

Efficient Credit Risk Simulation by Optimal Mean-Reversion Adjustment

HENRY SCHELLHORN*

October 21, 2002

ABSTRACT. Two different probability measures are of importance when calculating the risk of a large portfolio: the risk-neutral measure for pricing, and the real measure to project true earnings. When using Monte Carlo, the natural method is to conduct two different simulations, one in each probability measure. We advocate to simulate portfolios only in the true measure, and, for pricing, to adjust the results by the Radon-Nikodym derivative. This scheme leads to simulations that are up to twice faster and whose results are easier to interpret. Our results can also be useful in the calibration of the BGM/J Libor model. It is well-known that calibrating this model to caps and swaptions results in a dispersion of risk-neutral rates that is much higher than the plausible dispersion of real rates. Our scheme uses mean-reversion to correct the real rate distribution while having the advantage of preserving the volatility structure coming from the calibration process. In order to solve this problem we determined a new expression for the variance of the Radon-Nikodym derivative of Gaussian measures, which could be useful in other areas of finance.

1. INTRODUCTION

Two different probability measures are of importance when calculating the risk of a large portfolio: the measure used for pricing - be it risk-neutral, forward,..., and the real measure to project true earnings. We consider the case when the logarithm of the state variables, forward rate and credit spreads, are Gaussian in both measures, which is important in practical applications. This article focuses on the BGM/J Libor model. For that model, forward rates are approximately lognormal in the measure used for pricing, namely the forward measure; normality is a key advantage for a successful and intuitive calibration to caps and swaptions [3][4]. Likewise, a joint normal distribution in the true measure of the logarithm of forward rates is much easier to interpret than any other more sophisticated and statistically more correct distribution.

We note in passing that, whereas a standard methodology arose in inferring the forward measure, there is less of a consensus to infer the real measure. Although the academic literature favours to first infer the real measure and then adjust it with a market price of risk to obtain the forward measure, information often flows the other way round in practice. In many bank departments the key requirement is to do a correct pricing. To this effect, the forward measure is calibrated first, to market prices and/or historical correlations. To calculate earnings one or several (for what-if analysis) specific forms of the true measure are then derived from the forward measure. If rates or spreads distributions in the forward measure are deemed "plausible", the real measure is sometimes chosen to be the same as the forward measure. Whereas theoretically incorrect, this offers the advantage of facilitating communication with upper management, circumventing the need

*HEC, University of Lausanne, CH-1015 Lausanne, Switzerland. Tel + 41 21 692 33 52, fax + 41 21 692 34 35. Henry.Schellhorn@hec.unil.ch

to explain why, along a particular scenario, rates are different when calculating Value-at-Risk or Earnings-at-Risk.

When using Monte Carlo, the natural method to analyze risk is to conduct two different simulations if not calibrations, one in each probability measure. We advocate in this paper to simulate portfolios only in the true measure, and, for pricing, to adjust the results by the Radon-Nikodym derivative of the forward measure with respect to the true measure; the forward measure is calibrated first, and then the true measure, in such a way that the variance of the Radon-Nikodym derivative is minimal. This scheme offers four distinct advantages, as we explain below. The first one justifies conducting only one simulation, and the second one justifies using two measures.

First, simulations can be up to twice faster, in the case of "what-if analysis" the improvement in speed is much more important. This was the first reason why the author undertook this project: speed is a key issue in many Asset/Liability Management simulations. When the whole portfolio is simulated, the time required to calculate the cash flows for all instruments completely dwarfs the time required to generate rate forecasts [7]. However appealing theoretically, this approach is delicate to implement. The Radon-Nikodym derivative tends to grow rapidly, resulting in a higher variance of the measure of interest, and a degraded accuracy of Monte Carlo simulation. We will show in the next section how our adjustment is "optimal" in the sense that it minimizes the variance of the Radon-Nikodym derivative, thereby enhancing the accuracy of our simulations.

