

Background Material: Gaussian and Quasi-Gaussian Models

Leif Andersen
Bank of America Merrill Lynch

First version: August, 2013

This version: October, 2014

Abstract

In Monte Carlo simulation algorithms for CVA and regulatory capital, it is convenient to work with an interest rate model that has at least three properties: a) the model should be fast and allow for easy pathwise construction of the interest rate forward curves; b) the model should be easy to parametrize, both with historical and market-implied data; and c) the model should allow for multiple factors. In this small note, we outline some relevant results for the Quasi-Gaussian model class, demonstrating that it satisfies all three criteria. Our focus is mainly on the convenient Gaussian sub-class of models, but the path to more complex models is given. We give several prescriptions for how the models can be parametrized in practice, both in isolation and in combination with other models (such as a credit intensity model).

1 One-Factor Quasi-Gaussian Model

1.1 Notation and HJM Basics

Let $P(t, T)$ be the time t discount bond price, for a maturity of $T \geq t$. Also, define instantaneous forwards $f(t, T)$ by

$$f(t, T) = \frac{-\partial \ln P(t, T)}{\partial T}$$

or, equivalently,

$$P(t, T) = \exp \left(- \int_t^T f(t, u) du \right).$$

The short rate $r(t)$ is defined as $f(t, t)$, and the money-market account $\beta(t)$ is

$$\beta(t) = \exp \left(\int_0^t r(u) du \right).$$

The risk-neutral measure corresponding to using $\beta(t)$ as the numeraire is denoted \mathbb{Q} .

Suppose that the evolution of $P(t, T)$ is driven by a single Brownian motion. In this case, by standard results, in \mathbb{Q} ,

$$dP(t, T)/P(t, T) = r(t) dt - \sigma_P(t, T) dW(t), \tag{1}$$

where $W(t)$ is a one-dimensional \mathbb{Q} -Brownian motion, and $\sigma_P(t, T)$ a T -indexed volatility process adapted to $W(t)$. The famous HJM result shows that (1) is equivalent to $f(t, T)$ dynamics in \mathbb{Q} of the form

$$df(t, T) = \sigma_P(t, T)\sigma_f(t, T) dt + \sigma_f(t, T) dW(t), \quad (2)$$

where

$$\sigma_f(t, T) = \frac{\partial \sigma_P(t, T)}{\partial T}$$

or

$$\sigma_P(t, T) = \int_t^T \sigma_f(t, u) du. \quad (3)$$

1.2 Separability Condition

A significant practical drawback of the specification (2) is that it typically leads to infinite-dimensional dynamics where one needs to simultaneously evolve all points on the forward curve continuum. To avoid this, we look for special cases where the entire forward curve dynamics can be captured in a few Markov state variables. A starting point is to specialize to the separable volatility case where

$$\sigma_f(t, T) = a(t)b(T) \quad (4)$$

for some stochastic process $a(t)$ and some deterministic function $b(T)$ of the forward rate maturity. Combining (2), (3) and (4) leads to

$$df(t, T) = a(t)^2 b(T) \int_t^T b(u) du dt + a(t)b(T) dW(t),$$

or, in integral form,

$$f(t, T) = f(0, T) + b(T) \int_0^t a(s)^2 \int_s^T b(u) du ds + b(T) \int_0^t a(s) dW(s). \quad (5)$$

Let us define

$$x(t) = b(t) \int_0^t a(s)^2 \int_s^T b(u) du ds + b(t) \int_0^t a(s) dW(s)$$

and notice that, by Leibniz' rule,

$$\begin{aligned} dx(t) &= b'(t) \int_0^t a(s)^2 \int_s^T b(u) du ds dt + b(t)^2 \int_0^t a(s)^2 ds dt \\ &\quad + b'(t) \int_0^t a(s) dW(s) dt + b(t)a(t) dW(t). \end{aligned}$$

Rearranging this expression leads to (show this!)

$$dx(t) = (y(t) - \kappa(t)x(t)) dt + a(t)b(t) dW(t),$$

where we have defined

$$y(t) = b(t)^2 \int_0^t a(u)^2 du, \quad \kappa(t) = -\frac{b'(t)}{b(t)}.$$

Obviously then also

$$dy(t) = 2b(t)b'(t) \int_0^t a(u)^2 du dt + b(t)^2 a(t)^2 dt = (b(t)^2 a(t)^2 - 2\kappa(t)y(t)) dt.$$

