Saddlepoint Approximations for Stochastic Volatility Models

Michael Studer
Dept. Mathematatik, ETH Zürich, 8092 Zürich, CH.
E-mail: studerm@math.ethz.ch

It is well-known that the saddlepoint approximation can give a quite accurate approximation for the distribution of a random variable. We study Stochastic Volatility models (SV-models). Although the cumulant generating function of the marginals may not be analytically tractable for many SV-models or may even not exist, one can often quite easily compute the cumulant generating function of an approximation by using a stochastic Taylor expansion. We examine the applicability of this procedure for some explicit examples.

Keywords: Saddlepoint approximations, Stochastic Taylor expansion and Stochastic Volatility models

1 Introduction

The main objects of this work are continuous-time SV-models as proposed in Frey (1997) and Hofman *et al.* (1992). The evolution of some financial asset (stock price) is in these models described by the following two-dimensional SDE.

$$dS_t = a(t, S_t, v_t)S_t dt + \sigma(t, S_t, v_t)S_t dW_{1,t} \quad (1)$$

$$dv_t = b(t, S_t, v_t)dt + \eta_1(t, S_t, v_t)dW_{1,t} + \eta_2(t, S_t, v_t)dW_{2,t}, \quad (2)$$

where W_1 and W_2 are two independent standard Brownian motions on some probability space $(\Omega, \mathcal{F}, P)$. The filtration $(\mathcal{F}_t)_{t\geq 0}$ is the augmented filtration generated by the Brownian motions. There are several restrictions on the different terms of these two equations to fulfill for example existence and uniqueness of the solution. These can be looked up in Frey (1997) and Hofman *et al.* (1992), the former gives an excellent overview over previously discussed stochastic volatility models and related models.

Although these models seem to fit better to real-world conditions than the standard Black-Scholes model they have a major drawback, namely they are analytically less tractable. If we don't want to do time-consuming simulations, we have to use some approximations. The aim of this work is to show that a classical statistical tool namely the saddlepoint approximation can be used for solving some problems concerning SV-models. We use a stochastic Taylor expansion to approximate the marginals of the SV-models to ensure existence of the cumulant generating function. This will also help us to calculate the saddlepoint approximation. Unfortunately this procedure restricts the applicability of the methodology since the stochastic Taylor expansion may only be accurate for short time-horizons.

The paper is organized as follows. In Section 2 we give the basic technical background which is needed, and can be omitted if one has already met stochastic Taylor expansions and saddlepoint approximations. In Section 3 we first analyse the methodology in the well-known Black-Scholes model which is not a SV-model, but can nevertheless be used to highlight the pros and cons of our methodology. We then move on to SV-models and compare the analytical results with results obtained from simulations. As already mentioned, the methodology is only applicable for "short-time" problems. But this is exactly the case for most applications in risk management. We also give some easy methods to approximate shortfall probabilities and expected shortfall.

2 A saddlepoint approximation for Stochastic Volatility models

2.1 Classical Saddlepoint Approximations

The saddlepoint approximation was already developed in the fifties by Daniels (1954) for the average of independent and identically distributed random variables. We only want to outline the idea of the saddlepoint approximation for a density in the one-dimensional absolutely continuous case. The interested reader may look up all the necessary details in the excellent textbook of Jensen (1995). Then we give an overview over the formulas used in the sequel.

Let X be a one-dimensional random variable on $(\Omega, \mathcal{F}, \mathcal{P})$ with density f and moment generating function $M_X(\theta) = E[e^{\theta X}]$. The domain of M_X is

$\Theta = \{\theta \in \mathbb{R} : M_X(\theta) < \infty\}$. Let $K_X(\theta) = \log M_X(\theta)$ denote the cumulant generating function of X. One way to explain the idea of the saddlepoint approximation is by introducing the so-called Esscher or exponential tilting. To approximate the density of X in x, we are looking for a measure Q equivalent to P, such that x is the expectation of X under the measure Q, i.e.

$$\begin{aligned} \frac{dP}{d\lambda}(x) &= \frac{dP}{dQ}(x)\frac{dQ}{d\lambda}(x) = \left(\frac{dQ}{dP}(x)\right)^{-1}\frac{dQ}{d\lambda}(x), \\ \text{where } E_Q[X] &= x. \end{aligned} \tag{3}$$

The first term of the right hand side of (3) can be calculated, because we restrict ourselves to the class of exponentially tilted measures, i.e.

$$\frac{dQ}{dP}(x) = \frac{dP_\theta}{dP}(x) = e^{\theta x} M_X(\theta)^{-1} = e^{\theta x - K_X(\theta)}.$$

One easily checks that $K'(\theta) = E_\theta[X]$, where E_θ denotes the expectation under the measure P_θ. This means that we can calculate the saddlepoint approximation for all x such that there exists a $\theta_0 \in \Theta$ with $K'(\theta_0) = x$.
The second term of the right hand side of (3) is approximated from a "local" Edgeworth expansion (see Field and Ronchetti (1990) or Jensen (1995)) under the tilted measure P_{θ_0}. The Edgeworth expansion uses an expansion of the characteristic function of a random variable to determine an approximation of the density. In its most simple form the value of the tilted density in x is approximated by the density of a normal random variable with variance $\text{Var}_{\theta_0}(X)$, the variance of X under the measure P_{θ_0}, at the mean:

$$\frac{dP_{\theta_0}}{d\lambda}(x) \approx (2\pi K''(\theta_0))^{-1/2} \tag{4}$$

After introducing the idea of the saddlepoint approximation we give some well-known results. We formulate the saddlepoint approximation in the following set-up. Let $X_1, \ldots, X_n$ be independent and identically distributed random variables with $X_i \in \mathbb{R}^d$ and with a positive definite variance. Let $K(\theta) = \log E[\exp(\theta X_1)]$ be the cumulant generating function defined for $\theta \in \Theta$. We assume that $0 \in \text{int}\,\Theta$ and $\varphi \in L^p(\mathbb{R}^d)$ for some $p \geq 1$, where φ is the characteristic function of X_1. Let $\mu(\theta) = \frac{\partial}{\partial\theta}K(\theta)$ and $\Sigma(\theta) = \text{Var}_\theta(X) = \frac{\partial^2}{\partial\theta\partial\theta^T}K(\theta)$ denote the mean resp. the variance of X_1 under the tilted measure P_θ. The error terms we give below hold uniformly for all $\theta \in C$ where C is a compact subset of $\text{int}\Theta$.
We can calculate the saddlepoint approximation only for values of x such

that there exists a $\theta_0(x) \in \text{int}\Theta$ with $E_{\theta_0}[X] = x$. The density f_n of $\bar{X} = (X_1 + \cdots + X_n)/n$ can be written as

