Long-Time Trajectorial Large Deviations and Importance Sampling for Affine Stochastic Volatility Models

Zorana Grbac, David Krief, Peter Tankov

Long-time trajectorial large deviations and importance sampling for affine stochastic volatility models

Zorana Grbac[1] David Krief[2] Peter Tankov[3]

Abstract: *We establish a pathwise large deviations principle for affine stochastic volatility models introduced in (Keller-Ressel, 2011), and present an application to variance reduction for Monte Carlo computation of prices of path-dependent options in these models, extending the method developed in (Genin and Tankov, 2020) for exponential Lévy models. To this end, we apply an exponentially affine change of measure and use Varadhan's lemma, in the fashion of (Guasoni and Robertson, 2008) and (Robertson, 2010), to approximate the problem of finding the measure that minimises the variance of the Monte-Carlo estimator. We test the method on the Heston model with and without jumps to demonstrate its numerical efficiency.*

Key words: large deviations, Monte Carlo methods, importance sampling, affine stochastic volatility

MSC2010: 91G60, 60F10

1. Introduction

The aim of this paper is to develop efficient importance sampling estimators for prices of path-dependent options in affine stochastic volatility (ASV) models of asset prices. To this end, we establish pathwise large deviation results for these models, which are of independent interest.

An ASV model, studied in (Keller-Ressel, 2011), is a two-dimensional affine process (X, V) on $\mathbb{R} \times \mathbb{R}_+$ with special properties, where X models the logarithm of the stock price and V its instantaneous variance. This class includes many well studied and widely used models such as Heston stochastic volatility model (Heston, 1993), the model of (Bates, 1996), the stochastic volatility model of (Barndorff-Nielsen and Shephard, 2001) and time-changed Lévy models with independent affine time change. European options in affine stochastic volatility models may be priced by Fourier transform, but for path-dependent options explicit formulas are in general not available and Monte Carlo is often the method of choice. At the same time, Monte Carlo simulation of such processes is difficult and time-consuming: the convergence rates of discretization schemes are often low due to the irregular nature of coefficients of the corresponding stochastic differential equations. To accelerate Monte Carlo simulation, it is thus important to develop efficient variance-reduction algorithms for these models.

In this paper, we therefore develop an importance sampling algorithm for ASV models. Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a given probability space and denote by $\mathbb{E}[\cdot]$ the expectation with respect to $\mathbb{P}$. The importance sampling method is based on the following

[1]Université Paris Diderot, Paris, France, email:grbac@math.univ-paris-diderot.fr
[2]Université Paris Diderot, Paris, France, email: krief@math.univ-paris-diderot.fr
[3]ENSAE ParisTech, Palaiseau, France, email: peter.tankov@ensae.fr

identity, valid for any probability measure $\mathbb{Q}$, with respect to which $\mathbb{P}$ is absolutely continuous. Let P be a deterministic function of a random trajectory S, then

$$\mathbb{E}[P(S)] = \mathbb{E}^{\mathbb{Q}}\left[\frac{d\mathbb{P}}{d\mathbb{Q}}P(S)\right].$$

This allows one to define the importance sampling estimator

$$\widehat{P}_N^{\mathbb{Q}} := \frac{1}{N}\sum_{j=1}^{N}\left[\frac{d\mathbb{P}}{d\mathbb{Q}}\right]^{(j)} P(S_{\mathbb{Q}}^{(j)}),$$

where $S_{\mathbb{Q}}^{(j)}$ are i.i.d. sample trajectories of S under the measure $\mathbb{Q}$. For efficient variance reduction, one needs then to find a probability measure $\mathbb{Q}$ such that S is easy to simulate under $\mathbb{Q}$ and the variance

$$\mathbb{V}\mathrm{ar}_{\mathbb{Q}}\left[P(S)\frac{d\mathbb{P}}{d\mathbb{Q}}\right]$$

is considerably smaller than the original variance $\mathbb{V}\mathrm{ar}_{\mathbb{P}}[P(S)]$.

In this paper, following the work of (Genin and Tankov, 2020) in the context of Lévy processes, we define the probability $\mathbb{Q}$ using the exponentially affine measure change,

$$\frac{d\mathbb{P}_\Theta}{d\mathbb{P}} = \frac{e^{\int_{[0,T]} X_t\cdot\Theta(dt)}}{\mathbb{E}\left[e^{\int_{[0,T]} X_t\cdot\Theta(dt)}\right]},$$

where X is the first component of the ASV model (the logarithm of stock price) and Θ is a (deterministic) bounded signed measure on $[0,T]$. Such a choice is justified by several considerations. First, under the new measure the characteristic function of X remains analytically tractable; moreover, if Θ is supported by a finite number of points, X is piecewise affine under $\mathbb{P}_\Theta$ and is relatively easy to simulate. Second, for a class of payoffs possessing the concavity property, such a choice leads to asymptotically optimal variance reduction (see Theorem 5.4). For other payoffs it may be necessary to use more strongly path-dependent measure changes (e.g., with stochastic Θ), which can be obtained by approximating the solution of the HJB equation resulting from the minimization of the variance, see for example (Dupuis and Wang, 2004; Dupuis and Wang, 2007). However, such schemes may lead to higher computational complexity.

The optimal choice of Θ should minimize the variance of the estimator under $\mathbb{P}_\Theta$,

$$\mathbb{V}\mathrm{ar}_{\mathbb{P}_\Theta}\left(P(S)\frac{d\mathbb{P}}{d\mathbb{P}_\Theta}\right) = \mathbb{E}\left[P^2(S)\frac{d\mathbb{P}}{d\mathbb{P}_\Theta}\right] - \mathbb{E}\left[P(S)\right]^2.$$

The computation of this variance is in general as difficult as the computation of the option price itself. Following (Glasserman et al., 1999; Guasoni and Robertson, 2008; Robertson, 2010) and more recently (Genin and Tankov, 2020), we propose to compute the variance reduction measure Θ^* by minimizing the *proxy* for the variance computed using the theory of large deviations.

To this end, we establish a pathwise large deviation principle (LDP) for affine stochastic volatility models. A one dimensional LDP for X_t/t as $t\to\infty$ where X is the first component of an ASV model has been proven in (Jacquier et al., 2013). In this paper, we first establish an equivalent result for multiple dates and then we use the Dawson-Gärtner theorem to extend it to the trajectorial setting,

in the spirit of the pathwise LDP principles of (Léonard, 2000), but in the weaker topology of pointwise convergence.

The rest of the paper is structured as follows. In Section 2, we describe the model and recall certain useful properties of ASV processes. In Section 3, we recall some general results of large deviations theory. In Section 4, we prove a LDP for the trajectories of ASV processes. In Section 5, we develop the variance reduction method, using an asymptotically optimal change of measure obtained via the LDP shown in Section 4. In Section 6, we test the method numerically on several examples of options, some of which are path-dependent, in the Heston model with and without jumps.

Model Description

2. Model description

In this paper, we model the price of the underlying asset $(S_t)_{t\geq 0}$ of an option as $S_t = S_0\, e^{X_t}$, where we model $(X_t)_{t\geq 0}$ as an affine stochastic volatility process. We recall, from (Keller-Ressel, 2011) and (Duffie et al., 2003), the definition and some properties of ASV models.

Definition 2.1. An ASV model $(X_t, V_t)_{t\geq 0}$, is a stochastically continuous, time-homogeneous Markov process such that $\left(e^{X_t}\right)_{t\geq 0}$ is a martingale and

$$\mathbb{E}\left(e^{uX_t+wV_t}\middle| X_0 = x, V_0 = v\right) = e^{\phi(t,u,w)+\psi(t,u,w)\,v+u\,x}\,, \tag{2.1}$$

eq:LaplaceTransfo

for all $(t,u,w) \in \mathbb{R}_+ \times \mathbb{C}^2$.

eneralizedRiccati

Proposition 2.2. *The functions ϕ and ψ satisfy generalized Riccati equations*

$$\partial_t\phi(t,u,w) = F(u,\psi(t,u,w))\,, \qquad \phi(0,u,w) = 0, \tag{2.2a}$$

eq:GeneralizedRic

$$\partial_t\psi(t,u,w) = R(u,\psi(t,u,w))\,, \qquad \psi(0,u,w) = w, \tag{2.2b}$$

eq:GeneralizedRic

where F and R have the Lévy-Khintchine forms

$$\begin{aligned} F(u,w) &= \begin{pmatrix} u & w \end{pmatrix} \cdot \frac{a}{2} \cdot \begin{pmatrix} u \\ w \end{pmatrix} + b \cdot \begin{pmatrix} u \\ w \end{pmatrix} \\ &\quad + \int_{D\setminus\{0\}} \left(e^{xu+yw} - 1 - w_F(x,y) \cdot \begin{pmatrix} u \\ w \end{pmatrix} \right) m(dx,dy)\,, \\ R(u,w) &= \begin{pmatrix} u & w \end{pmatrix} \cdot \frac{\alpha}{2} \cdot \begin{pmatrix} u \\ w \end{pmatrix} + \beta \cdot \begin{pmatrix} u \\ w \end{pmatrix} \\ &\quad + \int_{D\setminus\{0\}} \left(e^{xu+yw} - 1 - w_R(x,y) \cdot \begin{pmatrix} u \\ w \end{pmatrix} \right) \mu(dx,dy)\,, \end{aligned}$$

where $D = \mathbb{R}\times\mathbb{R}_+$,

$$w_F(x,y) = \begin{pmatrix} \frac{x}{1+x^2} \\ 0 \end{pmatrix} \qquad \text{and} \qquad w_R(x,y) = \begin{pmatrix} \frac{x}{1+x^2} \\ \frac{y}{1+y^2} \end{pmatrix}$$

and (a,α,b,β,m,μ) satisfy the following conditions

- *a,α are positive semi-definite 2×2-matrices where $a_{12} = a_{21} = a_{22} = 0$.*
- *$b \in D$ and $\beta \in \mathbb{R}^2$.*
- *m and μ are Lévy measures on D and $\int_{D\setminus\{0\}} ((x^2+y)\wedge 1)\, m(dx,dy) < \infty$.*