Importantly, knowledge of the x and y processes may be used to compute all forward rates, as

$$\begin{aligned} f(t, T) &= f(0, T) + x(t) \frac{b(T)}{b(t)} + b(T) \int_0^t a(s)^2 \int_s^T b(u) du ds - b(T) \int_0^t a(s)^2 \int_s^t b(u) du ds \\ &= f(0, T) + x(t) \frac{b(T)}{b(t)} + b(T) \int_t^T b(u) \int_0^t a(s)^2 ds \\ &= f(0, T) + x(t) \frac{b(T)}{b(t)} + y(t) \frac{b(T)}{b(t)^2} \int_t^T b(u) du. \end{aligned} \quad (6)$$

1.3 Markov Representation

While (6) has a very convenient form, it does *not* necessarily imply a low-dimensional Markov representation of the forward curve. For that, we need the process $x(t)$ and $y(t)$ to be jointly Markov. The key step here is to assume that

$$a(t)b(t) = \sigma_f(t, t) = q(t, x(t), y(t)), \quad (7)$$

for some deterministic function q . If this is the case, we see that

$$dx(t) = (y(t) - \kappa(t)x(t)) dt + q(t, x(t), y(t)) dW(t), \quad x(0) = 0, \quad (8)$$

$$dy(t) = (q(t, x(t), y(t))^2 - 2\kappa(t)y(t)) dt, \quad y(0) = 0, \quad (9)$$

which is a regular 2-dimensional Markov SDE. Notice that $y(t)$ has no diffusion term and is often said to be *locally deterministic*.

Above, we set $b'(t)/b(t) = -\kappa(t)$. If we, without loss of generality, scale b such that $b(0) = 1$, then it follows that

$$b(t) = \exp\left(-\int_0^t \kappa(u) du\right).$$

Using this relation in (6), we get a more explicit prescription for how to reconstitute the full forward curve at time t from knowledge of $x(t)$ and $y(t)$:

$$f(t, T) = f(0, T) + e^{-\int_t^T \kappa(u) du} \{x(t) + y(t)G(t, T)\}, \quad G(t, T) = \int_t^T e^{-\int_t^u \kappa(s) ds} du. \quad (10)$$

In discount bond space, we integrate and get (show this!)

$$P(t, T) = P(0, T)/P(0, t) \exp \left(-x(t)G(t, T) - \frac{1}{2}y(t)G(t, T)^2 \right). \quad (11)$$

Let us summarize. IF a) the forward rate volatility $\sigma_f(t, T)$ factorizes as in (4); AND b) the volatility $\sigma_f(t, t)$ of the short rate $r(t) = f(t, t)$ is a deterministic function of the variables x and y ; THEN the evolution of the forward curve is fully characterized by the 2-dimensional vector-SDE (8)-(9). Given the state of x and y at time t , we can reconstitute the entire forward and discount bond curves by the simple formulas (10) and (11). The resulting model framework is known as a *quasi-Gaussian model*.

In the system (8)-(9), it is clear that the x -variable is the primary driver of yield curve evolution: not only is it the only variable directly affected by the Brownian motion, we also have¹

$$r(t) = f(t, t) = f(0, t) + x(t), \quad (12)$$

i.e., $x(t)$ determines directly how the short rate moves relative to its time 0 forward value. The role of the y -variable is secondary as its presence in (10) is largely to enforce a “convexity”-correction to preclude arbitrages between short- and long-dated discount bonds.