Second, as we will explain in more details later, separating the two measures has a clear advantage for calibration. To take a concrete example, it is well-known that calibrating the BGM/Libor model to caps and swaptions results in a dispersion of rates forecast, in the forward measure, that is much higher than the plausible dispersion of real rates in the US, because of the high skewness of the lognormal distribution. This is the reason why alternate models like [2] were designed to prevent a too rapid increase in risk-neutral rates; while widely used, for instance at Bank of America in the 1990s [8] this model is less practical to calibrate to caps and swaptions than the BGM/Libor. Again, we advocate to calibrate the forward measure first, and then to derive the true measure by a mean reversion adjustment, such that rates will still be lognormal in the true measure, but with a smaller dispersion. The volatility structure is chosen to be the same in both measures - note that this has the further advantage to reduce the excessively large number of degrees of freedom when calibrating the true measure. To summarize, the information we need to infer both measures is:

- caps and swaption prices
- credit spreads dynamics (in the risk-neutral measure)
- the a priori knowledge that future rates/spreads are lognormal in the real measure
- an educated guess about the maximal plausible value of the rate/ spread variance in the true measure.

No extra knowledge of the market price of risk, often difficult to obtain, is required, while the whole information from the market is completely integrated.

Third, results are easier to interpret in our scheme, because both measures are tightly linked by the minimization of the variance of the Radon-Nikodym, which acts as a measure of "distance" between two measures. Let us suppose we have 2 scenarios, H and L

resulting in future (say one year from now) rates in the forward measure, of 15% and 5%. When adding mean-reversion, the future rates in the real measure could be 12% in the H scenario and 8% in the L scenario. The point is that the H scenarios yield higher rate levels in both measures than the L scenarios. This is of particular interest in an ad hoc analysis, when practitioners isolate particular scenarios after a Monte Carlo run to better understand the results: a high rate scenario for pricing is a high rate scenario for earning analysis, and this across different horizons.

Fourth, there is a quasi-analytical solution for the optimal speed of mean-reversion in our scheme. By "quasi-analytical", we mean that the optimal solution is an analytical function of only one parameter. The value of this parameter solves a nonlinear equation, which is very easy to do numerically.

Finally, our lemma, albeit modest in scope, seems to be an unknown theoretical result. Whereas much attention has been given to the distribution of exponentials of functionals of Brownian motion [9] in the context of Asian option pricing, little has been done about the moments of a Radon-Nikodym derivative of one Gaussian measure with respect to another one. We suspect our result could be applied to the entropy minimization problems that are currently being studied in mathematical finance [6].

2. FORMULATION OF THE PROBLEM

We first remind the reader how the state variables of a BGM/J model with credit risk can be approximated by an Ornstein-Uhlenbeck process. We then expose the general simulation problem, irrespectively of the particular model for the state variables.

2.1. A BGM/J Model with Credit Risk. The BGM/J model of risk-free interest rates can be found in many textbooks [3][4]. For definiteness, we choose a model of $i = 1..m$ discrete forward rates f_i with initial values $f_{i,0}$. Each rates represents the rate of a discount bound that starts at time T_{i-1} and ends at time T_i . These forward rates therefore span the period starting now (at T_0) and ending at T_m . For simplicity, we suppose that $T_{i-1} - T_i = \tau$. We let W^F be n -dimensional Brownian motion in the forward measure, i.e., the measure where the $i + 1$ -th discount bound is a martingale between time T_{i-1} and T_i . Let $\sigma_i(t)$ be the deterministic volatility at time t of f_i , and $\rho_{ik}(t)$ be the deterministic correlation between forward rates i and k . Since usually $n < m$, the values of ρ_{ik} are completely determined by the values of coefficients b_{ij} for $j = 1..m$, which are such that:

$$\sum_{j=1}^d b_{ij}^2 = 1 \quad (1)$$

The system of SDEs for the risk-free forward rate is then.

$$\frac{df_i}{f_i} = -\sigma_i \sum_{k=i+1}^m \frac{\sigma_k \rho_{ik} f_k \tau}{1 + f_k \tau} dt + \sigma_i \sum_{j=1}^d b_{ij} dW_j^F \quad (2)$$

$$f_i(0) = f_{i,0} \quad (3)$$

When resorting to Monte Carlo simulation, this system needs to be discretized. To avoid cluttering the paper with new notation, we develop our results only in continuous time. The reader interested in discretization schemes is referred to [3].