Define the function

$$\chi(u) = \partial_w R(u,w)|_{w=0} = \alpha_{12}u + \beta_2 + \int_{D\setminus\{0\}} y\left(e^{xu} - \frac{1}{1+y^2}\right)\mu(dx,dy)\,.$$

In the rest of the paper, we shall impose the following standing assumption to ensure nondegeneracy of the model (dependence of the law of $(X_t)_{t\geq 0}$ on V_0), the martingale property of $(S_t)_{t\geq 0}$, where $S_t = S_0\, e^{X_t}$, and the existence of long-time limit of the functions ϕ and ψ, see (Keller-Ressel, 2011, Corollary 2.7 and Theorem 3.4).

ass1 **Assumption 1** (Nondegeneracy and martingale property)**.** The functions F, R and χ are such that there exists $u \in \mathbb{R}$ such that $R(u,0) \neq 0$, $F(1,0) = R(1,0) = 0$, $\chi(0) < 0$ and $\chi(1) < 0$.

In the following theorem, we compile several results of (Keller-Ressel, 2011), which describe the behaviour of the solutions to equations given in (2.2) as $t \to \infty$. The functions $w(u)$ and $\tilde{w}(u)$ denoting respectively the unique stable equilibrium and the unstable equilibrium (if there is one) of (2.2b), which are defined below in (1) and (2), will be used throughout the paper, as well as the function $h(u)$ defined in (2.5).

rResselProperties **Theorem 2.3.** *The following statements hold true.*

(1) *There exists a maximal interval $I \supseteq [0,1]$, and a unique function $w \in C(I) \cap C^1(\mathring{I})$ such that*

$$R(u,w(u)) = 0 \quad \text{for all} \quad u \in I,$$

$w(0) = w(1) = 0$, $w(u) < 0$ for all $u \in (0,1)$, $w(u) > 0$ for all $u \in I \setminus [0,1]$ and

$$\partial_w R(u,w(u)) < 0 \quad \text{for all } u \in \mathring{I}.$$

In other words, the function $w(u)$ is the unique stable equilibrium of (2.2b).

(2) *For $u \in I$, eq. (2.2b) admits at most one other equilibrium $\tilde{w}(u)$, which is unstable.*

(3) *For $u \in \mathbb{R}\setminus I$, eq. (2.2b) does not have any equilibrium.*

We denote $\mathcal{B}(u)$ the basin of attraction of the stable solution $w(u)$ of eq. (2.2b) and $J = \{u \in I : F(u,w(u)) < \infty\}$, the domain of $u \mapsto F(u,w(u))$. We have that

(4) *J is an interval such that $[0,1] \subseteq J \subseteq I$.*

(5) *For $u \in I$, $w \in \mathcal{B}(u)$ and $\Delta t > 0$, we have*

$$\psi\left(\frac{\Delta t}{\epsilon}, u, w\right) \underset{\epsilon\to 0}{\longrightarrow} w(u)\,. \tag{2.3}$$ eq:ConvergencePsi

(6) *For $u \in J$, $w \in \mathcal{B}(u)$ and $\Delta t > 0$,*

$$\epsilon\,\phi\left(\frac{\Delta t}{\epsilon}, u, w\right) \underset{\epsilon\to 0}{\longrightarrow} \Delta t\, h\,(u)\ , \tag{2.4}$$ eq:ConvergencePhi

where

$$h(u) := F(u,w(u)) = \lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[e^{uX_{1/\epsilon}}\right]. \tag{2.5}$$ def-h

(7) *For every $u \in I$, $0 \in \mathcal{B}(u)$.*

Definition 2.4. A convex function $f : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ with effective domain D_f is *essentially smooth* if

i. D_f° is non-empty;
ii. f is differentiable in D_f°;
iii. f is *steep*, that is, for any sequence $(u_n)_{n\in\mathbb{N}} \subset D_f^\circ$ that converges to a point in the boundary of D_f,

$$\lim_{n\to\infty} ||\nabla f(u_n)|| = \infty\,.$$

In the rest of the paper, to establish large deviations results, we shall make the following assumptions on the model.

ass:H **Assumption 2.** The function h, defined in (2.5), satisfies the following properties:
ass:Ha (1) There exists $u < 0$, such that $h(u) < \infty$.
ass:Hb (2) $u \mapsto h(u)$ is essentially smooth.

In (Jacquier et al., 2013), a set of sufficient conditions is provided for Assumption 2 to be verified:

Proposition 2.5 (Corollary 8 in (Jacquier et al., 2013))**.** *Let (X, V) be an ASV model satisfying Assumption 1. If either of the following conditions holds*

(i) The Lévy measure μ of R has exponential moments of all orders, F is steep and $(0,0), (1,0) \in D_F^\circ$.
(ii) (X, V) is a diffusion,

then function h is well defined, for every $u \in \mathbb{R}$ with effective domain J. Moreover h is essentially smooth and $\{0, 1\} \subset J^\circ$.

We now discuss the form of the basin of attraction of the unique stable solution of (2.2b).

lem:RConvex **Lemma 2.6.** *(Keller-Ressel, 2011, Lemma 2.2.)*
(a) F and R are proper closed convex functions on $\mathbb{R}^2$.
(b) F and R are analytic in the interior of their effective domain.
(c) Let U be a one-dimensional affine subspace of $\mathbb{R}^2$. Then $F|U$ is either a strictly convex or an affine function. The same holds for $R|U$.
(d) If $(u, w) \in D_F$, then also $(u, \eta) \in D_F$ for all $\eta \le w$. The same holds for R.

BasinOfAttraction **Lemma 2.7.** *Let $f : \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ be a convex function with either two zeros $w < \tilde{w}$, or a single zero w. In the latter case, we let $\tilde{w} = \infty$. Assume that there exists $y \in (w, \tilde{w})$ such that $f(y) < 0$. Then for every $x \in D_f$,*

$$\begin{cases} f(x) > 0\,, & \text{if } x < w \quad \text{or} \quad \tilde{w} < x\,, \\ f(x) < 0\,, & \text{if } x \in (w, \tilde{w})\,. \end{cases}$$

Proof. By convexity, for every $x \in D_f$ such that $x < w$,

$$\frac{y-w}{y-x} f(x) + \frac{w-x}{y-x} f(y) \ge f(w) = 0$$

and therefore $f(x) \ge -\frac{w-x}{y-w} f(y) > 0$. Furthermore, for every $x \in (w, y]$,

$$f(x) \le \frac{y-x}{y-w} f(w) + \frac{x-w}{y-w} f(y) < 0\,.$$

Let $s = \sup\{x \in D_f : f(x) < 0\}$. If f is continuous in s, then $\tilde{w} = s$ and for every $x > \tilde{w}$ in D_f, $f(x) \ge -\frac{\tilde{w}-x}{y-\tilde{w}} f(y) > 0$. If f is discontinuous in s however, then by convexity, $f(x) = +\infty$ for $x > s$. □

Proposition 2.8. *Let $u \in I$ and consider $w(u)$ the stable equilibrium of (2.2b), defined in Theorem 2.3 (1). Then the basin of attraction of $w(u)$ is $\mathcal{B}(u) = (-\infty, \tilde{w}(u)) \cap D_{R(u,\cdot)}$, where $\tilde{w}(u)$ is the unstable equilibrium of (2.2b), defined in Theorem 2.3 (2), and $\tilde{w}(u) = \infty$ when (2.2b) admits only one equilibrium.*

Proof. By Lemma 2.6, $w \mapsto R(u, w)$ is convex. Since $w(u)$ is a stable equilibrium, the hypotheses of Lemma 2.7 are verified. Therefore, $R(u, w) > 0$ for every $w < w(u)$, whereas $R(u, w) < 0$ for every $w \in D_{R(u,\cdot)}$ such that $w(u) < w < \tilde{w}(u)$. This implies that the solution of

$$\partial_t \psi(t, u, w) = R(u, \psi(t, u, w)) , \qquad \psi(0, u, w) = w \tag{2.6}$$

eq:REquationAlone

converges to $w(u)$ for every $w \in (-\infty, \tilde{w}(u)) \cap D_{R(u,\cdot)}$, whereas, if $w > \tilde{w}$, the solution of (2.6) diverges to ∞. □

3. Large deviations theory

eDeviationsTheory

In this section, we recall some useful classical results of the large deviations theory. We refer the reader to (Dembo and Zeitouni, 1998) for the proofs and for a broader overview of the theory.

thm:GartnerEllis

Theorem 3.1 (Gärtner-Ellis)**.** *Let $(X^\epsilon)_{\epsilon \in]0,1]}$ be a family of random vectors in $\mathbb{R}^n$ with associated measure μ_ϵ. Assume that for each $\lambda \in \mathbb{R}^n$, the logarithmic moment generating function, defined as the limit*

$$\Lambda(\lambda) := \lim_{\epsilon \to 0} \epsilon \log \mathbb{E}\left[e^{\frac{\langle \lambda, X^\epsilon \rangle}{\epsilon}} \right],$$

exists as an extended real number. Assume also that 0 belongs to the interior of $D_\Lambda := \{\lambda \in \mathbb{R}^n : \Lambda(\lambda) < \infty\}$. Denoting

$$\Lambda^*(x) = \sup_{\lambda \in \mathbb{R}^n} \langle \lambda, x \rangle - \Lambda(\lambda) , \quad x \in \mathbb{R}^n,$$

the following hold:

(a) For any closed set F,

$$\limsup_{\epsilon \to 0} \epsilon \log \mu_\epsilon(F) \leq - \inf_{x \in F} \Lambda^*(x) .$$

(b) For any open set G,

$$\liminf_{\epsilon \to 0} \epsilon \log \mu_\epsilon(G) \geq - \inf_{x \in G \cap \mathcal{F}} \Lambda^*(x) ,$$

where $\mathcal{F}$ is the set of exposed points of Λ^*, whose exposing hyperplane belongs to the interior of D_Λ.

