

**MARKET MODEL  
OF STOCHASTIC IMPLIED VOLATILITY  
WITH APPLICATION TO THE BGM MODEL**

ALAN BRACE, BEN GOLDYS, FIMA KLEBANER, AND ROB WOMERSLEY

ABSTRACT. Using a stochastic implied volatility method we show how to introduce smiles and skews into the BGM interest rate model.

CONTENTS

1.	Introduction	1
1.1.	Derivative formulae	3
2.	Dynamics of the implied volatility surface	5
3.	Different formulations	6
4.	Some properties of solutions	9
5.	Toy model	11
6.	Application to BGM	13
7.	Marginal Distributions	15
7.1.	Bachelier Model	16
7.2.	Derivatives	16
7.3.	Numerical results	17
	Appendix A. Ito-Venttsel formula	17
	References	17

1. INTRODUCTION

The aim of this paper is to present a new model for implied volatility. The main results and some properties of the model are announced without proofs. Those will be contained in the second part of this work, currently under preparation.

Suppose, as in [1], there are a full spectrum of zero coupon bonds  $P(t, T)$  maturing at all times  $T$  up to a finite horizon  $T^*$ , and let  $W_T(t)$  be Brownian motion under the forward measure  $\mathbb{P}_T$  located at maturity  $T$  (with corresponding numeraire  $P(t, T)$ ). Recall that the forwards  $L(t, T)$  over the interval  $[T, T_1]$ , where  $T_1 = T + \delta$ , are related to zero coupons via the relation

$$L(t, T) = \frac{1}{\delta} \left[ \frac{P(t, T)}{P(t, T_1)} - 1 \right].$$

---

WORKING PAPER S01-1, DEPARTMENT OF STATISTICS, UNIVERSITY OF NEW SOUTH WALES

From the arbitrage free dynamics of the zero coupon bonds,  $L(t, T)$  must be a positive martingale under the forward measure  $\mathbb{P}_{T_1}$  located at the end of  $[T, T_1]$ , so model it with the SDE

$$\frac{dL(t, T)}{L(t, T)} = \theta^*(t, T) dW_{T_1}(t), \quad (1.1)$$

where in general the volatility function  $\theta_t = \theta(t, T)$  is stochastic. The form of the SDE (1.1) under  $\mathbb{P}_{T_1}$  is similar to that of a stock under the spot measure  $\mathbb{P}$  in the standard Black-Scholes (BS) model when interest rates are zero. So to clarify ideas, first consider modelling stochastic implied volatility for a stock.

Following the notation and approach of Carr [4] but assuming interest rates are zero, under the spot arbitrage free measure  $\mathbb{P}$ , the numeraire will be unity and the underlying stock  $S_t$  is a martingale which we may assume satisfies the SDE

$$dS_t = S_t \theta_t dW_t^{(1)}, \quad (1.2)$$

where  $\theta_t$  is stochastic and  $dW_t$  is multi-dimensional Brownian motion under  $\mathbb{P}$ . Note that, with no loss of generality, we are taking all components of the instantaneous volatility vector  $\theta_t$  to be zero except the first.

The BS implied volatility convention says that if the time  $t$  (stochastic) implied volatility of an option exercising at time  $T$  with strike  $K$  is

$$\sigma_t = \sigma_t(T, K) = \sigma(t, T, K),$$

then the time  $t$  price of a call option will be

$$C_t = C(t, T, K) = \ell(S_t, \sigma_t(T, K), T - t; K), \quad (1.3)$$

where

$$\begin{aligned} \ell &= \ell(S, \sigma, \tau; K) = \text{SN}(h_1) - \text{KN}(h_2), \\ h_1 &= \frac{\ln \frac{S}{K}}{\sigma \sqrt{\tau}} + \frac{1}{2} \sigma \sqrt{\tau}, \quad h_2 = h_1 - \sigma \sqrt{\tau}. \end{aligned} \quad (1.4)$$

If the implied volatility is also a diffusion satisfying an SDE like

$$d\sigma_t = m_t(T, K, S_t, \theta_t, \sigma_t) dt + v_t(T, K, S_t, \theta_t, \sigma_t)^* dW_t = m_t dt + v_t^* dW_t, \quad (1.5)$$

then because the calls  $C_t$  must also be martingales under  $\mathbb{P}$ , it follows that the drift  $m_t$  and “volvol”  $v_t$  cannot be arbitrary, but must satisfy certain extra conditions. Those conditions will lead naturally to a system of SDEs for the implied volatility  $\sigma_t$ . The dependence of the volvol  $v_t(T, K, S_t, \theta_t, \sigma_t)$  on  $\sigma_t$  will be specified to get rid of some troublesome singularities.

We also suppose there are a full spectrum of call options available for all strikes  $K$  and all maturities  $T$  up to some horizon  $T^*$ . This assumption leads to two critical feedback conditions:

- The implied volatility  $\sigma_t(T, K)$  of the  $T$ -maturing call must remain finite at maturity, that is for  $t \leq T$

$$\sigma_t^2(T, K)(T - t) \geq 0 \quad \text{and} \quad \lim_{t \rightarrow T} \sigma_t^2(T, K)(T - t) = 0. \quad (1.6)$$

- The instantaneous volatility  $\theta_t$  of the underlying stock  $S_t$  must equal the implied volatility of the at-the-money option maturing immediately, that is

$$\theta_t = \sigma_t(t, S_t). \quad (1.7)$$

Let us emphasise importance of condition (1.6). It imposes a very severe restriction on the volvol process  $(v_t)$ . It becomes even more striking if we rewrite the SDE for the process  $(\sigma_t)$  in terms of a new process  $\xi_t = (T - t)\sigma_t^2$  (see Section 3 for details). In that case we end up with a stochastic differential equation for the process  $(\xi_t)$  with the initial condition  $\xi_0(T, K) = f(T, K)$  (say) and the terminal condition  $\xi_T(T, K) = 0$ . It is well known (see for example [5], that stochastic differential equations of this type need not have adapted solutions, unless the coefficients of this equation satisfy certain conditions. In our case, this fact is a source of mathematical difficulties but on the other hand it allows to obtain a closed system of equations with the coefficients which are determined intrinsically.

Returning to caps and caplets, two additional problems that must be tackled to integrate the above approach into the interest rate area are:

- How to approach a spectrum of caplets maturing at  $T$  and paying at  $T_1$  when the dynamics of each is specified under its own forward measure  $\mathbb{P}_{T_1}$ .
- How to use correlation to transfer feedback information from the immediately maturing caplet to later caplets.

**1.1. Derivative formulae.** Here are some formulae that will be required later, for the first and second partial derivatives ( $\partial_x$  stands for  $\frac{\partial}{\partial x}$  etc) of the BS call with respect to the underlying stock, strike and implied volatility. Starting with

$$\begin{aligned} \ell &= \ell(S, \sigma, \tau; K) = S\mathbf{N}(h_1) - K\mathbf{N}(h_2), \\ h_1 &= \frac{\ln \frac{S}{K}}{\sigma\sqrt{\tau}} + \frac{1}{2}\sigma\sqrt{\tau}, \quad h_2 = h_1 - \sigma\sqrt{\tau}, \\ \partial_S h_1 &= \partial_S h_2 = \frac{1}{S\sigma\sqrt{\tau}}, \quad \partial_K h_1 = \partial_K h_2 = -\frac{1}{K\sigma\sqrt{\tau}}, \\ \partial_\sigma h_2 &= \partial_\sigma h_1 - \sqrt{\tau}, \quad \partial_\sigma h_1 = -\frac{h_2}{\sigma}, \quad \partial_\sigma h_2 = -\frac{h_1}{\sigma}, \end{aligned}$$

where  $\mathbf{N}(\bullet)$  is the standard normal cumulative density function, and using

$$K\mathbf{N}'(h_2) = K\mathbf{N}'(h_1) \exp\left(h_1\sigma\sqrt{\tau} - \frac{1}{2}\sigma^2\tau\right) = S\mathbf{N}'(h_1),$$

the first partial derivatives of  $\ell$  with respect to  $S, \sigma$  and  $K$  are respectively

$$\begin{aligned} \partial_S \ell &= \mathbf{N}(h_1) + S\mathbf{N}'(h_1)\partial_S h_1 - K\mathbf{N}'(h_2)\partial_S h_2 = \mathbf{N}(h_1), \\ \partial_\sigma \ell &= S\mathbf{N}'(h_1)\partial_\sigma h_1 - K\mathbf{N}'(h_2)\partial_\sigma h_2 = \sqrt{\tau}K\mathbf{N}'(h_2) = \sqrt{\tau}S\mathbf{N}'(h_1), \\ \partial_K \ell &= S\mathbf{N}'(h_1)\partial_K h_1 - \mathbf{N}(h_2) - K\mathbf{N}'(h_2)\partial_K h_2 = -\mathbf{N}(h_2). \end{aligned}$$

Then, recalling that  $\mathbf{N}''(x) = -x\mathbf{N}'(x)$ , the second partial derivatives are

$$\begin{aligned}\partial_S^2 \ell &= \mathbf{N}'(h_1) \frac{\partial h_1}{\partial S} = \frac{1}{S\sigma\sqrt{\tau}} \mathbf{N}'(h_1), \\ \partial_K^2 \ell &= -\mathbf{N}'(h_2) \frac{\partial h_2}{\partial K} = \frac{1}{K\sigma\sqrt{\tau}} \mathbf{N}'(h_2) = \frac{S}{K^2\sigma\sqrt{\tau}} \mathbf{N}'(h_1), \\ \partial_\sigma^2 \ell &= \sqrt{\tau} S \mathbf{N}''(h_1) \frac{\partial h_1}{\partial \sigma} = \frac{S\sqrt{\tau}}{\sigma} h_1 h_2 \mathbf{N}'(h_1), \\ \partial_S \partial_K \ell &= \mathbf{N}'(h_1) \frac{\partial h_1}{\partial K} = -\frac{1}{K\sigma\sqrt{\tau}} \mathbf{N}'(h_1), \\ \partial_S \partial_\sigma \ell &= \mathbf{N}'(h_1) \frac{\partial h_1}{\partial \sigma} = -\frac{h_2}{\sigma} \mathbf{N}'(h_1), \\ \partial_K \partial_\sigma \ell &= -\mathbf{N}'(h_2) \frac{\partial h_2}{\partial \sigma} = \frac{h_1}{\sigma} \mathbf{N}'(h_2) = \frac{S h_1}{K\sigma} \mathbf{N}'(h_1).\end{aligned}$$

In addition we have the relation

$$\partial_\tau \ell = \frac{1}{2} \sigma^2 S^2 \partial_S^2 \ell = \frac{\sigma S}{2\sqrt{\tau}} \mathbf{N}'(h_1),$$

which comes from the BS partial differential equation, and holds for any option in the BS world.