$$f_n(x) = e^{n(K(\theta)-\theta x)} n^{d/2} \phi(0; \Sigma(\theta))(1 + O(n^{-1})), \tag{5}$$

where $\phi(x; \Sigma(\theta))$ denotes the density of a normal random variable with covariance matrix $\Sigma(\theta)$ and mean 0.
We can also calculate tail probabilities in the one-dimensional case in a similar way. We content ourselves by giving two tail formulas for the continuous case. The first one can be found in Jensen (1995, p. 27) and the second also in Jensen (1995, p. 79 ff) :

$$\begin{aligned} P[\bar{X} > x] &= \frac{e^{n(K(\theta)-\theta x)}}{\sqrt{n}\theta\sigma(\theta)} \Big(B_0(\lambda) + \tfrac{\zeta_3(\theta)}{6\sqrt{n}}(\theta) B_3(\lambda) + \tfrac{1}{n} \Big(\tfrac{\zeta_4(\theta)}{24} B_4(\lambda) \\ &\quad + \tfrac{\zeta_3^2}{72} B_6(\lambda) \Big) + O(B_0(\lambda) n^{-3/2}) \Big) . \qquad (6) \\ P[\bar{X} > x] &= (1 - \Phi(r)) + \phi(r) \left(\frac{1}{\lambda} - \frac{1}{r} \right) + O(n^{-3/2}), \qquad (7) \end{aligned}$$

The formulas (6), (7) are valid for $x \geq \mu(0)$. We denote by $B_0, \ldots, B_6$ the so-called Esscher functions (see Appendix C), $\Phi(r)$ and $\phi(r)$ denote the distribution function resp. the density function of a standard normal random variable, $r(y) = \sqrt{2n(\theta y - K(\theta))}$ and $\lambda = \sqrt{n}\,\theta\sigma(\theta)$. The formulas for $x < \mu(0)$ are very similar.
Formula (7) is called the Lugannani-Rice formula. It was derived in Lugannani and Rice (1980).

Remark. It is also possible to use other distributions than the normal as approximating densities in (4). One possible alternative is a Gamma-based approximation as discussed in Jensen (1995).
Although the saddlepoint approximation has its major applications for big n, we show that in the cases studied in this paper it is already a reasonable approximation for $n = 1$.

2.2 Stochastic Taylor expansion

As already mentioned, it may be impossible for models of the type (1), (2) to determine the cumulant generating function of the marginals which is a key instrument for the saddlepoint approximation. But often we are able to determine this function for some approximation of the marginals. As a key example, we consider a Hull-White type model discussed in Hull and White

(1987):

$$
\begin{aligned}
dS_t &= S_t\sqrt{v_t}\,dW_{1,t}, && (8)\\
dv_t &= \gamma v_t\,dt + \delta\rho v_t\,dW_{1,t} + \delta\sqrt{1-\rho^2}v_t\,dW_{2,t}, && (9)
\end{aligned}
$$

where $\delta > 0, -1 < \rho < 1$ and $\gamma,\ S_0,\ v_0 \in \mathbb{R}$. The inclusion of a drift component in the asset price does not change the problem significantly, but we want to capture correlation between the asset price and the volatility.
As the processes S_t and v_t stay positive almost surely, we can take the logarithm in both components and use the Itô formula (see also Lamberton and Lapeyre (1997)). This procedure, although not necessary, simplifies both the stochastic Taylor expansion and the saddlepoint approximation. Let R_t denote $\log S_t$ and $U_t = \log v_t$, then we have the following form for the resulting SDE:

$$
\begin{aligned}
dR_t &= -\frac{1}{2}\exp(U_t)\,dt + \exp(U_t/2)\,dW_{1,t}, && (10)\\
dU_t &= (\gamma - \delta^2/2)\,dt + \delta\rho\,dW_{1,t} + \delta\sqrt{1-\rho^2}\,dW_{2,t}. && (11)
\end{aligned}
$$

Following the notation of Kloeden and Platen (1992, p. 180 ff), by taking the hierarchical set Γ_2, we get the following truncated Taylor-Itô expansion for the transformed SDE (10), (11).

$$
\begin{aligned}
\tilde{R}_t &= \overbrace{\log S_0 - \frac{1}{2}(v_0 + \delta\rho\frac{\sqrt{v_0}}{2})t}^{x_0} + \overbrace{\sqrt{v_0}}^{x_1}\,W_{1,t} + \overbrace{\delta\rho\frac{\sqrt{v_0}}{4}}^{x_2}\,W_{1,t}^2\\
&\quad + \overbrace{\delta\sqrt{1-\rho^2}\frac{\sqrt{v_0}}{2}}^{x_3}\int_0^t W_{2,s}\,dW_{1,s}, && (12)\\
\tilde{U}_t &= \underbrace{\log v_0 + (\gamma - \frac{\delta^2}{2})t}_{o_0} + \underbrace{\delta\rho}_{o_1}\,W_{1,t} + \underbrace{\delta\sqrt{1-\rho}}_{o_2}\,W_{2,t}. && (13)
\end{aligned}
$$

Remark. The basic principle of the stochastic Taylor expansion is an iterative use of the Itô formula.

Now one can either compute the cumulant generating function of the marginals of this 2-dimensional process or of one of the components. By using a formula of Yor (1980) (see Appendix A) we get the following expression for the cumulant-generating functions:

$$\begin{aligned} K_{(\tilde{R}_t,\tilde{U}_t)}(\theta) &= g(\theta) - \log\frac{\sin b(\theta)}{b(\theta)} - \frac{1}{2}\log|\Sigma^{-1}(\theta)| + \frac{1}{2}u(\theta)^T\Sigma\, u(\theta), \quad (14)\\ \text{where } g(\theta) &= x_0\theta_1 + o_0\theta_2,\\ b(\theta) &= -\frac{x_3 t\theta_1}{2},\\ \Sigma^{-1}(\theta) &= \begin{bmatrix} b(\theta)\cot b(\theta) - 2u_3(\theta) & -a(\theta)\\ -a(\theta) & b(\theta)\cot b(\theta)\end{bmatrix}, \quad (15)\\ a(\theta) &= -b(\theta),\\ u(\theta) &= (u_1(\theta), u_2(\theta)) = ((x_1\theta_1 + o_1\theta_2)\sqrt{t}, o_2\sqrt{t}\,\theta_2),\\ u_3(\theta) &= x_2 t\theta_1. \end{aligned}$$

The function $K_{(\tilde{R}_t,\tilde{U}_t)}$ is C^∞ in a suitably chosen domain of $\mathbb{R}^2$ including 0 and similarly for $K_{\tilde{R}_t}$ and $K_{\tilde{U}_t}$ (see Appendix A).

$$\begin{aligned} K_{\tilde{R}_t}(\theta_1) &= K_{(\tilde{R}_t,\tilde{U}_t)}(\theta_1, 0)\\ &= g(\theta_1, 0) - \log\frac{\sin b(\theta_1)}{b(\theta_1)} - \frac{1}{2}\log|\Sigma^{-1}(\theta_1)|\\ &\quad + \frac{u_1(\theta_1,0)^2 b(\theta_1)\cot b(\theta_1)}{2|\Sigma^{-1}(\theta_1)|}, \quad (16)\\ K_{\tilde{U}_t}(\theta_2) &= g(0,\theta_2) + \frac{1}{2}(o_1^2 + o_2^2)\, t\,\theta_2^2. \quad (17) \end{aligned}$$

Remark. $K_{\tilde{U}_t}$ is just the cumulant generating function of a normal random variable, as it should be.