(c) If Λ is an essentially smooth, lower semi-continuous function, then μ_ϵ satisfies a LDP with good rate function Λ^*.

Remark 3.2. In our paper, the random variable X^ϵ will correspond to the value of the affine stochastic volatility process X computed at time $1/\epsilon$, and the limiting log-Laplace transform $\Lambda(\lambda)$ from the Gärtner-Ellis theorem therefore coincides with the function h defined in 2.5.

Definition 3.3. A partially ordered set $(\mathcal{P}, \leq)$ is called *right-filtering* if for every $i, j \in \mathcal{P}$, there exists $k \in \mathcal{P}$ such that $i \leq k$ and $j \leq k$.

Definition 3.4. A *projective system* $(\mathcal{Y}_j, p_{ij})_{i\leq j\in\mathcal{P}}$ on a partially ordered right-filtering set $(\mathcal{P},\leq)$ is a family of Hausdorff topological spaces $(\mathcal{Y}_j)_{j\in\mathcal{P}}$ and continuous maps $p_{ij}:\mathcal{Y}_j\to\mathcal{Y}_i$ such that $p_{ik}=p_{ij}\circ p_{jk}$ whenever $i\leq j\leq k$.

Definition 3.5. Let $(\mathcal{Y}_j, p_{ij})_{i\leq j\in\mathcal{P}}$ be a projective system on a partially ordered right-filtering set $(\mathcal{P},\leq)$. The *projective limit* of $(\mathcal{Y}_j, p_{ij})_{i\leq j\in\mathcal{P}}$, denoted $\mathcal{X}=\varprojlim\mathcal{Y}_j$, is the subset of topological spaces $\mathcal{Y}=\prod_{j\in\mathcal{P}}\mathcal{Y}_j$, consisting of all the elements $x=(y_j)_{j\in\mathcal{P}}$ for which $y_i=p_{ij}(y_j)$ whenever $i\leq j$, equipped with the topology induced by $\mathcal{Y}$. The *projective limit of closed subsets* $F_j\subseteq\mathcal{Y}_j$ are defined in the same way and denoted $F=\varprojlim F_j$.

Remark 3.6. The canonical projections of $\mathcal{X}$, i.e. the restrictions $p_j:\mathcal{X}\to\mathcal{Y}_j$ of the coordinate maps from $\mathcal{X}$ to $\mathcal{Y}_j$, are continuous.

thm:DawsonGartner

Theorem 3.7 (Dawson-Gärtner)**.** *Let $(\mathcal{Y}_j, p_{ij})_{i\leq j\in\mathcal{P}}$ be a projective system on a partially ordered right-filtering set $(\mathcal{P},\leq)$ and let (μ_ϵ) be a family of probabilities on $\mathcal{X}=\varprojlim\mathcal{Y}_j$, such that for any $j\in\mathcal{P}$, the Borel probability $\mu_\epsilon\circ p_j^{-1}$ on $\mathcal{Y}_j$ satisfies the LDP with good rate function Λ_j^*. Then μ_ϵ satisfies the LDP with good rate function*

$$\Lambda^*(x)=\sup_{j\in\mathcal{P}}\Lambda_j^*(p_j(x))\,.$$

thm:Varadhan

Theorem 3.8 (Varadhan's Lemma, version of (Guasoni and Robertson, 2008))**.** *Let $\mathcal{X}$ be a regular Hausdorff space, and let $(X^\epsilon)_{\epsilon\in]0,1]}$ be a family of $\mathcal{X}$-valued random variables, whose laws μ_ϵ satisfy a LDP with rate function Λ^*. If $\varphi:\mathcal{X}\to\mathbb{R}\cup\{-\infty\}$ is a function such that the set $\{\varphi>-\infty\}$ is open and φ continuous on this set and satisfies*

$$\limsup_{\epsilon\to0}\epsilon\,\log\mathbb{E}\left[\exp\left(\frac{\gamma\,\varphi(X^\epsilon)}{\epsilon}\right)\right]<\infty \tag{3.1}$$

varadhan.eq

for some $\gamma>1$, then

$$\lim_{\epsilon\to0}\epsilon\,\log\mathbb{E}\left[\exp\left(\frac{\varphi(X^\epsilon)}{\epsilon}\right)\right]=\sup_{x\in\mathcal{X}}\{\varphi(x)-\Lambda^*(x)\}\,.$$

4. Trajectorial large deviations for ASV models

sec:LDP-Affine

In this section, we prove a trajectorial LDP for (X_t) when the time horizon is large. Define, for $\epsilon\in(0,1]$ and $0\leq t\leq T$, the scaling $X_t^\epsilon=\epsilon X_{t/\epsilon}$. We proceed by proving first a LDP for X_t^ϵ in finite dimension, that we extend, in a second step to the whole trajectory of $(X_t^\epsilon)_{0\leq t\leq T}$.

4.1. Finite-dimensional LDP.

Let $\tau=\{0<t_1<\ldots<t_n=t\}$, by convention $t_0=0$, and define

$$\Lambda_{\epsilon,\tau}(\theta):=\log\mathbb{E}\left[e^{\sum_{k=1}^n\theta_k X_{t_k}^\epsilon}\right]\,,$$

for $\theta\in\mathbb{R}^n$. We start by formulating our main technical assumption. Recall that $w(u)$ and $\tilde{w}(u)$ denote, respectively, the stable and unstable equilibrium point of the Riccati equation (2.2b), defined in Theorem 2.3 (1) and (2).

ass:H2

Assumption 3. One of the following conditions is verified:

ass:H2a

(1) The interval support of F is $J=[u_-,u_+]$ and $w(u_-)=w(u_+)$.

ass:H2b (2) For every $u \in \mathbb{R}$, $\tilde{w}(\cdot) = \infty$, i.e, the generalized Riccati equations have only one (stable) equilibrium.

The above assumption may be seen as rather restrictive, however there are important models for which it is satisfied, such as the Heston model, with or without jumps, when there is no correlation between the Brownian motions driving the asset price and the volatility (see Remark 6.6). In the following Lemma we state a consequence of Assumption 3 which will be used hereafter.

lem:TildeW **Lemma 4.1.** *Let Assumption 3 hold true. For every $u_1, u_2 \in I$, $\tilde{w}(u_1) \geq w(u_2)$.*

Proof. If Assumption 3(2) holds, then the result is obvious. Assume then that it is Assumption 3(1) that holds. Since $u \mapsto w(u)$ is convex and $u \mapsto \tilde{w}(u)$ is concave (Keller-Ressel, 2011, Lemma 3.3), then for every $u_1, u_2 \in I$,

$$\tilde{w}(u_1) \geq \frac{u_+ - u_1}{u_+ - u_-}\tilde{w}(u_-) + \frac{u_1 - u_-}{u_+ - u_-}\tilde{w}(u_+) \geq w(u_-)\,,$$

while

$$w(u_2) \leq \frac{u_+ - u_2}{u_+ - u_-}w(u_-) + \frac{u_2 - u_-}{u_+ - u_-}w(u_+) = w(u_-)\,.$$

Therefore $\tilde{w}(u_1) \geq w(u_2)$ for every $u_1, u_2 \in I$. □

As a first step to apply Theorem 3.1, we prove the following result.

thm:LambdaTau **Theorem 4.2.** *Let $\theta \in \mathbb{R}^n$. If Assumption 3 holds, then*

$$\Lambda_\tau(\theta) := \lim_{\epsilon\to 0} \epsilon\, \Lambda_{\epsilon,\tau}(\theta/\epsilon) = \begin{cases} \sum_{j=1}^n (t_j - t_{j-1})\, h\left(\bar{\theta}_j\right) & \text{if } \bar{\theta}_j \in J\,,\ \forall j \\ \infty & \text{otherwise} \end{cases},$$

where $\bar{\theta}_j := \sum_{k=j}^n \theta_k$ and h is defined in (2.5).

