \cdot e^{\lambda (T-t)} \, dt + \sqrt{\delta} \cdot e^{2 \cdot \lambda (T-t)} \, dW \)

³The matrix \( \{l_{ij}(t)\} \) could be any rotation of the Cholesky decomposition of the matrix \( \{\Sigma_{ij}(t)\} \).
Let us continue to focus on the affine class. Let
\[ b_i(t, r_i) \overset{\text{def.}}{=} \eta_i(t) + \lambda_i(t) \cdot r_i \quad \forall i = 1, \ldots, \nu \]
and
\[ P(t, T) \overset{\text{def.}}{=} A(t, T) \cdot e^{-\sum_{i=1}^{\nu} r_i \cdot B^{(i)}(t, T)} \]
with \( A(T, T) = 1 \) and \( B^{(1)}(T, T) = \cdots = B^{(\nu)}(T, T) = 0 \). Substituting this form into the Black-Scholes PDE leads to the same simplified Riccati equations for the \( B \)'s,
\[ B^{(i)}(t, T) = \int_t^T e^{\int_{t'}^T \lambda_i(t') dt'} dT' \quad \forall i = 1, \ldots, \nu \]
and a similar result to (4) for \( A \).
\[ -\log A(t, T) = \int_t^T \sum_{i=1}^{\nu} \eta_i(t') \cdot B^{(i)}(t', T) \]
\[ \quad + \frac{1}{2} \cdot \sum_{i,j=1}^{\nu} \Sigma_{ij} \cdot B^{(i)}(t, T) \cdot B^{(j)}(t, T) \]
We can evaluate this in the time-homogeneous case.
\[ P(t, T) = \exp \left( \left( \sum_{i=1}^{\nu} \frac{\eta_i}{\lambda_i} + \frac{1}{2} \cdot \sum_{i,j=1}^{\nu} \frac{\Sigma_{ij}}{\lambda_i \cdot \lambda_j} \right) \cdot (T - t) \right) \]
\[ \quad - \sum_{i=1}^{\nu} \left( r_i + \frac{\eta_i}{\lambda_i} + \frac{1}{2} \cdot \sum_{j=1}^{\nu} \frac{\Sigma_{ij}}{\lambda_i + \lambda_j} \right) \cdot B^{(i)}(t, T) \]
\[ \quad + \frac{1}{2} \cdot \sum_{i,j=1}^{\nu} \frac{\Sigma_{ij}}{\lambda_i + \lambda_j} \cdot B^{(i)}(t, T) \cdot B^{(j)}(t, T) \]
(19)

3.3 Deterministic shift extension

Brigo-Mercurio point out that it is redundant to include all of the \( \eta \)'s in the specification. A more parsimonious and more flexible version is
\[ r \overset{\text{def.}}{=} \varphi(t) + \sum_{i=1}^{\nu} x_i \]
\[ dx_i \overset{\text{def.}}{=} \lambda_i \cdot x_i \cdot dt + \sum_{j=1}^{\nu} b_{ij} dW_j \quad \forall i = 1, \ldots, \nu \]
with \( \varphi(0) = r|_F_0 \) and all of the \( x \)'s initially zero. This provides
\[ P(t, T) = e^{-\int_t^T \varphi(t') dt'} \cdot \exp \left( \frac{1}{2} \cdot \sum_{i,j=1}^{\nu} \frac{\Sigma_{ij}}{\lambda_i \cdot \lambda_j} \cdot (T - t) \right) \]
\[ \quad - \sum_{i=1}^{\nu} \left( x_i + \frac{1}{\lambda_i} \cdot \sum_{j=1}^{\nu} \frac{\Sigma_{ij}}{\lambda_i + \lambda_j} \right) \cdot e^{\lambda_i \cdot (T - t) - 1} \]
\[ \quad + \frac{1}{2} \cdot \sum_{i,j=1}^{\nu} \frac{\Sigma_{ij}}{\lambda_i + \lambda_j} \cdot \frac{e^{\lambda_i \cdot (T - t) - 1}}{\lambda_i} \cdot \frac{e^{\lambda_j \cdot (T - t) - 1}}{\lambda_j} \]
(20)
where \( \varphi(\cdot) \) can be set to fit the initial term structure and the \( \frac{\nu(\nu+3)}{2} \) parameters to fit the initial volatility structure.

