

OPTIMAL CHANGES OF GAUSSIAN MEASURES, WITH AN APPLICATIONS TO RESIMULATION

Henry Schellhorn

Abstract. We derive optimality conditions and calculate approximate solutions to the problem of determining the optimal speed of mean reversion to be applied to a Gaussian state variable. The optimality criterion is the minimization of the variance of the Radon-Nikodym derivative of the measure "with mean-reversion" with respect to the measure "without mean-reversion" under constraints. We show that we can increase the speed of performing resimulation and sensitivity analysis in a Monte Carlo simulation. We apply our results to the pricing of real options and the pricing of interest-rate derivatives in the BGM/Libor model of interest rates.

Keywords: Monte Carlo simulation; change of measure.

Claremont Graduate University
School of Mathematical Sciences
710 N. College Ave
Claremont, CA 91711
Henry.Schellhorn@cgu.edu.

1. Introduction

We examine the problem of characterizing a "target", or "final" Gaussian probability measure by doing an optimal change of measure from an "initial" (completely characterized) Gaussian probability measure. In Monte Carlo simulation, it is often useful to resimulate the model using a different (final) probability measure when the integrand is expensive to compute. We thus advocate, rather than performing two independent simulations, to use the so-called *change of measure (CM) resimulation* scheme, whereby we calculate (scenario by scenario) the integrand only once in the initial measure and then multiply it by the Radon-Nikodym derivative of the final measure w.r.t. the initial measure to obtain the expected value of the integrand in the final measure.

This problem has several applications in computational finance, in both complete and incomplete markets. The traditional approach to calibrate option pricing models is to estimate a "physical", or "actual" measure by time-series analysis (see for instance Chan et al (1992) in the case of interest rate models), and then calibrate the market price of risk in such a way that (discounted) traded instruments are martingales in the risk-neutral measure; in this context, the initial measure is the physical measure, and the final measure is the risk-neutral measure. This paper also describes an application to the converse problem, where the initial measure is the risk-neutral measure, and the final measure is the physical measure; this problem has been less discussed in the literature, but is nevertheless quite important in practice. In incomplete markets, there are not enough traded instruments to fully determine a final (risk-neutral) measure, therefore a final measure is often defined

by changing optimally the initial (physical) measure or, viewed differently, by selecting an "optimal" measure from a class of probability measures (see e.g. Rouge and El Karaoui (2000) or Schweitzer (1996)).

We restrict our attention to probability measures where the state variable is a Gaussian diffusion, namely an Ornstein-Uhlenbeck process or an arithmetic Brownian motion, because their analytical tractability makes them quite important in practice. The final measure is the solution of an optimization problem, whereby the variance of the Radon-Nikodym derivative of the final measure with respect to the initial measure is minimized subject to constraints on some characteristics of the final measure. It is well-known that a Gaussian process is fully determined by its mean and autocovariance functions. Thus, our stylized analysis will focus on two different constraints: a constraint on the variance of the state variable at a terminal time, and a constraint on the (time-)average variance of the state variable. In our application section, we describe a case where the state variable in the final measure should have a lower variance than in the initial measure; as a result, the overall speed of mean-reversion should be higher in the final measure than in the initial measure; our methodology suggests the optimal shape of the speed of mean-reversion curve. The objective chosen, the minimization of the variance of the Radon-Nikodym derivative, is important in Monte Carlo simulation.

Our main contributions are the following. We first show that the variance of the Radon-Nikodym derivative is the exponential of the time-integral of the solution of an ordinary differential equation, thereby making our optimal control problem deterministic. We then explore the properties of this optimal control problem subject

to (i) a constraint on the average variance of the state variable in the final measure and (ii) a constraint on the terminal variance, and propose an approximation.

In the second part of our paper, we present two applications of our methodology to finance.

The first application is Monte Carlo resimulation of a real option model. In real options, the physical process is not traded, and its drift in the risk-neutral measure depends on the market price of risk. The initial measure is the physical measure, and the terminal measure is the risk-neutral measure.

The second application is Monte Carlo resimulation of the BGM/Libor model. In that problem the initial measure is the rolling forward measure, and the final measure is the physical measure. It is well-known that the calibration of the BGM model to caps and swaptions results in an implausibly high dispersion of interest rates (see e.g., Rebonato (1999)), while information is lacking to determine the variance of interest rates at all times in the physical measure. The constraint on the variance of interest rates in the physical measure (which has to be equal to the estimated value obtained from time-series analysis) and the minimization of the Radon-Nikodym derivative complete the determination of the optimal drift of interest rates in the physical measure. For this particular application, we compare our methodology with perhaps the simplest benchmark, namely selecting a constant speed of mean-reversion, and observe a significant albeit not drastic improvement. We did not compare our methodology to other methodologies for two reasons. The first one is that calculations become much more intensive: for more general changes of measure one needs to solve backward stochastic differential equations (see, e.g., Schweizer (1996)) instead of ordinary ones; besides, the variance can be calculated

only via simulation. Second, the resulting process (in the final measure) may have some undesirable properties that are hard to detect, as opposed to the simplicity of a Gaussian process.

2. Theory

Notation: The complete filtered probability space $(\Omega, \mathcal{F}, P^I)$ supports a Brownian motion W^I . We use the superscript I and F to refer to the probability measure, expectation operator, variance (Var) operator, and Brownian motion in the initial/final measure. When not shown otherwise, the expectation and variance operators are taken at time zero.

