Sample Path Large Deviations and Optimal Importance Sampling for Stochastic Volatility Models

Scott Robertson *†

October 19, 2009

Abstract

Sample path Large Deviation Principles (LDP) of the Freidlin-Wentzell type are derived for a class of diffusions which govern the price dynamics in common stochastic volatility models from Mathematical Finance. LDP are obtained by relaxing the non-degeneracy requirement on the diffusion matrix in the standard theory of Freidlin and Wentzell. As an application, a sample path LDP is proved for the price process in the Heston stochastic volatility model.

Using the sample path LDP for the Heston model, the problem is considered of selecting an importance sampling change of drift, for both the price and the volatility, which minimize the variance of Monte Carlo estimators for path dependent option prices. An asymptotically optimal change of drift is identified as a solution to a two dimensional variational problem. The case of the arithmetic average Asian put option is solved in detail.

Mathematics Subject Classification (2010): 60F10, 91G20, 91G60, 91B25
Mathematics Subject Classification (2000): 60F10, 65C05, 91B28

Keywords: Importance Sampling, Large Deviations, Monte Carlo Methods, Stochastic Volatility, Variance Reduction

1 Introduction

Sample path Large Deviations Principles (LDP) yield approximations for the probability that the path of a random process lies in a particular set of paths. The LDP results of Freidlin & Wentzell (1984) concern the family of diffusions

$$dX_t^\varepsilon = b(X_t^\varepsilon)dt + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)dW_t,$$

*Partially supported by the National Science Foundation under grants DGE-0221680 and DMS-0532390

†Department of Mathematical Sciences, Wean Hall 6113, Carnegie Mellon University, Pittsburgh, PA 15213 Email: scottrob@andrew.cmu.edu

for b, σ bounded, Lipschitz functions and small $\varepsilon > 0$. As $\varepsilon \downarrow 0$ LDP estimates are of the form

$$\lim_{\varepsilon \downarrow 0} \varepsilon \log P(X^\varepsilon \in A) = -\inf_{x \in A} I(x),$$

where A is a set of paths and I the rate function governing the sample path LDP for X. This paper derives sample path LDP[1] for a class of diffusions which govern the price dynamics in common stochastic volatility models from Mathematical Finance. In a typical stochastic volatility model, the asset price S and volatility v evolve according to

$$\begin{aligned} \frac{dS_t}{S_t} &= r dt + \sqrt{v_t} dW_t, \\ dv_t &= b(v_t)dt + \sigma(v_t)dB_t, \end{aligned} \tag{1.1}$$

where r is the interest rate and W, B are standard Brownian motions with $d\langle W, B\rangle_t = \rho dt$. By taking the orthogonal decomposition $W = \rho B + \bar{\rho} Z$ where Z is a standard Brownian motion independent of B and $\bar{\rho} = \sqrt{1 - \rho^2}$ it follows that the dynamics for S are governed by the two dimensional diffusion

$$\begin{aligned} dv_t &= b(v_t)dt + \sigma(v_t)dB_t, \\ dY_t &= \bar{\rho}\sqrt{v_t} dZ_t, \end{aligned} \tag{1.2}$$

in the sense that, up to a mild restriction upon σ, S is a continuous function of $\{v, Y\}$ on $[0, T]$[2].

Since the square root function is both degenerate and non locally Lipschitz the standard theory of Freidlin-Wentzell does not apply to the process in (1.2). Thus, in order to derive LDP for the price process S, it is first necessary to derive LDP for two dimensional diffusions of the form

$$\begin{aligned} dX_t &= b(X_t)dt + \sigma(X_t)dB_t, \\ dY_t &= f(X_t)dZ_t, \end{aligned} \tag{1.3}$$

where it is assumed a LDP holds for X but the function f is not required to be locally Lipschitz and non-degenerate on the state space of X. The results of Donati-Martin, Rouault, Yor & Zani (2004) show this assumption is valid in the model of Heston (1993). It is also valid in the model of Hull & White (1987).

If a LDP exists for $\{X, Y\}$ in (1.3) or correspondingly for $\{v, Y\}$ in (1.2) then, aside from some minor technicalities, a LDP for S follows via the Contraction Principle. The technicalities arise when introducing the parameter ε into the stochastic differential equation (SDE) for $\{S, v\}$. They concern integrability requirements upon v and must be handled on a case by case basis. Here, the details are worked out for the Heston (1993) model but the results extend to other models as well (e.g. Hull & White (1987)).

[1] When there is no risk of confusion, "LDP" will replace "sample path LDP"

[2] The restriction upon σ is that $\int_0^\cdot \rho\sqrt{v_s} dB_s$ is almost surely a continuous function of v on $[0, T]$

The use of LDP is widespread in Mathematical Finance. For an overview, see Pham (2007). With respect to sample path LDP and option pricing, the small time estimates for the transition kernel of a diffusion first obtained by Varadhan (1967) have been used to derive asymptotic formulae for the implied, local and effective volatilities near expiry for call options. For example, Avellaneda, Boyer-Olson, Busca & Friz (2003) and Avellaneda, Boyer-Olson, Busca & Friz (2002) consider index options in a local volatility model and Berestycki, Busca & Florent (2004) consider the effective volatility in a general stochastic volatility framework.

The application of sample path LDP considered here concerns importance sampling for pricing path dependent options where the underlying price evolves according to (1.1) under a given probability measure P. The option price takes the form

$$E_P\left[G(S)\right], \tag{1.4}$$

where G is a non-negative functional of the entire path of S on a time interval $[0,T]$. In most situations computing (1.4) requires Monte Carlo simulation: sampling paths of S according to P and taking averages. Importance sampling is a variance reduction technique based on the fact that if Q is a measure equivalent to P then

$$G(S)\frac{dP}{dQ} \tag{1.5}$$

is an unbiased Q estimator of (1.4). The Q variance of this estimator is

$$E_P\left[G^2(S)\frac{dP}{dQ}\right] - E_P\left[G(S)\right]^2. \tag{1.6}$$

Importance sampling seeks to identify the measure Q which minimizes (1.6) over an acceptable[3] class of equivalent measures, usually where each measure corresponds to a different shift in the average price path of S.

Only the first term in (1.6) varies with Q. Computing this term is at least as difficult as (1.4) and hence some approximation is necessary. One possible approximation is implied by Varadhan's Integral Lemma:

$$\log E_P\left[G^2(S)\frac{dP}{dQ}\right] \approx \sup\left(2\log G + \log\frac{dP}{dQ} - I\right), \tag{1.7}$$

where I is the rate function governing the LDP for the pair $\{v, Y\}$. An asymptotically optimal change of measure is thus identified as a solution to a min-max problem.

Glasserman, Heidelberger & Shahabuddin (1999) and Guasoni & Robertson (2008) study selecting an asymptotically optimal change of measure when the underlying price follows the Black-Scholes model where volatility is constant. Glasserman et al. (1999) discretize the time interval in order to use the finite dimensional LDP results of Cramér (1938) and Chernoff (1952). Guasoni &

[3] Q defined via $dQ/dP = G(S)/(E_P\left[G(S)\right])$ achieves zero variance but this choice of Q is not allowed since it requires knowledge of (1.4).

Robertson (2008) update the methodology to a continuous time setting, using the results of Schilder (1966).

The advantage of updating the methodology in Guasoni & Robertson (2008) to stochastic volatility models is that there are now changes of drift for both the asset price and volatility. Since there is control over the average volatility, an optimal drift may be selected for path-dependent options for which there is no obvious direction in which to move the asset price. However, without a LDP for $\{v, Y\}$ the justification for selecting an asymptotically optimal change of measure is baseless. If the pair $\{v, Y\}$ do satisfy a LDP the argument of Guasoni & Robertson (2008) holds in spirit, though the details are significantly more complicated.

This approach based upon Varadhan's Integral Lemma results in a deterministic change of drift and is motivated by its ease of implementation into the Monte Carlo simulation. An alternate importance sampling scheme using stochastic changes of drift and LDP approximations as $T - t \to 0$ to the value function

$$v(x,t) = E_P \left[G\left(S_s, t \leq x \leq T\right) \middle| S_t = x \right],$$

is given in Fournié, Lasry & Touzi (1997) and Pham (2007).