*Exercise 1.1.* Show that in the normal Bachelier model where

$$\begin{aligned}dS_t &= \theta_t dW_t^{(1)}, \quad C_t = \mathbf{E} \{ [S_T - K]^+ | \mathcal{F}_t \} = \sigma\sqrt{T-t} \Phi\left(\frac{S_t - K}{\sigma\sqrt{T-t}}\right), \\ \text{with } \Phi(x) &= \int_{-\infty}^x \mathbf{N}(u) du = \mathbf{N}'(x) + x\mathbf{N}(x), \\ \text{and } \ell &= \ell(S, \sigma, \tau; K) = \sigma\sqrt{\tau} \Phi(h), \quad h = \frac{(S - K)}{\sigma\sqrt{\tau}},\end{aligned}$$

the equivalent expressions for the first and second derivatives of  $\ell$  with respect to  $S$  and  $\sigma$  are

$$\begin{aligned}\partial_S \ell &= -\partial_K \ell = \mathbf{N}(h), \quad \partial_\sigma \ell = \sqrt{\tau} \Phi(h) - \sqrt{\tau} h \mathbf{N}(h) = \sqrt{\tau} \mathbf{N}'(h), \\ \partial_S^2 \ell &= -\partial_S \partial_K \ell = \partial_K^2 \ell = \frac{1}{\sigma\sqrt{\tau}} \mathbf{N}'(h), \quad \partial_\sigma^2 \ell = -\sqrt{\tau} \mathbf{N}''(h) \frac{h}{\sigma} = \frac{\sqrt{\tau}}{\sigma} h^2 \mathbf{N}'(h), \\ \partial_S \partial_\sigma \ell &= -\partial_K \partial_\sigma \ell = -\frac{h}{\sigma} \mathbf{N}'(h), \quad \partial_\tau \ell = \frac{1}{2} \sigma^2 \partial_S^2 \ell = \frac{1}{2} \frac{\sigma}{\sqrt{\tau}} \mathbf{N}'(h).\end{aligned}$$

## 2. DYNAMICS OF THE IMPLIED VOLATILITY SURFACE

Assuming that the drift  $m(\cdot)$  and volvol  $v(\cdot)$  in (1.5) are well behaved functions, applying Ito to (1.3) produces the following SDE for a call option:

$$\begin{aligned}
dC_t &= \partial_\tau C_t dt + \partial_S C_t dS_t + \frac{1}{2} \partial_S^2 C_t d\langle S \rangle_t + \partial_\sigma C_t d\sigma_t + \frac{1}{2} \partial_\sigma^2 C_t d\langle \sigma \rangle_t + \partial_S \partial_\sigma C_t d\langle S, \sigma \rangle_t, \\
&= -\frac{\sigma_t S_t}{2\sqrt{T-t}} \mathbf{N}'(h_1) dt + \mathbf{N}(h_1) dS_t + \frac{1}{2} \frac{1}{S_t \sigma_t \sqrt{T-t}} \mathbf{N}'(h_1) d\langle S \rangle_t \\
&\quad + \sqrt{T-t} S_t \mathbf{N}'(h_1) d\sigma_t + \frac{1}{2} \frac{S_t \sqrt{T-t}}{\sigma_t} h_1 h_2 \mathbf{N}'(h_1) d\langle \sigma \rangle_t - \frac{h_2}{\sigma_t} \mathbf{N}'(h_1) d\langle S, \sigma \rangle_t, \\
&= \mathbf{N}(h_1) S_t \theta_t dW_t^{(1)} + \sqrt{T-t} S_t \mathbf{N}'(h_1) v_t^* dW_t \\
&\quad + \sqrt{T-t} S_t \mathbf{N}'(h_1) \left[ m_t + \frac{\theta_t^2}{2\sigma_t(T-t)} - \frac{\sigma_t}{2(T-t)} + \frac{h_1 h_2 |v_t|^2}{2\sigma_t} - \frac{h_2 \theta_t v_t^{(1)}}{\sigma_t \sqrt{T-t}} \right] dt.
\end{aligned} \tag{2.1}$$

For this to be a martingale under the arbitrage free measure  $\mathbb{P}$ , its drift must be zero, and so

$$\begin{aligned}
m_t &= \frac{1}{2\sigma_t(T-t)} \left[ \sigma_t^2 - \theta_t^2 - (T-t) h_1 h_2 |v_t|^2 + 2\sqrt{T-t} h_2 \theta_t v_t^{(1)} \right], \\
&= \frac{1}{2\sigma_t(T-t)} \left\{ \begin{array}{l} \sigma_t^2 - \theta_t^2 + \left[ \frac{1}{4} \sigma_t^2 (T-t)^2 - \frac{\ln^2 \frac{K}{S_t}}{\sigma_t^2} \right] |v_t|^2 \\ - \left[ \frac{1}{2} \sigma_t (T-t) + \frac{\ln \frac{K}{S_t}}{\sigma_t} \right] 2\theta_t v_t^{(1)} \end{array} \right\},
\end{aligned} \tag{2.2}$$

and the arbitrage free dynamics of  $C_t$  becomes

$$dC_t = \mathbf{N}(h_1) S_t \theta_t dW_t^{(1)} + \sqrt{T-t} S_t \mathbf{N}'(h_1) v_t^* dW_t. \tag{2.3}$$

The only reasonable choice for the **dependence** of the volvol  $v_t$  on the implied volatility  $\sigma_t$  is now seen to be linear because that removes the troublesome singularities in (2.2). Setting

$$v_t = v_t(T, K, S_t, \theta_t, \sigma_t) = \sigma_t u_t(T, K, S_t) = \sigma_t u_t,$$

from (1.4), (1.5) and (2.2) the dynamics of  $S_t, \theta_t$  and  $\sigma_t$  are therefore determined by the non-linear set of equations

$$\begin{aligned}
dS_t &= S_t \theta_t dW_t^{(1)}, \quad \theta_t = \sigma(t, t, S_t), \quad \lim_{t \rightarrow T} \sigma_t(T, K) < \infty, \\
d\sigma_t &= \frac{1}{2\sigma_t(T-t)} \left[ \begin{array}{l} \sigma_t^2 + \frac{1}{4} \sigma_t^4 (T-t)^2 |u_t|^2 - \sigma_t^2 (T-t) \theta_t u_t^{(1)} \\ - \left| \theta_t + u_t \ln \frac{K}{S_t} \right|^2 dt \end{array} \right] dt + \sigma_t u_t^* dW_t,
\end{aligned} \tag{2.4}$$

and we now analyse this system further.

## 3. DIFFERENT FORMULATIONS

Our aim in this section is to list four different formulations of the stochastic implied volatility problem, so as to have the flexibility of choosing one that is most convenient to the problem at hand. In (2.4) we derived SDEs for the implied volatility  $\sigma_t(T, K)$  expressed in terms of the “absolute parameters”  $T$  and  $K$ . But implied volatility can also be expressed “relatively” as follows.

For given constants  $x > 0$  and  $y$ , define  $\eta$  by the equations

$$\eta_t(x, y) = \sigma_t(t + x, e^y S_t), \quad \sigma_t(T, K) = \eta_t\left(T - t, \ln \frac{K}{S_t}\right).$$

The correct interpretation of this transformation is that  $\eta_t(\cdot, \cdot)$  is the relative volatility surface at time  $t$  as seen by an observer moving with the stock. The parameters

$$x = T - t \geq 0, \quad y = \ln \frac{K}{S_t},$$

are, respectively, relative maturity ( $x$  becomes zero as options mature) and log-moneyness ( $y$  is zero at-the-money and negative for out-of-the-money puts or in-the-money calls), and the relative surface at time  $t$  is obtained by plotting  $\eta_t(x, y)$  against  $x$  and  $y$ . Note that because

$$\theta_t = \eta_t(0, 0),$$

a system of SDEs for  $\eta_t$  will also include one for the spot volatility  $\theta_t$ .

