The Use of Importance Sampling to Speed Up Stochastic Volatility Simulations

Stan Stilger*

June 6, 2012

Fouque and Tullie (2002) use importance sampling for variance reduction in stochastic volatility simulations. In this document, I replicate the numerical example comparing small noise expansion (SNE) and fast mean reversion (FMR) described in that paper using Matlab and NAG routines. The NAG's Mersenne Twister random number generator produced more stable results than many other random number generators.

1 Importance Sampling

Importance sampling is a popular variance reduction technique. Suppose we wish to estimate $\theta = \mathbb{E}_f[\phi(X)]$, where X has probability density f, and $\phi(\cdot)$ is a payoff function. Let g be another probability density such that $f(x) > 0 \to g(x) > 0$

$$\theta = \mathbb{E}_f[\phi(X)] = \int \phi(x) f(x) dx = \int \phi(x) \frac{f(x)}{g(x)} g(x) dx = \mathbb{E}_g\left[\phi(X) \frac{f(X)}{g(X)}\right]$$

Here, f and g are referred to as the original and importance sampling densities, and $\frac{f(x)}{g(x)}$ represents the likelihood ratio. It is the Radon-Nikodym derivative of the original measure with respect to the importance sampling measure. If density g can be chosen so that the random variable $\phi(X)\frac{f(X)}{g(X)}$ has small variance, then importance sampling can result in an efficient estimation of θ. To make the variance of $\phi(X)\frac{f(X)}{g(X)}$ small, $\frac{f(x)}{g(x)}$ should be inversely related to $\phi(x)$.

In the context of the Black-Scholes European call option, the importance sampling estimator is given by

$$\theta = \exp^{-rT} \mathbb{E}_g\left[\max\left\{S_0 \exp^{\left(r-\frac{\sigma^2}{2}\right)T+\sigma\sqrt{T}X} - K, 0\right\} \frac{f(x)}{g(x)}\right]$$

*Stan Stilger (przemyslaw.stilger@postgrad.mbs.ac.uk) is at Manchester Business School. I would like to thank Mike Croucher, Jean-Pierre Fouque, Eberhard Mayerhofer and Ser-Huang Poon for their helpful comments and suggestions.

where $\phi(X) = \max\left\{ S_0 \exp^{\left(r-\frac{\sigma^2}{2}\right)T+\sigma\sqrt{T}X} - K, 0 \right\}$ and $X \sim N(0,1)$. Thus, θ is the price of European call option, $\phi(X)$ is the payoff function and $\frac{f(x)}{g(x)}$ represents the likelihood ratio. The original density is

$$\ln(S_T) \stackrel{d}{=} \mathcal{N}\left(\ln(S_0) + (r - \frac{\sigma^2}{2})T, \sigma\sqrt{T} \right)$$

whereas the importance sampling density is given by

$$\ln(S_T) \stackrel{d}{=} \mathcal{N}\left(\ln(K) + \left(r - \frac{\sigma^2}{2}\right)T, \sigma\sqrt{T} \right)$$

and the likelihood ratio is

$$\frac{f(x)}{g(x)} = \exp\left((-2(\alpha - \beta)X + \alpha^2 - \beta^2)/(2\sigma^2 T) \right)$$

with $\alpha = \left(r - \frac{\sigma^2}{2}\right)T$ and $\beta = \ln\left(\frac{K}{S_0}\right) + \left(r - \frac{\sigma^2}{2}\right)T$. The change of measure is done by changing the drift of the stochastic process. Here, the drift of the stochastic process under the importance sampling measure has been shifted by $\ln\left(\frac{K}{S_0}\right)$, and all simulated paths end up at-the-money or in-the-money which is more efficient as some simulated paths under the original measure end up out-of-the-money.

2 Stochastic Volatility Application

Fouque and Tullie (2002) present a variance reduction scheme stochastic volatility simulations. In their setup, the price of a risky asset X_t, evolves according to the following SDE

$$dX_t = \mu X_t dt + \sigma(Y_t) X_t dW_t$$

where μ represents a constant mean return rate, $\sigma(Y_t)$ represents the volatility which is driven by another stochastic process Y_t which takes the following form

$$dY_t = \alpha(m - Y_t)dt + \beta(\rho dW_t + \sqrt{1-\rho^2} dZ_t). \tag{1}$$

The invariant distribution of Y_t is $\mathcal{N}\left(m, \frac{\beta^2}{2\alpha}\right)$. Let $v^2 = \frac{\beta^2}{2\alpha}$ and

$$dV_t = b(t, V_t)dt + a(t, V_t)d\eta_t$$

where $dV_t = \begin{pmatrix} X_t \\ Y_t \end{pmatrix}$, $b(t,V_t) = \begin{pmatrix} rx \\ \alpha(m - Y_t) \end{pmatrix}$, $a(t,V_t) = \begin{pmatrix} \sigma(Y_t)x & 0 \\ \nu\sqrt{2\alpha}\rho & \nu\sqrt{2\alpha\left(1-\rho^2\right)} \end{pmatrix}$, and $d\eta_t = \begin{pmatrix} W_t \\ Z_t \end{pmatrix}$. The price $P(t,V_t)$ of an European option at time t is then given by