The dynamics of the stochastic process y of interest are:

$$(2.1) \quad y(t) = x(t) + \alpha(t)$$

$$(2.2) \quad dx(t) = \sigma(t)dW^I(t)$$

$$(2.3) \quad x(0) = 0$$

where α and σ are deterministic functions of time. For simplicity, we choose x to be our state variable: we note that the average/terminal variance of y are the same as the ones of the variable x . Typically in finance y would be the logarithm of an asset price (please see the application section for more specific explanations), α its mean function (in the initial measure), and σ the volatility. In the final measure P^F the process W^F is Brownian motion, where:

$$dW^F(t) = dW^I(t) + \frac{a(t)x(t)}{\sigma(t)}dt$$

Once the speed of mean-reversion $a(t)$ is specified, P^F becomes fully specified. We now proceed to determine an expression for the second moment of the value at the horizon time T of the Radon-Nikodym derivative:

$$(2.4) \quad g(T) \equiv E^I \left[\frac{dP^F}{dP^I} \middle| \mathcal{F}_T \right]$$

Lemma 1:

$$(2.5) \quad E^I[g^2(T)] = \exp \left[\int_0^T \sigma^2(t) f(t) dt \right]$$

where:

$$(2.6) \quad \begin{aligned} \frac{df(t)}{dt} &= -\frac{a^2(t)}{\sigma^2(t)} + 4a(t)f(t) - 2\sigma^2(t)f^2(t) \\ f(T) &= 0 \end{aligned}$$

This lemma is standard, and can be proved by Ito's lemma. The technique of proof is similar to the calculation of the value of a discount bond in the Cox, Ingersoll, and Ross model (see e.g., Duffie (1996)). As Leventorksii (2004) points out, there is no "truly analytical" formula in that case, unless a and σ are constant. Various expansions exist for the solution of this problem though (see, e.g., Grasselli and Hurd (2003)).

Although both time-average variance and terminal variance constraints can be incorporated into the same optimal control problem, it is simpler for stylized analysis to consider both problems and optimality conditions separately. This also

enables us, in our result section, to decompose the effect of each constraint on the shape of the optimal control.

2.1. Average Variance Constraint (AVC) Problem. The *AVC problem* consists of selecting a so that, for a fixed constant A (the average variance times the horizon):

$$(2.7) \quad \min_a E^I[g^2(T)]$$

$$(2.8) \quad E^F\left[\int_0^T x^2(t)dt\right] \leq A$$

$$(2.9) \quad dx(t) = -a(t)x(t)dt + \sigma(t)dW^T(t)$$

$$(2.10) \quad x(0) = 0$$

THEOREM 1 *A necessary condition for a to solve the AVC problem is to set:*

$$(2.11) \quad a = \sigma^2\left[2f - \frac{zv}{y}\right]$$

where, for some value of $\lambda \geq 0$, the functions f, z, y, v solve the following boundary value problem:

$$(2.12) \quad \int_0^T v(t)dt \leq A$$

$$(2.13) \quad \frac{df}{dt} = -\frac{a^2}{\sigma^2} + 4af - 2\sigma^2 f^2 \quad f(T) = 0$$

$$(2.14) \quad \frac{dv}{dt} = -2av + \sigma^2 \quad v(0) = 0$$

$$(2.15) \quad \frac{dy}{dt} = \sigma^2 + 4y[\sigma^2 f - a] \quad y(T) = 0$$

$$(2.16) \quad \frac{dz}{dt} = \lambda + 2az \quad z(T) = 0$$

Proof. The control a solves (AVC) only if it minimizes the logarithm of $E^I[g^2(T)]$ under the same constraints. We define:

$$v(t) = E^F \left[\int_0^t x^2(u)du \right]$$

The Lagrange dual of the (AVC) problem is:

$$(2.17) \quad \max_{\lambda \geq 0} \mathcal{L}(\lambda)$$

$$(2.18) \quad \mathcal{L}(\lambda) = \min_a \int_0^T \sigma^2(t)f(t) + \lambda \int_0^T v(t)dt$$

$$(2.19) \quad \frac{df}{dt} = -\frac{a^2}{\sigma^2} + 4af - 2\sigma^2 f^2 \quad f(T) = 0$$

$$(2.20) \quad \frac{dv}{dt} = -2av + \sigma^2 \quad v(0) = 0$$

A necessary condition for a to solve the primal problem is that (λ, a) solve the Lagrange dual. A weaker condition is that a solves the inner minimization problem for a value of λ which is such that (2.8) holds. The inner minimization problem is then:

$$\max_a \int_0^T -\sigma^2(t)f(t) - \lambda v(t)dt$$

under (2.19) and (2.20). The Hamiltonian is:

$$(2.21) \quad H = -\sigma^2 f - \lambda v + y\left(-\frac{a^2}{\sigma} + 4af - 2\sigma^2 f^2\right) + z(-2av + \sigma^2)$$

Conditions (2.11), (2.13) to (2.16) follow from Pontryagin's maximum principle.

□

The advantage of this formulation is to replace a complicated stochastic control problem by a system of ordinary differential equations. Unfortunately the conditions of the theorem are not sufficient for a minimum: indeed the Hessian of the Hamiltonian is neither positive-definite nor negative-definite so the sufficient conditions of Mangasarian (1966) do not apply. Numerically, one is then reduced to searching all possible local minima. Since 3 terminal conditions are supplied versus only 1 initial condition, it may then appear superior to solve the boundary problem (2.11), (2.13) to (2.16) by the shooting method going backward, and then vary the free parameter $v(T)$ until the initial condition $v(0) = 0$ is met. Unfortunately, the initial value of the control $a(T)$ is then undefined, and all the approximations we used to fix this problem turned out to yield suboptimal solutions.