To get SDEs for  $\eta_t = \eta_t(x, y)$  we need to make the parameters  $T$  and  $K$  in  $\sigma_t = \sigma_t(T, K)$  into stochastic variables  $T_t$  and  $K_t$  like

$$T_t = t + x, \quad dT_t = dt$$

$$K_t = e^y S_t, \quad dK_t = e^y dS_t = e^y S_t \theta_t dW_t^{(1)},$$

and then rewrite the SDE (2.4) for  $\sigma_t$  using the Ito-Venttsel formula described in Appendix A. As we shall see, that will require the following partial derivatives:

$$\begin{aligned} \partial_T \sigma_t(T, K) &= \partial_T \eta_t\left(T - t, \ln \frac{K}{S_t}\right) = \partial_x \eta_t(x, y); \\ \partial_K \sigma_t(T, K) &= \partial_K \eta_t\left(T - t, \ln \frac{K}{S_t}\right) = \frac{1}{K} \partial_y \eta_t(x, y) = \frac{1}{e^y S_t} \partial_y \eta_t(x, y); \\ \partial_K^2 \sigma_t(T, K) &= \partial_K^2 \eta_t\left(T - t, \ln \frac{K}{S_t}\right) = \partial_K \left\{ \frac{1}{K} \partial_y \eta_t\left(T - t, \ln \frac{K}{S_t}\right) \right\}, \\ &= \frac{1}{K^2} \left\{ \partial_y^2 \eta_t(x, y) - \partial_y \eta_t(x, y) \right\} = \frac{1}{e^{2y} S_t^2} \left\{ \partial_y^2 \eta_t(x, y) - \partial_y \eta_t(x, y) \right\}. \end{aligned} \tag{3.1}$$

We are now in a position to obtain four systems of SDEs describing the evolution of the stochastic implied volatility surface in terms of:

- The absolute implied volatility  $\sigma_t = \sigma_t(T, K)$  as in (2.4).
- The square of the absolute implied volatility multiplied by time to maturity.

$$\xi_t = \xi_t(T, K) = \sigma_t^2(T - t) = \sigma_t^2(T, K)(T - t).$$

- The relative implied volatility  $\eta_t = \eta_t(x, y)$ .
- The square of the relative implied volatility multiplied by relative maturity

$$\zeta_t = \zeta_t(x, y) = \eta_t^2 x = \eta_t^2(x, y) x.$$

Repeating (2.4), the  $\sigma_t$ - formulation is

$$\begin{aligned}
 & dS_t = S_t \theta_t dW_t^{(1)} \quad (\text{spot SDE}) \\
 & d\sigma_t = \frac{1}{2\sigma_t(T-t)} \left[ \sigma_t^2 - \left| \theta_t + u_t \ln \frac{K}{S_t} \right|^2 \right] dt \\
 & \quad + \left[ \frac{1}{8} \sigma_t^3 (T-t) |u_t|^2 - \frac{1}{2} \sigma_t \theta_t u_t^{(1)} \right] dt + \sigma_t u_t^* dW_t, \\
 & \sigma_0(T, K) \quad \text{specified (initial condition),} \\
 & \begin{cases} \theta_t = \sigma(t, t, S_t) & \text{and} \\ \lim_{t \rightarrow T} \sigma_t(T, K) < \infty \end{cases} \quad (\text{feedback}).
 \end{aligned} \tag{3.2}$$

Multiplying the SDE (3.2) by  $2\sigma_t(T-t)$  and using

$$d\xi_t = d[\sigma_t^2(T-t)] = (T-t)2\sigma_t d\sigma_t + (T-t)\sigma_t^2 |u_t|^2 dt - \sigma_t^2 dt,$$

produces the  $\xi$ -formulation

$$\begin{aligned}
 & \xi_t = \xi_t(T, K) = \sigma_t^2(T-t) \quad (\text{definition}) \\
 & dS_t = S_t \theta_t dW_t^{(1)} \quad (\text{spot SDE}) \\
 & d\xi_t = \xi_t \left\{ \left[ 1 + \frac{1}{4} \xi_t \right] |u_t|^2 - \theta_t u_t^{(1)} \right\} dt - \left| \theta_t + u_t \ln \frac{K}{S_t} \right|^2 dt + 2\xi_t u_t^* dW_t \\
 & \xi_0(T, K) \quad \text{specified (initial condition)} \\
 & \begin{cases} \xi_t(t, K) = 0 & \text{and} \\ \theta_t^2 = \partial_T \xi_t(t, K) \end{cases} \quad (\text{feedback}).
 \end{aligned} \tag{3.3}$$

In the SDE (3.2) for  $\sigma_t$  set

$$\begin{aligned}
 T_t &= t + x, \quad dT_t = dt, \\
 K_t &= e^y S_t, \quad dK_t = e^y dS_t = e^y S_t \theta_t dW_t^{(1)},
 \end{aligned}$$

and apply the Ito-Venttsel TheoremA to get

$$\begin{aligned}
 & d\sigma_t(T_t, K_t) = d\sigma_t(t+x, e^y S_t) = d\eta_t(x, y) \\
 & dS_t = S_t \theta_t dW_t^{(1)}, \quad \theta_t = \sigma_t(t, S_t), \quad \lim_{t \rightarrow T} \sigma_t(T, K) < \infty, \\
 & d\sigma_t = \frac{1}{2\sigma_t(T-t)} \left[ \sigma_t^2 - \theta_t^2 - |v_t|^2 \ln^2 \frac{K}{S_t} - 2\theta_t u_t^{(1)} \ln \frac{K}{S_t} \right] dt \\
 & \quad + \left[ \frac{1}{8} \sigma_t^3 (T-t) |u_t|^2 - \frac{1}{2} \sigma_t \theta_t u_t^{(1)} \right] dt + \sigma_t u_t^* dW_t \\
 & + \left[ \partial_T \sigma_t(T, K) + \frac{1}{2} \partial_K^2 \sigma_t(T, K) e^{2y} S_t^2 \theta_t^2 + \partial_K \left( \sigma_t v_t^{(1)} \right) e^y S_t \theta_t \right] dt \\
 & \quad + \partial_K \sigma_t(T, K) e^y S_t \theta_t dW_t^{(1)}.
 \end{aligned}$$

To express these equations solely in terms of  $\eta_t(x, y)$  and  $x, y$ , assume  $u_t$  is expressed in terms of  $x$  and  $y$  rather than  $T$  and  $K$ , substitute

$$x = T - t, \quad y = \ln \frac{K}{S_t}, \quad d(\ln S_t) = \theta_t dW_t^{(1)} - \frac{1}{2} \theta_t^2 dt,$$

and apply the change of variable formulae (3.1). The  $\sigma$ -formulation then becomes the  $\eta$ -formulation

$$\begin{aligned}
dS_t &= S_t \theta_t dW_t^{(1)}, \quad (\text{spot SDE}) \\
d\eta_t &= \frac{1}{2\eta_t x} \left\{ \eta_t^2 - |\theta_t + yv_t|^2 \right\} dt + \left\{ \frac{1}{8}\eta_t^3 x |u_t|^2 - \frac{1}{2}\eta_t \theta_t u_t^{(1)} \right\} dt + \eta_t u_t^* dW_t \\
&\quad + \left\{ \partial_x \eta_t + \frac{1}{2}\theta_t^2 \partial_y^2 \eta_t + \theta_t \partial_y \left( \eta_t u_t^{(1)} \right) \right\} dt + \partial_y \eta_t d(\ln S_t) \\
\eta_0(x, y) &\text{ specified (initial condition)} \\
\begin{cases} \theta_t = \eta_t(0, 0) \\ \lim_{x \rightarrow 0} \eta_t(x, y) < \infty \end{cases} &\quad (\text{feedback}).
\end{aligned} \tag{3.4}$$

Similarly, applying Ito-Venttsel to the  $\xi$ -formulation (3.3) produces the  $\zeta$ -formulation

$$\begin{aligned}
\zeta_t &= \zeta_t(x, y) = \eta_t^2(x, y) x \quad (\text{definition}) \\
dS_t &= S_t \theta_t dW_t^{(1)} \quad (\text{spot SDE}) \\
d\zeta_t &= \zeta_t \left[ \left( 1 + \frac{1}{4}\zeta_t \right) |u_t|^2 - \theta_t u_t^{(1)} \right] dt - |\theta_t + yu_t|^2 dt \\
&\quad + \left[ \partial_x \zeta_t + \frac{1}{2}\theta_t^2 \partial_y^2 \zeta_t + 2\theta_t \partial_y \left( \zeta_t u_t^{(1)} \right) \right] dt + 2\zeta_t u_t^* dW_t + \partial_y \zeta_t d(\ln S_t) \\
\zeta_0(x, y) &\text{ specified (initial condition)} \\
\begin{cases} \zeta(t, 0, y) = 0 \quad \text{and} \\ \theta_t^2 = \partial_x \zeta_t(0, 0) \end{cases} &\quad (\text{feedback}).
\end{aligned} \tag{3.5}$$