1.1 Outline of Paper

The LDP results for (1.3) are given in Section 2. The only assumptions made on X and f are in Assumption 2.1: namely, that X satisfies a LDP and f is differentiable on the state-space of X. The main result is Theorem 2.2, which gives two variational conditions in equations (2.7) and (2.8) that allow application of the Approximately Continuous Contraction Principle (Dembo & Zeitouni 1998, Theorem 4.2.23) and hence are sufficient for a LDP to hold. Corollary 2.3 gives a single variational condition, (2.12), which implies the two conditions in Theorem 2.2, easing its implementation.

In Section 3, it is shown the price process in the Heston (1993) model satisfies a LDP. In the Heston model, the volatility v follows a CIR process and hence v satisfies a LDP via Donati-Martin et al. (2004, Theorem 1.3). Lemma 3.1 proves the LDP for $\{v, Y\}$ by invoking Corollary 2.3. The particular form the SDE for v, see (3.2) and (3.4), allows the price S to be viewed as an approximate contraction of $\{v, Y\}$. Proposition 3.2 proves S satisfies an LDP by showing the approximate part of the contraction disappears on a Large Deviations scale.

Section 4 outlines the methodology for selecting an asymptotically optimal importance sampling change of measure for pricing path dependent options in the Heston model. Through the lens of Girsanov's Theorem, the family of candidate measures correspond to shifts in both the volatility and the independent Brownian motion and hence are parameterized by two functions as in (4.4).

Asymptotic Optimality is formally defined in Definition 4.1 and a candidate asymptotically optimal measure is given by the solution to the variational problem in (4.9). Proposition 4.2 gives growth conditions under which Varadhan's Integral Lemma may be applied and Proposition 4.3 shows, in the case of bounded, continuous functionals, existence of maximizers to (4.9). Section 4 concludes by considering the arithmetic average Asian Put option. A numerical example is given to show the type of shifts corresponding to an asymptotically optimal measure.

2 Large Deviations Results

In accordance with (1.3), LDP are considered for the two-dimensional diffusion

$$\begin{aligned} dX_t^\varepsilon &= b(X_t^\varepsilon)dt + \sqrt{\varepsilon}\sigma(X_t^\varepsilon)dB_t, \\ dY_t^\varepsilon &= \sqrt{\varepsilon} f(X_t^\varepsilon)dZ_t, \end{aligned} \tag{2.1}$$

in the limit $\varepsilon \downarrow 0$ on a fixed time interval $[0, T]$ with $X_0^\varepsilon = x, Y_0^\varepsilon = 0$. The underlying probability space is $\left(C\left([0,T];\mathbb{R}^2\right), \mathcal{F}_T, P\right)$ where P is Wiener measure, defined on the completion of $\mathcal{B}\left(C\left([0,T];\mathbb{R}^2\right)\right)$ given by $\mathcal{F}_T$. The open sets are generated by the uniform norm. Under this setup, the coordinate mapping process (B, Z) is a standard Brownian Motion with respect to the filtration $\{\mathcal{F}_t\}_{t\in[0,T]}$, the usual augmentation of the natural filtration of (B, Z).

Let $E \subset \mathbb{R}$ be an open interval. It is assumed there is an $\varepsilon_0 > 0$ such that for $0 < \varepsilon < \varepsilon_0$ a strong solution to (2.1) exists and that

$$P\left(X_t^\varepsilon \in E; 0 \leq t \leq T\right) = 1. \tag{2.2}$$

For ease of exposition, LDP results will be stated for $\{X^\varepsilon, Y^\varepsilon\}$ and not for the measures P^ε induced by $\{X^\varepsilon, Y^\varepsilon\}$ under P for $0 < \varepsilon < \varepsilon_0$.

The classical result regarding sample path LDP is attributable to Freidlin & Wentzell (1984). Specified to (2.1) with $E = \mathbb{R}$, it states (Deuschel & Stroock 1989, Chapter 1.4) that if

i) For some $M > 0$ and all $x \in \mathbb{R}$:

$$0 < \sigma^2(x), f^2(x) < M(1 + x^2), \qquad b^2(x) \leq M(1 + x^2), \tag{2.3}$$

ii) For each $r > 0$ there is an $M_r > 0$ such that for $x, y \in \mathbb{R}, |x|, |y| < r$:

$$|\sigma^2(x) - \sigma^2(y)| + |f^2(x) - f^2(y)| + |b(x) - b(y)| \leq M_r|y - x|, \tag{2.4}$$

then $\{X^\varepsilon, Y^\varepsilon\}$ satisfy a LDP in $C\left([0,T];\mathbb{R}^2\right)$ with the good rate function

$$I^X(\phi, \psi) = \begin{cases} \frac{1}{2}\int_0^T \left(\left(\frac{\dot{\phi}_t - b(\phi_t)}{\sigma(\phi_t)}\right)^2 + \left(\frac{\dot{\psi}_t}{f(\phi_t)}\right)^2\right) dt & (\phi, \psi) \in \mathbb{H}_T^x \\ \infty & else \end{cases}, \tag{2.5}$$

where, specified to the case $d = 2$,

$$\mathbb{H}_T^x = \left\{\phi : \phi_t = x + \int_0^t \psi_s ds, \psi \in L_2\left([0,T];\mathbb{R}^d\right)\right\}. \tag{2.6}$$

This paper is concerned with relaxing the non-degeneracy and local-Lipschitz requirement on f. However, since the motivating example is where $f(x) = \sqrt{x}$ and X^ε is a positive process, it is desirable simply to assume that $\{X^\varepsilon\}$ satisfies a LDP and not necessarily that b, σ satisfy the requirements in (2.3) and (2.4). Therefore, in addition to the existence of a strong solution to (2.1) with X^ε satisfying (2.2), the following is assumed regarding X, f:

Assumption 2.1. *$\{X^\varepsilon\}$ satisfy a LDP on $C\left([0,T];\bar{E}\right)$ with good rate function I^X with domain*

$$DI^X = \left\{\phi \in C\left([0,T];\bar{E}\right) : I^X(\phi) < \infty\right\} \subset \mathbb{H}_T^x.$$

$f \in C(\bar{E};\mathbb{R})$ is differentiable in E.

Assumption 2.1 permits extensions of the Freidlin-Wentzell theory for $\{X^\varepsilon\}$ to cases where the SDE state space is not $\mathbb{R}$ or where $DI^X \neq \mathbb{H}_T^x$. Under Assumption 2.1, the following theorem gives two variational conditions sufficient for $\{X^\varepsilon, Y^\varepsilon\}$ to satisfy a LDP:

Theorem 2.2. *Let Assumption 2.1 hold. If*

i) For each $\alpha > 0$ there exists a $\beta(\alpha) > 0$ such that

$$\gamma(\alpha) \equiv \sup_{\{I^X(\phi)\leq\alpha\}} \left(\beta(\alpha)\int_0^T \dot{f}(\phi_t)^2\dot{\phi}_t^2 dt - I^X(\phi)\right) < \infty, \tag{2.7}$$

ii) $\forall \delta > 0$

$$\limsup_{m\uparrow\infty} \sup_{\{\phi\in DI^X\}} \left(\frac{-\delta^2 m^2}{2T^2\int_0^T \dot{f}(\phi_t)^2\dot{\phi}_t^2 dt} - I^X(\phi)\right) = -\infty, \tag{2.8}$$

Then $\{X^\varepsilon, Y^\varepsilon\}$ satisfy a LDP on $C\left([0,T];\bar{E}\times\mathbb{R}\right)$ with the good rate function

$$I^{X,Y}(\phi,\psi) = \inf\left\{I^X(\phi) + I^Z(\varphi) : \psi_\cdot = \int_0^\cdot f(\phi_t)\dot{\varphi}_t dt\right\}, \tag{2.9}$$

where

$$I^Z(\varphi) = \begin{cases} \frac{1}{2}\int_0^T \dot{\varphi}_t^2 dt & \varphi \in \mathbb{H}_T^0 \\ +\infty & else \end{cases}. \tag{2.10}$$

Furthermore, $\{Y^\varepsilon\}$ satisfies a LDP with the good rate function

$$I^Y(\psi) = \inf\left\{I^{X,Y}(\phi,\psi) : \phi \in C\left([0,T];\bar{E}\right)\right\}. \tag{2.11}$$

The following corollary shows that if the constants $\gamma(\alpha)$ and $\beta(\alpha)$ from (2.7) are uniform in α then condition (2.8) is automatically satisfied.