A reduced-form model is applied to analyze credit risk. We have $l = 1..L$ different credit ratings. Let $F_{i,0}^l$ be the initial forward curve for rating l . The logarithm of the forward spread s_i^l (which corresponds to rating l and period $[T_{i-1}, T_i]$) follows an Ornstein-Uhlenbeck process, that is:

$$\frac{ds_i^l}{s_i^l} = (\beta_i^l - \alpha_i^l s_i^l)dt + \sigma_i^l \sum_{j=1}^d b_{ij}^l dW_j^F \quad (4)$$

$$s_i^l(0) = F_{i,0}^l - f_{i,0} \quad (5)$$

where $\alpha_i^l, \beta_i^l, \sigma_i^l, b_{ij}^l$ are at most functions of time.

Whereas the logarithm of the spread is Gaussian, the logarithm of the risk-free forward rate is not quite Gaussian in the forward rate. However, the drift term is of proportional to σ^2 , against σ for the diffusion term, so with a fine maturity discretization (τ small) and normal volatility and forward rate levels, the drift term in (2) is negligible. This fact has been well documented in [3][4]. From now on we assume that this drift is zero, that is, the logarithm of the risk-free forward rate is Gaussian.

We suppose that the forward measure has already been inferred, that is, all parameters above have been calibrated. Let:

$$x = \begin{bmatrix} \log f \\ \log s^1 \\ \dots \\ \log s^L \end{bmatrix} \quad (6)$$

$$b^F = \begin{bmatrix} 0 \\ \beta^1 \\ \dots \\ \beta^L \end{bmatrix} \quad (7)$$

$$A^F = \begin{bmatrix} 0 & \dots & \dots & 0 \\ \dots & -A^1 & 0 & 0 \\ \dots & 0 & \dots & 0 \\ 0 & 0 & 0 & -A^L \end{bmatrix} \quad (8)$$

$$\Sigma = \begin{bmatrix} \sigma \otimes b_{..1} & \dots & \dots & \sigma \otimes b_{.d} \\ \sigma^1 \otimes b_{..1}^1 & \dots & \dots & \sigma^1 \otimes b_{.d}^1 \\ \dots & \dots & \dots & \dots \\ \sigma^d \otimes b_{..1}^d & \dots & \dots & \sigma^d \otimes b_{.d}^d \end{bmatrix} \quad (9)$$

where \otimes is the term-by-term Kronecker product of two vectors, and A^l is the diagonal matrix, with α_i^l in the diagonal. Our system is then

$$dx = (b^F - A^F x)dt + \Sigma dW^F \quad (10)$$

where Σ is a matrix with N rows and d columns.

2.2. The Simulation Problem. Let $\varsigma = \{T_1, \dots, T_I\}$ be the collection of cash flow dates of our portfolio, with T the last cash flow date. Let $C : C^2 \otimes \varsigma \rightarrow R$ be the cash flow mapping. In other terms, $C(x, T_i)$ is the cash flow occurring at time T_i , and, for fixed time, $C(\cdot, T_i)$ is a functional of our state variable x . Let $D : C^2 \otimes R \rightarrow R$ be the discount factor functional in the forward measure, i.e., $D(x, t)$ is the discount factor at time t . The price of the portfolio is:

$$V = E^F \left[\sum_{T_i \in \varsigma} D(x, T_i) C(x, T_i) \right] \quad (11)$$

where E^F is expectation in the forward measure at time 0. Risk managers are interested in knowing not only V but also the distribution:

$$P(C(x, t) \leq c) \quad (12)$$

for various times in the true measure of the cash flows C at various dates t . Our notation is the following: expectations operators, probabilities, and Brownian motion have a subscript F when in the forward measure, and no subscript at all in the true measure. A common form of "what if" analysis is to calculate (12) under different volatility assumptions. Quite loosely, let σ be some measure of overall state variable volatility. For each value of σ we are interested in calculating:

$$P^\sigma(C(x, t) \leq c) \quad (13)$$

where P^σ is the true probability measure when volatility is σ . For simplicity, we will

not mention what if analysis anymore to focus on the case of only 2 measures, but the reader should keep in mind that our results can be applied to more than 2 measures.