We decided instead to simplify the problem, that is, to reduce our "high-order" system into a "low-order" system. For a lucid exposition of these concepts, as well as an exposition of a different approximating procedure in optimal control we refer the reader to Sannutti (1968). Suppose that $\sigma^2 = O(\varepsilon^2)$. More specifically, we set

$$\sigma(t) = \varepsilon\sigma_1(t) + \dots$$

We do an asymptotic expansion of all the unknown functions:

$$a(t) = a_0(t) + \varepsilon a_1(t) + ..$$

$$f(t) = f_0(t) + \varepsilon f_1(t) + ..$$

$$v(t) = v_0(t) + \varepsilon v_1(t) + ..$$

where each of the terms is $O(1)$. We try the solution $a_0(t) = 0$, $v_0(t) = v_1(t) = 0$, $y_0(t) = y_1(t) = 0$. Separating the terms $O(1)$ and $O(\varepsilon)$, we obtain:

$$(2.22) \quad \frac{df_0}{dt} = -\frac{a_1^2}{\sigma_1^2}$$

$$(2.23) \quad \frac{df_1}{dt} = 4a_1 f_0$$

$$(2.24) \quad \frac{dv_2}{dt} = \sigma_1^2$$

The solution of equations (2.22) to (2.24) is clearly $O(1)$. We then calculate a first-order approximation of $E^I[g^2(T)]$:

$$(2.25) \quad \begin{aligned} \ln E^I[g^2(T)] &= \varepsilon^2 \int_0^T \sigma_1^2(t) f_0(t) dt + O(\varepsilon^3) \\ &= \varepsilon^2 \left[\int_0^T \sigma_1^2(u) du f_0(t) \right]_0^T - \varepsilon^2 \int_0^T \int_0^T \sigma^2(u) du \frac{df_0}{dt} \Big|_t dt + O(\varepsilon^3) \\ &= \varepsilon^2 \int_0^T \frac{\int_0^t \sigma^2(u) du}{\sigma^2(t)} a_1^2(t) dt + O(\varepsilon^3) \end{aligned}$$

where (2.25) follows from the boundary condition $f_0(T) = 0$, inherited from the boundary condition of (2.19). So far, we do not know the order of λ , so that

we keep in our "low order formulation" the full dynamics of v . We now define the approximated Lagrangian:

$$(2.26) \quad \mathcal{L}_{app}(\lambda) = \min_a \ln E^I[g_{app}^2(T)] + \lambda \int_0^T v(t) dt$$

where:

$$(2.27) \quad \ln E^I[g_{app}^2(T)] = \int_0^T \frac{\int_0^t \sigma^2(u) du}{\sigma^2(t)} a^2(t) dt$$

The approximated control problem consists then of (2.26), under (2.27) and (2.14). The Hamiltonian of the approximated problem becomes then:

$$H_{app}(v(t), a(t), t) = \frac{\int_0^t \sigma^2(u) du}{\sigma^2(t)} a^2(t) - \lambda v(t) + z(t)(-2a(t)v(t) + \sigma^2(t))$$

The optimal control of the approximated problem is:

$$(2.28) \quad a(t) = \frac{-\sigma^2(t)z(t)v(t)}{\int_0^t \sigma^2(u) du}$$

where v and z follow as before (2.14) and (2.16). An advantage of this formulation is that the approximated Hamiltonian is concave in v and t , and the Pontryagin optimality conditions are then sufficient.

We summarize our findings in an algorithm to calculate the optimal speed of mean-reversion. It uses a simple Euler scheme with time step Δt (a divisor of T) combined with the shooting method.

Algorithm AVC

Set $\lambda = a(0) = v(0) = 0$

Repeat

$$z(0) = 0$$

Repeat

For $t = \Delta t, \dots, T$

$$v(t) = v(t - \Delta t)(1 - 2a(t - \Delta t)\Delta t) + \sigma^2(t)\Delta t$$

$$z(t) = z(t - \Delta t)(1 + 2a(t - \Delta t)\Delta t) + \lambda\Delta t$$

$$a(t) = \frac{-\sigma^2(t)z(t)v(t)}{\mathbb{P}_{i=0}^{T/\Delta t} \sigma^2(i\Delta t)\Delta t}$$

Next t

Decrease $z(0)$ by some amount $\Delta z(0)$

Until $z(T)$ sufficiently close to zero

Increase λ by some amount $\Delta\lambda$

Until $\mathbb{P}_{i=0}^{T/\Delta t} v(i\Delta t)\Delta t$ sufficiently close to A

The Lagrange multiplier, λ represents the trade-off between a small variance of $g(T)$ and a small average variance of x . The higher λ the smaller the average variance of x (relatively to the variance of $g(T)$). In the case that interests us (namely where the average variance of x is higher than A for a null speed of mean reversion), λ is always positive. As a result z is negative, so that a and v are positive. Is this scheme able to reduce the average variance of x to any positive number? Clearly, the control:

$$(2.29) \quad a(t) = \frac{\sigma^2(t)}{2v(t)}$$

results in a null variance of x at all times, but also in an infinite variance of $g(T)$. Practically, for strictly positive final average variance of x the experimental answer seems to be negative. We cannot tell whether this is due to our approximation of the variance formula or to some intrinsic feature of the problem.

2.2. Terminal Variance Constraint (TVC) Problem. The *TVC problem* consists of selecting a so that, for a fixed constant M (the terminal variance times the horizon):

$$(2.30) \quad \min_a E^I[g^2(T)]$$

$$(2.31) \quad E^F[x^2(T)dt] = M$$

$$(2.32) \quad dx(t) = -a(t)x(t)dt + \sigma(t)dW^T(t)$$

$$(2.33) \quad x(0) = 0$$

The analog of theorem 1 follows.