Corollary 2.3. *If there exists $\beta > 0$ such that*

$$\gamma \equiv \sup_{\{\phi\in DI^X\}} \left(\beta\int_0^T \dot{f}(\phi_t)^2\dot{\phi}_t^2 dt - I^X(\phi)\right) < \infty, \tag{2.12}$$

then conditions (2.7) and (2.8) are satisfied.

2.1 Proofs of Large Deviations Results

The proof of Theorem 2.2 is close in spirit to the standard Freidlin & Wentzell (1984) proof. It exploits the Approximately Continuous Contraction Principle (Dembo & Zeitouni 1998, Theorem 4.2.23) and Schilder's Theorem (Schilder 1966) which identifies the rate function for the process $\{X^\varepsilon\}$ in (2.1) when $b = 0, \sigma = 1$.

Proof of Theorem 2.2. For each $m \in \mathbb{N}^+$ consider the partition of $[0, T]$, $\{t_i\}_{i=0,\dots,m}$ defined by $t_i = iT/m$. Define the function $g_m : C\left([0,T]; \bar{E} \times \mathbb{R}\right) \mapsto C\left([0,T]; \mathbb{R}\right)$ by

$$g_m(\phi, \varphi)_t = \sum_{i=1}^{m(t)} f(\phi_{t_{i-1}})(\varphi_{t_i} - \varphi_{t_{i-1}}) + f(\phi_{t_{m(t)}})\left(\varphi_t - \varphi_{t_{m(t)}}\right), \tag{2.13}$$

where $m(t) = \max\{j : t_j < t\}$. Assumption 2.1 implies g_m is a continuous map. Define the measurable map $g : DI^X \times \mathbb{H}_T^0 \mapsto \mathbb{H}_T^0$ by

$$g(\phi, \varphi)_t = \int_0^t f(\phi_s)\dot{\varphi}_s ds. \tag{2.14}$$

Let $Z^\varepsilon = \sqrt{\varepsilon} Z$. Schilder's Theorem yields an LDP for $\{Z^\varepsilon\}$ with the good rate function I^Z of (2.10). Since $\{X^\varepsilon\}$ and $\{Z^\varepsilon\}$ are independent, the pair $\{X^\varepsilon, Z^\varepsilon\}$ satisfy a LDP (Dembo & Zeitouni 1998, Chapter 4.2) on $C\left([0,T]; \bar{E} \times \mathbb{R}\right)$ with good rate function:

$$I^{X,Z}(\phi, \varphi) = I^X(\phi) + I^Z(\varphi). \tag{2.15}$$

Set $Y^{\varepsilon,m} = g_m(X^\varepsilon, Z^\varepsilon)$. For each m, the pair $\{X^\varepsilon, Y^{\varepsilon,m}\}$ is a continuous function of $\{X^\varepsilon, Z^\varepsilon\}$ and hence satisfies a LDP via the Contraction Principle. Since $\{X^\varepsilon\}$ is not changing with m, for the Approximately Continuous Contraction Principle to hold it suffices to show for each $\alpha > 0$

$$\limsup_{m\uparrow\infty} \sup_{\{I^{X,Z}(\phi,\varphi)\le\alpha\}} \sup_{0\le t\le T} |g_m(\phi,\varphi)_t - g(\phi,\varphi)_t| = 0, \tag{2.16}$$

and for each $\delta > 0$

$$\lim_{m\uparrow\infty} \limsup_{\varepsilon\downarrow 0} \varepsilon \log P\left(\sup_{0\le t\le T} |Y_t^{\varepsilon,m} - Y_t^\varepsilon| > \delta\right) = -\infty. \tag{2.17}$$

The requirement (2.16) is handled first. Let $\alpha > 0, \phi \in DI^X, \psi \in \mathbb{H}_T^0$. For any $0 \le a < b \le T$, Hölder's inequality yields

$$\left|\int_a^b \left(f(\phi_s) - f(\phi_a)\right)\dot{\psi}_s ds\right| \le \frac{b-a}{2}\left(\int_a^b \left(\dot{f}(\phi_s)\dot{\phi}_s\right)^2 ds + \int_a^b \left(\dot{\psi}_s\right)^2 ds\right). \tag{2.18}$$

Therefore

$$
\begin{aligned}
&|g(\phi,\varphi)_t - g_m(\phi,\varphi)_t| \\
&\qquad \leq \sum_{i=1}^{m(t)} \left| \int_{t_{i-1}}^{t_i} \left(f(\phi_s) - f(\phi_{t_{i-1}})\right) \dot{\varphi}_s ds \right| + \left| \int_{t_{m(t)}}^{t} \left(f(\phi_s) - f(\phi_{t_{m(t)}})\right) \dot{\varphi}_s ds \right| \\
&\qquad \leq \frac{T}{2m} \left(\int_0^T \left(\dot{f}(\phi_s)\dot{\phi}_s^2\right)^2 ds + \int_0^T \dot{\varphi}_s^2 ds \right).
\end{aligned}
$$

Using $\beta(\alpha), \gamma(\alpha)$ from (2.7) and the fact that $I^X(\phi) + I^Z(\varphi) \leq \alpha$

$$
\frac{T}{2m} \left(\int_0^T \left(\dot{f}(\phi_s)\dot{\phi}_s^2\right)^2 ds + \int_0^T \dot{\varphi}_s^2 ds \right) \leq \frac{T}{2m} \left(\frac{\gamma(\alpha) + \alpha}{\beta(\alpha)} + 2\alpha \right).
$$

Since the quantity on the right hand side goes to 0 uniformly in t, ϕ and ψ as $m \uparrow \infty$, the equality in (2.16) is satisfied.

Equation (2.17) is now handled. Fix $\delta > 0$ and set

$$
M_t^{\varepsilon,m} \equiv Y_t^{\varepsilon} - Y_t^{\varepsilon,m} = \sqrt{\varepsilon} \int_0^t h^m(X^{\varepsilon})_s dZ_s,
$$

where

$$
h^m(x_s) = f(x_s) - f(x_{t_{i-1}}), s \in [t_{i-1}, t_i). \tag{2.19}
$$

Extend $M^{\varepsilon,m}$ to $t > T$ by setting $h_t^m = K$ for some non-zero constant K. Thus, $M^{\varepsilon,m}$ is a continuous local martingale on $[0, \infty)$ with quadratic variation

$$
\langle M^{\varepsilon,m} \rangle_t = \varepsilon \int_0^t (h^m(X^{\varepsilon})_s)^2 ds,
$$

and so $\langle M^{\varepsilon,m} \rangle_t \uparrow \infty$ as $t \uparrow \infty$. Therefore, from the time change theorem for continuous local martingales (Karatzas & Shreve 1991, Theorem 3.4.6),

$$
M_t^{\varepsilon,m} \stackrel{d}{=} W_{\varepsilon \int_0^t (h^m(X^{\varepsilon})_s)^2 ds},
$$

where W is a standard Brownian motion under P. Conditioning on X^{ε}:

$$
\begin{aligned}
P\left[\sup_{0 \leq t \leq T} |M_t^{\varepsilon,m}| > \delta \middle| X^{\varepsilon} \right] &= P\left[\sup_{0 \leq t \leq \varepsilon \int_0^T (h^m(X^{\varepsilon})_s)^2 ds} |W_t| > \delta \middle| X^{\varepsilon} \right] \\
&\leq 4 \exp\left(\frac{1}{\varepsilon} \left(\frac{-\delta^2}{2 \int_0^T (h^m(X^{\varepsilon})_s)^2 ds} \right) \right),
\end{aligned}
$$

where the last inequality comes from Dembo & Zeitouni (1998, Lemma 5.2.1). Note that

$$
\frac{-\delta^2}{2 \int_0^T (h^m(X^{\varepsilon})_s)^2 ds} \tag{2.20}
$$

is a continuous, non-positive, functional of X^ε. Since $\{X^\varepsilon\}$ satisfies an LDP with the good rate function I^X, Varadhan's Integral Lemma (Dembo & Zeitouni 1998, Theorem 4.3.1) yields

$$\lim_{\varepsilon\downarrow 0}\varepsilon\log E_P\left[\exp\left(\frac{1}{\varepsilon}\left(\frac{-\delta^2}{2\int_0^T (h^m(X^\varepsilon)_s)^2 ds}\right)\right)\right]$$
$$= \sup_{\phi\in DI^X}\left(\frac{-\delta^2}{2\sum_{i=1}^m\int_{t_{i-1}}^{t_i}(f(\phi_s)-f(\phi_{t_{i-1}}))^2 ds} - I^X(\phi)\right),$$

where the last equality follows by substituting for h^m. For any $0\le a<b\le T$

$$\int_a^b (f(\phi_s)-f(\phi_a))^2\, ds \le (b-a)^2\int_a^b \left(\dot f(\phi_s)\dot\phi_s\right)^2 ds.$$