For large portfolios it is [7] much more time consuming to calculate C than x . As explained in the introduction, it is therefore advantageous to simulate C only in one measure and then calculate (11) (13) with a change a measure. The Radon-Nikodym derivative tends to grow rapidly, therefore results from a Monte Carlo simulation will be less accurate in the measure where results are adjusted by the Radon-Nikodym derivative. Also the additional error our scheme brings is more important when evaluating the tails of the true distribution than the middle. Since risk managers are more interested in accuracy in the tails of a distribution than in the middle, we recommend to simulate the true measure first, and then adjust the results for pricing. Pricing will be less accurate, but variance reduction techniques can be used to improve pricing, whereas they are hard to use when estimating (12)¹. The Radon-Nikodym derivative of the forward measure with respect to the true measure is written:

$$g(\omega, t) = E_t \left[\frac{dP^F}{dP} \right] \quad (14)$$

To summarize, the basic algorithm is the following.

1. Simulate $x(\omega)$ in the true measure, for scenario $\omega = 1.. \Omega$
2. Calculate the cash flow $C(x(\omega), T_i)$ for $\omega = 1.. \Omega, T_i \in \varsigma$

¹From the viewpoint of Monte Carlo, estimating the probability P consists in estimating a collection of expectations of an indicator function, one for each quantile c . Variance reduction techniques typically aim at optimizing simulations for only one of them

3. Calculate the discount factor $D(x(\omega), T_i)$ for $\omega = 1.. \Omega, T_i \in \varsigma$
4. Estimate (12) by sorting the results from step 2 for various times.
4. Estimate V by:

$$V_{\text{est}} = \frac{1}{\Omega} \sum_{\omega=1}^{\Omega} \sum_{T_i \in \varsigma} g(x(\omega), T_i) D(x(\omega), T_i) C(x(\omega), T_i) \quad (15)$$

The problem is now to infer the drift of our state variables in the real measure, since the volatility is assumed to be the same in both measures. We restrict our attention to drifts that are functions of the state variables, such that the resulting process is Gaussian in the true measure. We are then limited to determine the parameters A, b of the Ornstein-Uhlenbeck process:

$$dx = (b - Ax)dt + \Sigma dW \quad (16)$$

Often, the best predictor of future state variables in finance is their current value, so we take $b = 0$. Also, we now assume $b^F = 0, A^F = 0$. Taking different values would complicate only slightly the following calculations. In general Σ has more rows than columns, therefore to keep absolute continuity across measures A must be of the form:

$$A = \Sigma R \quad (17)$$

for some free matrix R . Therefore

$$dx = -\Sigma R x dt + \Sigma dW \quad (18)$$

Let:

$$\lambda = Rx. \quad (19)$$

Then:

$$g(t) = e^{\int_{s=0}^t \lambda(s) dW(s) - \frac{1}{2} \int_{s=0}^t |\lambda(s)|^2 ds} \quad (20)$$

The number of Monte Carlo scenarios required to obtain a prescribed order of accuracy is of order $O(\sqrt{\text{Var}(V_{\text{est}})})$. We need then to minimize the variance of our estimator or, equivalently, its second moment. Since

$$E(V_{\text{est}}^2) = E\left[\left(\sum_{T_i \in \varsigma} g(T_i) D(T_i) C(T_i)\right)^2\right] \quad (21)$$

$$\leq E[g^2(T)] E\left[\left(\sum_{T_i \in \varsigma} D(T_i) C(T_i)\right)^2\right] \quad (22)$$