Note that Brigo-Mercurio in their treatment for \( \nu = 2 \) suggest that the \( \lambda \)'s and most of the off-diagonal \( \Sigma \)'s are negative.

### 3.4 Multi-factor conditional risk-neutral density

Finite increments of the factors under the risk-neutral spot measure are multinormal.

\[
x(T)|\mathcal{F}_t \sim N(\mu(T - t), \Sigma(T - t))
\]  \hfill (21)

where

\[
\mu_i(T - t) \overset{\text{def.}}{=} (x_i|\mathcal{F}_t) \cdot e^{\lambda_i (T - t)}
\]

\[
\Sigma_{ij}(T - t) \overset{\text{def.}}{=} \Sigma_{ij} \cdot \frac{e^{(\lambda_i + \lambda_j) (T - t) - 1}}{\lambda_i + \lambda_j} \quad \forall i, j = 1, \ldots, \nu
\]

### 4 Forward Rate Analogs

The Heath-Jarrow-Morton framework is not so much a model for interest rates as a perspective for modeling interest rates. Given the importance of the bank account process, the short rate is a natural place to start modeling dynamics. But the multifactor approach we have just concluded suggests that it might be enlightening to consider the dynamics of the whole term structure. Heath-Jarrow-Morton discovered that an elegant approach to whole-curve dynamics is the volatility of the (instantaneous) forward rate.

\[
f(t, T) \overset{\text{def.}}{=} -\frac{\partial}{\partial T} \log P(t, T)
\]

\[
df(t, T) \overset{\text{def.}}{=} \alpha(t, T) \, dt + \sigma(t, T)^T \, dW
\]  \hfill (22)

### 4.1 Affine forward rates

For the single-factor time-homogeneous affine model with deterministic shift extension, (9), the forward rate is

\[
f(t, T) = (r - \varphi(t)) \cdot (1 + \lambda \cdot B(t, T) - \frac{1}{2} \cdot \gamma \cdot B^2(t, T)) + \\
\varphi(T) + \eta \cdot B(t, T) - \frac{1}{2} \cdot \delta \cdot B^2(t, T)
\]  \hfill (23)

with \( B \) as defined in (7). The corresponding volatility is

\[
\text{var}_t[\text{df}(t, T)] = \text{var}_t[\text{df}(t, t)]
\]

\[
\left( \frac{\cosh \left( \coth^{-1} \left( \sqrt{1 + \frac{2}{\lambda^2}} \right) \right)}{\cosh \left( \coth^{-1} \left( \sqrt{1 + \frac{2}{\lambda^2}} \right) - \frac{1}{2} \cdot \sqrt{1 + \frac{2}{\lambda^2}} \cdot (T - t) \right)} \right)^2
\]  \hfill (24)

The volatility term structure for this model looks reasonable. We have already
pointed out a general problem with single-factor models, in that all rate are perfectly correlated. A deeper problem with this representation of the affine model is that forward rates are not markovian; rather they all depend on the level of \( r = f(t, t) \) which cannot be uniquely determined by \( T - t \) and the level \( f(t, T) \) for \( T > t \). This is a major impediment to an implementation in the HJM framework.

### 4.2 Gaussian forward rates

In contrast, the gaussian short-rate model does lead to a markovian specification for the forward rate dynamics. In fact, Brigo-Mercurio cite claims that any markovian forward rate model is equivalent to a gaussian short-rate model.