$$P(t,v) = \mathbb{E}\left[e^{-r(T-t)}\phi(V_T)|V_t = v\right]$$

Next, Fouque and Tullie (2002) consider the following process

$$Q_t = exp\left(\int_0^T h(s,V_s)d\eta_s + \frac{1}{2}\int_0^T ||h(s,V_s)||^2 ds\right)$$

which is a positive martingale if $\mathbb{E}\left[Q_t^{-1}\right] = 1$. Then they define a new probability measure $\tilde{\mathbb{P}}$ equivalent to $\mathbb{P}$

$$\frac{d\tilde{\mathbb{P}}}{d\mathbb{P}} = (Q_T)^{-1}$$

By Girsanov Theorem, the process under the new measure, $\tilde{\mathbb{P}}$,

$$\tilde{\eta}_t = \eta_t + \int_0^t h(s,V_s)d\eta_s$$

is a standard Brownian motion. Processes V_t and Q_t can be rewritten in terms of $\tilde{\eta}_t$

$$\begin{aligned} dV_t &= \left(b(t,V_t) - a(t,V_t)h(t,V_t)\right)dt + a(t,V_t)d\tilde{\eta}_t \qquad (2) \\ Q_t &= exp\left(\int_0^T h(s,V_s)d\tilde{\eta}_s - \frac{1}{2}\int_0^T ||h(s,V_s)||^2 ds\right) \end{aligned}$$

With respect to this new measure, $P(t,V_t)$ can be written as

$$P(t,v) = \tilde{\mathbb{E}}\left[e^{-r(T-t)}\phi(V_T)Q_T|V_t = v\right]$$

Applying Ito's lemma to $P(t,V_t)Q_t$ and using Kolmogorov backward equation for $P(t,V_t)$ yields

$$d\left(P(t,V_t)Q_t\right) = Q_t\left(P(t,V_t)h(t,V_t) + a(t,V_t)^T\nabla P(t,V_t)\right)d\tilde{\eta}_t$$

Integrating $d\left(P(t,V_t)Q_t\right)$ from 0 to T gives

$$P(T,V_T)Q_T = P(0,V_0)Q_0 + \int_0^T Q_t\left(P(t,V_t)h(t,V_t) + a(t,V_t)^T\nabla P(t,V_t)\right)d\tilde{\eta}_t$$

This reduces to

$$\phi(V_T)Q_T = P(0,v) + \int_0^T Q_t\left(P(t,V_t)h(t,V_t) + a(t,V_t)^T\nabla P(t,V_t)\right)d\tilde{\eta}_t$$

Therefore, the variances of Monte Carlo estimators under $\tilde{\mathbb{P}}$ and $\mathbb{P}$ are given by

$$\begin{aligned} Var_{\tilde{\mathbb{P}}}\left(\phi(V_T)Q_T\right) &= \tilde{\mathbb{E}}\left[\int_0^T Q_t^2 ||P(t,V_t)h(t,V_t) + a(t,V_t)^T \nabla P(t,V_t)||^2 dt\right] \\ Var_{\mathbb{P}}\left(\phi(V_T)\right) &= \tilde{\mathbb{E}}\left[\int_0^T ||a(t,V_t)^T \nabla P(t,V_t)||^2 dt\right] \end{aligned}$$

Hence, the optimal choice of h for which the variance of $\phi(V_T)Q_T$ under $\tilde{\mathbb{P}}$ is

$$h(t,V_t) = -\frac{1}{P(t,V_t)} a(t,V_t)^T \nabla P(t,V_t) \tag{3}$$

In Fouque's and Tullie's work importance sampling is associated with $h(t,v)$ as it dictates the change of drift of the stochastic process given by equation (2). An appropriate choice of $h(t,v)$ leads to smaller variance for Monte Carlo estimator under $\tilde{\mathbb{P}}$ than under $\mathbb{P}$. Unfortunately, $h(t,v)$ can not be calculated directly from equation (3) as it depends on $P(t,v)$ which is an unknown. However, this equation gives some intuition about the possible choice of $h(t,v)$. To determine $h(t,v)$, Fouque and Tullie approximate $P(t,v)$ by its fast mean reversion expansion. Then they compare fast mean reversion with small noise expansion.