3 Risk measures and saddlepoint techniques

Let $X \in L^1(P)$ be the stochastic outcome of holding a position in a financial asset over some fixed period t. We tackle the task of computing the market risk of this position. The first task is the choice of an appropriate risk measure. We do not classify different risk measures but examine the two following measures to calculate market risk of a position (we neglect interest rates). We start with the popular Value-at-Risk (VaR). A possible mathematical definition of VaR proposed in Delbaen (1998) is the following:

$$VaR_\alpha(X) = -q_\alpha^+(X), \text{ where } q_\alpha^+(X) = \inf\{q|P[X \le q] > \alpha\}.$$

This risk measure can be found in the guidelines of the Basle Committee (1996a and 1996b) which is one reason for its popularity. Its simplicity and analytical tractability are certainly other reasons. But VaR lacks in most cases the subadditivity property which is required in the axiomatic approach of coherent risk measures of Artzner *et al.* (1999). Therefore we study also another risk measure called tail conditional expectation (TCE).

$$TCE_\alpha(X) = -E[X|X \leq -VaR_\alpha(X)]. \tag{18}$$

The definition is given in Artzner *et al.* (1999). By adapting a technique outlined in Rogers and Zane (1998) we can also approximate tail conditional expectation, which is a coherent risk measure under some conditions (see Delbaen (1998)).

3.1 Tail conditional expectation for log-return models

We examine models where the asset price dynamics S_t are given through specification of the dynamics for the log-return $X_t = \log S_t$. We suppose that X_t has a density f_{X_t} and the cumulant-generating function $K_{X_t} = \log M_{X_t}$ exists in an open interval containing 0 and 1. We fix the desired level α and set $a = -VaR_\alpha(S)$, then we have the following (we drop the subscript t):

$$\begin{aligned} E[S|S \leq a] &= E[e^X|e^X \leq a] = E[e^X|X \leq \log a] \\ &= \int_{-\infty}^{\log a} e^x \frac{f(x)}{P[X \leq \log a]}\, dx = \frac{M_X(1)}{P[X \leq \log a]} \int_{-\infty}^{\log a} g(x)\, dx, \end{aligned}$$

where $g(x) = \exp(x) f(x)/M_X(1)$ is just the density of X under an exponentially tilted measure P_1 with $dP_1 = \exp(X - K_X(1))dP$. So the cumulant-generating function of g, K_g can easily be derived from K_X:

$$K_g(u) = K_X(u+1) - K_X(1).$$

Finally we put all together and get the following easy expression:

$$E[S|S \leq a] = M_X(1) \frac{P_1[X \leq \log a]}{P[X \leq \log a]}. \tag{19}$$

Now we can use a tail formula to approximate tail conditional expectation.

Remark. This method is used in Rogers and Zane (1998) to price options for models with log-returns given by Lévy processes. Unfortunately the approximation is in this case too bad for small time horizons so that especially for risk management purposes we can not use it.

3.2 Tail conditional expectation for negative random variables

We can also use saddlepoint approximations where it is not appropriate to work with log-returns. Suppose $X \leq 0$, set $a = -VaR_\alpha(S)$, then we can approximate TCE in the following way:

$$\begin{aligned} -E[X|X \leq a] &= -\frac{1}{P[X \leq a]} \int_{-\infty}^{a} x\, f(x)\, dx = \frac{E[-X]}{P[X \leq a]} \int_{-a}^{\infty} \underbrace{\frac{x\, f(-x)}{E[-X]}}_{h(x)} dx \\ &= E[-X] \frac{P_h[-X \geq -a]}{P[X \leq a]} = E[-X] \frac{P_h[X \leq a]}{P[X \leq a]} \end{aligned} \tag{20}$$

The only remaining task is calculating the cumulant generating function of the random variable Z having density h.

$$\begin{aligned} M_h(u) &= \int_0^{\infty} e^{ux} h(x)\, dx = \frac{M_X(-u)}{E[-X]} \int_0^{\infty} x\, \underbrace{\frac{e^{ux} f(-x)}{M_X(-u)}}_{g(x)} dx \\ &= \frac{M_X(-u)}{E[-X]} E_g[-X] = \frac{M_X'(-u)}{M_X'(0)} \qquad \text{, as} \\ M_g(\tilde{u}) &= \frac{M_X(u + \tilde{u})}{M_X(u)} \end{aligned} \tag{21}$$

Hence by taking logarithms we have the expression for the cumulant generating function of the conditional distribution and can apply the tail formulas. It is obvious how to extend this method to random variables which are bounded from above.

4 Applications for risk management

The outlined methodology is perfectly matched for short-time risk management. Not only can one easily determine shortfall probabilities via the Lugannani-Rice formula (7), which was already used for option pricing in Rogers and Zane or other tail formulas, e.g. (6), but we are also able to approximate the whole distribution which allows us to determine other functionals. We compare the results derived from the approximation with the exact results in the Black-Scholes case and with results derived from simulations in the Hull-White case.

4.1 The Black-Scholes model

We consider the following well-known model for the evolution of the asset price:

$$dS_t = \mu S_t \, dt + \sigma S_t \, dW_t. \tag{22}$$

By the same "trick" as in the Hull-White case, i.e. taking the logarithm of the asset price, the resulting process becomes Brownian motion. This does not yield a convincing example. The saddlepoint approximation is exact in this case as pointed out in Daniels (1980). So let C_t denote a call option in the Black-Scholes case. We compare the exact distribution of C_t for t days with the saddlepoint approximation and it is even possible to calculate easily the exact distribution of the stochastic Taylor expansion, which is just a non-central chi-square distribution.