*Exercise 3.1.* In the Bachelier model define moneyness by

$$y = K - S_t,$$

and also take the volvol  $v_t$  to be linear in the implied volatility, that is

$$v_t = u_t \sigma_t.$$

Show that the equivalent formulations for  $\sigma_t, \xi_t, \eta_t$  and  $\zeta_t$  have

$$\begin{aligned}
d\sigma_t &= \frac{1}{2\sigma_t(T-t)} \left[ \sigma_t^2 - |\theta_t + (K - S_t) u_t|^2 \right] dt + \sigma_t u_t^* dW_t, \\
d\xi_t &= \xi_t |u_t|^2 dt - |\theta_t + (K - S_t) u_t|^2 dt + 2\xi_t u_t^* dW_t, \\
d\eta_t &= \frac{1}{2\eta_t x} \left[ \eta_t^2 - |\theta_t + yu_t|^2 \right] dt + \eta_t u_t^* dW_t \\
&\quad + \left[ \partial_x \eta_t + \frac{1}{2}\theta_t^2 \partial_y^2 \eta_t + \theta_t \partial_y \left( \eta_t u_t^{(1)} \right) \right] dt + \partial_y \eta_t dS_t, \\
d\zeta_t &= \zeta_t |u_t|^2 dt - |\theta_t + yu_t|^2 dt + 2\zeta_t u_t^* dW_t \\
&\quad + \left[ \partial_x \zeta_t + \frac{1}{2}\theta_t^2 \partial_y^2 \zeta_t + 2\theta_t \partial_y \left( \zeta_t u_t^{(1)} \right) \right] dt + \partial_y \zeta_t dS_t.
\end{aligned}$$

*Remark 3.2.* Considering the  $\xi$ -formulations in the Bachelier and Black-Scholes models

$$\begin{aligned}
d\xi_t &= \xi_t |u_t|^2 dt - |\theta_t + (K - S_t) u_t|^2 dt + 2\xi_t u_t^* dW_t \quad (\text{Bachelier}) \\
d\xi_t &= \xi_t \left\{ \left[ 1 + \frac{1}{4}\xi_t \right] |u_t|^2 - \theta_t u_t^{(1)} \right\} dt - \left| \theta_t + u_t \ln \frac{K}{S_t} \right|^2 dt + 2\xi_t u_t^* dW_t \quad (\text{BS}),
\end{aligned}$$

one is struck by the similarity between the two models. The main difference is the unpleasant non-linear term

$$\frac{1}{4}\xi_t^2 |u_t|^2 dt$$

in the drift of the Black-Scholes equation.that will cause us some problems.

*Remark 3.3.* Let us comment on the feedback condition which was introduced in Section 3. Note that on the boundary  $x = 0$ , the condition

$$\zeta(t, 0, y) = 0 \Rightarrow \partial_y \zeta_t(0, y) = \partial_y^2 \zeta_t(0, y) = \partial_y \left( \zeta_t v_t^{(1)} \right) = 0$$

for all  $y$ , so that the SDE for  $\zeta_t(0, y)$  on the “leading edge”  $x = 0$  reduces to

$$d\zeta_t(0, y) = -|\theta_t + yu_t|^2 dt + \partial_x \zeta_t(0, y) dt.$$

But for  $\zeta_t(0, y)$  to remain zero  $\forall t \geq 0$  the increment  $d\zeta_t(0, y)$  itself must also be zero, which means

$$\partial_x \zeta_t(0, y) = |\theta_t + yu_t|^2, \quad \forall y.$$

In terms of  $\eta_t$ , the boundary conditions at  $x = 0$  are

$$\eta_t(0, y) = |\theta_t + yu_t| < \infty,$$

which we might have obtained directly by asking that the drift in (3.4) not have a singularity at  $x = 0$ . Various skews and smiles of the familiar “upward hook” type can be obtained by changing the correlation which shifts the vertex of the underlying parabola. The initial volatility surface must of course satisfy this equation, which gives useful information about how  $u_t$  varies with  $y$  near the boundary  $x = 0$  (remember also that  $u_t(0, y)$  can depend on  $y$ ).

#### 4. SOME PROPERTIES OF SOLUTIONS

Let  $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$  be a filtered probability space with the filtration satisfying the usual conditions. We assume that this space carries a two-dimensional Wiener process  $(W_t) = (W_t^{(1)}, W_t^{(2)})$ . In this section we will consider the equation

$$\begin{aligned} d\zeta_t &= \zeta_t \left( |u_t|^2 - \theta_t u_t^{(1)} \right) dt + 2\zeta_t u_t^* dW_t - \left( |\theta_t + yu_t|^2 + \frac{1}{2}\theta_t^2 \right) dt \\ &+ \left( \frac{\partial \zeta_t}{\partial x} + \frac{1}{2}\theta_t^2 \frac{\partial^2 \zeta_t}{\partial y^2} + 2\theta_t \frac{\partial \left( \zeta_t u_t^{(1)} \right)}{\partial y} \right) dt + \theta_t \frac{\partial \zeta_t}{\partial y} dW_t^{(1)} + \frac{1}{4}\zeta_t^2 dt, \end{aligned} \quad (4.1)$$

where  $\theta_t^2 = \partial_x \zeta_t(0, 0)$ ,

$$\zeta_0(x, y) = f(x, y), \quad \text{and} \quad \zeta_t(0, y) = 0.$$

This equation has a solution only if the process  $(u_t) = (u_t^{(1)}, u_t^{(2)})$  is chosen in a special way. We will start with a simpler problem. Namely, we will study equation (4.1) without the boundary condition  $\zeta_t(0, y) = 0$  and with the process  $(u_t)$  given in advance and such that

**Hypothesis 4.1.** *The process  $(u_t) = (u_t(x, y))$  is adapted for each  $(x, y)$ , continuous in  $(t, x, y)$ . Moreover, we assume that the process  $\partial_y u_t^{(1)}(x, y)$  is well defined and continuous in  $(t, x, y)$ .*

By a local solution to (4.1) we mean a process  $\{\zeta_t(x, y) : x \geq 0, y \in \mathbb{R}\}$  and a stopping time  $\tau$  such that the following holds.

1. The process  $(\zeta_t(x, y))$  is continuous in  $(t, x, y) \in [0, \tau) \times \mathbb{R}_+ \times \mathbb{R}$ .
2. For each  $t < \tau$  the function  $\zeta_t(\cdot, \cdot) \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$  and the processes

$$\frac{\partial \zeta_s}{\partial x}(x, y), \quad \frac{\partial \zeta_s^2}{\partial y^2}(x, y) \quad \text{and} \quad \frac{\partial \zeta_s}{\partial y}(x, y)$$

are continuous in  $s < \tau$  for each  $(x, y) \in \mathbb{R}_+ \times \mathbb{R}$ .

3. For each  $f \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$  and  $(x, y) \in \mathbb{R}_+ \times \mathbb{R}$  we have for  $t < \tau$

$$\begin{aligned} \zeta_t = f &+ \int_0^t \zeta_s \left( \frac{1}{4} \zeta_s + |u_s|^2 - \theta_s u_s^{(1)} \right) ds + \int_0^t \theta_s \partial_y \zeta_s dW_s^{(1)} + \int_0^t 2\zeta_s u_s^* dW_s \\ &+ \int_0^t \left( \partial_x \zeta_s + \frac{1}{2} \theta_s^2 \partial_y^2 \zeta_s + 2\theta_s \partial_y (\zeta_s u_s^{(1)}) \right) ds - \int_0^t \left( |\theta_s + y u_s|^2 + \frac{1}{2} \theta_s^2 \right) ds, \end{aligned}$$

where the dependence on  $(x, y)$  is suppressed, wherever possible.

**Theorem 4.2.** *Assume Hypothesis 4.1. Then for each  $f \in C^{1,2}(\mathbb{R}_+ \times \mathbb{R})$  there exists a unique local solution to (4.1).*

**Lemma 4.3.** *Assume Hypothesis 4.1. Let*

$$M_t = \int_0^t \theta_s dW_s^{(1)},$$

and

$$\begin{aligned} N_s^t(x, y) &= 2 \int_s^t u_r^*(x + t - r, y - M_r) dW_r - \int_s^t |u_r(x + t - r, y - S_r)|^2 dr \\ &+ \int_s^t \left( \frac{1}{4} \zeta_r(x + t - r, y - M_r) - \theta_r u_r^{(1)}(x + t - r, y - M_r) \right) dr. \end{aligned}$$

Then

$$\zeta_t(x, y) = X_t(x, y + M_t),$$

where

$$X_t(x, y) = e^{N_0^t(x, y)} f(t + x, y) - \int_0^t e^{N_s^t(x, y)} \left( |\theta_s + (y - S_s) u_s(x + t - s, y - S_s)|^2 + \frac{1}{2} \theta_s^2 \right) ds.$$

*Exercise 4.4.* In the Bachelier Model denote

$$N_t = \int_0^t 2u_s^* dW_s - \int_0^t |u_s|^2 ds.$$

Show that the process  $\xi$  if exists, must satisfy an integral equation

$$\xi_t = e^{N_t} f(T, K) - e^{N_t} \int_t^T e^{-N_s} |\theta_s - (K - S_s) u_s|^2 ds,$$

and therefore the feedback condition takes the form

$$f(T, K) = \int_0^T e^{-N_s} |\theta_s - (K - S_s) u_s|^2 ds.$$

Finally, show that the process  $(\xi_t)$  exists for all times provided the process  $(u_t)$  is locally bounded in  $t, T, K \geq 0$ ,

**Proposition 4.5.** *Assume Hypothesis 4.1. Moreover, assume that the process  $(u_t)$  is chosen in such a way that  $\zeta_t(0, y) = 0$  for all  $t < \tau$  and  $y \in \mathbb{R}$ . Then the unique local solution  $(\zeta_t)$  is strictly positive: for each  $x \geq 0, y \in \mathbb{R}$  and  $t < \tau$  we have  $\zeta_t(x, y) \geq 0$ .*