1.4 Example: Gaussian Model with Constant Mean Reversion

As a useful (and very common) special case of the setup above, let us assume that the deterministic mean reversion speed $\kappa(t)$ defined above is just a positive constant κ . Also, let us assume that the volatility function $q(t, x, y)$ only depends on time t , $q(t, x, y) = \sigma_r(t)$. In this case, it follows that

$$G(t, T) = \int_t^T e^{-\kappa(u-t)} du = \frac{1}{\kappa} e^{\kappa t} [-e^{-\kappa u}]_t^T = \frac{1 - e^{-\kappa(T-t)}}{\kappa}.$$

Also, the y -process simplifies to

$$dy(t) = (\sigma_r(t)^2 - 2\kappa y(t)) dt$$

which is now fully (rather than just locally) deterministic. To solve for $y(t)$, set $z(t) = e^{2\kappa t} y(t)$, we see that

$$dz(t) = e^{2\kappa t} dy(t) + 2\kappa e^{2\kappa t} y(t) dt = \sigma_r(t) e^{2\kappa t} dt$$

¹This can be seen by setting $t = T$ in (10).

whereby $y(t)$ is

$$y(t) = e^{-2\kappa t} z(t) = \int_0^t \sigma_r(u)^2 e^{2\kappa(u-t)} du.$$

The process for $x(t)$ reduces to

$$dx(t) = (y(t) - \kappa x(t)) dt + \sigma_r(t) dW(t)$$

which is a Gaussian process of the Ornstein-Uhlenbeck type. This process is tractable, and the distribution of $x(t)$ is known in closed form. Specifically, for given $x(t)$ the distribution of $x(t + \delta)$ is Gaussian with mean

$$E(x(t + \delta)|x(t)) = e^{-\kappa\delta} x(t) + \int_t^{t+\delta} e^{-\kappa(t+\delta-s)} y(s) ds$$

and variance

$$\text{Var}(x(t + \delta)|x(t)) = \int_t^{t+\delta} \left(e^{-\kappa(t+\delta-s)} \sigma_r(s) \right)^2 ds.$$

These relations are useful in constructing a bias-free simulation scheme for paths of $x(t)$. Since $x(t)$ is Gaussian (and has no lower bound), we notice that there will inevitably be paths where forward rates become negative. This feature is actually not unrealistic these days, as negative spot and forward rates have been observed in several countries during the last few years.

1.5 Parametrization of Gaussian model

The model in Section 1.4 requires specification of two quantities: a) the mean reversion κ ; and b) the volatility function $\sigma_r(t)$. From (12), it is clear that

$$\begin{aligned} dr(t) &= \frac{\partial f(0, t)}{\partial t} dt + dx(t) = \frac{\partial f(0, t)}{\partial t} dt + (y(t) - \kappa x(t)) dt + \sigma_r(t) dW(t) \\ &= (\theta(t) - \kappa r(t)) dt + \sigma_r(t) dW(t) \end{aligned}$$

where

$$\theta(t) = \frac{\partial f(0, t)}{\partial t} + y(t) + \kappa f(0, t).$$

As such, the function $\sigma_r(t)$ can be identified as the volatility function for the short rate $r(t)$, i.e., as the volatility of the short end of the forward curve. Often this volatility is observable through option quotes on short-dated options. On the other hand, the choice of $q(t, x, y) = \sigma_r(t)$ and $b(T) = e^{-\kappa T}$ must (through (7)) mean that $a(t) = \sigma_r(t)e^{\kappa t}$ and therefore

$$\sigma_f(t, T) = a(t)b(T) = \sigma_r(t)e^{-\kappa(T-t)}.$$

In other words, the mean reversion κ controls the way volatility decays along the forward curve, a decay that is, for instance, observable in the way caplet volatilities decline as a function of maturity.

1.6 Other Models

Besides Gaussian models, a variety of quasi-Gaussian models are used in practice. A common class involves setting

$$q(t, x(t), y(t)) = \sigma(t)r(t)^\beta = \sigma(t)(x(t) + f(0, t))^\beta.$$

In this case the distribution of forward rates are not known in closed form, and any simulation scheme would generally involve a numerical discretization of (8)-(9). Notice that if $\beta = 0$, the model reduces to the one in Section 1.4 above.

2 Two-Factor Gaussian Models

One drawback of the model specification in Section 1 is the fact that the forward curve is driven only by a single Brownian motion; as a consequence, the forward curve will tend to move in largely parallel fashion. For instruments (or instrument portfolios²) that have sensitivity to yield curve twists, one-factor models are not realistic. While the quasi-Gaussian model above can be developed in full generality for an arbitrary number of Brownian motions, the notation gets somewhat unwieldy so we just refer to Chapter 13 in [1] for the results. For our needs here, we specialize to a two-factor Gaussian model³, a setup that, despite its apparent simplicity, still has a number of subtleties.