THEOREM 2 *A necessary condition for a to solve the TVC problem is to set:*

$$(2.34) \quad a = 2\sigma^2\left[f - \frac{z}{y}v\right]$$

where, for some value of z_T , the functions f, z, y, v solve the following boundary value problem:

$$(2.35) \quad \frac{df}{dt} = -\frac{a^2}{\sigma^2} + 4af - 2f^2\sigma^2 \quad f(T) = 0$$

$$(2.36) \quad \frac{dv}{dt} = -2av + \sigma^2 \quad v(T) = M$$

$$(2.37) \quad \frac{dy}{dt} = \sigma^2 + 4y[\sigma^2 f - a] \quad y(T) = 0$$

$$(2.38) \quad \frac{dz}{dt} = 2az \quad z(T) = z_T$$

Solving the TVC problem is easier than solving the AVC problem because the Lagrange multiplier λ disappears. However, the equations are badly-conditioned, in the sense that both z and y increase quite fast with time, which results quickly in numerical errors when calculating (2.34). We resorted to the same approximation as before, i.e., using (2.27), which results in the same suboptimal control (2.28).

Algorithm TVC

Choose Δz_T and $\Delta z(0)$ so that $\Delta z_T > \Delta z(0)$

Set $a(0) = v(0) = z_T = 0$

Repeat

$$z(0) = 0$$

Repeat

For $t = \Delta t, \dots, T$

$$v(t) = v(t - \Delta t)(1 - 2a(t - \Delta t)\Delta t) + \sigma^2(t)\Delta t$$

$$z(t) = z(t - \Delta t)(1 + 2a(t - \Delta t)\Delta t)$$

$$a(t) = \frac{-\sigma^2(t)z(t)v(t)}{\mathbb{P} \sum_{i=0}^{T/\Delta t} \sigma^2(i\Delta t)\Delta t}$$

Next t

Decrease $z(0)$ by some amount $\Delta z(0)$

Until $z(T)$ sufficiently close to z_T

Decrease z_T by some amount Δz_T

Until $v(T)$ sufficiently close to M .

3. Results

3.1. Average Variance Constraint Problem. We define the *relative average variance of $x(T)$* (in the final measure) as the ratio of A divided by the cumulated variance that we would obtain if $a = 0$ (that is, $\int_0^T \int_0^u \sigma^2(u) du dt$). We report in figures 2 and 3, as a function of the relative average variance of $x(T)$ the value of $Var^I[g(T)]$ for our suboptimal control obtained from the algorithm (AVC) with $\Delta t = 0.01$. In figure 2 volatility is constant, while in figure 3 volatility is time-dependent. Note that, for relevance we report the "true" variance of $g(T)$ and not the approximation (i.e., the exponential of (2.27)). We also compare our suboptimal control to the most naive control, that is, the lowest constant speed of mean reversion such that (2.8) is met. In all cases our suboptimal control beats the constant speed of mean reversion control. We observe in all these results that the suboptimal control $a(t)$ follows a slightly downward trend.

3.2. **Terminal Variance Constraint Problem.** We define the *relative terminal variance of $x(T)$* (in the final measure) as the ratio of M divided by the terminal variance that we would obtain if $a = 0$, that is $\int_0^T \sigma^2(t) dt$. We report in figures 4 and 5, as a function of the relative terminal variance of $x(T)$ the value of $Var^I[g(T)]$ for our suboptimal control obtained from the algorithm (TVC) with $\Delta t = 0.01$. In figure 4 volatility is constant, while in figure 5 volatility is time-dependent. As before, we report the "true" variance of $g(T)$ and not the approximation (i.e., the exponential of (2.27)). As before, we also compare our suboptimal control to the control corresponding to the lowest constant speed of mean reversion such that (2.8) is met. In all cases our suboptimal control beats the constant speed of mean reversion control, albeit more moderately than in the average variance case. For the terminal variance ratio and horizons chosen, the terminal variance turns out to be not significantly lower than the maximum variance of x across its path.

4. Application to Real Options

In this section the initial measure is the physical measure and the final measure is the risk-neutral measure. We suppose that there are two inputs to the model:

- the underlying variable, Y which is not traded
- a European option which payoff H at time T depends only on $Y(T)$; the observed market price of this claim, namely H_0 , is universally accepted as being "right".

The dynamics of Y are:

$$y(t) = \ln Y(t) - \ln(Y(0)) + \frac{1}{2} \int_0^t \sigma^2(u) du$$

$$dy(t) = \mu dt + \sigma(t) dW^I(t)$$

The goal is to manage a portfolio of various other contingent claims on Y . For simplicity of exposition we assume that all cash flows occur at maturity, so that we can write $h(Y)$, for this single (path-dependent) cash-flow. Observe that h is a functional. Portfolio managers are interested, among other things, in calculating, by Monte Carlo simulation:

- the price of the portfolio
- the distribution of $h(Y)$ in the physical measure.

We describe in the appendix a data model corresponding to this problem.

It is well-known (see e.g. Dixit and Pindyck (1994)) that the rate of return of Y is not equal to the risk-free interest rate r in the risk-neutral measure. Here we select a mean-reverting market price of risk, that is, we let:

$$x(t) = y(t) + \mu t$$

$$dx = -a(t)x(t)dt + \sigma(t)dW^F(t)$$

Note that, on top of the mean-reverting component we could have added a constant market price of risk, but this would have resulted in a more complicated expression for $E^I[g^2(T)]$ than (2.5). By assumption:

$$(4.1) \quad H_0 = e^{-rT} E^I[g(T)H]$$

Since Y is geometric Brownian motion, the Black-Scholes formula shows that constraint (4.1) can be expressed as a relationship between the observed implied volatility and the model implied volatility, namely, it can be translated into constraint (2.31). The TVC algorithm can then be used to determine a shape of $a(t)$ that best accelerates Monte Carlo resimulation, as we will show shortly. We first show the full algorithm.