*Proof.* Let  $(\zeta_t)$  be a local solution to (4.1) and let

$$dS_t = \theta_t S_t dW_t^{(1)}, \quad t < \tau.$$

It follows from Section 3 that for  $t < \tau$

$$\zeta_t(t, x) = \xi_t(t + x, e^y S_t),$$

and therefore it is enough to show that  $\xi_t(T, K) \geq 0$  for  $t < \tau$ . Let

$$L_t = L_t(T, K) = \int_0^t 2u_s^* dW_s - \int_0^t \left( \left(1 - \frac{1}{4}\xi_s\right) |u_s|^2 + \theta_s u_s^{(1)} \right) ds.$$

Then

$$\xi_t(T, K) = e^{L_t} f(T, K) - e^{L_t} \int_0^T e^{-L_s} \left| \theta_s + u_s \ln \frac{K}{S_s} \right|^2 ds.$$

Since  $\xi_T(T, K) = 0$  we find that

$$f(T, K) - \int_0^t e^{-L_s} \left| \theta_s + u_s \ln \frac{K}{S_s} \right|^2 ds = \int_t^T e^{-L_s} \left| \theta_s + u_s \ln \frac{K}{S_s} \right|^2 ds.$$

Hence

$$\xi_t(T, K) = e^{L_t} \int_t^T e^{-L_s} e^{-L_s} \left| \theta_s + u_s \ln \frac{K}{S_s} \right|^2 ds > 0,$$

for all  $t < \tau$ . □

## 5. TOY MODEL

Start with the  $\zeta$ -formulation in the Bachelier model (see Exercises 1.1 and 3.1), and assume the volvol  $u_t$  is stochastic but independent of both  $x$  and  $y$ :

$$d\zeta_t = \zeta_t |u_t|^2 dt - |\theta_t + y u_t|^2 dt + \partial_y \zeta_t \theta_t dW_t^{(1)} + 2\zeta_t u_t^* dW_t \quad (5.1)$$

$$+ \left[ \partial_x \zeta_t + \frac{1}{2} \theta_t^2 \partial_y^2 \zeta_t + 2\theta_t \partial_y \left( \zeta_t u_t^{(1)} \right) \right] dt,$$

$$\zeta_0(x, y) = f(x, y), \quad (\text{initial condition})$$

$$\partial_x \zeta_t(0, y) = |\theta_t + y u_t|^2. \quad (\text{feedback})$$

If the initial implied volatility surface  $\zeta_0(x, y) = f(x, y)$  is quadratic in  $y$  then clearly  $\zeta_t(x, y)$  will remain quadratic in  $y$  because all terms in the SDE and both initial and feedback conditions appearing in (5.1) are linear in  $\zeta$  or quadratic in  $y$ .

So suppose the initial implied volatility surface is given by

$$\eta_0^2 = \eta_0^2(x, y) = \frac{(1 - e^{-2\lambda x})}{2\lambda x} |\theta_0 + y u_0|^2,$$

$$\text{or } \zeta_0 = \zeta_0(x, y) = f(x, y) = \frac{1}{2\lambda} (1 - e^{-2\lambda x}) |\theta_0 + y u_0|^2,$$

For values of  $\theta_0$  around 20%,  $u_0$  between 20% and 40% and  $\lambda$  about .2, such initial surfaces are semi-believable. Setting

$$M_t = \int_0^t \theta_t dW_t^{(1)}, \quad N_t = \int_0^t 2u_s^* dW_s - \int_0^t |u_s|^2 ds,$$

a solution to (5.1) for all  $x \geq 0$  and  $y \in \mathbb{R}$  is

$$\zeta_t(x, y) = e^{N_t} f(t+x, y+M_t) - e^{N_t} \int_0^t e^{-N_s} |\theta_s + (y+M_t-M_s)u_s|^2 ds. \quad (5.2)$$

Differentiating with respect to  $x$

$$\frac{\partial}{\partial x} \zeta_t(x, y) = e^{N_t} \frac{\partial}{\partial x} f(t+x, y+M_t) = e^{N_t} e^{-2\lambda(t+x)} |\theta_0 + (y+M_t)u_0|^2,$$

and so the feedback condition at  $x=0$  gives

$$\frac{\partial}{\partial x} \zeta_t(0, y) = |\theta_t + yu_t|^2 = e^{N_t-2\lambda t} |\theta_0 + (y+M_t)u_0|^2,$$

or

$$\theta_t^2 + 2y\theta_t u_t^{(1)} + y^2 |u_t|^2 = e^{N_t-2\lambda t} \left[ \begin{array}{l} (\theta_0^2 + 2\theta_0 u_0^{(1)} M_t + |u_0|^2 M_t^2) \\ + 2y (\theta_0 u_0^{(1)} + |u_0|^2 M_t) + y^2 |u_0|^2 \end{array} \right].$$

Comparing coefficients of powers of  $y$  yields

$$\begin{aligned} |u_t| &= e^{\frac{1}{2}N_t - \lambda t} |u_0|, \\ \theta_t &= e^{\frac{1}{2}N_t - \lambda t} |\theta_0 + M_t u_0|, \\ u_t^{(1)} &= e^{\frac{1}{2}N_t - \lambda t} \frac{(\theta_0 u_0^{(1)} + |u_0|^2 M_t)}{|\theta_0 + M_t u_0|}, \\ e^{\frac{1}{2}N_t} &= \mathcal{E} \left( \int_0^t u_s^* dW_s \right), \quad M_t = \int_0^t \theta_t dW_t^{(1)}. \end{aligned} \quad (5.3)$$

Re-expressing the equation for  $|u_t|$  as

$$\begin{aligned} |u_t| &= |u_0| \mathcal{E} \left( \int_0^t [u_s^* dW_s - \lambda dt] \right), \\ d|u_t| &= |u_t| [u_s^* dW_s - \lambda dt], \end{aligned}$$

and introducing the new Brownian motion  $\widetilde{W}_t = \int_0^t \frac{u_s^* dW_s}{|u_s|}$ , gives

$$d|u_t| = |u_t| \left[ |u_t| d\widetilde{W}_t - \lambda dt \right] = |u_t|^2 \left[ d\widetilde{W}_t - \frac{\lambda}{|u_t|} dt \right],$$

which in turn, after applying Girsanov ( $d\widehat{W}_t = d\widetilde{W}_t - \frac{\lambda}{|u_t|} dt$ ), has form

$$d|u_t| = |u_t|^2 d\widehat{W}_t,$$

which does not explode in finite time.

In this example it is relatively easy to simulate the equations (5.3): simply increment the processes  $M_t$  and  $N_t$ , calculate the values of  $\theta_t$  and  $u_t = \left( u_t^{(1)}, u_t^{(2)} \right)^*$  from the feedback equations, and increment again.

## 6. APPLICATION TO BGM

To apply the above results to interest rates, first focus on what will be martingales and under what measures in a stochastic volatility version of BGM. As mentioned in the introduction, from [1] the Libor forward rates  $L(t, T)$  must be positive martingales under  $\mathbb{P}_{T_1}$  and can be taken to satisfy SDEs like

$$\frac{dL(t, T)}{L(t, T)} = \theta_t^*(T) dW_{T_1}(t),$$

where the  $\theta_t(T)$  are stochastic. We emphasise that here the maturity dependent volatilities  $\theta_t(T)$  are vectors, unlike the spot volatility  $\theta_t$  used for stocks above.

The **Black convention** for quoting cap prices in the presence of a volatility smile or skew is similar to that for stocks. The volatility in the Black cap formula, which adds component caplet values, is adjusted to produce the correct price. To analyse further, return to the standard lognormal BGM model in which  $\theta_t(T)$  is deterministic. The present value  $\text{Cpl}_t(T)$  of a caplet struck at  $\kappa$ , maturing at  $T$ , and paying at  $T_1 = T + \delta$  is given by the Black caplet formula

$$\begin{aligned} \text{Cpl}_t(T) &= P(t, T_1) \ell\{L(t, T), \sigma_t(T, \kappa), T - t; \kappa\}, \\ \sigma_t(T, \kappa) &= \sqrt{\frac{1}{(T-t)} \int_t^T |\theta_s(T)|^2 ds}. \end{aligned} \quad (6.1)$$

Suppose we can break the cap skew down into a caplet skew by distributing the cap prices at different strikes into Black caplet prices in such a way that the corresponding caplet volatility profile  $\sigma_t(T, \kappa)$  plotted against  $T$  and  $\kappa$ , is reasonably smooth (this step could very well involve some heroic numerical work). Then (after interpolation, if necessary) we can assume that for all maturities  $T$  we have an initial implied volatility surface  $\sigma_0 = \sigma_0(T, \kappa)$  which can be input as a start parameter.