Let us consider the multi-factor time-homogeneous affine model with deterministic shift extension in (20).

\[
f(t, T) = \varphi(T) + \sum_{i=1}^{\nu} x_i \cdot e^{\lambda_i (T-t)}
\]

so

\[
df(t, T) = \sum_{i,j=1}^{\nu} \frac{e^{\lambda_i (T-t)} - 1}{\lambda_i} \cdot \Sigma_{ij} \cdot e^{\lambda_j (T-t)} - 1 \cdot dt
\]

\[
+ \sum_{i,j=1}^{\nu} e^{\lambda_i (T-t)} \cdot I_{i,j} dW_j
\]
This is clearly markovian: instantaneous increments in the forward rates are determined solely by the model parameters and the term $T - t$. The noise for the forward rate has a particularly simple form.

$$
\frac{\text{cov}_t [df(t, T_1), df(t, T_2)]}{dt} = \sum_{i,j=1}^{\nu} \Sigma_{ij} \cdot e^{\lambda_i (T_1 - t) + \lambda_j (T_2 - t)}
$$

(25)

### 4.3 Drift condition

In terms of the HJM form, (22), we have

$$
\sigma_i(t, T) = \sum_{j=1}^{\nu} e^{\lambda_j (T - t)} \cdot l_{ji}, \quad \forall i = 1, \ldots, \nu
$$

$$
\alpha(t, T) = \sum_{i,j=1}^{\nu} e^{\lambda_i (T - t)} \cdot \frac{1}{\lambda_i} \cdot \Sigma_{ij} \cdot e^{\lambda_j (T - t)}
$$

for this model.

The main HJM result is

$$
\alpha(t, T) = \sigma^T(t, T) \cdot \int_t^T \sigma(t, T') \, dT'
$$

(26)

which we can verify holds.

### 5 Pure-Discount Calls

#### 5.1 Forward Measure

The stochastic discount factor is a path-dependent quantity representing how much money one would need on deposit today to meet a unit liability in the future.
Figure 5: typical two-factor gaussian forward rate correlation surface

Since the bank balance $B(t)$ grows at the risk-free rate $r(t)$, we have

$$dB(t) = r(t) \cdot B(t) \, dt$$

and hence

$$D(t, T) = \frac{B(t)}{B(T)} = e^{-\int_t^T r(T') \, dT'}$$

The arbitrage-free value of a pure-discount bond is just the risk-neutral expected value of the stochastic discount factor:

$$P(t, T) = E_t D(t, T)$$

(27)

Similarly, the value of a European call with exercise date $T > t$ and strike price $X$ on an underlying pure-discount bond with maturity $S > T$ is

$$V_{call}(t, T, S, X) = E_t \left[ D(t, T) \cdot (E_T [D(T, S)] - X)^+ \right]$$

(28)

This expectation would seem to present challenges. We could abandon this effort and go back to the PDE approach from above; but a significant technical advance attributed to Geman provides a way forward.

If we change our unit of account (“numéraire”) to $P(t, T)$ and our risk-neutral probability measure from $\mathbb{Q}$ to $\mathbb{Q}^T$ with Radon-Nikodym derivative

$$\frac{d\mathbb{Q}_t^T}{d\mathbb{Q}_t} = \frac{B(t)}{B(T)} \cdot \frac{P(T, T)}{P(t, T)}
\begin{align*}
&= \frac{D(t, T)}{P(t, T)}
\end{align*}$$

(29)
we can re-write (28) as

\[ V_{call}(t, T, S, X) = P(t, T) \cdot \mathbb{E}_t^T \left( P(T, S) - X \right)^+ \]

The expectation may be manageable using this trick. The new challenge is to work out the density of \( P(t, S) \) under the new so-called forward measure.

Note that \( Q^T \) is termed the forward measure because in general

\[ E_t \left[ D(t, T) \cdot H(T) \right] = P(t, T) \cdot E_t^T H(T) \]

and hence

\[
E_t^T r(T) = \frac{1}{P(t, T)} \cdot E_t \left[ r(T) \cdot e^{-\int_t^T r(T') \, dt'} \right] \\
= \frac{1}{P(t, T)} \cdot E_t \left[ -\frac{\partial}{\partial T} e^{-\int_t^T r(T') \, dt'} \right] \\
= -\frac{1}{P(t, T)} \cdot \frac{\partial P(t, T)}{\partial T} \\
= f(t, T)
\]

That is, the forward rate for \( T \) is simply the expected value of the spot rate under the risk-neutral \( T \)-forward measure.