Algorithm CM Resimulation

- I. Calculate the optimal $a(t)$ using the TVC algorithm.*
- II. Simulate $Y(\omega)$ for each scenario $\omega = 1..\Omega$ in the initial measure*
- III. Calculate the portfolio cash flows $h(Y(\omega))$*
- IV. Calculate the empirical distribution of the cash flows $h(Y(\omega))$ in the initial measure*
- V. Calculate the Radon-Nikodym derivative $g(T, \omega)$ for each scenario $\omega = 1..\Omega$*
- VI. Calculate the estimator of market value $V_{\text{CM}} = \frac{1}{\Omega} e^{-rT} \mathbb{P}_{\omega=1} g(T, \omega) h(Y(\omega))$*

By comparison, a traditional resimulation algorithm would be:

Algorithm Traditional Resimulation

- 1. Calculate a constant a so that (2.31) is met with equality*
- 2. Simulate $Y(\omega)$ for each scenario $\omega = 1..\Omega$ in the initial measure*
- 3. Calculate the portfolio cash flows $h(Y(\omega))$*
- 4. Calculate the empirical distribution of the cash flows $h(Y(\omega))$ in the initial measure*
- 5. Simulate $Y(\omega')$ for each scenario $\omega' = 1..\Omega$ in the final measure*

6. Calculate the estimator of market value $V_{\text{TRAD}} = \frac{1}{\Omega} e^{-rT} \mathbb{P}_{\omega^0=1} h(Y(\omega'))$

The CM Resimulation algorithm should be applied only when step III is more time-consuming than step II. We reported in an earlier paper (Schellhorn and Kidani (2000)) that, for the pricing of large portfolios of mortgage-backed securities, calculating the sum of the (discounted) payoffs $h(Y)$ can take more than 100 more time than simulating the state variable and its Radon-Nikodym derivative g .

4.1. Selection of an Objective to Minimize. By definition,

$$E^I[V_{CM}^2] = e^{-2rT} (E^I[g^2(T)]E^I[h^2(Y)] + Cov^I[g^2(T), h^2(Y)])$$

The sign of the covariance depends on the cash flow h . For an uncorrelated cash flow, we have:

$$(4.2) \quad Var^I[V_{CM}] = e^{-2rT} (E^I[g^2(T)]E^I[h^2(Y)] - E^F[h(Y)]^2)$$

It is well-known that the error in Monte Carlo simulation is proportional to the variance of the estimator. Therefore (4.2) shows that, on average (across all possible cash flows), minimizing $E^I[g^2(T)]$ results in accelerating the calculation time of V_{CM} .

There is a completely different reason why minimizing the variance of g is appealing for real option pricing, or, more generally, in incomplete markets. It is well-known that (see e.g., Duffie and Richardson (1991), Schweizer (1996)) when not all claims are attainable, the minimum variance martingale measure should be used

as a pricing measure when the objective is to minimize the $\mathcal{L}^2(P^I)$ - approximation error between the payoff of a non-attainable claim, and the value of its "best hedge". There are however some non-trivial differences between the definition of a minimum variance martingale measure and our definition, which makes this result delicate to translate to our model. We decided to leave this task for future research.

We also note that entropy minimization (see e.g., Rouge and El Karoui (2001)) seems to be gaining in popularity in the incomplete market literature compared to variance-minimization. Roughly speaking, minimization of the entropy of a measure corresponds to maximizing the expected value of an exponential utility function, whereas finding the minimum variance martingale measure corresponds to maximizing the expected value of a quadratic utility function. However, unlike the variance of (the Radon-Nikodym of) a measure, it seems difficult to relate the entropy of a measure to the accuracy of Monte Carlo resimulation.

5. Application to the BGM/Libor Model

The BGM/Libor model is currently one of the most widely used models for the pricing of interest rate options. In a one-factor BGM/Libor model, forward rates F_i for a loan between period T_i and $T_{i+1} \equiv T_i + \tau$ (with $i = 1..m$) follow the system of SDE:

$$(5.1) \quad \frac{dF_i}{F_i} = -\sigma_i \sum_{k=i+1}^n \frac{\sigma_k F_k \tau}{1 + F_k \tau} dt + \sigma_i dW^F$$

$$(5.2) \quad F_i(0) = F_{i,0}$$

where W^F is Brownian motion in the measure used for pricing, namely the rolling forward measure. The drift term in (5.1) is in practice very small, so that forward rates are approximately lognormal. Normality (of the logarithm of the forward rate) is a key advantage for a successful and intuitive calibration to caps and swaptions, as explained for instance in Rebonato (1999),(2002)). Likewise, a joint normal distribution in the physical measure of the logarithm of forward rates is much easier to interpret than any other more sophisticated and statistically more correct distribution.

Although the academic literature favours to first infer the physical measure and then adjust it with a market price of risk (see Heath et al (1992)) to obtain the rolling forward measure, information often flows the other way round in practice. In many bank departments the key requirement is to do a correct pricing, in the sense that the risk-neutral pricing formula applied to caps and swaptions returns the observed market prices. To this effect, the rolling forward measure is calibrated first to the prices of caps and swaptions and/or historical correlations. The initial measure is then the rolling forward measure. The final measure is then either the physical measure (to calculate *Value-at-Risk*, i.e. the probability distribution of future portfolio value), or some measure derived from either the rolling forward or the physical measure, to perform sensitivity analysis (e.g., what happens to Value-at-Risk when the average volatility changes by 1%, 2%, 5%).