Proof. Since Assumption 3 holds, then, by Lemma 4.1, $w(\bar{\theta}_{j+1}) \in \mathcal{B}(\bar{\theta}_j)$ for every j. Assume first that $\bar{\theta}_j \in J$ for every j. Using the Markov property and eq. (2.1), we obtain

$$\begin{aligned}
\Lambda_\tau(\theta) &= \lim_{\epsilon\to 0} \epsilon \log\left(\mathbb{E}\left[e^{\sum_{j=1}^n \theta_j X_{t_j/\epsilon}}\right]\right) \\
&= \lim_{\epsilon\to 0} \epsilon \log\left(\mathbb{E}\left[e^{\sum_{j=1}^{n-1} \theta_j X_{t_j/\epsilon}}\, \mathbb{E}\left(e^{\bar{\theta}_n X_{t_n/\epsilon}} \middle| X_{t_{n-1}/\epsilon}, V_{t_{n-1}/\epsilon}\right)\right]\right) \\
&= \lim_{\epsilon\to 0} \epsilon\, \phi\left(\frac{t_n - t_{n-1}}{\epsilon},\, \bar{\theta}_n,\, 0\right) \\
&\qquad + \epsilon \log\left(\mathbb{E}\left[e^{\sum_{j=1}^{n-2} \theta_j X_{t_j/\epsilon} + \bar{\theta}_{n-1} X_{t_{n-1}/\epsilon} + \psi\left(\frac{t_n - t_{n-1}}{\epsilon}, \bar{\theta}_n, 0\right) V_{t_{n-1}/\epsilon}}\right]\right).
\end{aligned}$$

Since $\bar{\theta}_n \in J$ and $0 \in \mathcal{B}(\bar{\theta}_n)$, eqs. (2.3) and (2.4) apply and

$$\begin{aligned}
\Lambda_\tau(\theta) = \lim_{\epsilon\to 0} \epsilon \log\Bigg(\mathbb{E}\Bigg[&e^{\sum_{j=1}^{n-2} \theta_j X_{t_j/\epsilon} + \bar{\theta}_{n-1} X_{t_{n-1}/\epsilon} + \psi\left(\frac{t_n - t_{n-1}}{\epsilon}, \bar{\theta}_n, 0\right) V_{t_{n-1}/\epsilon}}\Bigg]\Bigg) \\
&+ (t_n - t_{n-1})\, h(\bar{\theta}_n)\,.
\end{aligned}$$

Using the fact that $\bar{\theta}_j \in J$ and $w(\bar{\theta}_{j+1}) \in \mathcal{B}(\bar{\theta}_j)$ for every j, we can iterate the procedure to obtain

$$\Lambda_\tau(\theta) = \sum_{j=1}^n (t_j - t_{j-1})\, h\left(\bar{\theta}_j\right) + \lim_{\epsilon\to 0} \epsilon\, \psi\left(\frac{t_1 - t_0}{\epsilon},\, \bar{\theta}_1,\, w\left(\bar{\theta}_2\right)\right) V_0 + \epsilon \sum_{k=1}^n \theta_k X_0$$

$$= \sum_{j=1}^{n} (t_j - t_{j-1})\, h\left(\bar{\theta}_j\right) \ . \tag{4.1}$$ eq:LambdaN

Assume now that there exists k such that $\bar{\theta}_k \notin J$. Without loss of generality, we take the largest such k. Following the same procedure, we find

$$\Lambda_\tau(\theta) = \lim_{\epsilon \to 0} \epsilon \log \left(\mathbb{E}\left[e^{\sum_{j=1}^{k-2} \theta_j X_{t_j/\epsilon} + \bar{\theta}_{k-1} X_{t_{k-1}/\epsilon} + \psi\left(\frac{t_k - t_{k-1}}{\epsilon}, \bar{\theta}_k, w(\bar{\theta}_{k+1})\right)} V_{t_{k-1}/\epsilon} \right] \right)$$
$$+ \epsilon\, \phi\left(\frac{t_k - t_{k-1}}{\epsilon}, \bar{\theta}_k, w(\bar{\theta}_{k+1})\right) + \sum_{j=k+1}^{n} (t_j - t_{j-1})\, h(\bar{\theta}_j) \ .$$

Noting that $\phi(\cdot, u, w)$ explodes in finite time for $u \notin J$ then finishes the proof. □

We now proceed to the finite-dimensional large deviations result. Let us denote $J^n := D_{\Lambda_\tau}$.

FiniteDimensional **Theorem 4.3.** *Let $(X_t^\epsilon)_{t \geq 0,\, \epsilon \in (0,1]}$ and $\tau = \{t_1, ..., t_n\}$ as previously. Let Assumptions 2 and 3 hold true. Then $(X_{t_1}^\epsilon, ..., X_{t_n}^\epsilon)$ satisfies a LDP on $\mathbb{R}^n$ with good rate function*

$$\Lambda_\tau^*(x) = \sup_{\bar{\theta} \in J^n} \left\{ \sum_{j=1}^{n} \bar{\theta}_j (x_j - x_{j-1}) - \sum_{j=1}^{n} (t_j - t_{j-1})\, h\left(\bar{\theta}_j\right) \right\} \ ,$$

where h is defined in (2.5).

Proof. By Assumption 2(1), there exists $u \in J$ such that $u < 0$, which implies that $[u, 1] \subset J$ and therefore 0 is in the interior of J^n. Theorem 4.2 implies that the limit

$$\Lambda_\tau(\theta) = \lim_{\epsilon \to 0} \epsilon\, \Lambda_{\epsilon,\tau}(\theta/\epsilon) = \begin{cases} \sum_{j=1}^{n} (t_j - t_{j-1})\, h\left(\bar{\theta}_j\right) & \text{if } \bar{\theta}_j \in J \ , \ \forall j \\ \infty & \text{otherwise} \end{cases} \ ,$$

where $\bar{\theta}_j := \sum_{k=j}^{n} \theta_k$, exists as an extended real number.

Since, by Assumption 2(2), h is essentially smooth and lower semi-continuous, then so is Λ_τ. Theorem 3.1 then applies and $(X_{t_1}^\epsilon, ..., X_{t_n}^\epsilon)$ satisfies a LDP, on $\mathbb{R}^n$, with good rate function

$$\Lambda_\tau^*(x) = \sup_{\theta \in \mathbb{R}^n} \left\{ \theta^\top x - \Lambda_\tau(\theta) \right\} \ .$$

Furthermore, with the convention $x_0 = 0$, and letting $\theta_j = \bar{\theta}_j - \bar{\theta}_{j+1}$ for $j = 1, \ldots, n-1$ and $\theta_n = \bar{\theta}_n$, we have,

$$\begin{aligned} \Lambda_\tau^*(x) &= \sup_{\theta \in \mathbb{R}^n} \left\{ \theta^\top x - \Lambda_\tau(\theta) \right\} \\ &= \sup_{\bar{\theta} \in J^n} \left\{ \sum_{j=1}^{n} \sum_{k=j}^{n} \theta_k (x_j - x_{j-1}) - \sum_{j=1}^{n} (t_j - t_{j-1})\, h\left(\bar{\theta}_j\right) \right\} \\ &= \sup_{\bar{\theta} \in J^n} \left\{ \sum_{j=1}^{n} \bar{\theta}_j (x_j - x_{j-1}) - \sum_{j=1}^{n} (t_j - t_{j-1})\, h\left(\bar{\theta}_j\right) \right\} \ , \end{aligned}$$

which finishes the proof. □

4.2. **Infinite-dimensional LDP.**

4.2.1. *Extension of the LDP.* We now extend the LDP to the whole trajectory of $(X_t^\epsilon)_{0\leq t\leq T}$ on $\mathcal{F}([0,T],\mathbb{R}) := \{x : [0,T] \to \mathbb{R},\, x_0 = 0\}$, the set of all functions from $[0,T]$ to $\mathbb{R}$ that vanish at 0, by proving the following general lemma.

lem:DawsonGartner

Lemma 4.4. *Let $(\mathcal{P},\leq)$ be the partially ordered right-filtering set*

$$\mathcal{P} = \bigcup_{n=1}^{\infty} \{(t_1,...,t_n)\ ,\ 0 \leq t_1 \leq ... \leq t_n \leq T\}$$

ordered by inclusion. We consider, on $(\mathcal{P},\leq)$, the projective system $(\mathcal{Y}_j, p_{ij})_{i\leq j\in\mathcal{P}}$ defined by $\mathcal{Y}_j = \mathbb{R}^{\#j}$ and $p_{ij} : \mathcal{Y}_j \to \mathcal{Y}_i$ the natural projection on shared times. Assume that for any $j = \{t_1,...,t_n\}$, the finite-dimensional process $(X^\epsilon_{t_1},...,X^\epsilon_{t_n})$ satisfies a large deviation property with good rate function Λ^*_j. Then the family $(X_t^\epsilon)_{0\leq t\leq T}$ satisfies a large deviation property on $\mathcal{X} = \mathcal{F}([0,T],\mathbb{R})$ equipped with the topology of pointwise convergence, with good rate function

$$\Lambda^*(x) = \sup_{j\in\mathcal{P}} \Lambda^*_j(p_j(x))\ ,$$

where $p_\tau(x) = (x_{t_1},...,x_{t_n})$ is the canonical projection from $\mathcal{X}$ to $\mathcal{Y}_\tau$.

Proof. Let μ^ϵ be the probability measure generated by $(X_t^\epsilon)_{0\leq t\leq T}$ on $\mathcal{X}$. Then, by hypothesis, for any $j \in \mathcal{P}$, $\mu^\epsilon \circ p_j^{-1}$ satisfies a LDP with good rate function Λ^*_τ. The result then follows from Theorem 3.7. □

ensionalExistence

Theorem 4.5. *Assume that Assumptions 2 and 3 hold. Then $(X_t^\epsilon)_{0\leq t\leq T}$ satisfies a LDP on $\mathcal{F}([0,T],\mathbb{R})$ equipped with the topology of point-convergence, as $\epsilon \to 0$, with good rate function*

$$\Lambda^*(x) = \sup_\tau \Lambda^*_\tau(x)\ ,$$

where the supremum is taken over the discrete ordered subsets of the form $\tau = \{t_1,...,t_n\} \subset [0,T]$.