2.1 Statement of Model

First, let now $W(t) = (W_1(t), W_2(t))^\top$ be a two-dimensional Brownian motion; as always, the components W_1 and W_2 are assumed uncorrelated. The forward rate process in \mathbb{Q} is now

$$df(t, T) = \sigma_f(t, T)^\top \int_0^T \sigma_f(t, u) du + \sigma_f(t, T)^\top dW(t)$$

where $\sigma_f(t, T)$ is a two-dimensional vector process. To arrive at the two-factor Gaussian Markov model, set

$$\sigma_f(t, T) = a(t)b(T) \tag{13}$$

where $a(t)$ is a deterministic 2×2 matrix of the form⁴

$$a(t) = \begin{pmatrix} \sigma_{11}(t)e^{\kappa_1 t} & 0 \\ \sigma_{21}(t)e^{\kappa_1 t} & \sigma_{22}(t)e^{\kappa_2 t} \end{pmatrix}$$

and $b(T)$ is a two-dimensional vector of the form

$$b(T) = \begin{pmatrix} e^{-\kappa_1 T} \\ e^{-\kappa_2 T} \end{pmatrix}.$$

²A classical example is a short-term interest rate asset portfolio hedged with long-term loans.

³Two-factor Gaussian models have been specified in numerous different ways in the literature; see Chapter 12 in [1] for the details. Our specification here is, we feel, the most intuitive.

⁴We use a lower-diagonal matrix, as a Cholesky decomposition can always be employed to generate such a matrix.

Equation (13) is the 2-D equivalent of the 1-D condition (4) and can be expected to lead to a Markov representation of the forward curve. Cranking through the same steps as in Sections 1.4 leads to⁵

$$f(t, T) = f(0, T) + M(t, T)^\top (x(t) + y(t)G(t, T)) \quad (14)$$

where

$$M(t, T) = \begin{pmatrix} e^{-\kappa_1(T-t)} \\ e^{-\kappa_2(T-t)} \end{pmatrix}, \quad G(t, T) = \int_t^T M(t, u) du,$$

and $x(t) = (x_1(t), x_2(t))^\top$ satisfies

$$dx(t) = (y(t)\mathbf{1} - \kappa x(t)) dt + \sigma_x(t)^\top dW(t) \quad (15)$$

with $\mathbf{1}$ being the vector $(1, 1)^\top$ and

$$\kappa = \begin{pmatrix} \kappa_1 & 0 \\ 0 & \kappa_2 \end{pmatrix}, \quad \sigma_x(t) = \begin{pmatrix} \sigma_{11}(t) & 0 \\ \sigma_{21}(t) & \sigma_{22}(t) \end{pmatrix}.$$

In addition, $y(t)$ is here the 2×2 deterministic matrix

$$y(t) = B(t) \left(\int_0^t a(u)^\top a(u) du \right) B(t), \quad (16)$$

with

$$B(t) = \text{diag}(b(t)) = \begin{pmatrix} e^{-\kappa_1 t} & 0 \\ 0 & e^{-\kappa_2 t} \end{pmatrix}.$$

It is easy to verify that these equations will reproduce those of Section 1.4 of $\sigma_{21}(t) = \sigma_{22}(t) = 0$, i.e., when the 2-D model is specialized to a 1-D model.

The 2-D generalization of (12) is easily seen (from (14)) to be

$$r(t) = f(0, t) + x_1(t) + x_2(t) \quad (17)$$

which highlights the fact that the curve is now being pushed by two state-variables, x_1 and x_2 . We notice that these variables are generally correlated, as (15) implies that the correlation coefficient is

$$\rho_x(t) = \frac{\sigma_{21}(t)\sigma_{22}(t)}{\sqrt{\sigma_{11}(t)^2 + \sigma_{21}(t)^2} \sqrt{\sigma_{22}(t)^2}}. \quad (18)$$

⁵See Chapter 13 in [1].