Now suppose, following on from our work on stocks above, that the implied volatilities  $\sigma_t = \sigma_t(T, \kappa)$  satisfy SDEs of the form

$$\begin{aligned} d\sigma_t &= m_t(T, \kappa, L(t, T), \theta_t, \sigma_t) dt + \sigma_t u_t(T, \kappa, L(t, T), \theta_t)^* dW_{T_1}(t), \\ &= m_t dt + \sigma_t u_t^* dW_{T_1}(t) \end{aligned}$$

under the forward measures  $\mathbb{P}_{T_1}$ , and that caplet values are given by (6.1) with  $\sigma_t = \sigma_t(T, \kappa)$  now stochastic. Because the caplets  $\text{Cpl}_t(T)$  are assets, their present values divided by the numeraire  $P(t, T_1)$  must be martingales under the forward measure  $\mathbb{P}_{T_1}$ . That is, for all positive  $T$  and  $\kappa$ , the expression

$$\ell\{L(t, T), \sigma_t(T, \kappa), T - t; \kappa\}$$

must be a martingale under the  $\mathbb{P}_{T_1}$  forward measure. Similarly to (3.3), that leads to the system

$$\begin{aligned} \xi_t &= \xi_t(T, \kappa) = \sigma_t^2(T - t), \quad dL(t, T) = L(t, T) \theta^*(t, T) dW_{T_1}(t), \\ d\xi_t &= \xi_t \left\{ \left[ 1 + \frac{1}{4} \xi_t \right] |u_t|^2 - \theta_t u_t^{(1)} \right\} dt - \left| \theta_t + u_t \ln \frac{\kappa}{L(t, T)} \right|^2 dt + 2\xi_t u_t^* dW_{T_1}(t), \\ \xi_0(T, \kappa) &= T \sigma_0^2(T, \kappa) = f(T, \kappa) \quad \text{specified (initial condition),} \\ \xi_T(T, \kappa) &= 0 \quad \forall T \quad \text{(feedback).} \end{aligned} \quad (6.2)$$

The Libor forward rate volatility  $\theta_t(T)$  and the implied volvol  $u_t(T, \kappa)$  must now be linked into a feedback loop, otherwise the system (6.2) for  $L(t, T)$  and its derivative caplet volatilities will be under-specified. Assuming  $\theta_t(T)$  and  $u_t(T, \kappa)$  are well defined, a formal solution to (6.2) is

$$\begin{aligned}\xi_t(T, \kappa) &= e^{N_t(T, \kappa)} f(T, \kappa) - e^{N_t(T, \kappa)} \int_0^t e^{-N_s(T, \kappa)} \left| \theta_s + u_s(T, \kappa) \ln \frac{\kappa}{L(s, T)} \right|^2 ds, \\ N_t(T, \kappa) &= \int_0^t 2 u_s^*(T, \kappa) dW_{T_1}(s) - \left\{ \left[ 1 - \frac{1}{4} \xi_s(T, \kappa) \right] |u_s(T, \kappa)|^2 + \theta_s u_s^{(1)}(T, \kappa) \right\} ds.\end{aligned}$$

The feedback condition  $\xi_T(T, \kappa) = 0$  for all maturities  $T$  implies

$$\begin{aligned}f(T, \kappa) &= \int_0^T e^{-N_s(T, \kappa)} \left| \theta_s(T) + u_s(T, \kappa) \ln \frac{\kappa}{L(s, T)} \right|^2 ds, \quad \text{and} \\ \partial_T f(T, \kappa) &= e^{-N_T(T, \kappa)} \left| \theta_T + u_T(T, \kappa) \ln \frac{\kappa}{L(T, T)} \right|^2 \\ &+ \left[ \int_0^t \partial_T e^{-N_s(T, \kappa)} \left| \theta_s + u_s(T, \kappa) \ln \frac{\kappa}{L(s, T)} \right|^2 \right]_{t=T}.\end{aligned}$$

which means

$$\begin{aligned}\partial_T \xi_t(T, \kappa) &= \xi_t(T, \kappa) \partial_T N_t(T, \kappa) \\ &+ e^{N_t(T, \kappa)} \partial_T \left\{ f(T, \kappa) - \int_0^t e^{-N_s(T, \kappa)} \left| \theta_s + u_s(T, \kappa) \ln \frac{\kappa}{L(s, T)} \right|^2 ds \right\}, \\ \text{and } [\partial_T \xi_t(T, \kappa)]_{t=T} &= e^{N_T(T, \kappa)} \left\{ e^{-N_T(T, \kappa)} \left| \theta_T + u_T(T, \kappa) \ln \frac{\kappa}{L(T, T)} \right|^2 \right\}, \\ &= \left| \theta_T + u_T(T, \kappa) \ln \frac{\kappa}{L(T, T)} \right|^2.\end{aligned}$$

In other words

$$\xi_T(T, \kappa) = 0 \Rightarrow [\partial_T \xi_t(T, \kappa)]_{t=T} = \left| \theta_T(T) + u_T(T, \kappa) \ln \frac{\kappa}{L(T, T)} \right|^2. \quad (6.3)$$

Putting  $\kappa = L(T, T)$  in (6.3) yields the **Libor volatility link**

$$|\theta_T(T)|^2 = [\partial_T \xi_t(T, L(T, T))]_{t=T} \quad \text{or} \quad |\theta_t(t)| = \sigma_t(t, L(t, t)), \quad (6.4)$$

which can be extended to include  $\theta_t(T)$  at later maturities using correlation information. Suppose that from historical data analysis with a standard BGM model, we have constructed a deterministic vector volatility function  $\gamma_t(T)$  which reflects the correlation structure we would like our model to exhibit. Namely, that the instantaneous correlation at time  $t$  between the  $T_i$  and  $T_j$  Libor forward rates is

$$\rho_t(T_i, T_j) = \frac{\gamma_t^*(T_i) \gamma_t(T_j)}{\|\gamma_t(T_i)\| \|\gamma_t(T_j)\|}.$$

For the forward rate volatility vector try

$$\theta_t(T) = \gamma_t(T) \psi_t, \quad (6.5)$$

where  $\psi_t$  is a scalar stochastic variable free to be determined by feedback (the deterministic vector  $\gamma_t(T)$  is of course already fully specified from the historical data analysis). From (6.4)

$$\psi_t = \frac{|\theta_t(t)|}{|\gamma_t(t)|} = \frac{\sigma_t(t, L(t, t))}{|\gamma_t(t)|},$$

and so

$$\theta_t(T) = \frac{\sigma_t(t, L(t, t))}{|\gamma_t(t)|} \gamma_t(T).$$

Moreover

$$\begin{aligned} \frac{d\langle L(\cdot, T) \rangle_t}{L^2(t, T)} &= |\theta_t(T)|^2 dt = \psi_t^2 |\gamma_t(T)|^2 dt, \\ \frac{d\langle L(\bullet, T_i), L(\bullet, T_j) \rangle_t}{L(t, T_i) L(t, T_j)} &= \theta_t^*(T_i) \theta_t(T_j) dt = \psi_t^2 \gamma_t^*(T_i) \gamma_t(T_j) dt, \end{aligned}$$

which returns the required instantaneous correlation

$$\frac{\theta_t^*(T_i) \theta_t(T_j)}{\|\theta_t(T_i)\| \|\theta_t(T_j)\|} = \rho_t(T_i, T_j).$$

The **volvol link** for  $u_t(T, \kappa)$  can be specified in a similar, but somewhat looser, fashion.. The spot volvol  $u_t(t, \kappa)$  is largely determined by the feedback condition (6.3), although (as in the stock case) there is still considerable freedom to engineer its dependence on the strike  $\kappa$ , and its distribution into components. For later maturities,  $u_t(T, \kappa)$  can be specified in term of its spot value  $u_t(t, \kappa)$  so as to exhibit the same sort of decay with respect to maturity  $T$  that is seen in historical data.

Hence the **BGM  $\xi$ -formulation**

$\begin{aligned} \xi_t &= \xi_t(T, K) = \sigma_t^2(T-t) \quad (\text{definition } \xi) \\ \theta_t(T) &= \psi_t \gamma_t(T) \quad (\text{volatility of forwards}) \\ dL(t, T) &= L(t, T) \theta_t^*(T) dW_{T_1}(t) \quad (\text{forward dynamics}) \\ d\xi_t &= \xi_t \left\{ \left[ 1 + \frac{1}{4} \xi_t \right]  u_t ^2 - \theta_t u_t^{(1)} \right\} dt - \left  \theta_t + u_t \ln \frac{\kappa}{L(t, T)} \right ^2 dt + 2\xi_t u_t^* dW_{T_1}(t) \\ \xi_0(T, \kappa) &= T \sigma_0^2(T, \kappa) = f(T, \kappa) \quad \text{specified} \quad (\text{initial condition}) \\ \left\{ \begin{array}{l} \xi_T(T, \kappa) = 0 \\ [\partial_T \xi_t(T, \kappa)]_{t=T} = \left  \theta_T(T) + u_T(T, \kappa) \ln \frac{\kappa}{L(T, T)} \right ^2 \end{array} \right. & \quad (\text{feedback}). \end{aligned}$	(6.6)
---	-------

## 7. MARGINAL DISTRIBUTIONS

Let  $p^*(S, T)$  denote the marginal distribution, so the price  $C$  of a European Call with strike  $K$  and expiry  $T$  can be written as

$$C(K, T) = \int_0^\infty \max(S - K) p^*(S, T) dS.$$

Following Breeden and Litzenberger [3] the marginal distribution can be recovered by

$$\begin{aligned} C(K, T) &= \int_K^\infty (S - K)p^*(S, T)ds \\ \frac{\partial C(K, T)}{\partial K} &= - \int_K^\infty p^*(S, T)dS \\ \frac{\partial^2 C(K, T)}{\partial K^2} &= p^*(K, T). \end{aligned}$$

**7.1. Bachelier Model.** In the Bachelier model

$$C(\sigma, K, t, T) = \sigma\sqrt{(T-t)}\Phi(h) \tag{7.1}$$

where

$$h = \frac{S - K}{\sigma\sqrt{T-t}},$$

$$\Phi(u) = \int_{-\infty}^x \mathbf{N}(u)du = x\mathbf{N}(x) + \mathbf{N}'(x),$$

and  $\mathbf{N}$  is the standard normal cumulative density function.