Thus, since $t_i - t_{i-1} = \frac{T}{m}$

$$\frac{-\delta^2}{2\sum_{i=1}^m\int_{t_{i-1}}^{t_i}(f(\phi_s)-f(\phi_{t_{i-1}}))^2 ds} \le \frac{-m^2\delta^2}{2T^2\int_0^T \dot f(\phi_s)^2\dot\phi_s^2 ds}.$$

This leaves the inequality

$$\limsup_{\varepsilon\downarrow 0}\varepsilon\log P\left[\sup_{0\le t\le T}|Y^{\varepsilon,m}-Y^\varepsilon|>\delta\right] \le \sup_{\phi\in DI^X}\left(\frac{-m^2\delta^2}{2T^2\int_0^T \dot f(\phi_s)^2\dot\phi_s^2 ds} - I^X(\phi)\right).$$

But, because of (2.8)

$$\limsup_{m\uparrow\infty}\sup_{\phi\in DI^X}\left(\frac{-m^2\delta^2}{2T^2\int_0^T \dot f(\phi_s)^2\dot\phi_s^2 ds} - I^X(\phi)\right) = -\infty,$$

and the equality in (2.17) holds finishing the proof.

□

Proof of Corollary 2.3. If (2.12) holds then clearly (2.7) holds. Furthermore, for each $\phi\in DI^X$

$$\frac{-\delta^2 m^2}{2T^2\int_0^T \dot f(\phi_t)^2\dot\phi_t^2 dt} - I^X(\phi) = \frac{-\delta^2 m^2}{2T^2\int_0^T \dot f(\phi_t)^2\dot\phi_t^2 dt} \pm \beta\int_0^T \dot f(\phi_t)^2\dot\phi_t^2 dt - I^X(\phi)$$
$$\le \frac{-\delta^2 m^2}{2T^2\int_0^T \dot f(\phi_t)^2\dot\phi_t^2 dt} - \beta\int_0^T \dot f(\phi_t)^2\dot\phi_t^2 dt + \gamma.$$

For $A,B>0$, on $x>0$

$$-\frac{A}{x} - Bx \le -2\sqrt{AB}.$$

Thus

$$\frac{-\delta^2 m^2}{2T^2\int_0^T \dot f(\phi_t)^2\dot\phi_t^2 dt} - I^X(\phi) \le -\frac{\delta m\sqrt{2\beta}}{T} + \gamma.$$

Since this bound is uniform, taking the limit at $m\uparrow\infty$ yields the result. □

3 The Heston Stochastic Volatility Model

In this section, Theorem 2.2 is used to establish a LDP for the price process in the stochastic volatility model of Heston (1993). In the model, the asset price and volatility dynamics are

$$\begin{aligned} \frac{dS_t}{S_t} &= rdt + \sqrt{v_t}dW_t. \\ dv_t &= \kappa\left(\theta - v_t\right)dt + \xi\sqrt{v_t}dB_t, \\ d\langle W, Z\rangle_t &= \rho dt. \end{aligned} \tag{3.1}$$

It is assumed $\kappa, \theta, \xi > 0$ and $-1 < \rho < 1$. To ensure the strict positivity of v, it is required that $\kappa\theta > \xi^2/2$. The Brownian motion W is decomposed $W = \rho B + \bar{\rho} Z$ where B and Z are independent Brownian motions and $\bar{\rho} = \sqrt{1-\rho^2}$. S_0 and v_0 are assumed constant. For each $0 < \varepsilon < 1$ the corresponding SDE is

$$\begin{aligned} \frac{dS^\varepsilon_t}{S^\varepsilon_t} &= rdt + \rho\sqrt{\varepsilon}\sqrt{v^\varepsilon_t}dB_t + \bar{\rho}\sqrt{\varepsilon}\sqrt{v^\varepsilon_t}dZ_t, \\ dv^\varepsilon_t &= \kappa\left(\theta - v^\varepsilon_t\right)dt + \xi\sqrt{\varepsilon}\sqrt{v^\varepsilon_t}dB_t. \end{aligned} \tag{3.2}$$

Setting

$$Y^\varepsilon_\cdot = \int_0^\cdot \sqrt{\varepsilon}\sqrt{v^\varepsilon_t}dZ_t, \tag{3.3}$$

$S^\varepsilon = S^\varepsilon\left(v^\varepsilon, Y^\varepsilon\right)$ where

$$S^\varepsilon(\phi,\psi)_t = S_0 \exp\left(rt - \frac{\varepsilon}{2}\int_0^t \phi_s ds + \frac{\rho}{\xi}\left(\phi_t - v_0 - \kappa\int_0^t \left(\theta - \phi_s\right)ds\right) + \bar{\rho}\psi_s\right) \tag{3.4}$$

is a continuous, ε-dependent functional on $C\left([0,T]; \mathbb{R}_+ \times \mathbb{R}\right)$. To prove $\{S^\varepsilon\}$ satisfy a LDP it suffices to prove

i) The pair $\{v^\varepsilon, Y^\varepsilon\}$ satisfy a LDP with a good rate function,

ii) $\{\log S^\varepsilon(v^\varepsilon, Y^\varepsilon)\}$ is exponentially equivalent (Dembo & Zeitouni 1998, Definition 4.2.10) to $\{\log S(v^\varepsilon, Y^\varepsilon)\}$ for

$$S(\phi,\psi)_t = S_0 \exp\left(rt + \frac{\rho}{\xi}\left(\phi_t - v_0 - \kappa\int_0^t \left(\theta - \phi_s\right)ds\right) + \bar{\rho}\psi_s\right). \tag{3.5}$$

Indeed, if $i)$ holds true then $\{\log S(v^\varepsilon, Y^\varepsilon)\}$ satisfies a LDP with good rate function via the Contraction Principle. If $ii)$ holds true then $\{\log S^\varepsilon(v^\varepsilon, Y^\varepsilon)\}$ satisfies a LDP from via Dembo & Zeitouni (1998, Theorem 4.2.13). The sample path LDP for $\{S^\varepsilon(v^\varepsilon, Y^\varepsilon)\}$ then follows from the Contraction Principle.