5.1. Selection of a Constraint. Calibrating the BGM/Libor model to caps and swaptions results in a dispersion of rates forecast, in the rolling forward measure, that is much higher than the plausible dispersion of physical rates in the US,

because of the high skewness of the lognormal distribution. This is one of the reasons why alternate models like Hull and White (1993), where rates are Gaussian when they are large and lognormal when they are small, were designed to prevent a too rapid increase in risk-neutral rates; while widely used, for instance at Bank of America in the 1990s (Williams (1999)) this model is less practical to calibrate to caps and swaptions than the BGM/Libor model. To summarize, we advocate to calibrate the rolling forward measure first, and then to derive the physical measure by a mean reversion adjustment, such that rates will still be lognormal in the physical measure, but with a smaller dispersion. Depending on how the "smaller dispersion" constraint is specified, the optimal speed of mean reversion solves either problem (TVC), or problem (AVC), where the state variable $y(t) = \log(F_i(t)) - \log(F_{i,0})$ for some well-chosen forward rate F_i .

5.2. Results. Our own experience showed that a good fit is obtained to cap prices in the US and UK when volatility takes the form:

$$(5.3) \quad \sigma(t) = \sigma_0(1 - 0.8 \exp(-2t) - mt)$$

This corresponds to the stylized cap curve observed in Rebonato (2002) p. 232: an initial very steep portion, a plateau area, followed by a rapid decline. For simplicity of exposition, we suppose that all forwards have the same instantaneous volatility, that is:

$$(5.4) \quad \sigma_i(t) = \sigma(t)$$

This results in our usual model (2.2), with:

$$(5.5) \quad x(t) \cong \ln(F_i(t)) - \ln(F_i(0)) + \frac{1}{2} \int_0^t \sigma^2(u) du$$

where the approximation comes from the fact that we ignore the drift in (5.1).

We applied the algorithm TVC to the model above, with $\Delta t = 0.015$ and $T = 15$. We calculated the optimal speed of mean-reversion for each volatility curve, and then calculated $E^I[g^2(T)]$ according to the formula in lemma 1. In each case we set up the constant M so as to achieve a 40% terminal variance reduction (i.e., the ratio of M over $\int_0^T \sigma^2(t) dt$ equals 60%). We also compared our results with the value of $E^I[g^2(T)]$ obtained when we apply the constant value of a that results in the same terminal variance. Figure 6 shows that in all cases we obtain lower variances for our suboptimal speed of mean reversion, compared to a constant speed of mean reversion, but the effect is more pronounced for steeper declines in the volatility curve.

In the rest of this section we consider the volatility function corresponding to $m = 0.06$ in (5.3). From figure 6, we see (assuming again that (4.2) holds) that a CM resimulation scheme with constant a would need 86% more scenarios than a CM resimulation scheme with suboptimal $a(t)$, for the same accuracy. The former would then be completely ineffective.

We now compare the CM resimulation scheme with the traditional resimulation scheme on a real example. Comparing to section 4, the roles of the initial and the terminal measure are reversed, since here the initial measure is the measure where we calculate market value. We would therefore change steps IV and VI to:

IV. Calculate the estimator of market value "in the initial measure"

VI. Estimate the empirical distribution "in the final measure"

In this particular example, instead of comparing the empirical distributions, we compare the estimator of the variance (in the physical measure) of an at the money floorlet, that is, an instrument with cash flow:

$$h(F_i(T, \omega)) = \max(K - F_i(T, \omega), 0)$$

for $T = T_i = 1.5$, and $K = F_i(0, \omega) = 0.05$. Since we have only one cash flow, the adequate pricing measure is the forward measure, where (5.5) holds exactly. Our estimators are then, for a given batch of scenarios b :

$$(5.6) \quad V_{CM}(b) = \frac{1}{\Omega} \sum_{\omega=1}^{\Omega} g(\omega) h^2(F_i(T, \omega)) - \frac{1}{\Omega} \sum_{\omega=1}^{\Omega} g(\omega) h(F_i(T, \omega)) \quad \#_2$$

$$(5.7) \quad V_{TRAD}(b) = \frac{1}{\Omega} \sum_{\omega^0=1}^{\Omega} h^2(F_i(T, \omega')) - \frac{1}{\Omega} \sum_{\omega=1}^{\Omega} h(F_i(T, \omega')) \quad \#_2$$

where again scenarios ω are sampled from the forward measure, and scenarios ω' are sampled from the physical measure, and $\Omega = 100$. We compared these estimators across $b = 1..100$ batches and obtained an empirical standard deviation of:

$$\text{stdev}(V_{CM}(b)) = 5.51 * 10^{-6}$$

$$\text{stdev}(V_{TRAD}(b)) = 5.49 * 10^{-6}$$

In this particular example, the CM resimulation scheme would therefore be twice faster than the traditional resimulation scheme (assuming again, that most of the computing time is spent evaluating the cash flow h).

6. Conclusion

We derive optimality conditions and calculate approximate solutions to the problem of determining the optimal speed of mean reversion to be applied to a Gaussian state variable. We show that we can increase the speed of resimulation and sensitivity analysis in a Monte Carlo simulation. In this article, we take the case of finance simulations, but our result can apply to many other simulation problems, potentially even to the numerical solution via finite differences of parabolic partial differential equations.

7. Acknowledgements

We thank for their helpful comments Hans Ulrich Gerber, Julien Hugonnier, Yuri Kabanov, and Maria Semenova. All errors are ours. This research has been carried out within the project on Credit Risk of the National centre of Competence in Research "Financial Valuation and Risk Management" (NCCR FINRISK). The NCCR FINRISK is a research program supported by the Swiss National Science Foundation.