Figure 1 shows the relative error of the approximation. The solid line shows

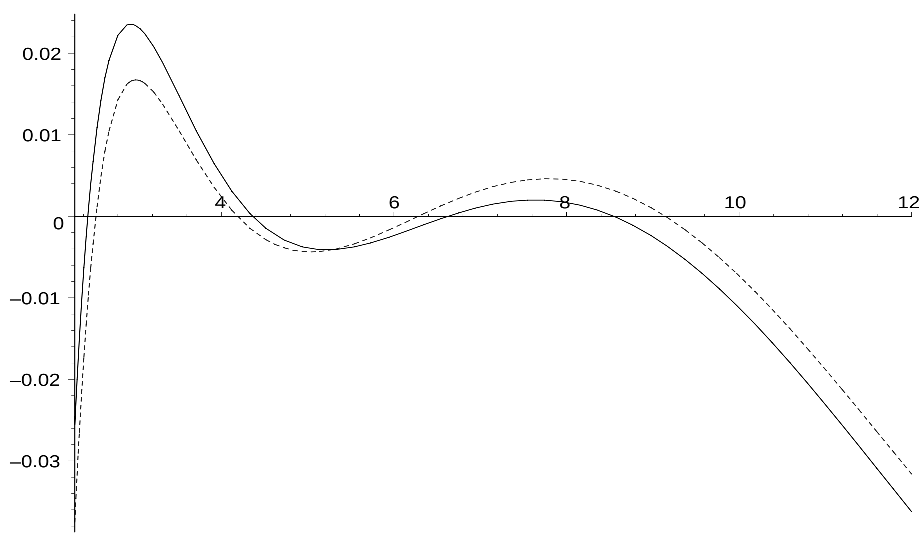

Figure 1: Relative error for the density of C_t in the Black-Scholes Case. The parameter values are $S_0 = K = 100$, $\sigma = 0.2$, $r = 0$ and $T = 125$ days. We consider a time horizon $t = 5$ days.

the relative error of a saddlepoint approximation (5). The dashed line shows the relative error of the exact distribution of the stochastic Taylor expansion. The interval shown captures over 99.7% of the mass of the exact distribution.

It is quite easy to approximate the density of a call in the Black-Scholes model, but in more complex models the cumulant generating function may be too complicated to determine the whole distribution. So instead of solving many equations of the type $K'(\theta) = x$, we might only be interested in shortfall probabilities. By using a tail formula we can easily calculate approximate shortfall probabilities. We compare them in Figure 2 for the case of the call with the exact shortfall probabilities. The solid line represents the relative error for the tail formula (6) in the upper tail, the dashed line for formula (7) respectively. Although we underestimate the probability of a big up move, the relative error goes only up to about 3%.

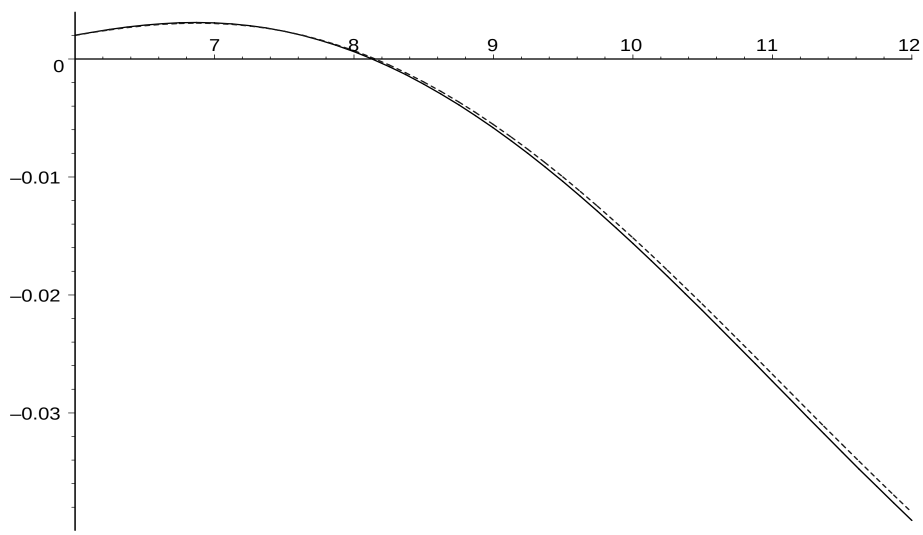

Figure 2: Comparison of the relative error for the shortfall probabilities using tail formulas (6) and (7) (dashed line) for the call in the Black-Scholes model.

As outlined in the previous section we are also able to calculate TCE by the adapted saddlepoint approximation (20). The TCE approximation for the underlying would be exact, as we are dealing with normal returns. So we only examine the TCE for the call. In Figure 3 we compare the relative error of using (6) or (7) (dashed line) in formula (20) to the exact TCE. The approximated TCE matches almost perfectly the true value. The errors in the nominator and denominator of (20) and (21) cancel very nicely.

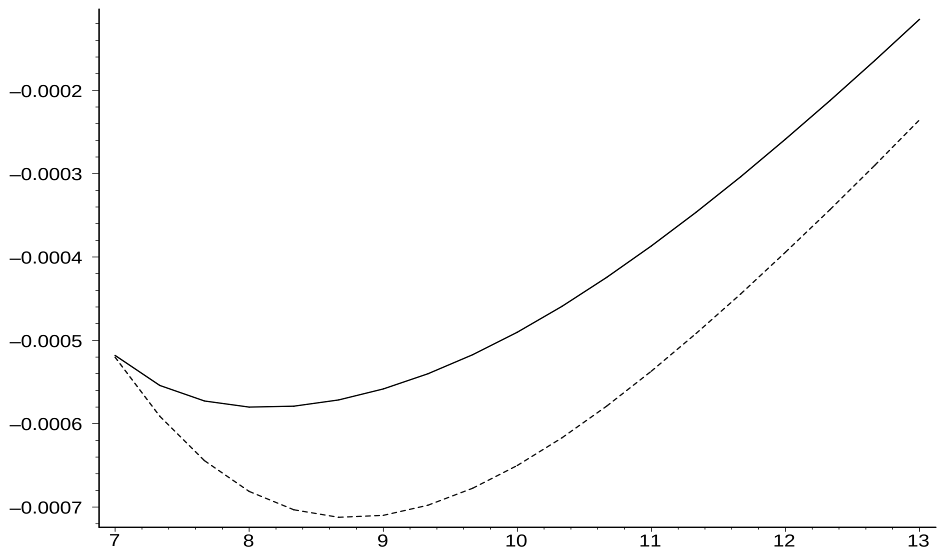

Figure 3: Comparison of the relative error for the approximated TCE for a call in the Black-Scholes model using tail formulas (6) and (7) (dashed line) in formula (20). The probability of the value of the call rising above 13 is 0.01%!