We now write $Z = \sum_{T_i} D(T_i) C(T_i)$ for clarity, and ρ_n for correlation between g^2 and Z^{4n} . We then use Hölder's inequality repeatedly:

$$E(V_{\text{est}}^2) \leq E[g^2(T)]^{1-\frac{1}{2n}} E[g^2(T)Z^{4n}]^{\frac{1}{2n}} \quad (23)$$

$$\leq E[g^2(T)]^{1-\frac{1}{2n}} (E[g^2(T)]E[Z^{4n}] - \rho_n \text{Var}[g^2(T)]^{\frac{1}{2}} \text{Var}[Z^{4n}]^{\frac{1}{2}})^{\frac{1}{2n}} \quad (24)$$

$$\leq E[g^2(T)]^{1-\frac{1}{2n}} (E[g^2(T)]E[Z^{4n}] + \text{Var}[g^2(T)]^{\frac{1}{2}} \text{Var}[Z^{4n}]^{\frac{1}{2}})^{\frac{1}{2n}} \quad (25)$$

$$\leq E[g^2(T)]^{1-\frac{1}{2n}} E[g^2(T)]^{\frac{1}{2n}} E[Z^{4n}]^{\frac{1}{2n}} (1 + \frac{\text{Var}[g^2(T)]^{\frac{1}{2}} \text{Var}[Z^{4n}]^{\frac{1}{2}}}{E[g^2(T)]E[Z^{4n}]})^{\frac{1}{2n}} \quad (26)$$

For finite variance of $g^2(T)$ the last term goes to one when n goes to infinity. We can then write, for large n

$$E(V_{\text{est}}^2) \leq E[g^2(T)]E[(\sum_{T_i} D(T_i)C(T_i))^{4n}]^{\frac{1}{2n}} \quad (27)$$

Accuracy is at worst inversely proportional to the square root of the second moment of $g(T)$. The goal is therefore to minimize $E[g^2(T)]$.

As explained earlier, mean-reversion is introduced in the true measure with the goal of limiting the dispersion of x , i.e., some linear form of the variance-covariance matrix of x . However, for controllability reasons, we are restricted to limiting the dispersion of a vector x' of dimension $d < N$, which is a linear transformation of the vector x :

$$x' = BRx \quad (28)$$

for some invertible matrix B . We will see in the next paragraph what is a good choice for BR . The dynamics of x' are:

$$dx' = -BR\Sigma B^{-1}x'dt + BR\Sigma dW \quad (29)$$

and we can write:

$$\lambda = B^{-1}x' \quad (30)$$

Using the new notation:

$$\Sigma' = BR\Sigma \quad (31)$$

$$A' = BR\Sigma B^{-1} \quad (32)$$

Let $e(E)$ be the vector (matrix) of ones in dimension N . The *control problem* becomes then:

$$\text{Min}_{A'(t)} E_0 \left[e^{r \int_0^T \lambda(t) dW(t) - \int_0^T |\lambda(t)|^2 dt} \right] \quad (33)$$

$$E_0 \left[\int_{t=0}^T (x')^t E x' dt \right] \leq M \quad (34)$$

$$dx' = -A'x'dt + \Sigma' dW \quad (35)$$

$$\lambda = (\Sigma')^{-1} A' x' \quad (36)$$

3. SOLUTION OF THE OPTIMAL CONTROL PROBLEM

The main difficulty in this paper was to find a better formula for the objective (34). If λ were non-stochastic, the following result would be trivial.

Lemma 1. Under (29)(30), $E_0[e^{\int_{t=0}^T 2\lambda(t)dW(t) - \int_{t=0}^T |\lambda(t)|^2 dt}] = e^{-E_0[\int_{t=0}^T |\lambda(t)|^2 dt]}$

Proof: See appendix.

Theorem 2. When Σ' is symmetric, the solution of the optimal control problem is, for some positive constant α :

$$A = \sqrt{\alpha}\Sigma' \quad (37)$$

Proof: Within this proof, we drop primes on our variables for clarity. The optimizer of the exponential of a function is the same as the optimizer of the function itself. The same fact carries over to functionals. Let α be the Lagrange multiplier of (34). The problem becomes then to minimize $J(0)$, where

$$J(t) = E_0\left[\int_{s=t}^T \lambda^t \lambda ds + \alpha x^t E x ds\right] \quad (38)$$

$$= E_0\left[\int_{s=t}^T x^t A^t \Sigma^{-2} A x + \alpha x^t E x ds\right] \quad (39)$$