Proof. The result is a direct application of Lemma 4.4. □

4.2.2. *Calculation of the rate function.* We finally calculate the rate function in Theorem 4.5.

ionalRateFunction

Theorem 4.6. *The rate function of Theorem 4.5 is*

$$\Lambda^*(x) = \int_0^T h^*(\dot{x_t^{ac}})\, dt + \int_0^T \mathcal{H}\left(\frac{d\nu}{d|\nu|}(t)\right) d|\nu|\ ,$$

where

$$h^*(y) := \sup_{\theta\in J} \{\theta y - h(\theta)\}\ , \qquad \mathcal{H}(y) := \lim_{\epsilon\to 0} \epsilon\, h^*(y/\epsilon)\ ,$$

h is defined in (2.5), $\dot{x^{ac}}$ is the derivative of the absolutely continuous part of x, ν is the finite signed measure which is the singular component of dx with respect to the Lebesgue measure, $|\nu|$ is the total variation measure of ν and $\frac{d\nu}{d|\nu|}$ is the Radon-Nikodym derivative of ν with respect to its total variation measure.

Proof. By Theorem 4.3, for every $x \in \mathcal{F}([0,T],\mathbb{R})$,

$$\sup_\tau \Lambda^*_\tau(x) = \sup_\tau \sup_{\bar\theta\in J^{\#\tau}} \sum_{j=1}^{\#\tau} \bar\theta_j(x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1})h(\bar\theta_j)$$

$$= \sup_{\tau} \sup_{\xi \in C([0,T],J)} \sum_{j=1}^{\#\tau} \xi_{t_j}(x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1})h(\xi_{t_j})$$

$$= \sup_{\xi \in C([0,T],J)} \sup_{\tau} \sum_{j=1}^{\#\tau} \xi_{t_j}(x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1})h(\xi_{t_j}) \,.$$

The second line follows from the first line since one can always find a continuous function ξ such that $\xi_{t_i} = \bar{\theta}_i$ for $i = 1, \ldots, n$. Since we have assumed that there exists $u < 0$ in J, then if x has infinite variation, we immediately find that $\Lambda^*(x) = \infty$. Assume therefore that x has finite variation. We wish to show that

$$\sup_{\xi \in C([0,T],J)} \sup_{\tau} \sum_{j=1}^{\#\tau} \xi_{t_j}(x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1})h(\xi_{t_j})$$

$$= \sup_{\xi \in C([0,T],J)} \int_0^T \xi_t dx_t - \int_0^T h(\xi_t)dt \,.$$

Notice that

$$\sup_{\xi \in C([0,T],J)} \sup_{\tau} \sum_{j=1}^{\#\tau} \xi_{t_j}(x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1})h(\xi_{t_j})$$

$$\geq \sup_{\xi \in C([0,T],J)} \limsup_{\tau} \sum_{j=1}^{\#\tau} \xi_{t_j}(x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1})h(\xi_{t_j})$$

$$= \sup_{\xi \in C([0,T],J)} \int_0^T \xi_t dx_t - \int_0^T h(\xi_t)dt \,.$$

To prove the other inequality, we use the following construction. Fix τ and let $\xi \in C([0,T],J)$. Let also $\epsilon > 0$ such that $\epsilon < \min(t_j - t_{j-1})$ and define $\xi^{\epsilon,\tau}$ as

$$\xi_t^{\epsilon,\tau} = \begin{cases} \xi_{t_{j-1}} + \frac{t - t_{j-1}}{\epsilon}\left(\xi_{t_j} - \xi_{t_{j-1}}\right) & \text{if } t \in [t_{j-1},\, t_{j-1} + \epsilon] \,, \\ \xi_{t_j} & \text{if } t \in [t_{j-1} + \epsilon,\, t_j] \,. \end{cases}$$

Then

$$\left| \sum_{j=1}^{\#\tau} \xi_{t_j}(x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1})h(\xi_{t_j}) - \int_0^T \xi_t^{\epsilon,\tau} dx_t + \int_0^T h(\xi_t^{\epsilon,\tau})dt \right|$$

$$= \left| \sum_{j=1}^{\#\tau} (\xi_{t_j} - \xi_{t_{j-1}}) \int_{t_{j-1}}^{t_{j-1}+\epsilon} \left(1 - \frac{t - t_{j-1}}{\epsilon}\right) dx_t + \int_{t_{j-1}}^{t_{j-1}+\epsilon} h(\xi_t^{\epsilon,\tau}) - h(\xi_{t_j})dt \right|$$

$$\leq \sum_{j=1}^{\#\tau} \left|\xi_{t_j} - \xi_{t_{j-1}}\right| \left| \int_{t_{j-1}}^{t_{j-1}+\epsilon} \left(1 - \frac{t - t_{j-1}}{\epsilon}\right) dx_t \right|$$

$$+ 2\epsilon \max\left\{ |h(\xi)| \,:\, \xi \in [\xi_{t_{j-1}}, \xi_{t_j}] \right\}$$

$$\leq \sum_{j=1}^{\#\tau} \left|\xi_{t_j} - \xi_{t_{j-1}}\right| \mu_x\big(]0,\epsilon]\big) + 2\epsilon \max\left\{|h(\xi)| \,:\, \xi \in [\xi_{t_{j-1}}, \xi_{t_j}]\right\} \underset{\epsilon\to 0}{\to} 0\,,$$

where μ_x is the measure associated with x. Hence

$$\sup_{\xi\in C([0,T],J)} \sup_{\tau} \sum_{j=1}^{\#\tau} \xi_{t_j}(x_{t_j} - x_{t_{j-1}}) - (t_j - t_{j-1})h(\xi_{t_j})$$
$$\leq \sup_{\xi\in C([0,T],J)} \int_0^T \xi_t dx_t - \int_0^T h(\xi_t) dt$$

and

$$\Lambda^*(x) = \sup_{\xi\in C([0,T],J)} \int_0^T \xi_t dx_t - \int_0^T h(\xi_t)\, dt\,.$$

We will now use (Rockafellar, 1971, Thm. 5.) to obtain the result. Since x has finite variation, the measure dx_t is regular. Using the notation of (Rockafellar, 1971), in our case the multifunction D is the constant multifunction $t \mapsto D(t) = J$. Therefore D is fully lower semi-continuous. Furthermore, since $[0,1] \subset J$, the interior of $D(t)$ is non-empty. The set $[0,T]$ is compact with no non-empty open sets of measure 0 and for every u in the interior of J, and $V \in [0,T]$ open,

$$\int_V |h(u)|\, dt \leq T|h(u)| < \infty\,.$$

(Rockafellar, 1971, Thm. 5.) then implies that

$$\sup_{\xi\in C([0,T],J)} \int_0^T \xi_t\, dx_t - \int_0^T h(\xi_t)\, dt = \int_0^T h^*(\dot{x_t^{ac}})\, dt + \int_0^T \mathcal{H}\left(\frac{d\nu}{d|\nu|}(t)\right) d|\nu|\,,$$

where

$$h^*(y) = \lim_{\epsilon\to 0} \sup_{\theta\in J} \{\theta y - h(\theta)\}\,, \qquad \mathcal{H}(y) = \lim_{\epsilon\to 0} \epsilon\, h^*(y/\epsilon)\,,$$

and $\dot{x}^{ac}$ and ν have been defined in the statement of the theorem. □

Remark 4.7. In particular, the proof of Theorem 4.6 shows that, if x does not belong to V_r, the set of trajectories $x : [0,t] \to \mathbb{R}$ with bounded variation, then $\Lambda^*(x) = \infty$.

VarianceReduction

5. Variance reduction

Denote $P(S)$ the payoff of an option on $(S_t)_{0\leq t\leq T}$. The price of an option is generally calculated as the expectation $\mathbb{E}[P(S)]$ under a certain risk-neutral measure $\mathbb{P}$. For any equivalent measure $\mathbb{Q}$, the price of the option can be written

$$\mathbb{E}[P(S)] = \mathbb{E}^{\mathbb{Q}}\left[P(S)\frac{d\mathbb{P}}{d\mathbb{Q}}\right]\,.$$

The variance of $P(S)$ is

$$\mathbb{V}\mathrm{ar}_{\mathbb{P}}\left(P(S)\right) = \mathbb{E}\left[P^2(S)\right] - \mathbb{E}\left[P(S)\right]^2\,,$$

whereas

$$\mathbb{V}\mathrm{ar}_{\mathbb{Q}}\left(P(S)\frac{d\mathbb{P}}{d\mathbb{Q}}\right) = \mathbb{E}^{\mathbb{Q}}\left[P^2(S)\left(\frac{d\mathbb{P}}{d\mathbb{Q}}\right)^2\right] - \left(\mathbb{E}^{\mathbb{Q}}\left[P(S)\frac{d\mathbb{P}}{d\mathbb{Q}}\right]\right)^2$$
$$= \mathbb{E}\left[P^2(S)\frac{d\mathbb{P}}{d\mathbb{Q}}\right] - \mathbb{E}\left[P(S)\right]^2 .$$

We can therefore choose $\mathbb{Q}$ in order to reduce the variance of the random variable, whose expectation gives the price of the option.