4.2 Stochastic Volatility models

We examine the Hull-White model (10), (11). We are now dealing with a two-dimensional saddlepoint approximation and as the cumulant generating function is more complicated than in the Black-Scholes case it is harder to solve an equation of the type $K'(\theta) = x$. But nevertheless it is still possible to approximate the joint density of $(\tilde{R}_t, \tilde{U}_t)$. But as the true density in the Hull-White model is unknown we have to simulate according to (10), (11) to compare the results obtained from the approximations.
We simulate 40000 paths of (10), (11) up to a time horizon t by the Euler scheme (see Kloeden and Platen (1992 p. 305 ff)). Before we perform some statistical tests we first make some simple graphical comparisons. We divide $\mathbb{R}^2$ in many small rectangles. We calculate the approximate measure of each rectangle induced by $(\tilde{R}_t, \tilde{U}_t)$. In Figure 4 we put a "0" on each rectangle having approximate expected frequency for 40000 independent trials of $(\tilde{R}_t, \tilde{U}_t)$ less than 1. For all the other rectangles we set a "+" sign, where the expected frequency is bigger than the observed frequency from 40000 simulations, a "–" sign otherwise. The time horizon chosen is again t =5 days.

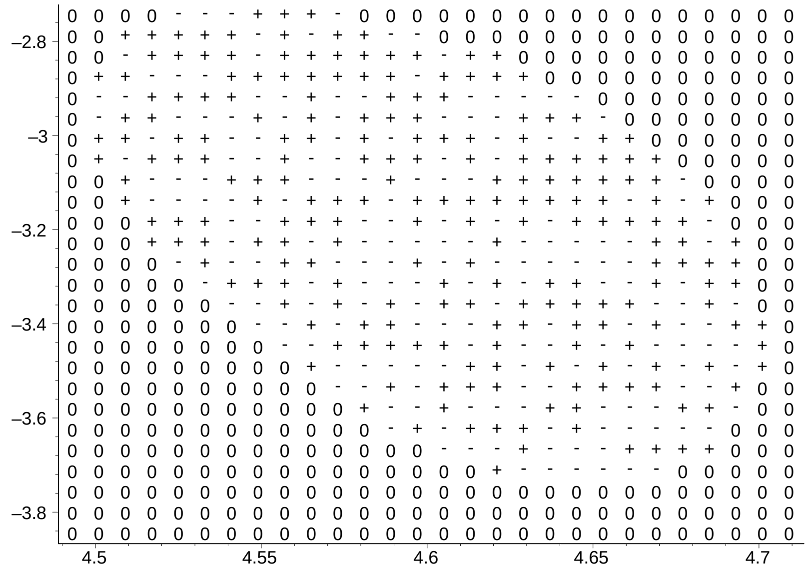

Figure 4: Graphical description of the difference of the approximation and the results from simulations for the joint density of $(\tilde{R}_t, \tilde{U}_t)$, $t = 5$ days.

We also compare the marginal density of the approximation for the logarithmic returns with a non-parametric density estimation from 40000 simulated points $x_1, \ldots, x_{40000}$. In Figure 5 we include over 99.9% of the points received by simulations and compare the relative error of the approximating density with the non-parametric density estimation and a simple normal approximation, i.e. we use a normal density where we match the first two moments of $\tilde{R}_t$.

In Table 1 we compare the (0.001, 0.005, 0.01, 0.05, 0.1) quantiles obtained by simulations and the corresponding upper quantiles of the approximation by saddlepoint approximation and by the simple normal approximation. In Table 2 we calculate approximate TCE for the 5 day time horizon, i.e. we use (19) for $\exp(\tilde{R}_t)$. Again we compare the relative error with respect to the results from the simulations and also include results from a simple normal approximation.

In Tables 1 and 2 the first column gives the chosen level α, the second column gives the result obtained from the simulations, i.e. quantile or TCE_α, the third column describes the relative error from the results of the saddlepoint approximation in percent and finally the fourth column gives the quotient of the relative error from the saddlepoint approximation and the relative error from the simple normal approximation.

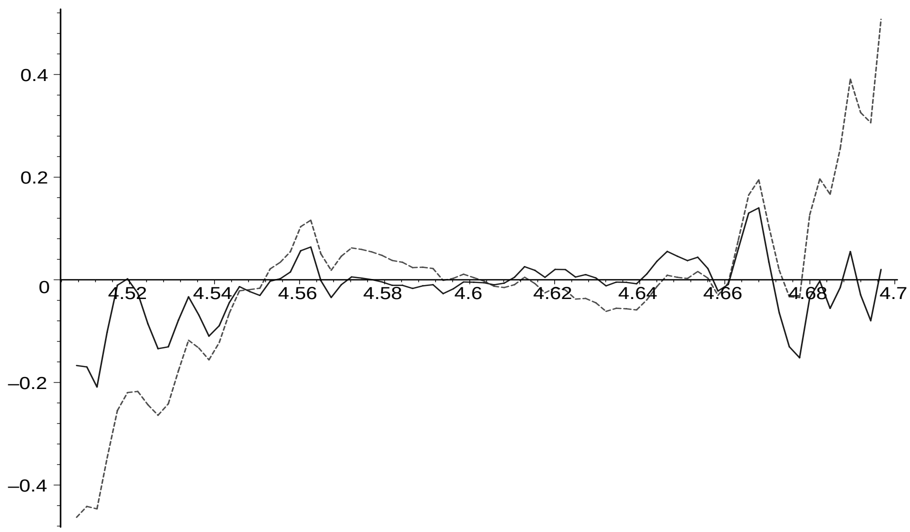

Figure 5: Relative error for the marginal density of X_5 in the Hull-White model. The dashed line represents the error from a simple normal approximation.

These tables suggest that we can approximate these risk measures in the Hull-White case almost perfectly. Although the saddlepoint approximation is especially in the quantile case and in the lower tail for the TCE much better than the simple normal approximation, already the simple normal approximation works very well. But this fact is due to the special form of these two measures. If we approximate $E[(q^+_{\alpha_i}(S_5) - S_5)_+]$ for $\alpha_1 = 0.01$ and $\alpha_2 = 0.05$ we get the following results. We divide the simulations into two parts of 20000 trials each and then estimate the quantity $E[(q^+_{\alpha_i}(S_5) - S_5)_+]$ so we get intervals $[a_i, b_i]$ for $i = 1, 2$. For both quantiles the results from the saddlepoint approximation lie in the corresponding interval whereas the simple normal approximation is about 4% smaller than a_2 and even 20% smaller than a_1. The following statistical tests also suggest that we make a significant improvement by using a saddlepoint approximation.