We can use Itô’s Lemma and the Black-Scholes PDE to determine the process for the underlying under the original measure.

\[
dP(t, S) = r \cdot P(t, S) \, dt + \sum_{i,k=1}^{\nu} P_{ri}(t, S) \cdot l_{ik} \, dW_k
\]

Under the original measure, the original (vector) brownian motion \( W \) is a martingale. Under the forward measure, it will have a drift. We can subtract off that drift from the brownian motion and add it to the drift for the underlying. In particular, from Girsanov’s theorem we can determine that

\[
dW_k = \sum_{j=1}^{\nu} \frac{P_{rj}(t, T)}{P(t, T)} \cdot l_{jk} \, dt + dW_k^T
\]

for \( k = 1, \ldots, \nu \) where \( W^T \) is a (vector) brownian motion under the \( T \)-forward measure. So

\[
\frac{dP(t, S)}{P(t, S)} = \left( r + \sum_{i,j=1}^{\nu} \frac{P_{ri}(t, S) \cdot \Sigma_{ij} \cdot P_{rj}(t, T)}{P(t, T)} \right) \, dt \\
+ \sum_{i,j=1}^{\nu} \frac{P_{ri}(t, S)}{P(t, T)} \cdot l_{ij} \, dW_j^T
\]

### 5.2 Single-Factor Affine Models

For a single-factor affine model, we have

\[ P(t, T) = A(t, T) \cdot e^{-r \cdot B(t, T)} \]
and
\[ dr = (\eta(t) + \lambda(t) \cdot r) \, dt + \sqrt{\delta(t) + \gamma(t) \cdot r} \, dW \]
under the original measure. Under the \( T \)-forward measure, we can use the definition (29) to find that the process is still affine, but transformed:
\[
\begin{align*}
\eta^T(t) &= \eta(t) - B(t, T) \cdot \delta(t) \\
\lambda^T(t) &= \lambda(t) - B(t, T) \cdot \gamma(t) \\
\delta^T(t) &= \delta(t) \\
\gamma^T(t) &= \gamma(t)
\end{align*}
\]
To evaluate
\[ V_{\text{call}}(t, T, S, X) = P(t, T) \cdot E_t^T \left( A(T, S) \cdot e^{-r(T) - B(T, S)} - X \right)^+ \tag{30} \]
we need the density of \( r(T) \) under the \( T \)-forward risk-neutral measure. Repeating the development from \( \S 2.4 \) for the conditional density of the short-rate under the time-homogeneous model, we can still work with
\[
x = 4 \cdot \frac{\lambda}{\gamma} \left( r + \frac{\delta}{\gamma} \right)
\]
and
\[
\nu = 4 \cdot \frac{\lambda}{\gamma} \left( \frac{\eta}{\lambda} - \frac{\delta}{\gamma} \right)
\]
but now the process for \( x \) under the forward measure is
\[
\frac{dx}{dt} = (\nu + (\lambda - \gamma \cdot B(t, T)) \cdot x) \, dt + 2 \cdot \sqrt{x} \, dW_T
\]
Similarly \( x = \sum_{i=1}^{\nu} y_i^2 \) where now
\[
dy_i = \frac{\lambda - \gamma \cdot B(t, T)}{2} \cdot y_i \, dt + dW_i^T \quad \text{for} \quad i = 1, \ldots, \nu
\]
so\(^5\)
\[
y_i(T)|\mathcal{F}_t \sim \mathcal{N} \left( \sqrt{-\frac{x}{\nu}} \cdot B_i(t, T), B(t, T) \right) \quad \text{for} \quad i = 1, \ldots, \nu
\]
and finally
\[ r(T)|\mathcal{F}_t \sim \frac{2}{\nu} \cdot B(t, T) \cdot \chi_{\nu}^2(c^T) - \frac{\delta}{\gamma} \]
under the \( T \)-forward measure, with non-centrality parameter
\[
c = \frac{4}{\gamma} \cdot \left( r + \frac{\delta}{\gamma} \right) \cdot \left( \frac{1}{B(t, T)} + \lambda - \frac{\gamma}{2} \cdot B(t, T) \right) \tag{31}
\]
We can use this density to evaluate the expectation in (30).
\[
V_{\text{call}}(t, T, S, X) = P(t, T) \cdot \int_0^\infty \left( A(T, S) \cdot e^{-B(T, S) \cdot \left( \frac{2}{\nu} \cdot B(t, T) \cdot x - \frac{\delta}{\gamma} \right)} - X \right) \cdot f_{\chi_{\nu}^2(c)}(x) \, dx
\]
\(^5\)Consider \( d \left( y_i \cdot e^{\frac{1}{2} B_i^T \cdot \chi_{\nu}^2(c^T)} \right) \).
where \( u \) represent at-the-money:

\[
\begin{align*}
  u &= \frac{\delta}{2} \cdot B(T, S) - \log \frac{X}{A(T, S)} \\
  &= \frac{\delta}{2} \cdot B(T, S) - \log X - \frac{\gamma}{4} \cdot B(T, S)
\end{align*}
\] (32)

The second term in the result is proportional to the CDF of a non-central chi-square. The first term is as well, once we apply the identity

\[
  e^{\frac{1-b}{2} \cdot x} \cdot f_{X^2}(c) (x) = e^{\frac{1-b}{2} \cdot c} \cdot b^{1 - \frac{c}{2}} \cdot f_{X^2}(c) (b \cdot x)
\]

for \( b > 0 \) with

\[
  b = 1 + \frac{\gamma}{2} \cdot B(t, T) \cdot B(T, S)
\] (33)

we can arrive at the classic solution for the value of pure-discount call under the affine model,

\[
  V_{\text{call}}(t, T, S, X) = P(t, S) \cdot F_{\chi^2}(\hat{\nu})(b \cdot u) - P(t, T) \cdot X \cdot F_{\chi^2}(\hat{\nu})(u)
\] (34)

5.3 Single-Factor Gaussian Model

The gaussian model is a special case of the affine model with \( \gamma = 0 \). Under the \( T \)-forward measure, the process for the short rate with time-homogeneous parameters is

\[
dr = (\eta + \lambda \cdot r - \delta \cdot B(t, T)) \ dt + \sqrt{\delta} \ dW_T
\]

as before, consider increments of the process

\[
r(t) \cdot e^{\lambda (T-t)}
\]

We can use this to evaluate the expectation in (30).

\[
  V_{\text{call}}(t, T, S, X) = P(t, T) \cdot \int_{-\infty}^{\log \frac{A(T, S)}{B(T, S)}} \left( A(T, S) \cdot e^{-B(T, S) \cdot x} - X \right) \cdot f_{r(T)}(x) \ dx
\]

As above, the relevant identity about the density is

\[
e^{-b \cdot x} \cdot f_{N(\mu, \sigma^2)}(x) = e^{-b \left( \mu - \frac{1}{2} \sigma^2 \cdot b \right)} \cdot f_{N(\mu, \sigma^2)}(x + \sigma^2 \cdot b)
\]

with \( b = B(T, S) \).

Finally, we get a result for pure-discount call under the Gaussian model exactly analogous to the Black-Scholes formula for a GBM asset:

\[
  V_{\text{call}}(t, T, S, X) = P(t, S) \cdot F_{\text{B}(0, 1)} \left( \frac{\log \frac{P(t, S)}{P(t, T)} \cdot X + \frac{1}{2} \sigma^2 \cdot (T - t)}{\sigma \cdot \sqrt{T - t}} \right) - P(t, T) \cdot X \cdot F_{\text{B}(0, 1)} \left( \frac{\log \frac{P(t, S)}{P(t, T)} \cdot X - \frac{1}{2} \sigma^2 \cdot (T - t)}{\sigma \cdot \sqrt{T - t}} \right)
\] (35)
with

\[ \sigma = B(T, S) \cdot \sqrt{\delta \cdot \frac{e^{2\lambda(T-t)} - 1}{2 \cdot \lambda \cdot (T - t)}} \]  

(36)