If  $\sigma$  does not depend on the strike  $K$  then

$$\frac{\partial^2 C}{\partial K^2} = \frac{\mathbf{N}'(h)}{\sigma\sqrt{T-t}}.$$

In our model  $\sigma = \sigma(K, T)$  depends on the strike  $K$  (and other variables). Noting that

$$\xi = \sigma^2(T-t),$$

the initial form for  $\sigma$  is implied by the initial volatility surface  $\xi(x, y) = f(x, y)$  where  $x = T-t$  is the maturity and  $y = K - S$  is the moneyness.

**7.2. Derivatives.** To use the Breeden and Litzenberger result we need

$$\frac{\partial^2 C(\sigma(K), K, 0, T)}{\partial K^2}.$$

As  $K = S + y$  it does not matter if we differentiate w.r.t  $y$  or  $K$ . Using  $\xi = \sigma^2x$  and

$$\begin{aligned} \frac{\partial \sigma}{\partial y} &= \frac{1}{2\sigma\sqrt{x}} \frac{\partial f}{\partial y} \\ \frac{\partial^2 \sigma}{\partial y^2} &= \frac{-1}{4\sigma^3x^2} \left( \frac{\partial f}{\partial y} \right)^2 + \frac{1}{2\sigma x} \frac{\partial^2 f}{\partial y^2}. \end{aligned}$$

Also

$$\frac{\partial h}{\partial y} = \frac{-1}{\sigma\sqrt{x}} + \frac{y}{\sigma^2\sqrt{x}} \frac{\partial \sigma}{\partial y}.$$

Writing  $C$  for  $C(\sigma(K), y, x, T)$  the Bachelier formula (7.1) eventually gives

$$\begin{aligned} \frac{\partial C}{\partial K} &= -\mathbf{N}(h) + \left( \frac{C + y\mathbf{N}(h)}{\sigma} \right) \frac{\partial \sigma}{\partial y} \\ \frac{\partial^2 C}{\partial K^2} &= \frac{\mathbf{N}'(h)}{\sigma\sqrt{x}} - \frac{2y\mathbf{N}'(h)}{\sigma^2\sqrt{x}} \frac{\partial \sigma}{\partial y} + \frac{y^2\mathbf{N}'(h)}{\sigma^3\sqrt{x}} \left( \frac{\partial \sigma}{\partial y} \right)^2 + \left( \frac{C + y\mathbf{N}(h)}{\sigma} \right) \frac{\partial^2 \sigma}{\partial y^2}. \end{aligned}$$

The marginal distribution implied by the initial volatility surface is then obtained by setting  $t = 0$ , so  $x = T$ , and  $y = K - S_0$ .

**7.3. Numerical results.** In the simulations the initial volatility surface is given by

$$f(x, y) = \frac{1 - e^{-2\lambda x}}{2\lambda} q(y)$$

where

$$q(y) = \theta_0^2 + 2\rho\theta_0\nu y + \nu^2 y^2,$$

where  $\nu$  is the volvol,  $\lambda$  controls the flattening out as maturity  $x$  increases, and  $\rho$  is the correlation in the quadratic dependence of the surface on moneyness  $y$ . Thus initially at  $t = 0$

$$\sigma^2(K) = \frac{1}{T} \left[ \frac{1 - e^{-2\lambda T}}{2\lambda} \theta_0^2 + 2\rho\theta_0\nu(K - S_0) + \nu^2(K - S_0)^2 \right].$$

The Bachelier model was simulated over 5 years using 400 time steps. The parameters in the initial volatility surface were  $\nu = 20\%$ ,  $\lambda = 0.25$ ,  $\theta_0 = 2\%$  and different values for the correlation  $\rho$ . Figures 1, 2 and 3 show the initial volatility surface and the empirical (using 100,000 simulations) and analytic marginal distributions at 2.5 and 5 years for  $\rho = 0, -0.7, 0.5$  respectively. The agreement in all cases is remarkable.

#### APPENDIX A. ITO-VENTTSEL FORMULA

To write down an SDE for  $\eta_t = \eta(t, x, y)$  we will need a generalization of the Ito-Venttsel formula as derived in [6] and [8].

**Theorem A.1.** *Let  $W_t$  be multi-dimensional Brownian motion. Suppose  $F(t, u)$  is twice differentiable with respect to the parameter  $u$  and satisfies the SDE*

$$dF(t, u) = A(t, u) dt + B^*(t, u) dW_t.$$

If  $u_t$  satisfies the SDE

$$du_t = C(t, u_t) dt + D^*(t, u_t) dW_t,$$

then an SDE for  $F(t, u_t)$  is

$$\begin{aligned} dF(t, u_t) &= A(t, u_t) dt + B^*(t, u_t) dW_t \\ &+ \frac{\partial}{\partial u} F(t, u_t) du_t + \frac{1}{2} \frac{\partial^2}{\partial u^2} F(t, u_t) |D(t, u_t)|^2 dt \\ &+ \frac{\partial}{\partial u} B^*(t, u_t) D(t, u_t) dt. \end{aligned}$$

#### REFERENCES

- [1] Brace A, Gatarek D & Musiela M (1997) "The market model of interest rate dynamics" *Math Finance* 7, p127-154.
- [2] Balland P & Hughston LP (2000) "Pricing and hedging with a sticky-delta smile" *April Risk Conference, Paris*
- [3] Breeden and Litzenberger (1978) "Prices of Contingent claims implied in option prices", *Journal of Business* 51, 621-651.
- [4] Carr P (2000) "A survey of preference free option valuation with stochastic volatility" *April Risk Conference, Paris*.

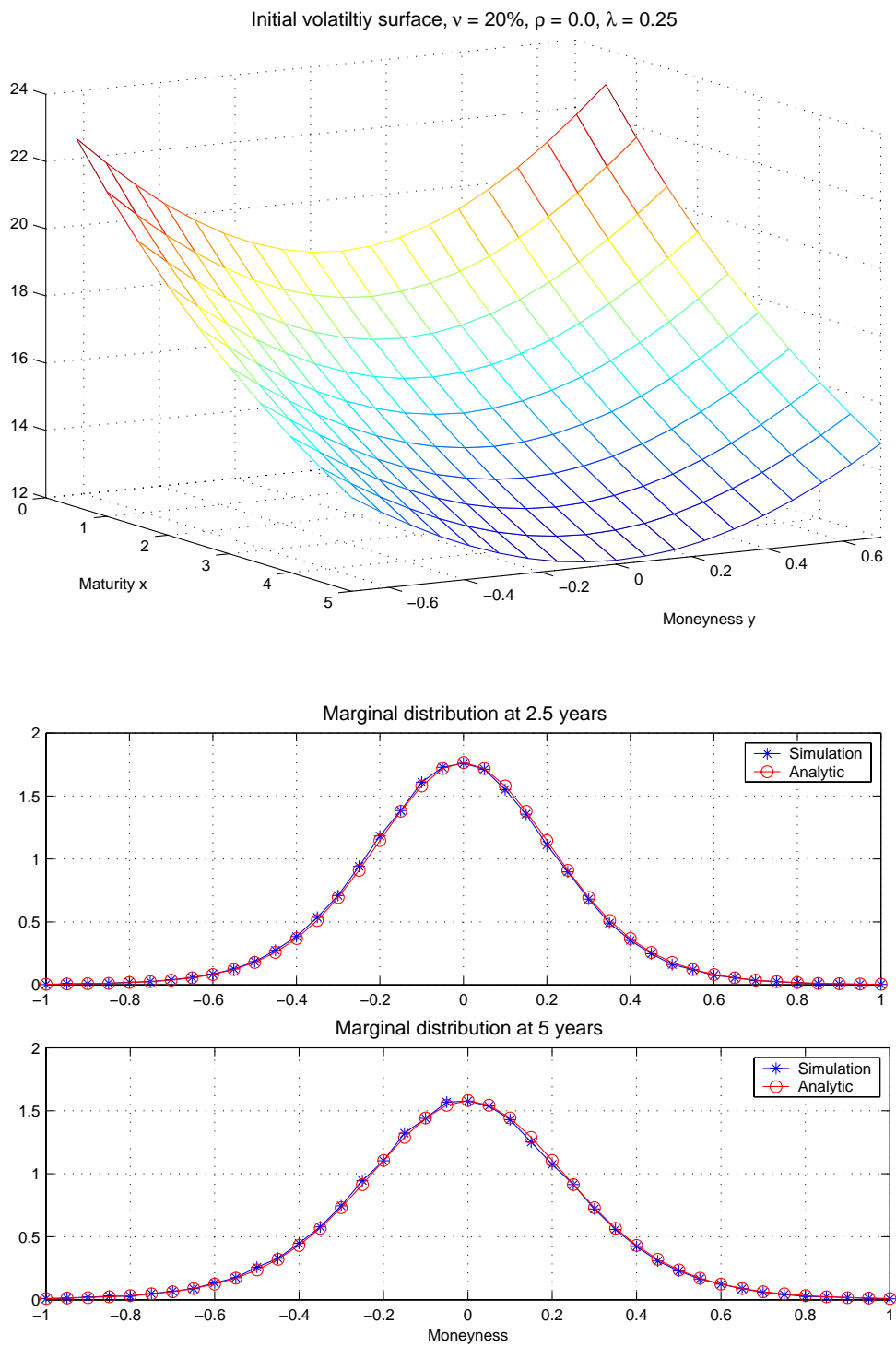
- [5] Nualart D and Pardoux E (1991) “Boundary value problems for stochastic differential equations, *Ann. Probab.* 19, 1118–1144
- [6] Rozovskii BL (1973) “On the Ito-Venttsel formula” *Vestnik Moskovskogo Universiteta, Matematika Vol 28 No1 p26-32.*
- [7] Schonbucher PJ (1998) “A market model for stochastic implied volatility” *Working paper, University Bonn.*
- [8] Venttsel AD (1965) “On the equations of the theory of conditional Markov processes” *Teoriya veroyatn i ee primenen, X No 2 p390-393.*