The first test is a simple chi-square goodness-of-fit test. We divide $\mathbb{R}^2$ in many small rectangles. We join all rectangles with approximated expected frequency for 40000 trials less then 1. In our example we have 406 regions. The chi-square statistic has a value of 405.7 so the P-value is 48%.
Another test is based on the integrated square error $\int (f_n(x) - f(x))^2\, a(x)\, dx$

level	emp. VaR	Rel. error of appr. in %	Comparison with normal appr.
0.001	91.076	0.105	0.203
0.005	92.568	0.099	0.283
0.01	93.300	0.090	0.339
0.05	95.284	0.063	0.714
0.1	96.358	0.022	- 4.200
0.9	103.57	0.029	1.000
0.95	105.75	- 0.011	1.046
0.99	106.59	-0.047	- 0.417
0.995	107.31	- 0.075	-0.533
0.999	108.75	- 0.055	- 0.207

Table 1: Comparison of VaR in the Hull-White model.

level	emp. TCE	Rel. error of appr. in %	Comparison with normal appr.
0.001	90.409	- 0.142	- 0.299
0.005	91.696	- 0.040	- 0.100
0.01	92.330	- 0.012	- 0.033
0.05	94.061	- 0.004	- 0.021
0.1	94.970	- 0.036	- 0.315
0.9	104.92	- 0.086	- 1.125
0.95	105.80	- 0.095	- 0.833
0.99	107.56	- 0.158	- 0.944
0.995	108.19	- 0.157	- 0.708
0.999	109.64	- 0.255	- 1.077

Table 2: Comparison of TCE in the Hull-White model.

(ISE), where a is a bounded integrable function whose set of discontinuities has 2-dimensional Jordan content zero, f_n is a non-parametric estimation of f. This ISE is under some assumptions (which are fulfilled in our example) asymptotically normal (see Bickel and Rosenblatt (1973) or Appendix B). So we take the approximating density and calculate the integrated square error with respect to the density estimation. For the joint density we have a P-value of 53%. These tests suggest no significant difference for the case of the joint distribution.
We can do exactly the same tests for the marginal distributions. For the marginal distribution of the log-stock, the corresponding P-values are 13% for the chi-square test (91 regions) and 38% for the ISE based test.
For the marginal distribution of the log squared volatility which is in fact a normal distribution we get the following P-values. We have 85% for the chi-square test and 69% for the ISE based test. So we have no significant statistical differences for all performed tests on a 5%-level. If we use the same tests for the case of the simple normal approximation the P-values for the joint distribution and the log-stock are much smaller than 1%.

5 Conclusion

We have shown that saddlepoint techniques yield excellent approximations to the shape of the distribution in a stochastic volatility set-up. We have further tested the method to many more examples in the case of market risk management; in many cases we obtained fast and accurate answers. The combination of stochastic Taylor expansion and saddlepoint approximations may therefore add a useful analytic tool to quantitative finance. Further investigation is needed to see how the method can be adapted to longer periods, beyond 10 days, say.

Appendix

A Calculation of the Characteristic function

We want to calculate the cumulant generating function of the two-dimensional random variable $(\tilde{R}_t, \tilde{U}_t)$ where $\tilde{R}_t$ and $\tilde{U}_t$ are given by (12) and (13). As a first step we adapt a result outlined in Yor (1980) which was originally presented in Lévy (1950). For completeness we give a proof of the result. We follow quite closely the proof in Yor (1980).

Lemma A.1 *Let (X_t, Y_t) be a two-dimensional standard Brownian motion starting at* 0. *For all $b \in \mathbb{R} \setminus \{0\}$:*

$$E\left[e^{b\int_0^1 (X_u\,dY_u - Y_u\,dX_y)}\Big| X_1 = x, Y_1 = y\right] = \frac{b}{\sin b} e^{(\frac{x^2+y^2}{2})(1-b\cot b)}. \tag{23}$$

Proof.

$$\int X_u\,dY_u - Y_u\,dX_u = \int X'_u\,dY'_u - Y'_u\,dX'_u$$

where $\binom{X'}{Y'} = O\binom{X}{Y}$ with O being an orthogonal transformation.

$$\int X_u\,dY_u - Y_u\,dX_u = \int \left\langle A\begin{pmatrix}X\\Y\end{pmatrix}, d\begin{pmatrix}X\\Y\end{pmatrix}\right\rangle$$

$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and $O^TAO = A$.

On the other hand the law of $\binom{X}{Y}$ is invariant under all orthogonal transformations. Because of these two remarks we have that

$$E\left[e^{b\int_0^1 (X_u\,dY_u - Y_u\,dX_y)}\big|_{X_1=x, Y_1=y}\right] = E\left[e^{b\int_0^1 (X_u\,dY_u - Y_u\,dX_y)}\big|_{X_1^2+Y_1^2=x^2+y^2}\right].$$

We are now considering the following three processes

$$\rho_t = \sqrt{X_t^2 + Y_t^2} \tag{24}$$

$$\beta_t = \int_0^t \frac{X_u\,dX_u + Y_u\,dY_u}{\rho_u} \tag{25}$$

$$\gamma_t = \int_0^t \frac{X_u\,dY_u - Y_u\,dX_u}{\rho_u}. \tag{26}$$

By inspection of the quadratic variations and the quadratic covariation of β and γ, we see that β and γ are one-dimensional orthogonal Brownian motions. By a result in Yor (1979) we know that the natural filtrations of the processes β and ρ coincide and conclude that ρ is independant of γ. Hence

$$\begin{aligned} E\left[e^{b\int_0^1 (X_u\,dY_u - Y_u\,dX_y)}\big|_{X_1^2+Y_1^2=x^2+y^2}\right] &= E\left[e^{b\int_0^1 \rho_u\,d\gamma_u}\big|\rho_1 = \rho\right] \\ &= E\left[e^{-\frac{b^2}{2}\int_0^1 \rho_u^2\,du}\big|\rho_1 = \rho\right] \\ &= \frac{b}{\sin b} e^{(\frac{x^2+y^2}{2})(1-b\cot b)}. \end{aligned}$$

The first two equalities follow from the remarks above while the last follows from the fact that we are dealing with a Brownian bridge and so we can apply Corollary 3.3 from Revuz and Yor (1991, p. 430). □

With the help of this lemma we can prove formula (14).