Regarding $\{v^\varepsilon, Y^\varepsilon\}$, Donati-Martin et al. (2004, Theorem 1.3) prove that $\{v^\varepsilon\}$ satisfies a LDP in $C_{v_0}\left([0,T]; \mathbb{R}_+\right)$ with the good rate function

$$I^v\left(\phi\right) = \begin{cases} \frac{1}{2}\int_0^T \frac{\left(\dot{\phi}_t - \kappa(\theta - \phi_t)\right)^2}{\xi^2 \phi_t}dt & \phi \in DI^v \\ +\infty & else \end{cases}, \tag{3.6}$$

where

$$DI^v = \left\{ \phi \in C_{v_0}\left([0,T];\mathbb{R}_+\right) : \frac{\dot\phi - \kappa(\theta-\phi)}{\sqrt{\phi}} \in L^2[0,T] \right\}. \tag{3.7}$$

The following Lemma uses Corollary 2.3 to prove $\{v^\varepsilon, Y^\varepsilon\}$ satisfies a LDP:

Lemma 3.1. *$\{v^\varepsilon\}$ and $f(x) = \sqrt{x}$ satisfy Assumption 2.1. Furthermore, for any $\beta < \frac{2}{\xi^2}$*

$$\sup_{\phi\in DI^v}\left(\beta\int_0^T \frac{\dot\phi_t^2}{4\phi_t}dt - I^v(\phi)\right) < \infty,$$

and hence by Corollary 2.3, $\{v^\varepsilon, Y^\varepsilon\}$ satisfy a LDP with the good rate function

$$I^{v,Y}(\phi,\psi) = \inf\left\{ I^v(\phi) + I^Z(\varphi) : \psi_\cdot = \int_0^\cdot \sqrt{\phi_s}\dot\varphi_s ds \right\}. \tag{3.8}$$

The following proposition affirms the exponential equivalence of $\{\log S^\varepsilon(v^\varepsilon, Y^\varepsilon)\}$ and $\{\log S(v^\varepsilon, Y^\varepsilon)\}$ and hence the existence of a LDP for $\{S^\varepsilon\}$:

Proposition 3.2. *$\{\log S^\varepsilon(v^\varepsilon, Y^\varepsilon)\}$ and $\{\log S(v^\varepsilon, Y^\varepsilon)\}$ are exponentially equivalent. Therefore, $\{S^\varepsilon\}$ satisfies a LDP with the good rate function*

$$I^S(\varphi) = \inf\{I^{v,Y}(\phi,\psi) : \varphi = S(\phi,\psi)\}, \tag{3.9}$$

where $I^{v,Y}$ is from (3.8) *and $S(\phi,\psi)$ is from* (3.5).

3.1 Proofs of Heston Model Results

Proof of Lemma 3.1. It is first shown that Assumption 2.1 holds with $E = (0,\infty)$. $f(x) = \sqrt{x}$ clearly satisfies Assumption 2.1. That $\{v^\varepsilon\}$ satisfies a LDP on $C([0,T];\mathbb{R}_+)$ follows by setting $I^v = \infty$ off of $C_{v_0}([0,T];\mathbb{R}_+)$ (Dembo & Zeitouni 1998, Lemma 4.1.5). To show $DI^v \subset \mathbb{H}_T^{v_0}$, let $\phi \in DI^v$. Since

$$\dot\phi_t^2 = \xi^2\phi_t\left(\frac{\dot\phi_t - \kappa(\theta-\phi_t)}{\xi\sqrt{\phi_t}}\right)^2 + 2\kappa\left(\theta-\phi_t\right)\dot\phi_t - \kappa^2\left(\theta-\phi_t\right)^2,$$

the non-negativity of ϕ implies

$$\int_0^T \dot\phi_t^2 dt \le 2\xi^2\left(\sup_{0\le t\le T}\phi_t\right) I^v(\phi) + \kappa\left(\theta - v\right)^2,$$

and hence $\phi \in \mathbb{H}_T^{v_0}$. As for the "Furthermore" statement: for any $0 < \beta < \frac{2}{\xi^2}$, $\phi \in DI^v$

$$\begin{aligned}\beta\int_0^T \frac{\dot\phi_t^2}{4\phi_t}dt - I^v(\phi) &= \left(\frac{\beta}{4} - \frac{1}{2\xi^2}\right)\int_0^T \frac{\dot\phi_t^2}{\phi_t}dt + \int_0^T\left(\frac{\kappa\left(\theta-\phi_t\right)\dot\phi_t}{\xi^2\phi_t} - \frac{\kappa^2(\theta-\phi_t)^2}{2\xi^2\phi_t}\right)dt \\ &\le \frac{\kappa}{\xi^2}\left(\theta\log\phi_T - \phi_T - \theta\log v_0 + v_0\right).\end{aligned}$$

On $x > 0$

$$\theta \log x - x \leq \theta \log \theta - \theta,$$

from whence it follows that

$$\sup_{\phi \in DI^v} \beta \int_0^T \frac{\dot{\phi}_t^2}{4\phi_t} dt - I^v(\phi) < \infty.$$

□

Proof of Proposition 3.2. It suffices to prove for any $\delta > 0$,

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \log P \left[\sup_{0 \leq t \leq T} \left| \log S^\varepsilon \left(v^\varepsilon, Y^\varepsilon\right)_t - \log S \left(v^\varepsilon, Y^\varepsilon\right)_t \right| > \delta \right] = -\infty. \quad (3.10)$$

The positivity of v^ε implies

$$\sup_{0 \leq t \leq T} \left| \log S^\varepsilon \left(v^\varepsilon, Y^\varepsilon\right)_t - \log S \left(v^\varepsilon, Y^\varepsilon\right)_t \right| = \frac{\varepsilon}{2} \int_0^T v_s^\varepsilon ds.$$

By Markov's inequality, for $M > 0$

$$P \left[\int_0^T v_s^\varepsilon ds > \frac{2\delta}{\varepsilon} \right] \leq \exp \left(-\frac{2M\delta}{\varepsilon^2} \right) E_P \left[\exp \left(\frac{M}{\varepsilon} \int_0^T v_s^\varepsilon ds \right) \right].$$

Thus, (3.10) will hold if there exists an $M > 0$ such that

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \log E_P \left[\exp \left(\frac{M}{\varepsilon} \int_0^T v_s^\varepsilon ds \right) \right] < \infty. \quad (3.11)$$

Jensen's inequality implies

$$E_P \left[\exp \left(\frac{M}{\varepsilon} \int_0^T v_s^\varepsilon ds \right) \right] \leq \frac{1}{T} \int_0^T E_P \left[\exp \left(\frac{MT}{\varepsilon} v_s^\varepsilon \right) \right] ds.$$

The process v^ε (Glasserman 2004, Chapter 3.4) has marginal distributions

$$v_s^\varepsilon \stackrel{d}{=} \frac{\xi^2 \varepsilon \left(1 - e^{-\kappa s}\right)}{4\kappa} X,$$

where X is a non-central chi-square random variable with

$$\frac{4\kappa\theta}{\xi^2\varepsilon}, \qquad \frac{4\kappa e^{-\kappa s} v_0}{\xi^2 \varepsilon \left(1 - e^{-\kappa s}\right)},$$

degrees of freedom and non-centrality parameter respectively. Thus (3.11) will follow by choosing, for any $0 < \alpha < 1/2$,

$$M = \frac{4\kappa\alpha}{\xi^2 T}.$$

□

4 Asymptotically Optimal Importance Sampling

In this section, Proposition 3.2 is used to outline a methodology for selecting an asymptotically optimal change of measure when pricing path dependent options in the Heston stochastic volatility model. For a detailed explanation of the method see Glasserman et al. (1999) or Guasoni & Robertson (2008).

Assume that the asset price S evolves according to (3.1) and G is a non-negative continuous functional on $C([0,T];\mathbb{R}_+)$. As stated in the introduction, the goal is to minimize

$$E_P\left[G^2(S)\frac{dP}{dQ}\right], \tag{4.1}$$

where Q ranges over a class of equivalent measures. Here, the class is defined as follows: with DI^v from (3.7) define the map $u : DI^v \mapsto L^2[0,T]$ via

$$u(f)_t = \frac{\dot{f}_t - \kappa\left(\theta - f_t\right)}{\xi\sqrt{f_t}}. \tag{4.2}$$

From Donati-Martin et al. (2004, Sections 5 and 6.1) it follows that u is a bijection. Define:

$$\begin{aligned} \tilde{\mathbb{H}}_T^{v_0} &= \left\{f \in DI^v : u(f) \in AC[0,T]\right\}, \\ \tilde{\mathbb{H}}_T^{0} &= \left\{g \in \mathbb{H}_T^0 : \dot{g} \in AC[0,T]\right\}. \end{aligned} \tag{4.3}$$

For each $f \in \tilde{\mathbb{H}}_T^{v_0}$, $g \in \tilde{\mathbb{H}}_T^0$ and $\varepsilon > 0$ an equivalent measure $Q^{f,g}(\varepsilon)$ is determined via:

$$\frac{dQ^{f,g}}{dP}(\varepsilon) = \mathcal{E}\left(\frac{1}{\sqrt{\varepsilon}}\left(\int_0^\cdot u(f)_t dB_t + \int_0^\cdot \dot{g}_t dZ_t\right)\right)_T. \tag{4.4}$$

Using Girsanov's Theorem it is seen that under $Q^{f,g}(\varepsilon)$, at $\varepsilon = 0$, v^ε and $\sqrt{\varepsilon}Z$ from (3.2) are equal to f and g respectively.