As discussed in the introduction, we follow (Genin and Tankov, 2020) by considering the class of exponentially affine transforms, defined as follows:

$$\frac{d\mathbb{P}_\Theta}{d\mathbb{P}} = \frac{e^{\int_0^T X_t\,\Theta(dt)}}{\mathbb{E}\left[e^{\int_0^T X_t\,\Theta(dt)}\right]},$$

where Θ belongs to M, the set of signed measures on $[0,T]$. Denoting $H(X) = \log P\left(S_0\, e^X\right)$, the optimization problem writes

$$\inf_{\Theta\in M} \mathbb{E}\left[\exp\left(2H(X) - \int_0^T X_t\,\Theta(dt) + \mathcal{G}_1(\Theta)\right)\right], \tag{5.1}$$

eq:OptimizationPr

where

$$\mathcal{G}_\epsilon(\Theta) := \epsilon \log \mathbb{E}\left[e^{\frac{1}{\epsilon}\int_0^T X_t^\epsilon\,\Theta(dt)}\right].$$

The optimization problem (5.1) cannot be solved explicitly. We therefore choose to solve the problem asymptotically using the two following lemmas. Denote $\bar{M}$ the set of measures $\Theta \in M$ with support on a finite set of points. We first give a lemma that characterizes the behaviour of $\mathcal{G}_\epsilon(\Theta)$ as $\epsilon \to 0$, for $\Theta \in \bar{M}$ as this will be sufficient for the cases that we will consider in Section 6 (see Prop. 5.5). We stress the fact that although Lemmas 5.1 and 5.2 are proved for $\Theta \in \bar{M}$, the resulting asymptotic proxy for the variance, and the resulting candidate importance sampling measure are well defined, and may be used for any $\Theta \in M$, provided that they lead to a sufficient reduction in the variance of the estimator.

lem:Gepsilon

Lemma 5.1. *If Assumption 3 holds, then for any measure $\Theta \in \bar{M}$, such that for every $t \in [0,T]$, $\Theta([t,T]) \in J$, we have*

$$\lim_{\epsilon\to 0} \mathcal{G}_\epsilon(\Theta) = \int_0^T h(\Theta([t,T]))\,dt .$$

where h is defined in (2.5).

Proof. Denote $\tau = \{t_1, ..., t_n\}$, the support of Θ. We then obtain

$$\lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[e^{\frac{1}{\epsilon}\int_0^T X_t^\epsilon\,\Theta(dt)}\right] = \lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[e^{\frac{1}{\epsilon}\sum_{j=1}^n X_{t_j}^\epsilon\,\Theta(\{t_j\})}\right]$$
$$= \sum_{j=1}^n (t_j - t_{j-1})\, h\left(\Theta(\{t_j\})\right)$$
$$= \int_0^T h(\Theta([t,T]))\,dt$$

by applying Theorem 4.2 to $\theta = (\Theta(\{t_1\}), ..., \Theta(\{t_n\}))$. □

Next, we give a result that characterizes the behaviour of the variance minimization problem (5.1) where X has been replaced by X^ϵ as $\epsilon \to 0$.

lem:LimitFunction

Lemma 5.2. *Let $\Theta \in \bar{M}$ such that $-\Theta([t,T]) \in J^\circ$ for every $t \in [0,T]$. Let Assumptions 2 and 3 hold true and assume furthermore that $H : \mathcal{F}([0,T],\mathbb{R}) \to \mathbb{R}$ is bounded from above by a constant C and continuous on D the set of functions $x \in V_r$, such that $H(x) > -\infty$, with respect with to the pointwise convergence topology. Then*

$$\lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[\exp\left(\frac{2H(X^\epsilon) - \int_0^T X_t^\epsilon\, \Theta(dt) + \mathcal{G}_\epsilon(\Theta)}{\epsilon}\right)\right] = \sup_{x\in D}\left\{2H(x) - \int_0^T x_t \Theta(dt) - \Lambda^*(x)\right\} + \int_0^T h(\Theta([t,T]))\, dt\,.$$

Proof. First note that, by Lemma 5.1,

$$\lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[\exp\left(\frac{2H(X^\epsilon) - \int_0^T X_t^\epsilon\, \Theta(dt) + \mathcal{G}_\epsilon(\Theta)}{\epsilon}\right)\right] = \lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[\exp\left(\frac{2H(X^\epsilon) - \int_0^T X_t^\epsilon\, \Theta(dt)}{\epsilon}\right)\right] + \int_0^T h(\Theta([t,T]))\, dt\,.$$

We therefore just need to prove that

$$\lim_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[\exp\left(\frac{2H(X^\epsilon) - \int_0^T X_t^\epsilon\, \Theta(dt)}{\epsilon}\right)\right] = \sup_{x\in D}\left\{2H(x) - \int_0^T x_t \Theta(dt) - \Lambda^*(x)\right\}.$$

Denote $\varphi : \mathcal{F}([0,T],\mathbb{R}) \to \mathbb{R}$ the function $\varphi(x) = 2H(x) - \int_0^T x_t\, \Theta(dt)$. Since H is assumed to be continuous and Θ has support on τ, φ is continuous. Let us show the integrability condition of Theorem 3.8. For every $\gamma > 0$

$$\begin{aligned}
&\limsup_{\epsilon\to 0} \epsilon\, \log \mathbb{E}\left[\exp\left(\frac{\gamma\, \varphi(X^\epsilon)}{\epsilon}\right)\right] \\
&\quad = \limsup_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[\exp\left(\frac{2\gamma H(X^\epsilon) - \gamma \int_0^T X_t^\epsilon\, \Theta(dt)}{\epsilon}\right)\right] \\
&\quad \le 2\gamma C + \limsup_{\epsilon\to 0} \epsilon \log \mathbb{E}\left[e^{\frac{1}{\epsilon}\int_0^T X_t^\epsilon\, d(-\gamma\Theta)_t}\right]\,.
\end{aligned}$$

Since $-\Theta([t,T]) \in J^\circ$ for every $t \in [0,T]$, there exists $\gamma > 1$ such that $-\gamma\Theta([t,T])$ remains in J for every t. Therefore Lemma 5.1 applies and

$$\limsup_{\epsilon\to 0} \epsilon\, \log \mathbb{E}\left[\exp\left(\frac{\gamma\, \varphi(X^\epsilon)}{\epsilon}\right)\right] \le 2\gamma C + \int_0^T h(-\gamma\, \Theta([t,T]))\, dt < \infty\,.$$

Theorem 3.8 then applies and yields the result. □

In view of Lemma 5.2 we suggest to compute the candidate variance reduction parameter by minimizing over $\Theta \in M$ the expression

$$\sup_{x \in V_r} \left\{ 2H(x) - \int_0^T x_t \, \Theta(dt) - \Lambda^*(x) \right\} + \int_0^T h(\Theta([t,T])) \, dt \, . \tag{5.2}$$ express

It is then natural to ask, how close the corresponding measure change will be to the optimal one which minimizes the variance of the Monte Carlo estimator over all possible measure changes. Varadhan's lemma allows to define a notion of asymptotic optimality, which provides a partial answer to this question. Consider a family of importance sampling measures $(\mathbb{Q}(\varepsilon))_{\varepsilon>0}$. By Jensen's inequality, for all $\varepsilon > 0$,

$$\mathbb{E}^{\mathbb{Q}(\varepsilon)} \left[\left(e^{\frac{1}{\varepsilon} H(X^\varepsilon)} \frac{d\mathbb{P}}{d\mathbb{Q}(\varepsilon)} \right)^2 \right] \geq \mathbb{E} \left[e^{\frac{1}{\varepsilon} H(X^\varepsilon)} \right]^2$$

A family of importance sampling measures changes is called *asymptotically optimal* if, for this family, the log-scale decay rates of the above expressions are the same. In other words, the asymptotically optimal measure change does at least as well as any other measure change at the logarithmic scale of large deviations.

optvar.def **Definition 5.3.** Let $(\mathbb{Q}(\varepsilon))_{\varepsilon>0}$ be a family of importance sampling measure changes. We say that $(\mathbb{Q}(\varepsilon))_{\varepsilon>0}$ is asymptotically optimal if

$$\lim_{\varepsilon \downarrow 0} \varepsilon \log \mathbb{E}^{\mathbb{Q}(\varepsilon)} \left[\left(e^{\frac{1}{\varepsilon} H(X^\varepsilon)} \frac{d\mathbb{P}}{d\mathbb{Q}(\varepsilon)} \right)^2 \right] = 2 \lim_{\varepsilon \downarrow 0} \varepsilon \log \mathbb{E} \left[e^{\frac{1}{\varepsilon} H(X^\varepsilon)} \right] .$$

The theorem below follows immediately from Theorem 8 of (Genin and Tankov, 2020) and shows that in the case of concave payoffs, the computation of the minimizer of (5.2) is simplified and we have asymptotic optimality.

thm:Equivalence **Theorem 5.4.** *Let H be concave and assume that the set $\{x \in V_r : H(x) > -\infty\}$ is non-empty and contains a constant element. Assume furthermore that H is continuous on this set with respect to the topology of pointwise convergence, that h is lower semi-continuous with open and bounded effective domain and that there exists a $\lambda > 0$ such that h is complex-analytic on $\{z \in \mathbb{C} : |Im(z)| < \lambda\}$. Then*

$$\inf_{\Theta \in M} \sup_{x \in V_r} \left\{ 2H(x) - \int_0^T x_t \Theta(dt) - \Lambda^*(x) + \int_0^T h(\Theta([t,T])) \, dt \right\} \tag{5.3}$$ line1.eq

$$= 2 \inf_{\Theta \in M} \left\{ \widehat{H}(\Theta) + \int_0^T h(\Theta([t,T])) \, dt \right\} , \tag{5.4}$$ line2.eq

where

$$\widehat{H}(\Theta) = \sup_{x \in V_r} \left\{ H(x) - \int_0^T x_t \, \Theta(dt) \right\} .$$

Furthermore, if the measure Θ^* minimises the left-hand side of the above equation, it also minimises the right-hand side. Finally, if the payoff functional and the measure Θ^* satisfy the assumptions of Lemma 5.2 then the importance sampling measure corresponding to Θ^* is asymptotically optimal.