P_{SNE} is a small noise expansion of $P(t,v)$ given by the Black-Scholes formula

$$P_{SNE} = xN(d_1) - Ke^{-r(T-t)}N(d_2)$$

with $d_1 = \frac{\ln(\frac{x}{K}) + (r + \frac{1}{2}\sigma^2(y))(T-t)}{\sigma(y)\sqrt{T-t}}$ and $d_2 = d_1 - \sigma(y)\sqrt{T-t}$. Here h takes the following form

$$h = -\frac{1}{P_{SNE}} \begin{pmatrix} x\sigma(y)\frac{\partial P_{SNE}}{\partial x} \\ 0 \end{pmatrix} \tag{4}$$

P_{FMR} is a fast mean reversion expansion of $P(t,v)$ derived from the following equation

$$P_{FMR} = P_{BS(\bar{\sigma})} - (T-t)\left(V_2 x^2 \frac{\partial^2 P_{BS(\bar{\sigma})}}{\partial x^2} + V_3 x^3 \frac{\partial^3 P_{BS(\bar{\sigma})}}{\partial x^3}\right) \tag{5}$$

where $P_{BS(\bar{\sigma})}$ is given by

$$xN(d_1) - Ke^{-r(T-t)}N(d_2)$$

with $d_1 = \frac{\ln(\frac{x}{K}) + (r + \frac{1}{2}\bar{\sigma}^2)(T-t)}{\sigma(y)\sqrt{T-t}}$, $d_2 = d_1 - \bar{\sigma}\sqrt{T-t}$ and V_2, V_3 are parameters calibrated from the implied volatility skew. In this case, h takes the following form

$$h = -\frac{1}{P_{FMR}} \begin{pmatrix} x\sigma(y)\frac{\partial P_{FMR}}{\partial x} \\ 0 \end{pmatrix} \tag{6}$$

In general, Small Noise Expansion works best when the mean reversion rate is low

($\alpha \to 0$), whereas Fast Mean Reversion works best when the mean reversion rate is high ($\alpha \to \infty$).

3 Numerical Implementation

Following the importance sampling schemes described in the previous section, I employ small noise expansion and fast mean reversion to price a European call using the following parameters:

$$
\begin{aligned}
r = 0.1,& \quad \sigma(y) = exp(y) \\
m = -2.6,& \ \ \nu = 1, \ \ \rho = -0.3 \\
X_0 = 110,& \quad exp(y_0) = 0.00983 \\
K = 100,& \quad T = 1, \quad \alpha = 1
\end{aligned}
$$

where r is the risk-free rate, ρ is the correlation in equation (1). As for the fast mean reversion, I use order zero expansion, where (5) reduces to $P_{BS(\bar{\sigma})}$.

In order to apply importance sampling, I calculate h according to equation (4). For FMR it is given by equation (6). Implementation is done in Matlab and using the following NAG routines:

- s30ab - Black-Scholes-Merton option pricing formula with greeks
- g05kf - initialization function for pseudo random number generator using
- g05sk - pseudo random number generator

Figures 1 and 2 plot the results of basic Monte Carlo (MC), small noise expansion (SNE) and fast mean reversion (FMR). The numerical example shows that both schemes converge well compared to the plain Monte Carlo simulations, while the Fast Mean Reversion performs slightly better the Small Noise Expansion based on the speed of price convergence and with a small variance. According to Fouque and Tullie (2002) Fast Mean Reversion outperforms Small Noise Expansion even when the mean reversion rate is 1.

References

[1] Fouque J.P. and T. Tullie (2002) Variance Reduction for Monte Carlo Simulation in a Stochastic Volatility Environment. Quantitative Finance, 2(1):24-30.

Figure 1: Price

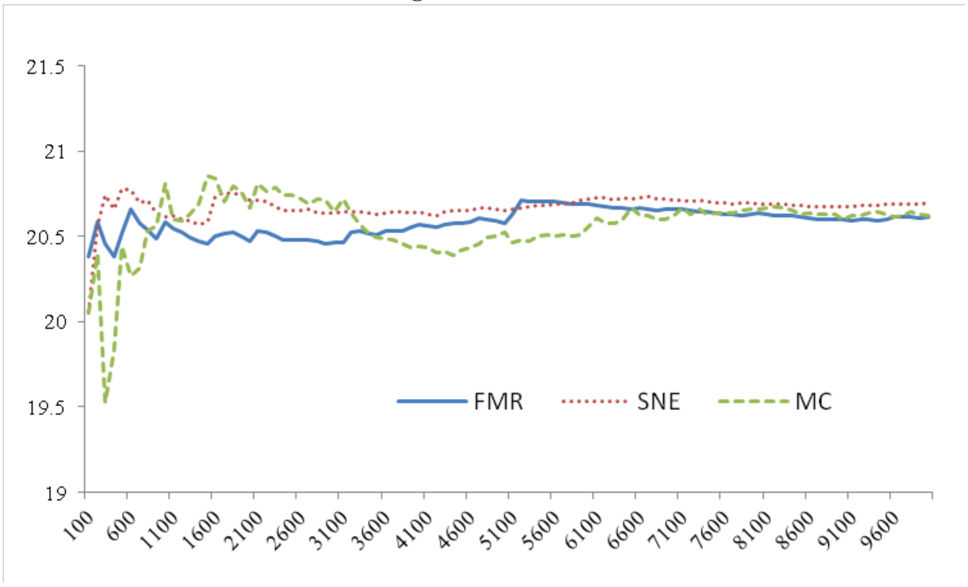

Figure 2: Variance