Let $V(t)$ be the variance-covariance matrix of $x(t)$. By Ito's lemma, we have:

$$\frac{dV}{dt} = -AV - VA + \Sigma^2 \quad (40)$$

Writing \otimes for the Kronecker (term by term) product of two matrices, we have:

$$E[x^t A^t \Sigma^{-2} A x] = e^t A^t \Sigma^{-2} A \otimes V e \quad (41)$$

Using the principle of optimality in dynamic programming, we have:

$$0 = \text{Min}_{R(t)} e^t [A^t \Sigma^{-2} A \otimes V + \alpha V + \frac{\partial J}{\partial V} \otimes \frac{dV}{dt}] e + \frac{\partial J}{\partial t} \quad (42)$$

$$= \text{Min}_{R(t)} e^t [A^t \Sigma^{-2} A \otimes V + \alpha V + \frac{\partial J}{\partial V} \otimes (-AV - VA + \Sigma^2)] e + \frac{\partial J}{\partial t} \quad (43)$$

Differentiating with respect to A yields:

$$0 = \Sigma^{-2}AV - \frac{\partial J}{\partial V}V \quad (44)$$

We choose:

$$\frac{\partial J}{\partial V} = \Sigma^{-2}A \quad (45)$$

Going back to (43) and separating the time term and the V term yields:

$$\frac{\partial J}{\partial t} = -e^t \Sigma \Sigma^t e \quad (46)$$

$$0 = e^t [A^t \Sigma^{-2}A \otimes V + \alpha V + \Sigma^{-2}A \otimes (-AV - VA)]e \quad (47)$$

The right handside of (47) becomes, when $A = \sqrt{\alpha} \Sigma^{-1}$

$$e^t [A^t \Sigma^{-2}A \otimes V + \alpha V - 2\Sigma^{-2}A \otimes AV]e = \sum_{ij} 2\alpha V_{ij} - 2e^t \alpha (\Sigma^{-1} \otimes \Sigma V)e \quad (48)$$

$$= \sum_{ij} 2\alpha V_{ij} - 2\alpha \sum_{ij} (\Sigma^{-1})_{ij} \sum_k \Sigma_{ik} V_{kj} \quad (49)$$

$$= \sum_{ij} 2\alpha V_{ij} - 2\alpha \sum_{jk} \sum_i (\Sigma^{-1})_{ji} \Sigma_{ik} V_{kj} \quad (50)$$

$$= \sum_{ij} 2\alpha V_{ij} - 2\alpha \sum_{jk} V_{kj} \quad (51)$$

$$= 0 \quad (52)$$

Q.E.D.

In light of this theorem, we see what is a good choice for the matrix BR of the previous section to keep our control simple. It is the matrix that renders Σ' as symmetric as possible while making the left side of (34) a positive linear form in $E[x'_i x'_j]$. In other terms, it could be the solution of the program, for some positive constants l, d :

$$\min e^t (BR \Sigma - \Sigma^t R^t B^t) e \quad (53)$$

$$l \leq (BR)_{ij} \leq d \quad (54)$$

We have not checked if the control problem is convex, in which case the constant α would be the solution of the dual problem. In practice, it is easy to see that (34) is always tight, and α can be calculated by univariate search as the minimum value that satisfies (34).

Example. In the monivariate case, we take:

$$\sigma(t) = 0.1 + 0.3(1 - e^{-t/5}) \quad (55)$$

$$T = 10 \quad (56)$$

$$M = 0.267 \quad (57)$$

We compare our optimal control, to the "naive" control to select a constant speed of mean-reversion. The lowest value of a that satisfies (34) is $a = 2.8$. This results in:

$$E[g^2(T)] = 9.277 * 10^{10} \quad (58)$$

The optimal control $a = \sigma(t)$ results in:

$$E[g^2(T)] = 1.3 \quad (59)$$

We now compare our methodology, to simulate cash flows only in the true-measure, and then estimate V according to (15), to the normal method, to perform two independent simulations, one in the true measure, and one in the forward measure to calculate V . The cost of estimating (12) is the same in both methods. In order to reach the same accuracy in the second method to estimate V , we would need however 30% more scenarios. The total processing time of our method is therefore roughly two-thirds of the processing time of the normal method.