8. Appendix: Data Model

The data model in figure 1 refers to the problem described in section 4, namely simulation of real options. We model the problem as a practitioner would face it, deliberately satisfying mathematical rigor. For simplicity, we do not describe obvious variables, such as time t and scenario ω . The field `Constraint_Id` can take two values:

- average variance
- terminal variance

In other terms, the final measure can be determined via either the average variance problem or the terminal variance problem¹. The field `Constraint_Val` takes the numerical value A (in case `Constraint_Id`=average variance) or M (in case `Constraint_Id`=terminal variance).

Fields in grey ("key fields") uniquely determine each record in these tables.

9. References

Chan, K.C., Karolyi, A., Longstaff, F. and Saunders, A. (1992). An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance*, 47, 1209-1227.

Dixit, A. and Pindyck, R. (1994). *Investment under uncertainty*. Princeton University Press.

Duffie, D. and Richardson, H. (1991). Mean-variance hedging in continuous-time. *The Annals of Applied Probability*, 1(1), 1-15.

¹We discard the zero probability case where both problems return the same solution, namely $a(t)$.

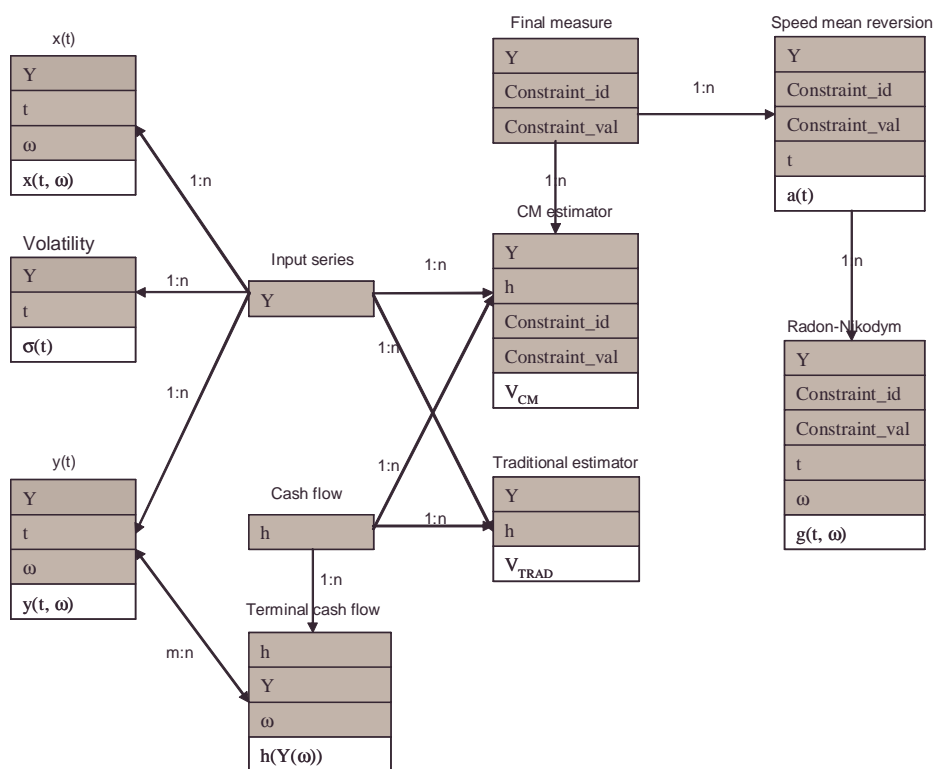


Figure 1. Data Model.

Duffie, D. (1996). *Dynamic Asset Pricing Theory*. Princeton University Press.

Grasselli, M. and Hurd, T.(2003). Wiener chaos and the Cox-Ingersoll-Ross model. Working paper. McMaster University.

Heath D., Jarrow, R. and Morton, A.(1992). Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica*, 60, 77-106.

Hull J., and White, A. (1993). Bond option pricing based on a model for the evolution of bond prices. *Advances in Futures and Options Research*, 6,1-13.

Levendorskii, S. (2004). Pseudo-diffusions and quadratic term structure models. Working paper, University of Texas at Austin.

Mangasarian, O.L. (1966). Sufficient conditions for the optimal control of nonlinear systems, *SIAM Journal of Control*, 2, 139-52.

Rebonato, R. (1999). *Volatility and Correlation*. Chichester. John Wiley & Sons.

Rebonato, R. (2002). *Modern Pricing of Interest-Rate Derivatives: The LIBOR Market Model and Beyond*. Princeton University Press.

Rouge, R. and El Karaoui, N. (2000). Pricing via utility maximization and entropy. *Mathematical Finance*, 10, 259-276.

Sannutti, P. (1968). Singular perturbation method in the theory of optimal control. Report R-379. University of Illinois, Urbana-Champaign.

Schellhorn, H. and Kidani, F. (2000). Variance reduction techniques for large scale risk management. In *Monte Carlo and Quasi-Monte Carlo Methods 1998*. Niederreiter and Spanier eds. Berlin.

Schweitzer, M.(1996). Approximation pricing and the variance-optimal martingale measure. *The Annals of Probability*, 24(1), 206-236.

Williams, M.(1999). Private communication.