Finite-dimensional dependence. A simple example of a functional H which is continuous in the topology of pointwise convergence is the situation when H depends on a path x only through its values at a finite number of points $x_{t_1}, ..., x_{t_n}$. In this case, our results hold under less stringent assumptions, and we state them as a separate proposition. The asymptotically optimal variance reduction measure is also supported by the points $\{t_1, \dots, t_n\}$. To simplify notation, we denote x_{t_i} by x_i and we introduce a function $H_\tau : \mathbb{R}^n \to \mathbb{R} \cup \{-\infty\}$ such that $H(x) = H_\tau(x_1, \dots, x_n)$.

MeasureIsDiscrete

Proposition 5.5. *Let $\tau = \{t_1, ..., t_n\}$ and let $H(x) = H_\tau(x_1, \dots, x_n)$. Assume that the set $\{x \in \mathbb{R}^n : H_\tau(x) > -\infty\}$ is nonempty and that H_τ is concave and continuous on this set. Let Assumptions 2 and 3 hold true, let the effective domain J of h be bounded and contain a neighborhood of 0, and assume that $h(x)$ is lower semicontinuous on $\overline{J}$. Define*

$$\widehat{H}_\tau(\bar\theta) = \sup_{x\in\mathbb{R}^n} \left\{ H_\tau(x) - \sum_{i=1}^n \bar\theta_i (x_i - x_{i-1}) \right\}.$$

Then, there exists $\bar\theta^*$ which minimizes

$$F(\bar\theta) := \widehat{H}_\tau(\bar\theta) + \sum_{i=1}^n (t_i - t_{i-1}) h(\bar\theta_i), \tag{5.5}$$

minim.eq

and the family of importance sampling measures defined by

$$\frac{d\mathbb{P}^\epsilon}{d\mathbb{P}} = \frac{e^{\sum_{i=1}^n (X^\epsilon_{t_i} - X^\epsilon_{t_{i-1}})\bar\theta^*_i}}{\mathbb{E}\left[e^{\sum_{i=1}^n (X^\epsilon_{t_i} - X^\epsilon_{t_{i-1}})\bar\theta^*_i}\right]}$$

is asymptotically optimal.

Proof. Let $\{\bar\theta^{(k)}\}_{k\geq 1} \subseteq J^n$ be a minimizing sequence for (5.5). Since J is bounded, this sequence has a subsequence $\{\bar\theta^{(k_m)}\}_{m\geq 1}$, converging to $\bar\theta^* \in \overline{J}^n$. As a supremum over a family of linear functions $\widehat{H}_\tau$ is lower semicontinuous. By the lower semicontinuity of h, it follows that F is lower semicontinuous as well and $\bar\theta^*$ is a minimizer of (5.5) on $\overline{J}^n$. Moreover, since H_τ is bounded and $h(0) = 0$, $F(0) < \infty$ and thus also $F(\bar\theta^*) < \infty$. Assume that $\bar\theta^* \in \partial(J^n)$, let $\rho \in (0,1)$ and let k be such that $\bar\theta^*_k \in \partial J$. Then, $\rho\bar\theta^* \in \mathring{J}^n$ and, using the convexity of $\widehat{H}_\tau$ and h, we have the following estimate

$$\begin{aligned} F(\rho\bar\theta^*) &= F(\bar\theta^*) + \widehat{H}_\tau(\rho\bar\theta^*) - \widehat{H}_\tau(\bar\theta) + \sum_{i=1}^n (t_i - t_{i-1})\{h(\rho\bar\theta^*_i) - h(\bar\theta^*_i)\} \\ &\leq F(\bar\theta^*) + (1-\rho)\{\widehat{H}_\tau(0) - \widehat{H}_\tau(\bar\theta^*)\} \\ &\quad + (1-\rho) \sum_{i=1,\dots,n; i\neq k} (t_i - t_{i-1})\{h(0) - h(\bar\theta^*_i)\} - (t_k - t_{k-1})(1-\rho)\bar\theta_k \nabla h(\rho\bar\theta_k) \end{aligned}$$

To fix the ideas, assume that $\bar\theta_k > 0$. Then, by the essential smoothness of h, $\nabla h(\rho\bar\theta_k) \to +\infty$ as $\rho \to 1$, so that there exists $\rho < 1$ with $F(\rho\bar\theta^*) < F(\bar\theta^*)$. This is in contradiction with the assumption that $\bar\theta^*$ is the minimizer of F, and therefore $\bar\theta^* \in \mathring{J}^n$. In view of this condition, and Lemma 5.1, assumption (3.1) holds for the

function ϕ given by

$$x \in \mathbb{R}^n \mapsto \phi(x) := 2H_\tau(x) - \sum_{i=1}^n \bar{\theta}_i^*(x_i - x_{i-1}),$$

with the convention $x_0 = 0$.

Define the measure $\Theta := \sum_{i=1}^n \theta_i \delta_{t_i}$, where δ_{t_i} is the Dirac measure at t_i, $\theta_i = \bar{\theta}_i^* - \bar{\theta}_{j+1}^*$ for $j = 1, \ldots, n-1$ and $\theta_n = \bar{\theta}_n^*$. By Varadhan's lemma (Theorem 3.8) applied with the LDP of Theorem 4.3 and Lemma 5.1 applied with the measure Θ, it follows that

$$\begin{aligned}
&\lim_{\epsilon\to 0} \epsilon \log \mathbb{E}^{\mathbb{P}^\epsilon}\left[\left(e^{\frac{1}{\epsilon}H_\tau(X^\epsilon)}\frac{d\mathbb{P}}{d\mathbb{P}^\epsilon}\right)^2\right] \\
&= \lim_{\epsilon\to 0} \epsilon \log \left\{ \mathbb{E}\left[\exp\left(\frac{2H_\tau(X^\epsilon) - \sum_{i=1}^n (X_{t_i}^\epsilon - X_{t_{i-1}}^\epsilon)\bar{\theta}_i^*}{\epsilon}\right)\right] \right. \\
&\qquad \left. \times \mathbb{E}\left[\exp\left(\frac{\sum_{i=1}^n (X_{t_i}^\epsilon - X_{t_{i-1}}^\epsilon)\bar{\theta}_i^*}{\epsilon}\right)\right] \right\} \\
&= \sup_{x\in\mathbb{R}^n} \left\{ 2H_\tau(x) - \sum_{i=1}^n (x_{t_i} - x_{t_{i-1}})\bar{\theta}_i^* - \Lambda_\tau^*(x) \right\} + \sum_{i=1}^n (t_i - t_{i-1}) h(\bar{\theta}_i^*) \\
&= \sup_{x\in\mathbb{R}^n} \inf_{\bar{\xi}\in J^n} \left\{ 2H_\tau(x) - \sum_{i=1}^n (x_{t_i} - x_{t_{i-1}})(\bar{\theta}_i^* + \bar{\xi}_i) + \sum_{i=1}^n (t_i - t_{i-1})(h(\bar{\xi}_i) + h(\bar{\theta}_i^*)) \right\}
\end{aligned}$$

In view of our assumptions (in particular, the boundedness of J), the inf and the sup may be exchanged, see Proposition VI.2.3 in (Ekeland and Temam, 1999), so that the above is equal to the following:

$$\begin{aligned}
&\inf_{\bar{\xi}\in J^n} \sup_{x\in\mathbb{R}^n} \left\{ 2H_\tau(x) - \sum_{i=1}^n (x_{t_i} - x_{t_{i-1}})(\bar{\theta}_i^* + \bar{\xi}_i) + \sum_{i=1}^n (t_i - t_{i-1})(h(\bar{\xi}_i) + h(\bar{\theta}_i^*)) \right\} \\
&= \inf_{\bar{\xi}\in J^n} \left\{ 2\widehat{H}_\tau\left(\frac{\bar{\xi} + \bar{\theta}^*}{2}\right) + \sum_{i=1}^n (t_i - t_{i-1}) h(\bar{\xi}_i) + \sum_{i=1}^n (t_i - t_{i-1}) h(\bar{\theta}_i^*) \right\} \\
&= 2\left\{ \widehat{H}_\tau\left(\bar{\theta}^*\right) + \sum_{i=1}^n (t_i - t_{i-1}) h(\bar{\theta}_i^*) \right\}
\end{aligned}$$

The last line follows because, on the one hand, by convexity of h and by definition of $\bar{\theta}^*$,

$$\begin{aligned}
&\inf_{\bar{\xi}\in J^n} \left\{ 2\widehat{H}_\tau\left(\frac{\bar{\xi} + \bar{\theta}^*}{2}\right) + \sum_{i=1}^n (t_i - t_{i-1}) h(\bar{\xi}_i) + \sum_{i=1}^n (t_i - t_{i-1}) h(\bar{\theta}_i^*) \right\} \\
&\geq \inf_{\bar{\xi}\in J^n} \left\{ 2\widehat{H}_\tau\left(\frac{\bar{\xi} + \bar{\theta}^*}{2}\right) + 2\sum_{i=1}^n (t_i - t_{i-1}) h\left(\frac{\bar{\xi}_i + \bar{\theta}_i^*}{2}\right) \right\} \\
&= 2 \inf_{\bar{\xi}\in J^n} \left\{ \widehat{H}_\tau\left(\bar{\xi}\right) + \sum_{i=1}^n (t_i - t_{i-1}) h\left(\bar{\xi}_i\right) \right\} = F(\bar{\theta}^*),
\end{aligned}$$

and on the other hand, choosing $\bar{\xi} = \bar{\theta}^*$, we have that

$$\inf_{\bar\xi\in J^n}\left\{2\widehat H_\tau\left(\frac{\bar\xi+\bar\theta^*}{2}\right)+\sum_{i=1}^n(t_i-t_{i-1})h(\bar\xi_i)+\sum_{i=1}^n(t_i-t_{i-1})h(\bar\theta^*_i)\right\}$$
$$\le 2\widehat H_\tau\left(\bar\theta^*\right)+2\sum_{i=1}^n(t_i-t_{i-1})h(\bar\theta^*_i).$$