Proposition A.1 *The cumulant generating function* $K_{(\tilde{R}_t,\tilde{U}_t)}(\theta_1,\theta_2)$ *of the two-dimensional random variable* $(\tilde{R}_t,\tilde{U}_t)$ *given by (12) and (13) exists* $\forall\,\theta_2\in\mathbb{R}$ *and* $\theta_1\in]-arccot(-x_1), arccot(x_2)[$ *with* $0<-x_1<x_2$. *And it has the form (14).*

Proof. Let us first recall the definitions of $\tilde{R}_t$ and $\tilde{U}_t$. We leave out the deterministic terms x_0, o_0. Let A denote the set $\{W_{1,t}=w_1\sqrt{t}, W_{2,t}=w_2\sqrt{t}\}$. Then

$$\begin{aligned}\tilde{R}_t &= x_1W_{1,t}+x_2W_{1,t}^2+x_3\int_0^t W_{2,s}\,dW_{1,s},\\ \tilde{U}_t &= o_1W_{1,t}+o_2W_{2,t}.\end{aligned}$$

To use Lemma A.1 we apply Itô's lemma:

$$\int_0^t Y_u\,dX_u=\frac{1}{2}X_tY_t-\frac{1}{2}\int_0^1 (X_u\,dY_u-Y_u\,dX_u) \tag{27}$$

Now we first calculate the conditional expectation with respect to A,

$$E[e^{\theta_1X_t+\theta_2\tilde{U}_t}|A]=E[e^{f(W_{1,t},W_{2,t})+\alpha\int_0^t W_{2,s}\,dW_{1,s}}|A]=$$

$$E[e^{f(\sqrt{t}X_1,\sqrt{t}Y_1)+\alpha t\int_0^1 Y_s\,dX_s}|A]=E[e^{f(\sqrt{t}X_1,\sqrt{t}Y_1)+\frac{\alpha t}{2}X_1Y_1-\frac{\alpha t}{2}\int_0^1(Y_s\,dX_s-X_s\,dY_s)}|A]=$$

$$e^{f(w_1,w_2)}\frac{b(\theta_1)}{\sin(b(\theta_1))}e^{(\frac{w_1^2+w_2^2}{2})(1-b(\theta_1)\cot b(\theta_1))},$$

where $X_u=\frac{1}{\sqrt{t}}W_{1,tu}$, $Y_u=\frac{1}{\sqrt{t}}W_{2,tu}$ $f(w_1,w_2)=(\theta_1x_1+\theta_2o_1)w_1+\theta_2o_2w_2+\theta_1w_1^2$.
For the third equality we use (27), while the fourth equality follows from Lemma A.1. We integrate over $\mathbb{R}^2$, $\binom{X_1}{Y_1}$ has a two-dimensional standard normal density and A$=\{X_1=w_1,Y_1=w_2\}$, thus

$$\begin{aligned}&E[e^{\theta_1\tilde{R}_t+\theta_2\tilde{U}_t}]=\frac{b(\theta_1)}{2\pi\sin b(\theta_1)}\int\int e^{f(w_1,w_2)-\frac{w_1^2+w_2^2}{2}b(\theta_1)\cot b(\theta_1)}\,dw_1\,dw_2=\\ &\frac{b(\theta_1)}{2\pi\sin b(\theta_1)}\int\int e^{<\hat{\theta},w>-\frac{1}{2}w^T\Sigma^{-1}w}\,dw_1\,dw_2=\\ &\frac{b(\theta_1)}{\sin b(\theta_1)}\frac{1}{|\Sigma^{-1}(\theta)|}\frac{|\Sigma^{-1}(\theta)|}{2\pi}\int\int e^{<u(\theta),w>-\frac{1}{2}w^T\Sigma^{-1}(\theta)w}\,dw_1\,dw_2.\end{aligned}$$

The result now follows easily from the well-known formula for the moment generating function of a multivariate normal.

The calculations are possible as long as $\Sigma^{-1}(\theta)$ stays positive definite. So we have to look at the equation $|\Sigma^{-1}(\theta)| = 0$.

$$|\Sigma^{-1}(\theta)| = b(\theta)^2 \cot^2(b(\theta)) - 2u_3(\theta)b(\theta)\cot(b(\theta)) - a(\theta)^2,$$

b, a, and u_3 are of the form a constant times θ. So the value of $|\Sigma^{-1}(\theta)|$ in 0 is 1. We know further that $|\Sigma^{-1}(\theta)|_{b(\theta)=\pm\frac{\pi}{2}} < 0$. By continuity there exist $I_1 < 0$, $I_2 > 0$ such that I_1 is the first time $|\Sigma^{-1}(I_1)|$ hits 0 coming from 0 on the negative half-line and I_2 respectively on the positive half-line. They can now be calculated. We just remark that in our case $a(\theta) = -b(\theta)$ and assume that they are not identically 0:

$$\begin{aligned}
c_1^2\theta^2\cot^2(c_1\theta) - 2c_1c_2\theta^2\cot(c_1\theta) - c_1^2\theta^2 &= 0 \Leftrightarrow \\
c_1^2\cot^2(c_1\theta) - 2c_1c_2\cot(c_1\theta) - c_1^2 &= 0 \text{ set } x = \cot c_1\theta \\
c_1^2x^2 - 2c_1c_2x - c_1^2 &= 0 \Leftrightarrow \\
x_{1,2} &= c_2 \pm \sqrt{c_2^2 + c_1^2}.
\end{aligned}$$

So the smallest roots are given by $I_1 = -\text{arccot}(-x_1)$ and $I_2 = \text{arccot}(x_2)$.□

B A quadratic measure of deviation

In this appendix, we give the necessary background for a test used in Section 4. The interested reader finds in Bickel and Rosenblatt (1973) and Rosenblatt (1975) further details.

Our goal is to test whether a density approximation is close to the true density of some random vector.
Let $X_1, \ldots, X_n$ denote iid random two-vectors with continuous density function. By choosing a bounded weight function w with finite support one can estimate the density from $X_1, \ldots, X_n$.

$$f_n(x) = \frac{1}{nb(n)^2}\sum_{j=1}^{n} w\left(\frac{x - X_j}{b(n)}\right), \tag{28}$$

where $b(n)$ is some reasonably chosen bandwidth such that $b(n) \downarrow 0$ and $nb(n)^2 \to \infty$ as $n \to \infty$. We make the following assumptions:

- a_1: w is bounded and zero outside $[-\frac{1}{2}, \frac{1}{2}]^2$ and $\int w\,dx = 1$,
- a_2: f is bounded and positive on $\mathbb{R}^2$ and $f \in C^2(\mathbb{R}^2)$,
- a_3: $w(u) = w(-u)$,
- a_4: the function a is bounded, integrable and its set of discontinuities has 2-dimensional Jordan content zero.