4. APPENDIX: PROOF OF LEMMA

For notational simplicity, we prove only the monivariate case. Let $M_t = \int_0^t \lambda dW_s$. By proposition (3.8) p.143 in [4], we have the Wiener chaos decomposition of $g(T)$:

$$g(T) = 1 + \sum_{n=0}^{\infty} \frac{1}{n!} \int_{t=0}^T L_t^n dM_t \quad (60)$$

where

$$L^0 = 1 \quad (61)$$

$$L^n = n \int L_t^{n-1} dM_t \quad (62)$$

Calculating explicitly the first terms yields:

$$L_t^1 = \int_{s=0}^t \lambda_s dW_s \quad (63)$$

$$L_t^2 = \int_{s=0}^t \int_{u=0}^s 2\lambda_u dW_u \lambda_s dW_s \quad (64)$$

We have then:

$$g(T) = 1 + \int_{t=0}^T \lambda_t \left(1 + \int_{s=0}^t \lambda_s \left(1 + \frac{1}{2} \int_{u=0}^s 2\lambda_u (1 + \dots) dW_u\right) dW_s\right) dW_t \quad (65)$$

Or, in differential form,

$$dg(t) = -\lambda_t \left(1 + \int_{s=0}^t \lambda_s \left(1 + \frac{1}{2} \int_{u=0}^s 2\lambda_u (1 + \dots) dW_u\right) dW_s\right) dW_t \quad (66)$$

By Ito's lemma and the martingale property of g we have $E_0[d(g^2)_t] = E_0[(dg_t)^2]$. We now compute the latter:

$$E_0[(dg_t)^2] = E_0\left[\left\{\lambda_t \left(1 + \int_{s=0}^t \lambda_s \left(1 + \frac{1}{2} \int_{u=0}^s 2\lambda_u (1 + \dots) dW_u\right) dW_s\right) dW_t\right\}^2\right] \quad (67)$$

$$= E_0[\lambda_t^2 E\left[\left\{\left(1 + \int_{s=0}^t \lambda_s \left(1 + \frac{1}{2} \int_{u=0}^s 2\lambda_u (1 + \dots) dW_u\right) dW_s\right)\right\}^2 \middle| x_t\right] dt] \quad (68)$$

Let us write $H_s = \lambda_s \left(1 + \int_{u=0}^s \lambda_u (1 + \dots) dW_u\right)$, and $T_t = \langle x, x \rangle_t$, which is deterministic. The p.d.f of x_t being identical to the p.d.f. of $W(T_t)$, we have:

$$E[f(z) E\left[\left\{\int_{s=0}^t H_s dW_s\right\}^2 \middle| x_t = z\right]] = E[f(z) E\left[\left\{\int_{s=0}^t H_s dW_s\right\}^2 \middle| W(T_t) = z\right]] \quad (69)$$

Because of mean-reversion $T_t \leq t$, so we have:

$$\begin{aligned} E\left[\left\{\int_{s=0}^t H_s dW_s\right\}^2 \middle| W(T_t)\right] &= E\left[\left\{\int_{s=0}^{T_t} H_s dW_s\right\}^2 \middle| W(T_t)\right] + \\ &E\left[\int_{s=0}^{T_t} H_s dW_s \int_{s=T_t}^t H_s dW_s \middle| W(T_t)\right] \\ &+ E\left[\left\{\int_{s=T_t}^t H_s dW_s\right\}^2 \middle| W(T_t)\right] \end{aligned} \quad (70)$$

Let $B(s) = W(s) - W(T_t)$. The distribution of the Brownian bridge (formula 8.5 in [1])

is:

$$\{B_t - T_t B(T_t), 0 \leq t \leq T_t\} \stackrel{d}{=} \{B_t, 0 \leq t \leq T_t | B(T_t) = 0\} \quad (71)$$

In other terms,

$$\{W_t - W(T_t), 0 \leq t \leq T_t\} \stackrel{d}{=} \{W_t - W(T_t), 0 \leq t \leq T_t | W(T_t)\} \quad (72)$$