Set $F = \log G$ and let S^ε be defined as in (3.2). For $Q^{f,g}(\varepsilon)$, the quantity in (4.1) is

$$E_P\left[\exp\left(\frac{2}{\varepsilon}F(S^\varepsilon)\right)\left(\frac{dQ^{f,g}}{dP}(\varepsilon)\right)^{-1}\right], \tag{4.5}$$

at $\varepsilon = 1$. The small-noise approximation to (4.5) is

$$L(f,g) = \limsup_{\varepsilon\downarrow 0}\varepsilon \log E_P\left[\exp\left(\frac{2}{\varepsilon}F(S^\varepsilon)\right)\left(\frac{dQ^{f,g}}{dP}(\varepsilon)\right)^{-1}\right]. \tag{4.6}$$

An asymptotically optimal pair $\{f,g\}$ is thus defined as:

Definition 4.1 (Asymptotically Optimal)**.** $\{f,g\}$ *is* asymptotically optimal *if it is a solution to the problem*

$$\min_{f\in\tilde{\mathbb{H}}_T^{v_0}, g\in\tilde{\mathbb{H}}_T^0} L(f,g).$$

In order to compute $L(f,g)$, it is necessary to use Varadhan's Integral Lemma. The restrictions $f \in \tilde{\mathbb{H}}_T^{v_0}$ and $g \in \tilde{\mathbb{H}}_T^0$ ensure, using the stochastic integration by parts formula, that

$$\varepsilon \log \frac{dQ^{f,g}}{dP}(\varepsilon)$$

can be defined path-wise as a continuous functional of $\{B^\varepsilon, Z^\varepsilon\}$ where $B^\varepsilon = \sqrt{\varepsilon}B$ and $Z^\varepsilon = \sqrt{\varepsilon}Z$. However, using the notation of Section 3 it follows that

$$2F(S^\varepsilon) = 2F\left(S^\varepsilon\left(v^\varepsilon, Y^\varepsilon\right)\right)$$

is an epsilon dependent continuous functional of the pair $\{v^\varepsilon, Y^\varepsilon\}$ which may take the value $-\infty$. Furthermore, $\{v^\varepsilon\}$ is not a continuous function of $\{B^\varepsilon\}$. In spite of this, the following Proposition shows that Varadhan's Integral Lemma is still applicable:

Proposition 4.2. *If there exists a $\gamma > 1$ such that*

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \log E_P\left[\exp\left(\frac{2\gamma}{\varepsilon}F(S^\varepsilon)\right)\right] < \infty, \tag{4.7}$$

then for all $f \in \tilde{\mathbb{H}}_T^{v_0}, g \in \tilde{\mathbb{H}}_T^0$

$$\begin{aligned} L(f,g) = \sup_{\phi \in \mathbb{H}_T^{v_0}, \varphi \in \mathbb{H}_T^0} & 2F\left(S\left(\phi, \int \sqrt{\phi}\dot{\varphi}\right)\right) + \frac{1}{2}\int_0^T \left(u(f)_t - u(\phi)_t\right)^2 dt \\ & + \frac{1}{2}\int_0^T \left(\dot{g}_t - \dot{\varphi}_t\right)^2 dt - 2\left(I^v(\phi) + I^Z(\varphi)\right). \end{aligned} \tag{4.8}$$

In the variational problem on the right hand side of (4.8) the functions f and g only appear in the middle two terms and these terms are non-negative for all f, g, ϕ and φ. Since the goal is to minimize $L(f,g)$, Proposition 4.2 suggests the following method for finding an asymptotically optimal pair $\{f,g\}$. First, solve the variational problem

$$\sup_{\phi \in \mathbb{H}_T^{v_0}, \varphi \in \mathbb{H}_T^0} F\left(S\left(\phi, \int \sqrt{\phi}\dot{\varphi}\right)\right) - I^v(\phi) - I^Z(\varphi). \tag{4.9}$$

Next, if an optimal $\hat{\phi}$, $\hat{\varphi}$ exist and are in $\tilde{\mathbb{H}}_T^{v_0}$, and $\tilde{\mathbb{H}}_T^0$ respectively set $f = \hat{\phi}, g = \hat{\varphi}$. Lastly, check if

$$L(f,g) \leq 2F\left(S\left(f, \int \sqrt{f}\dot{g}\right)\right) - 2\left(I^v(f) + I^Z(g)\right). \tag{4.10}$$

If this is the case, then the pair $\{f,g\}$ is asymptotically optimal. When successful, this method has the advantage in that it replaces solving a min-max variational problem with solving two maximization problems. However, there is no guarantee the solutions to the min-max and max-min problems will coincide: for example, this will never be the case if the solution to (4.9) is not unique. Furthermore, for many functionals of interest (e.g. Arithmetic Asian options), $F(S(\phi,\varphi))$ is not concave and hence the general theory on when the order of the min and max can be switched does not apply.

Assuming an asymptotically optimal pair $\{f,g\}$ is found, Monte Carlo simulation is run under the measure $Q^{f,g}=Q^{f,g}(1)$ where the price and volatility dynamics are

$$\begin{aligned}\frac{dS_t}{S_t} &= (r+\sqrt{v_t}\,(\rho u(f)_t+\bar{\rho}\dot{g}_t))\,dt+\rho\sqrt{v_t}d\hat{B}_t+\bar{\rho}\sqrt{v_t}d\hat{Z}_t,\\ dv_t &= (\kappa\theta-\kappa v_t+\xi u(f)_t\sqrt{v_t})\,dt+\xi\sqrt{v_t}d\hat{B}_t,\end{aligned} \tag{4.11}$$

where $(\hat{B},\hat{W})$ is a standard $Q^{f,g}$ Brownian Motion.

If condition (4.10) holds for a particular choice of $f\in\tilde{\mathbb{H}}_T^{v_0}$, $g\in\tilde{\mathbb{H}}_T^{0}$ then $Q^{f,g}(\varepsilon)$ is asymptotically optimal when compared against all families of equivalent measures $Q(\varepsilon)$ and not just those parameterized via (4.4). Indeed, using Jensen's inequality and the fact $\{S^\varepsilon\}$ satisfies a LDP

$$\begin{aligned}\liminf_{\varepsilon\downarrow 0}\varepsilon\log & E_{Q(\varepsilon)}\left[\left(G(S^\varepsilon)\frac{dP}{dQ}(\varepsilon)\right)^{2/\varepsilon}\right]\\ &\geq 2\liminf_{\varepsilon\downarrow 0}\varepsilon\log E_P\left[G\left(S^\varepsilon\right)^{1/\varepsilon}\right]\\ &\geq \sup_{\phi\in\mathbb{H}_T^{v_0},\varphi\in\mathbb{H}_T^0} 2F\left(S\left(\psi,\int\sqrt{\phi}\dot{\varphi}\right)\right)-2\left(I^v(\phi)+I^Z(\varphi)\right)\\ &= L(f,g).\end{aligned}$$

The primary goal when solving the variational problem in (4.9) is to obtain the $\hat{\phi}$ and $\hat{\varphi}$. Thus, it is of interest to know when maximizers exist. In the case when F is bounded from above the following proposition shows that the variational problems in (4.8) and (4.9) do admit solutions.