Theorem B.1 *Under assumptions* $a_1 - a_4$*, if* $nb(n)^2 \to \infty$*,* $b(n) = o(n^{-\frac{1}{5}})$ *for* $n \to \infty$*,*

$$b(n)^{-1}\left(nb(n)^2\int (f_n(x)-f(x))^2 a(x)\,dx - \int f(x)a(x)\,dx\int w(u)^2\,du\right) \tag{29}$$

is asymptotically normally distributed with mean zero and variance

$$2w^{*4}(0)\int a(x)^2 f(x)^2 dx, \tag{30}$$

where $w^{*n}(x)$ *denotes the n-th convolution of* w*.*

Proof. see Rosenblatt (1975)

Corollary B.1 *Suppose one considers the one-dimensional analogue of the result obtained in Theorem B.1 with* $a'_1 - a'_4$ *the corresponding one-dimensional assumptions. If* $nb(n)^2 \to \infty$ *and* $b(n) = o(n^{-\frac{2}{9}})$ *for* $n \to \infty$*.*

$$b(n)^{-\frac{1}{2}}\left(nb(n)\int (f_n(x)-f(x))^2 a(x)\,dx - \int f(x)a(x)\,dx\int w(u)^2\,du\right) \tag{31}$$

is asymptotically normally distributed with mean zero and variance (30).

Remark. It is straightforward to use the results from this section as possible statistical tests whether the density approximation yields an accurate answer.

C Esscher functions

In formula (6) we use certain non-standard functions. According to Jensen (1995, p. 23) we call them Esscher functions. They appear naturally as inversion integrals when we approximate tail probabilities. We include them here to make the paper as self-contained as possible. For $\lambda > 0$ we define $B_k(\lambda)$, $k = 0, 1, \ldots$, by

$$B_k(\lambda) = \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-\frac{t^2}{2}}\frac{(it)^k}{1+it/\lambda}\,dt. \tag{32}$$

One can prove an easy lemma to connect the Esscher functions with the standard normal distribution function.

Lemma C.1

$$B_k(\lambda) = (-1)^k \frac{d^k}{dx^k} \left(\lambda e^{\frac{\lambda^2}{2}+\lambda x} \Big(1 - N(\lambda + x) \Big) \right) \Big|_{x=0} \tag{33}$$

Proof. Jensen (1995, p. 24) □

The lemma allows us to write down the Esscher functions in a more convenient way.

$$\begin{aligned}
B_0(\lambda) &= \lambda \exp\left(\frac{\lambda^2}{2}\right)(1 - N(\lambda)) \\
B_1(\lambda) &= -\lambda \left(B_0(\lambda) - (2\pi)^{-\frac{1}{2}} \right) \\
B_2(\lambda) &= \lambda^2 \left(B_0(\lambda) - (2\pi)^{-\frac{1}{2}} \right) \\
B_3(\lambda) &= -\left(\lambda^3 B_0(\lambda) - (\lambda^3 - \lambda^2)(2\pi)^{-\frac{1}{2}} \right) \\
B_4(\lambda) &= \lambda^4 B_0(\lambda) - (\lambda^4 - \lambda^2)(2\pi)^{-\frac{1}{2}} \\
B_5(\lambda) &= -\left(\lambda^5 B_0(\lambda) - (\lambda^5 - \lambda^3 + 3\lambda)(2\pi)^{-\frac{1}{2}} \right) \\
B_6(\lambda) &= \lambda^6 B_0(\lambda) - (\lambda^6 - \lambda^4 + 3\lambda^2)(2\pi)^{-\frac{1}{2}} \\
&\vdots \\
&\mathit{etc.}
\end{aligned}$$

Acknowledgements

The author would like to thank Professor P. Embrechts and P. Cheridito for the careful reading of earlier versions of this paper.

References

[1] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath. Coherent measures of risk. *Math. Finance*, 9(3):203–228, 1999.

[2] P.J. Bickel and M. Rosenblatt. On some global measures of the deviations of density function estimates. *Ann. Statist.*, 1:1071–1095, 1973.

[3] H.E. Daniels. Saddlepoint approximations in statistics. *Ann. Math. Statistics*, 25:631–650, 1954.

[4] H.E. Daniels. Exact saddlepoint approximations. *Biometrika*, 67:53–58, 1980.

[5] F. Delbaen. Coherent risk measures on general probability spaces. ETHZ Preprint, November 1998.

[6] C. Field and E. Ronchetti. *Small Sample Asymptotic*. Institute of Mathematical Statistics, Hayward, California, 1990.

[7] R. Frey. Derivative asset analysis in models with level-dependent and stochastic volatility. Discussion Paper No. B-401 Universiät Bonn, January 1997.

[8] N. Hofman, E. Platen, and M. Schweizer. Option pricing under incompleteness and stochastic volatility. *Mathematical Finance*, 2:153–187, 1992.

[9] J. Hull and A. White. The pricing of options on assets with stochastic volatilities. *Journal of Finance*, 42(2):281–300, 1987.

[10] J. L. Jensen. *Saddlepoint Approximations*. Oxford University Press, 1995.

[11] P. E. Kloeden and E. Platen. *Numerical Solutions of Stochastic Differential Equations*. Springer, 1992.

[12] D. Lamberton and B. Lapeyre. *Introduction to stochastic calculus applied to finance*. Chapman and Hall, 1997.

[13] P. Lévy. Wiener's random function, and other laplacian random functions. Second Symposium of Berkeley. Probability and Statistics, 1950. 171-186.

[14] R. Lugannani and S. Rice. Saddlepoint approximations for the distribution of the sum of independant random variables. *Adv. Appl. Probab.*, 12:475–490, 1980.

[15] Basle Committee on Banking Supervision. Amendment to the capital accord to incorporate market risk. Basle, January 1996.

[16] Basle Committee on Banking Supervision. Overview of the amendment to the capital accord to incorporate market risk. Basle, January 1996.

[17] D. Revuz and M. Yor. *Continuous Martingales and Brownian Motion*. Springer, Berlin, 1991.

[18] L. C. G. Rogers and O. Zane. Saddle-point approximations to option prices. Preprint, University of Bath, 1998.

[19] M. Rosenblatt. A quadratic measure of deviation of two-dimensional density estimates and a test of independence. *Ann. Statist.*, 3(1):1–14, 1975.

[20] M. Yor. Les filtrations de certaines martingales du mouvement brownien dans $\mathbb{R}^n$. *Séminaire de Probabilités XIII, 1977/78*. Lect. Notes in Math., 721:427-440, Springer, 1979.

[21] M. Yor. Remarques sur un formule de Paul Lévy. *Séminaire de Probabilités XIV, 1978/79*. Lect. Notes in Math., 784:343-346, Springer, 1980.