10. Figures

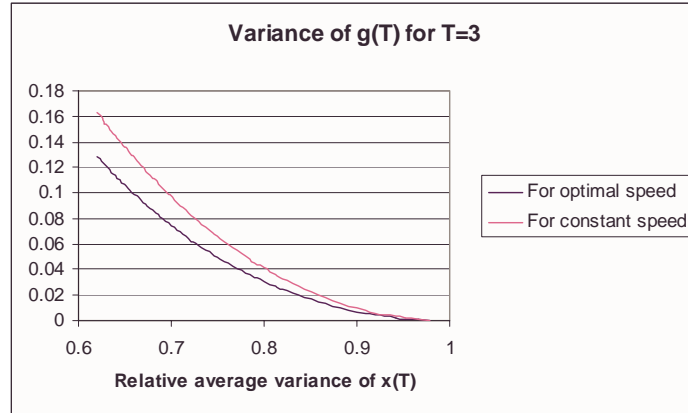


Figure 2. Variance of g as a function of the ratio of A over the cumulated variance of x in the uncontrolled case ($a = 0$). The volatility is $\sigma(t) = 0.2$.

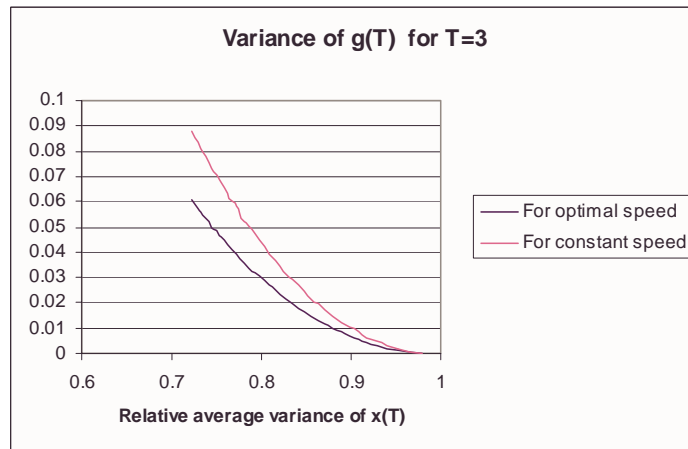


Figure 3. Variance of g as a function of the ratio of A over the cumulated variance of x in the uncontrolled case ($a = 0$). The volatility is $\sigma(t) = 0.2(1 + 0.2 \cos(\frac{t}{4}))$.

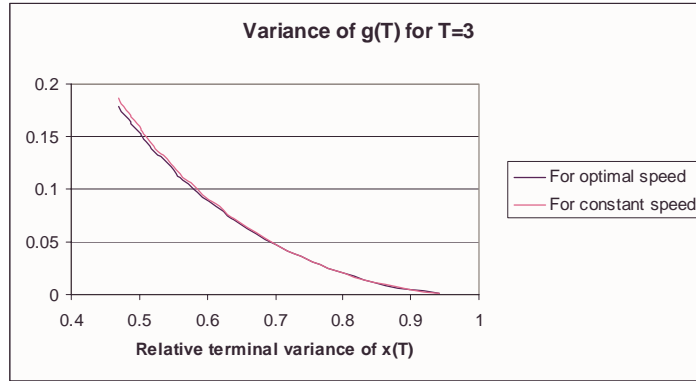


Figure 4. Variance of g as a function of the ratio of M over the terminal variance of x in the uncontrolled case ($a = 0$). The volatility is $\sigma(t) = 0.2$.

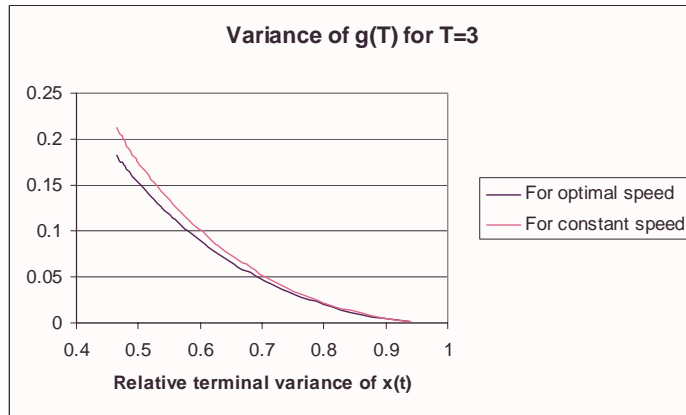


Figure 5. Variance of g as a function of the ratio of M over the terminal variance of x in the uncontrolled case ($a = 0$). The volatility is $\sigma(t) = 0.2(1 + 0.2 \cos(\frac{t}{4}))$.

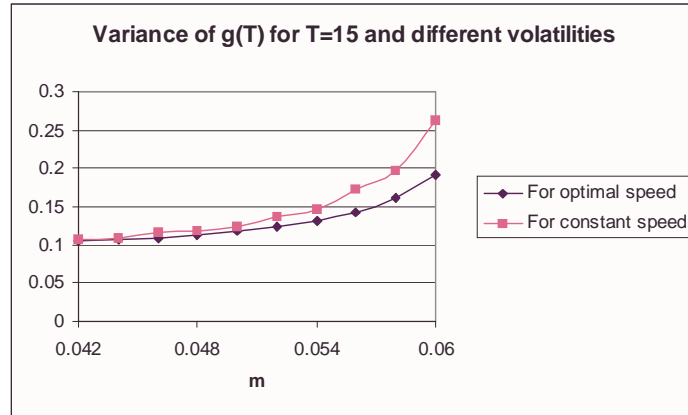


Figure 6. This figure represents the variance of $g^2(T)$ corresponding to a 40% terminal variance reduction (i.e., the ratio of terminal variance of $x(a(t) \neq 0)$ over terminal variance of $x(a(t) = 0)$ equals 60%). The horizontal axis corresponds to different volatility functions $\sigma = 0.35(1 - 0.8 \exp(-2t) - mt)$.