The first term to the right of (70) is equal to

$$E\left[\left\{\int_{s=0}^{T_t} H_s dW_s\right\}^2 | W(T_t)\right] = E\left[\left\{\int_{s=0}^{T_t} H_s d(W_s - W(T_t))\right\}^2 | W(T_t)\right] \quad (73)$$

$$= E\left[\int_{s=0}^{T_t} E[H_s^2 | W(T_t)] d(W_s - W(T_t))^2\right] \quad (74)$$

$$= \int_{s=0}^{T_t} E[H_s^2 | W(T_t)] ds \quad (75)$$

The second term is zero, and the third term is:

$$E\left[\left\{\int_{s=T_t}^t H_s dW_s\right\}^2 | W(T_t)\right] = E\left[\int_{s=T_t}^t H_s^2 ds | W(T_t)\right] \quad (76)$$

Therefore (70) can be rewritten:

$$E\left[\left\{\int_{s=0}^t H_s dW_s\right\}^2 | W(T_t)\right] = \int_{s=0}^t E[H_s^2 | W(T_t)] ds \quad (77)$$

And (68) becomes, with the help of (69):

$$E_0[(dg_t)^2] = E_0[\lambda_t^2 dt \{1 + [\int_{s=0}^t E_0[H_s^2 | x_t] ds\}] \quad (78)$$

$$= E_0[\lambda_t^2 dt \{1 + \int_{s=0}^t \lambda_s^2 (1 + \int_{u=0}^s \lambda_u (1 + \dots) dW_u)^2 ds\}] \quad (79)$$

$$= E_0[\lambda_t^2 dt (1 + \sum_{n=1}^{\infty} \int_{t_{n-1}=0}^t \dots \int_{t_0=0}^{t_1} \prod_{i=0}^{n-1} \lambda^2(t_i) dt_i)] \quad (80)$$

Writing $v(t) = E_0[x^2(t)]$ we obtain:

$$E_0[d(g^2)_t] = E_0[\lambda^2(t)]dt(1 + \sum_{n=1}^{\infty} \int_{t_{n-1}=0}^t \dots \int_{t_0=0}^{t_1} \Pi_{i=0}^{n-1} E_0[\lambda^2(t_i)]dt_i) \quad (81)$$

Let

$$h(t) = (1 + \sum_{n=1}^{\infty} \int_{t_{n-1}=0}^t \dots \int_{t_0=0}^{t_1} \Pi_{i=0}^{n-1} E_0[\lambda^2(t_i)]dt_i) \quad (82)$$

Since:

$$\frac{dh}{dt} = E_0[\lambda^2(t)]h(t) \quad (83)$$

We have:

$$h(t) = \exp\left[\int_{s=0}^t E_0[\lambda^2(s)]ds\right] \quad (84)$$

Finally,

$$E_0[d(g^2)_t] = E_0[\lambda^2(t)] \exp\left[\int_{s=0}^t E_0[\lambda^2(s)]ds\right]dt \quad (85)$$

$$= \frac{d}{dt} \exp\left[\int_{s=0}^t E_0[\lambda^2(s)]ds\right]dt \quad (86)$$

Therefore

$$E_0[g^2(T)] = \exp\left[\int_{s=0}^T E_0[\lambda^2(s)]ds\right] \quad (87)$$

5. REFERENCES

- [1] Durrett, R. Probability: Theory and Examples. Brooks and Cole. 1991
- [2] Hull J, and White. Bond Option Pricing Based on a Model for the Evolution of Bond Prices. *Advances in Futures and Options Research*, Vol 6 (1993) 1-13.
- [3] Rebonato, R. Modern Pricing of Interest-Rate Derivatives: The LIBOR Market Model and Beyond. Princeton University Press, 2002.
- [4] Rebonato, R. Volatility and Correlation. John Wiley & Sons, 1999.

[5] Revuz, D., and Yor. Continuous Martingales and Brownian Motion. Springer, 1991.

[6] Rouge, R., and El Karoui. Pricing via Utility Maximization and Entropy. *Mathematical Finance* 10 (2000), 259-276.

[7] Schellhorn, H. and Kidani. Variance Reduction Techniques for Large Scale Risk Management. In *Monte Carlo and Quasi-Monte Carlo Methods 1998*. Niederreiter and Spanier eds. Springer.

[8] Williams, M. Private communication, 1999.

[9] Yor, M. *Exponential Functionals of Brownian Motion and Related Processes*. Springer, 2001.