2.2 Reparameterization

To emphasize the correlation structure of the 2-D model, we may write

$$dx(t) = (y(t)\mathbf{1} - \kappa x(t)) dt + \sigma_x^*(t)^\top dW^*(t) \quad (19)$$

where

$$\sigma_x^*(t) = \begin{pmatrix} \sqrt{\sigma_{11}(t)^2 + \sigma_{21}(t)^2} & 0 \\ 0 & |\sigma_{22}(t)| \end{pmatrix} \triangleq \begin{pmatrix} \sigma_1(t) & 0 \\ 0 & \sigma_2(t) \end{pmatrix}$$

and the elements of $W^*(t) = (W_1^*(t), W_2^*(t))^\top$ have instantaneous correlation of $\rho_x(t)$. This form of the x -dynamics is more intuitive and reduces the specification of the model to the five quantities $\rho_x(t)$, $\sigma_1(t)$, $\sigma_2(t)$, κ_1 , κ_2 . As in Section 1.5, we wish to link these quantities to statistics of the forward curve. For this, we note from (14) and (19) that

$$df(t, T) = O(dt) + \begin{pmatrix} \sigma_1(t)e^{-\kappa_1(T-t)} \\ \sigma_2(t)e^{-\kappa_2(T-t)} \end{pmatrix}^\top dW^*(t) \quad (20)$$

whereby we can easily demonstrate that

$$\text{Var}_t(df(t, T)) = \sigma_1(t)^2 e^{-2\kappa_1(T-t)} b(t, T, T), \quad (21)$$

$$\text{corr}(df(t, T_1), df(t, T_2)) = \rho(t, T_1, T_2) = \frac{b(t, T_1, T_2)}{\sqrt{b(t, T_1, T_1)b(t, T_2, T_2)}}, \quad (22)$$

with

$$b(t, T_1, T_2) = 1 + \rho_x(t) \frac{\sigma_2(t)}{\sigma_1(t)} \left(e^{-(\kappa_2 - \kappa_1)(T_1 - t)} + e^{-(\kappa_2 - \kappa_1)(T_2 - t)} \right) + \left(\frac{\sigma_2(t)}{\sigma_1(t)} \right)^2 e^{-(\kappa_2 - \kappa_1)(T_1 + T_2 - 2t)}.$$

2.3 Practical Parameterization

In practice, we would normally insist that the correlation structure in (22) is stationary, in the sense that

$$\text{corr}(df(t, T_1), df(t, T_2)) = g(T_1 - t, T_2 - t)$$

for some function g . Evidently this requires that $\rho_x(t)$ as well as the ratio $\sigma_2(t)/\sigma_1(t)$ are set to constants, whereby we ultimately have only one free time-dependent function to specify, along with the four constants κ_1 , κ_2 , ρ_x , and $v = \sigma_2(t)/\sigma_1(t)$. We notice that in this case the correlation between the short rate $r(t)$ and the “perpetual” forward $f(t, \infty)$ becomes (assuming, without loss of generality, that $\kappa_1 < \kappa_2$)

$$\rho_\infty = \rho(t, t, \infty) = \frac{1 + \rho_x v}{\sqrt{1 + 2\rho_x v + v^2}}, \quad (23)$$

which is independent of mean reversions. Empirical examination of the correlation between short and long rates can often be used to establish $\rho(t, t, \infty)$, which in turn will help us estimate v and/or ρ_x .

In choosing mean reversions κ_1 and κ_2 , it is common to set $\kappa_1 = 0$ and $\kappa_2 > 0$ which ensures that the (very) long end of the forward curve $f(t, \infty)$ is non-random⁶. In particular, we see from (20) that when $\kappa_1 = 0$ and T is large,

$$df(t, T) \approx O(dt) + \sigma_1(t)dW_1^*(t) = O(dt) + dx_1(t). \quad (24)$$

Notice that $x_1(t)$ can thus be identified directly as the driver of the long end of the forward curve, such that

$$\sigma_l(t) = \sigma_1(t)$$

will effectively be the volatility of long-dated forward rates.