Proposition 4.3. *Assume that F is bounded from above and let $f\in\tilde{\mathbb{H}}_T^{v_0}$, $g\in\tilde{\mathbb{H}}_T^{0}$. Then, the variational problems in* (4.8) *and* (4.9) *each admit maximizers.*

4.1 The Arithmetic Average Asian Put Option

Consider the arithmetic average Asian put option:

$$G(S)=\left(K-\frac{1}{T}\int_0^T S_t dt\right)^+. \tag{4.12}$$

Since G is bounded from above, Propositions 4.2 and 4.3 hold. To solve (4.9), it is convenient to make the transformation $\dot{\varphi}=\sqrt{\phi}h$ for $h\in L^2[0,T]$. This leaves the variational problem

$$\sup_{\phi\in\mathbb{H}_T^{v_0},h\in L^2[0,T]} F\left(S\left(\phi,\int\phi h\right)\right)-\left(I^v(\phi)+\frac{1}{2}\int_0^T\phi_t h_t^2 dt\right). \tag{4.13}$$

This transformation is allowed because for $\phi\in DI^v$

$$\inf_{0\le t\le T}\phi_t \ge v_0\exp\left(-\frac{v_0}{\theta}-\frac{\xi^2}{\kappa\theta}I^v(\phi)\right)>0. \tag{4.14}$$

For functionals of the form

$$F(\phi)=\int_0^T f(\phi_t,\dot{\phi}_t)dt,$$

where $f(x,y)$ is a known smooth function and ϕ is sufficiently regular, define $DF(\phi)$ via

$$DF(\phi)_t=\partial_x f(\phi_t,\dot{\phi}_t)-\frac{d}{dt}\partial_y f(\phi_t,\dot{\phi}_t).$$

With this notation, the Euler Lagrange equation for (4.13) is

$$\begin{aligned} h_t+\bar{\rho}\frac{\frac{1}{T}\int_t^T S_s ds}{K-\frac{1}{T}\int_0^T S_s ds}&=0,\\ DI^v(\phi)_t&=\frac{1}{2}h_t^2+\frac{\rho\kappa}{\xi\bar{\rho}}h_t-\frac{\rho}{\xi\bar{\rho}}\dot{h}_t, \end{aligned} \tag{4.15}$$

where $S_s=S(\phi,\int\phi h)_s$.

4.1.1 Numerical Example

For the arithmetic average Asian put option, the following parameter values are considered [4].

$$\begin{aligned} &\kappa=2,\theta=0.09,\xi=0.2,v_0=0.04,\\ &r=0.05,T=1,S_0=50,K=30,\rho=-0.5. \end{aligned}$$

For these parameter values, $\hat{\phi},\hat{\varphi}$ from (4.9) satisfy the inequality in (4.10) and hence are asymptotically optimal. Figure 1 shows the deterministic price and volatility path under the optimal change of drift. For comparison, also included is the deterministic volatility path under the original measure P.

The interpretation behind Figure 1 is that since under the original measure P the Put Option is out of the money, in order to bring the option more into the money either the "average" price path must come down or the volatility must go up. Under the asymptotically optimal measure it is seen that both of these shifts are taking place.

[4]The model parameter values are taken to approximately correspond with those in Heston (1993).

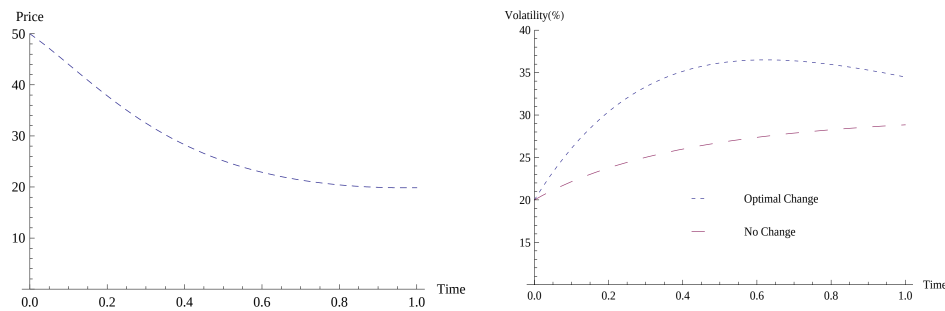

Figure 1: Deterministic Price and Volatility Paths under the Asymptotically Optimal change of drift. Also included in the volatility plot is the deterministic volatility path under the original measure. The asset price is in dollars and time horizon in years. Parameter values are $\kappa = 2, \theta = 0.09, \xi = 0.2, v_0 = 0.04, r = 0.05, T = 1, S_0 = 50, K = 30, \rho = -0.5$.

4.2 Proof of Importance Sampling Results

The proof of Proposition 4.2 requires the following extension of Varadhan's Integral Lemma which will be proved at the end of this section.

Lemma 4.4. *Let $\mathcal{X}$ and $\mathcal{Y}$ be two metric spaces. Suppose $\{\omega^\varepsilon\}$ satisfies a LDP with good rate function $I : \mathcal{X} \mapsto [0, \infty]$. Let $\Lambda : \mathcal{X} \mapsto \mathcal{Y}$ be a continuous map. For all $\varepsilon > 0$, let $\Lambda^\varepsilon : \mathcal{X} \mapsto \mathcal{Y}$ be measurable such that $\{\Lambda^\varepsilon(\omega^\varepsilon)\}$ is exponentially equivalent to $\{\Lambda(\omega^\varepsilon)\}$. Let $\Phi : \mathcal{Y} \mapsto [-\infty, \infty)$ and $\Psi : \mathcal{X} \mapsto \mathbb{R}$ be continuous. If there exists a $\gamma > 1$ such that*

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \log E_P \left[\exp \left(\frac{\gamma}{\varepsilon} \left(\Phi \left(\Lambda^\varepsilon(\omega^\varepsilon) \right) + \Psi(\omega^\varepsilon) \right) \right) \right] < \infty,$$

then

$$\lim_{\varepsilon \downarrow 0} \varepsilon \log E_P \left[\exp \left(\frac{1}{\varepsilon} \left(\Phi \left(\Lambda^\varepsilon(\omega^\varepsilon) \right) + \Psi(\omega^\varepsilon) \right) \right) \right] = \sup_{x \in \mathcal{X}} \left(\Phi(\Lambda(x)) + \Psi(x) - I(x) \right).$$

Proof of Proposition 4.2. Let $\omega^\varepsilon = \{B^\varepsilon, v^\varepsilon, Z^\varepsilon, Y^\varepsilon\}$. Since the Euclidean distance in $\mathbb{R}^d$ compares the coordinates separately, by combining the argument in Donati-Martin et al. (2004, Sections 5 and 6.1) with the argument used to prove Theorem 2.2 it follows that $\{\omega^\varepsilon\}$ satisfies a LDP with good rate function

$$I^\omega(\eta, \phi, \varphi, \psi) = \begin{cases} I^v(\phi) + I^Z(\varphi) & (\eta, \phi, \varphi, \psi) \in DI^\omega \\ \infty & else \end{cases},$$

where

$$DI^\omega = \left\{ (\eta, \phi, \varphi, \psi) : \eta, \varphi \in \mathbb{H}_T^0, \phi \in DI^v, u(\phi) = \dot{\eta}, \psi = \int \sqrt{\phi} \dot{\varphi} \right\}.$$

Let γ be as in (4.7) and let $1 < \delta < \gamma$. For any $f \in \tilde{\mathbb{H}}_T^{v_0}$, $g \in \tilde{\mathbb{H}}_T^0$, Hölder's inequality with $p = \gamma/\delta$, $q = \gamma/(\gamma - \delta)$ yields

$$\limsup_{\varepsilon \downarrow 0} \varepsilon \log E_P \left[\exp\left(\frac{\delta}{\varepsilon} \left(2F(S^\varepsilon) - \varepsilon \log\left(\frac{dQ^{f,g}}{dP}(\varepsilon) \right) \right) \right) \right] < \infty.$$

Thus, using Proposition 3.2 and the stochastic integration by parts formula, Lemma 4.4 applies with

$$\mathcal{X} = C\left([0,T]\,; \mathbb{R} \times \mathbb{R}_+ \times \mathbb{R}^2\right), \qquad \mathcal{Y} = C\left([0,T]\,;\mathbb{R}\right), \tag{4.16}$$

and for $x = (\eta, \phi, \varphi, \psi) \in \mathcal{X}$, $y \in \mathcal{Y}$:

$$\Lambda(x) = \log S(\phi, \psi), \qquad \Lambda^\varepsilon(x) = \log S^\varepsilon(\phi, \psi), \qquad \Phi(y) = 2F(e^y),$$
$$\Psi(x) = -u(f)_T \eta_T - \dot{g}_T \varphi_T + \int_0^T \left(\dot{u}(f)_t \eta_t + \frac{1}{2} u(f)_t^2 + \ddot{g}_t \varphi_t + \frac{1}{2} \dot{g}_t^2 \right) dt. \tag{4.17}$$