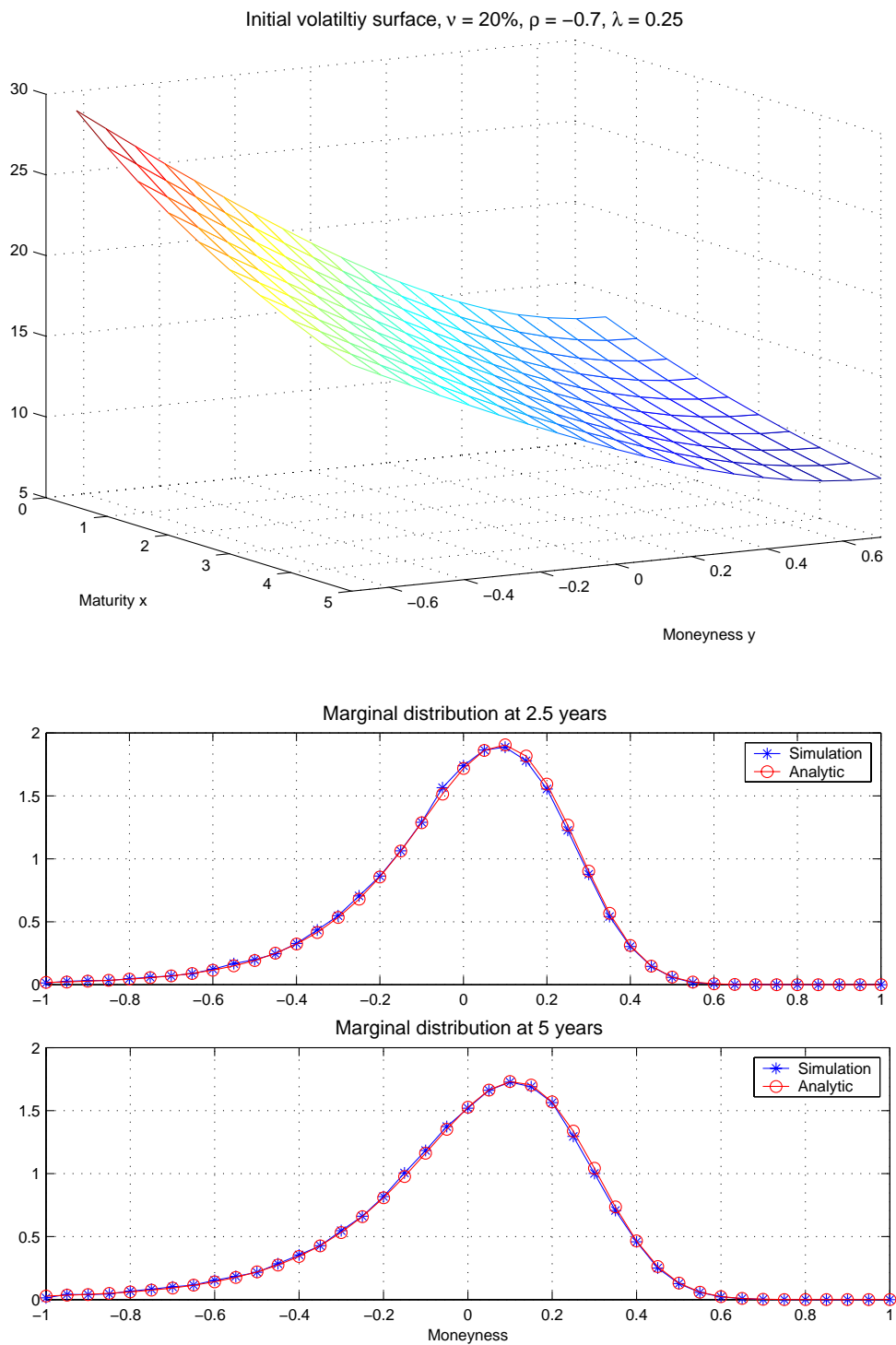
NATIONAL AUSTRALIA BANK AND FINANCIAL MATHEMATICAL MODELLING AND ANALYSIS  
*E-mail address:* [abrace@ozemail.com.au](mailto:abrace@ozemail.com.au)

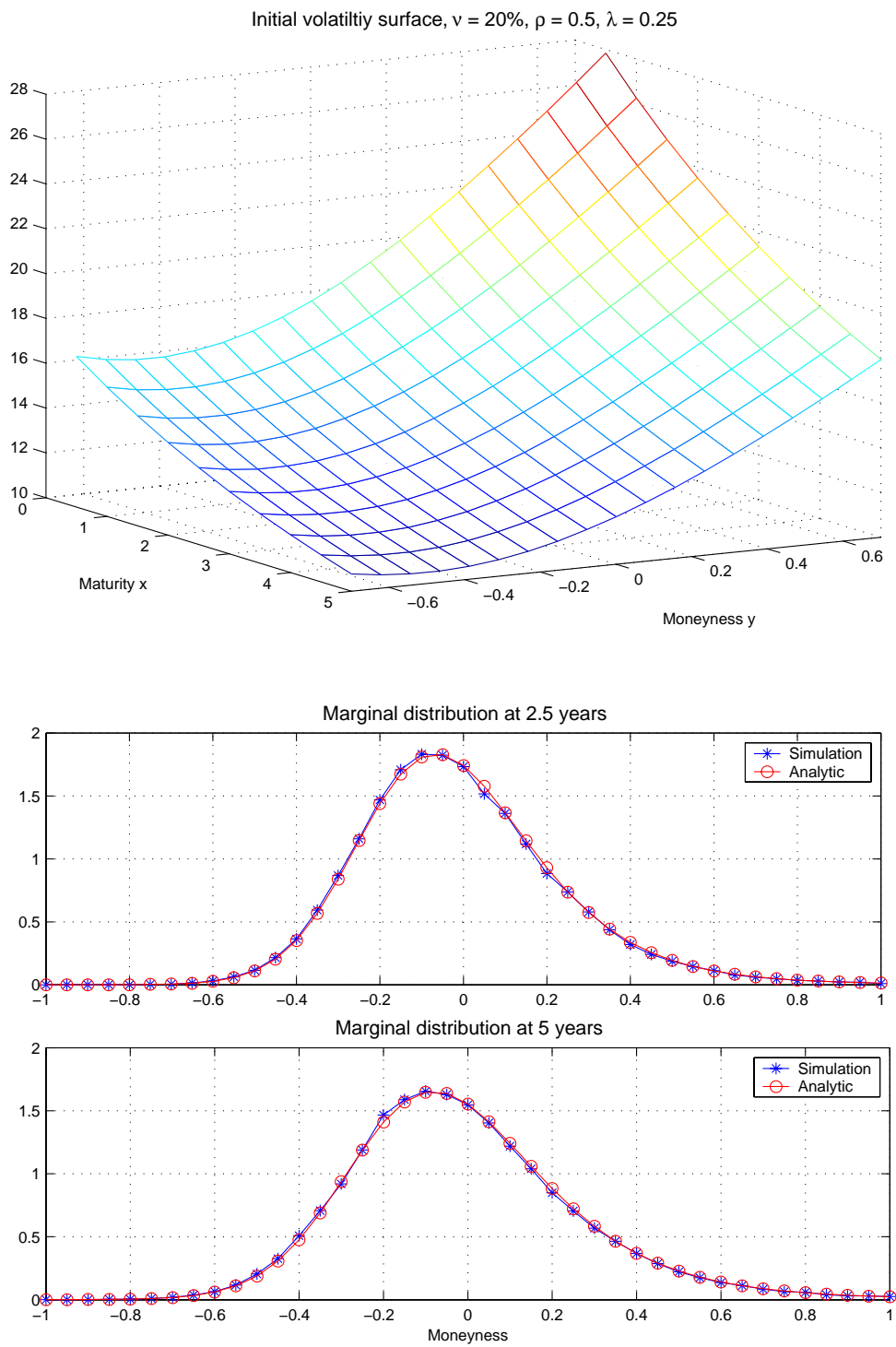
SCHOOL OF MATHEMATICS, THE UNIVERSITY OF NEW SOUTH WALES, SYDNEY, NSW 2052, AUSTRALIA  
*E-mail address:* [B.Goldys@unsw.edu.au](mailto:B.Goldys@unsw.edu.au)

DEPARTMENT OF MATHEMATICS AND STATISTICS, UNIVERSITY OF MELBOURNE, PARKEVILLE, VIC 3052,  
AUSTRALIA  
*E-mail address:* [F.Klebaner@ms.unimelb.edu.au](mailto:F.Klebaner@ms.unimelb.edu.au)

SCHOOL OF MATHEMATICS, THE UNIVERSITY OF NEW SOUTH WALES, SYDNEY 2052, AUSTRALIA  
*E-mail address:* [R.Womersley@unsw.edu.au](mailto:R.Womersley@unsw.edu.au)

FIGURE 1. Initial volatility surface and marginal distributions,  $\rho = 0$

FIGURE 2. Initial volatility surface and marginal distributions,  $\rho = -0.7$

FIGURE 3. Initial volatility surface and marginal distributions,  $\rho = 0.5$