Finally, observe that the short rate process becomes

$$dr(t) = O(dt) + \sigma_1(t)dW_1^*(t) + \sigma_2(t)dW_2^*(t)$$

which implies that the total short rate volatility is

$$\sigma_r(t) = \sqrt{\sigma_1(t)^2 + \sigma_2(t)^2 + 2\rho_x\sigma_1(t)\sigma_2(t)} = \sigma_l(t)\sqrt{1 + 2\rho_x v + v^2} = \sigma_l(t)\frac{1 + \rho_x v}{\rho_\infty}.$$

In particular, we find that the ratio of short and long forward rate volatilities is

$$c = \frac{\sigma_l(t)}{\sigma_r(t)} = \frac{\rho_\infty}{1 + \rho_x v},$$

or

$$\rho_x v = \frac{\rho_\infty}{c} - 1. \quad (25)$$

Inserting this into (23) yields

$$v = \sqrt{c^{-2} - 1 - 2\left(\frac{\rho_\infty}{c} - 1\right)}. \quad (26)$$

Based on the results above, we can use the following model parameterization strategy:

1. Estimate empirically the correlation $\rho_\infty > 0$ between the short and the long end of the forward curve.
2. Estimate empirically the ratio $c < 1$ of short and long forward rate volatilities.
3. Specify the short-term volatility term structure $\sigma_r(t)$ to match short term option prices.
4. Specify a mean reversion κ_2 ; set $\kappa_1 = 0$.
5. Compute v from (26) and ρ_x from (25); set $\sigma_l(t) = c\sigma_r(t)$.

⁶If both mean reversions are positive, all volatility will decay away in the large- T limit.

6. Set $\sigma_1(t) = \sigma_l(t)$ and set $\sigma_2(t) = v\sigma_1(t)$.

In Step 4, determination of a meaningful mean reversion parameter κ_2 can be done in numerous ways. An easy approach is to use this parameter to match as well as possible the decay in volatility from $\sigma_r(t)$ to $\sigma_l(t)$ as the forward rate maturity is increased. For this we can rely on the following (from (21))

$$\text{Var}_t(df(t, T)) = \sigma_r(t)^2 \frac{1 + v^2 e^{-2\kappa_2(T-t)} + 2\rho_x v e^{-\kappa_2(T-t)}}{1 + v^2 + 2\rho_x v}. \quad (27)$$

2.4 Usage for Hybrids

As formulated in Section 2.1 above, the 2-D Gaussian model is characterized by the two (Markov) Gaussian variables $x_1(t)$ and $x_2(t)$. We have shown that the correlation between these drivers (and their volatilities) determine the correlation between long and short forward rates on the forward rate curve. Suppose now, however, that an additional financial variable $Z(t)$ is introduced and needs to be correlated with interest rates. For this, it is normally easier to link x_1 and x_2 to specific points on the forward curve, which then may be correlated to $Z(t)$ directly.

To illustrate how this idea can work, assume as in Section 2.3 that $\kappa_1 = 0$ (but $\kappa_2 > 0$). Suppose also that we have estimated a correlation of ρ_l between Z and the long end of the forward curve (represented by $f(t, \infty)$). According to (24), this simply requires that Z is correlated to $x_1(t)$ with a correlation coefficient of

$$\rho_{Z,1} = \rho_l.$$

If, in addition, we have measured that Z is correlated to short-term rates (represented by $r(t)$) with a correlation coefficient of ρ_s , then (17) shows that we need the correlation between Z and $x_1(t) + x_2(t)$ to equal ρ_s . If $\sigma_Z(t)$ is the volatility of Z , we therefore need the correlation $\rho_{Z,2}$ between Z and x_2 to satisfy

$$\rho_s \sigma_Z(t) \sigma_r(t) = \rho_{Z,1} \sigma_Z(t) \sigma_1(t) + \rho_{Z,2} \sigma_Z(t) \sigma_2(t)$$

or

$$\begin{aligned} \rho_{Z,2} &= \frac{\rho_s \sqrt{\sigma_1(t)^2 + \sigma_2(t)^2 + 2\rho_x \sigma_1(t) \sigma_2(t)} - \rho_{Z,1} \sigma_1(t)}{\sigma_2(t)} \\ &= \rho_s \sqrt{v^2 + 1 + 2\rho_x v} - \rho_{Z,1} v \end{aligned}$$

where ρ_x and $v = \sigma_2(t)/\sigma_1(t)$ were defined earlier. This completes the correlation specification. As always, a check for positive definiteness of the total correlation matrix (for x_1 , x_2 , and Z) is required.

References

- [1] Andersen, L. and V. Piterbarg (2010), *Interest Rate Modeling*, Atlantic Financial Press.