Therefore,

$$L(f,g) = \sup_{x \in \mathcal{X}} \left(\Phi(\Lambda(x)) + \Psi(x) - I^\omega(x) \right).$$

The equivalence of this with the right hand side of (4.8) follows via the identifications in (4.17); the fact that u from (4.2) is a bijection between DI^v and $L^2[0,T]$; using the regular integration by parts formula on Ψ for $\eta, \varphi \in \mathbb{H}_T^0$; and noting that for $\phi \in DI^v$

$$I^v(\phi) = \frac{1}{2} \int_0^T u(\phi)_t^2 dt. \tag{4.18}$$

□

Proof of Proposition 4.3. Let

$$v : L^2[0,T] \mapsto DI^v \subset C_{v_0}\left([0,T]\,;\mathbb{R}_+\right)$$

be the inverse of u. It will be shown that for any $f \in \mathbb{H}_T^{v_0}$, $g \in \mathbb{H}_T^0$ and $M > 0$ the variational problem

$$\sup_{x,y \in L^2[0,T]} 2F\left(S\left(v(x), \int \sqrt{v(x)} y \right) \right) - M \int_0^T \left(u(f)_t + x_t\right)^2 dt$$
$$- M \int_0^T \left(\dot{g}_t + y_t\right)^2 dt + \int_0^T \left(u(f)_t^2 + \dot{g}_t^2 \right) dt, \tag{4.19}$$

admits a maximizer. Since u is a bijection and (4.18) holds, setting $M = 1/2$ will yield a maximizer to (4.8) and setting $M = 1$, $g = 0$ and

$$f_t = \theta + e^{-\kappa t}\left(v_0 - \theta\right),$$

will yield a maximizer to (4.9). From Donati-Martin et al. (2004, Sections 5 and 6.1) it follows that v is weakly continuous and there exist constants $C, D > 0$ such that

$$\sup_{0\leq t\leq T} v(x)_t \leq C + D\int_0^T x_t^2 dt.$$

This, combined with the continuity of F and S imply

$$F\left(S\left(v(x), \int \sqrt{v(x)}y\right)\right)$$

is weakly continuous in (x, y). Since

$$M\int_0^T (u(f)_t + x_t)^2\, dt + M\int_0^T (\dot{g}_t + y_t)^2\, dt$$

is weakly lower semi-continuous in (x, y) and tends to ∞ as $\|x\|, \|y\| \uparrow \infty$ the existence of a maximizer follows for bounded F (Reed & Simon 1972, Theorem S.6).

□

Proof of Lemma 4.4. The exponential equivalence of $\{\Lambda^\varepsilon(\omega^\varepsilon)\}$ and $\{\Lambda(\omega^\varepsilon)\}$ implies the following two facts:

i) For any $r > 0, \delta > 0$ and $x \in \mathcal{X}$,

$$\begin{aligned}\liminf_{\varepsilon\downarrow 0} \varepsilon \log & P\left(\omega^\varepsilon \in B(x,r), \Lambda^\varepsilon(\omega^\varepsilon) \in B\left(\Lambda(\omega^\varepsilon), \delta\right)\right) \\ &= \liminf_{\varepsilon\downarrow 0} \varepsilon \log P\left(\omega^\varepsilon \in B(x,r)\right).\end{aligned}$$

ii) For bounded measurable functions $G : \mathcal{X} \times \mathcal{Y} \mapsto \mathbb{R}$, open $A \subset \mathcal{X}$ and $\delta > 0$,

$$\begin{aligned}\limsup_{\varepsilon\downarrow 0} \varepsilon \log E_P & \left[\exp\left(\frac{1}{\varepsilon}G(\omega^\varepsilon, \Lambda^\varepsilon(\omega^\varepsilon))\right) 1_{\{\omega^\varepsilon\in A\}}\right] \\ &= \limsup_{\varepsilon\downarrow 0} \varepsilon \log E_P \left[\exp\left(\frac{1}{\varepsilon}G(\omega^\varepsilon, \Lambda^\varepsilon(\omega^\varepsilon))\right) 1_{\{\omega^\varepsilon\in A\}} 1_{\{\Lambda^\varepsilon(\omega^\varepsilon)\in B(\Lambda(\omega^\varepsilon),\delta)\}}\right].\end{aligned}$$

Therefore, the proof for real valued Φ follows by mimicking the proofs in Dembo & Zeitouni (1998, Lemmas 4.3.4, 4.3.6, 4.3.8). The extension to when Φ may take the value $-\infty$ follows from Guasoni & Robertson (2008, Lemma A.5).

□

References

Avellaneda, M., Boyer-Olson, D., Busca, J. & Friz, P. (2002), 'Reconstruction of volatility, pricing index options using the steepest-descent approximation', *Risk* **15**(10), 87–91.

Avellaneda, M., Boyer-Olson, D., Busca, J. & Friz, P. (2003), 'Application of large deviation methods to the pricing of index options in finance', *C. R. Math. Acad. Sci. Paris* **336**(3), 263–266.

Berestycki, H., Busca, J. & Florent, I. (2004), 'Computing the implied volatility in stochastic volatility models', *Comm. Pure Appl. Math.* **57**(10), 1352–1373.

Chernoff, H. (1952), 'A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations', *Ann. Math. Statistics* **23**, 493–507.

Cramér, H. (1938), 'Sur un nouveau théorème-limite de la théorie des probabilités', *Actualités Scientifiques et Industrielles* (736), 5–23.

Dembo, A. & Zeitouni, O. (1998), *Large deviations techniques and applications*, Vol. 38 of *Applications of Mathematics (New York)*, second edn, Springer-Verlag, New York.

Deuschel, J.-D. & Stroock, D. W. (1989), *Large deviations*, Vol. 137 of *Pure and Applied Mathematics*, Academic Press Inc., Boston, MA.

Donati-Martin, C., Rouault, A., Yor, M. & Zani, M. (2004), 'Large deviations for squares of Bessel and Ornstein-Uhlenbeck processes', *Probab. Theory Related Fields* **129**(2), 261–289.

Fournié, E., Lasry, J. M. & Touzi, N. (1997), Monte Carlo methods for stochastic volatility models, *in* 'Numerical methods in finance', Publ. Newton Inst., Cambridge Univ. Press, Cambridge, pp. 146–164.

Freidlin, M. I. & Wentzell, A. D. (1984), *Random perturbations of dynamical systems*, Vol. 260 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*, Springer-Verlag, New York. Translated from the Russian by Joseph Szücs.

Glasserman, P. (2004), *Monte Carlo methods in financial engineering*, Vol. 53 of *Applications of Mathematics (New York)*, Springer-Verlag, New York. Stochastic Modelling and Applied Probability.

Glasserman, P., Heidelberger, P. & Shahabuddin, P. (1999), 'Asymptotically optimal importance sampling and stratification for pricing path-dependent options', *Math. Finance* **9**(2), 117–152.

Guasoni, P. & Robertson, S. (2008), 'Optimal importance sampling with explicit formulas in continuous time', *Finance Stoch.* **12**(1), 1–19.

Heston, S. L. (1993), 'A closed-form solution for options with stochastic volatility with applications to bond and currency options', *The Review of Financial Studies* **6**(2), 327–343.

Hull, J. & White, A. (1987), 'The pricing of options on assets with stochastic volatilities', *The Journal of Finance* **42**(2), 281–300.

Karatzas, I. & Shreve, S. E. (1991), *Brownian motion and stochastic calculus*, Vol. 113 of *Graduate Texts in Mathematics*, second edn, Springer-Verlag, New York.

Pham, H. (2007), Some applications and methods of large deviations in finance and insurance, *in* 'Paris-Princeton Lectures on Mathematical Finance 2004', Vol. 1919 of *Lecture Notes in Math.*, Springer, Berlin, pp. 191–244.

Reed, M. & Simon, B. (1972), *Methods of modern mathematical physics. I. Functional analysis*, Academic Press, New York.

Schilder, M. (1966), 'Some asymptotic formulas for Wiener integrals', *Trans. Amer. Math. Soc.* **125**, 63–85.

Varadhan, S. R. S. (1967), 'Diffusion processes in a small time interval', *Comm. Pure Appl. Math.* **20**, 659–685.