By another application of Varadhan's lemma (here the boundedness of H_τ suffices to ensure its applicability), exchanging the inf and the sup by the same argument as above, we have:

$$\begin{aligned}
&\lim_{\epsilon\to0}\epsilon\log\mathbb{E}\left[\exp\left(\frac{H_\tau(X^\epsilon)}{\epsilon}\right)\right]=\sup_{x\in\mathbb{R}^n}\{H_\tau(x)-\Lambda^*_\tau(x)\}\\
&=\sup_{x\in\mathbb{R}^n}\inf_{\bar\theta\in J^n}\left\{H_\tau(x)-\sum_{j=1}^n\bar\theta_j(x_j-x_{j-1})+\sum_{j=1}^n(t_j-t_{j-1})h(\bar\theta_j)\right\}\\
&=\inf_{\bar\theta\in J^n}\sup_{x\in\mathbb{R}^n}\left\{H_\tau(x)-\sum_{j=1}^n\bar\theta_j(x_j-x_{j-1})+\sum_{j=1}^n(t_j-t_{j-1})h(\bar\theta_j)\right\}\\
&=\inf_{\bar\theta\in J^n}\left\{\widehat H_\tau(\bar\theta)-\sum_{j=1}^n(t_j-t_{j-1})h(\bar\theta_j)\right\}\\
&=F(\bar\theta^*),
\end{aligned}$$

which shows the asymptotic optimality of $\bar\theta^*$ as per Definition 5.3.

□

6. Numerical examples

NumericalExamples

In this section, we apply the variance reduction method to several examples. We first prove a result for options on the average value of the underlying over a finite set of points.

prop:HHat

Proposition 6.1. *Let $\tau=\{t_1,...,t_n\}$ and consider an option with log-payoff*

$$H(x)=\log\left(K-\frac{S_0}{n}\sum_{j=1}^n e^{x_j}\right)_+. \tag{6.1}$$

asianpayoff.eq

Then H satisfies the assumptions of Proposition 5.5 and for any $\bar\theta\in\mathbb{R}^n$,

$$\widehat H(\bar\theta)=\log\left(\frac{K}{1-\sum_{l=1}^n\theta_l}\right)-\sum_{m=1}^n\theta_m\log\left(\frac{-\theta_m\,n\,K/S_0}{1-\sum_{l=1}^n\theta_l}\right) \tag{6.2}$$

eq:HHat

where we use the notation $\theta_j=\bar\theta_j-\bar\theta_{j+1}$ for $j=1,\dots,n-1$ and $\theta_n=\bar\theta_n$.

Proof. Let us first show that the assumptions of Proposition 5.5 are satisfied. The concavity of H follows from Lemma 10 in (Genin and Tankov, 2020), see the examples in section 4 of that paper; the payoff is clearly bounded from above and continuous on the set where it is finite.

Let us now turn to the computation of $\widehat H$. For this payoff,

$$H(x)-\sum_{i=1}^n\bar\theta_i(x_i-x_{i-1})=\log\left(K-\frac{S_0}{n}\sum_{j=1}^n e^{x_j}\right)_+-\sum_{j=1}^n\theta_jx_j\,.$$

When the option is out or at the money, the log-payoff is $-\infty$. Assume that x is such that $H(x) > -\infty$ and differentiate with respect to x_j. We obtain

$$0 = \partial_{x_j} \left\{ \log\left(K - \frac{S_0}{n} \sum_{l=1}^n e^{x_l} \right) - \sum_{l=1}^n x_l \theta_l \right\} = \frac{-\frac{S_0}{n} e^{x_j}}{K - \frac{S_0}{n} \sum_{l=1}^n e^{x_l}} - \theta_j \, .$$

Therefore the x that maximises $H(x) - \sum_{j=1}^n \theta_j x_j$ satisfies

$$\frac{e^{x_j}}{\theta_j} = -n \frac{K}{S_0} + \sum_{l=1}^n e^{x_l} = -n \frac{K}{S_0} + \frac{e^{x_j}}{\theta_j} \sum_{l=1}^n \theta_l \, ,$$

for every j. Therefore

$$x_j = \log\left(\frac{-\theta_j \, n \, K/S_0}{1 - \sum_{l=1}^n \theta_l} \right) \, .$$

Inserting x_j in the value of $H(x) - \sum_{j=1}^n \theta_j x_j$, we obtain the result. □

6.1. **European and Asian put options in the Heston model.** Consider the Heston model (Heston, 1993)

$$\begin{aligned} dX_t &= -\frac{V_t}{2} \, dt + \sqrt{V_t} \, dW_t^1 \, , & X_0 &= 0 \\ dV_t &= \lambda(\mu - V_t) \, dt + \zeta \sqrt{V_t} \, dW_t^2 \, , & V_0 &> 0 \\ d\langle W^1, W^2 \rangle_t &= \rho \, dt \, , \end{aligned} \tag{6.3}$$

eq:HestonDynamics

where W^1, W^2 are standard $\mathbb{P}$-Brownian motions. The Laplace transform of (X_t, V_t) is

$$\mathbb{E}\left(e^{uX_t + wV_t} \right) = e^{\phi(t,u,w) + \psi(t,u,w) V_0 + uX_0} \, ,$$

where ϕ, ψ satisfy the Riccati equations

$$\begin{aligned} \partial_t \phi(t,u,w) &= F(u, \psi(t,u,w)) & \phi(0,u,w) &= 0 \\ \partial_t \psi(t,u,w) &= R(u, \psi(t,u,w)) & \psi(0,u,w) &= w \end{aligned} \tag{6.4}$$

eq:RiccatiEquatio

for $F(u,w) = \lambda \mu w$ and

$$R(u,w) = \frac{\zeta^2}{2} w^2 + \zeta \rho \, uw - \lambda w + \frac{1}{2}(u^2 - u) \, .$$

A standard calculation shows that the solution of the Riccati equations (6.4) is

$$\begin{aligned} \psi(t,u,w) &= \frac{1}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u \right) - \frac{\gamma}{\zeta^2} \frac{\tanh\left(\frac{\gamma}{2} t\right) + \eta}{1 + \eta \tanh\left(\frac{\gamma}{2} t\right)} \\ \phi(t,u,w) &= \mu \frac{\lambda}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u \right) t - 2\mu \frac{\lambda}{\zeta^2} \log\left(\cosh\left(\frac{\gamma}{2} t\right) + \eta \sinh\left(\frac{\gamma}{2} t\right) \right) \, , \end{aligned} \tag{6.5}$$

eq:RiccatiSolutio

where $\gamma = \gamma(u) = \zeta \sqrt{\left(\frac{\lambda}{\zeta} - \rho u\right)^2 + \frac{1}{4} - \left(u - \frac{1}{2}\right)^2}$ and $\eta = \eta(u,w) = \frac{\lambda - \zeta\rho u - \zeta^2 w}{\gamma(u)}$. Furthermore, for the Heston model, the function h (asymptotic Laplace exponent of X, see Equation (2.4)) is given by

$$h(u) = \begin{cases} \mu \frac{\lambda}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u \right) - \mu \frac{\lambda}{\zeta^2} \gamma(u), & \text{if} \quad \left(\frac{\lambda}{\zeta} - \rho u \right)^2 + \frac{1}{4} - \left(u - \frac{1}{2} \right)^2 > 0 \\ +\infty, & \text{otherwise.} \end{cases} \tag{6.6}$$

eq:HestonH

Remark 6.2. The function h is the log-Laplace transform of the normal inverse Gaussian process (Barndorff-Nielsen, 1997) which is complex-analytic on a strip around the real axis.

The following proposition describes the effect of the time dependent Esscher transform on the dynamics of the Heston model.

tonDynamicsPTheta

Proposition 6.3. *Let $\tau = \{t_1, ..., t_n\}$ and $\mathbb{P}_{\bar\theta}$ the measure given by*

$$\frac{d\mathbb{P}_{\bar\theta}}{d\mathbb{P}} = \frac{e^{\sum_{j=1}^{n} \bar\theta_j \left(X_{t_j} - X_{t_{j-1}}\right)}}{\mathbb{E}\left[e^{\sum_{j=1}^{n} \bar\theta_j \left(X_{t_j} - X_{t_{j-1}}\right)}\right]} .$$

Under $\mathbb{P}_{\bar\theta}$, the dynamics of the $\mathbb{P}$-Heston process (X_t, V_t) becomes

$$\begin{aligned} dX_t &= \left(\bar\theta_{n_t} + \zeta\rho\,\Psi\left(\tau_t - t, \bar\theta_{n_t}, ..., \bar\theta_n\right) - \frac{1}{2}\right) V_t\,dt + \sqrt{V_t}\,d\tilde W_t^1\,, \quad X_0 = 0 \\ dV_t &= \tilde\lambda_t\left(\tilde\mu_t - V_t\right)dt + \zeta\sqrt{V_t}\,d\tilde W_t^2\,, \qquad V_0 = V_0 \\ d\left\langle \tilde W^1, \tilde W^2\right\rangle_t &= \rho\,dt\,, \end{aligned} \tag{6.7}$$

eq:HestonDynamics

where $\tilde W$ is 2-dimensional correlated $\mathbb{P}_{\bar\theta}$-Brownian motion, $n_t = \inf\{k \in \mathbb{N} : t_k \geq t\}$, $\tau_t = \inf\{s \in \tau\,:\, s \geq t\}$, Ψ is defined iteratively as

$$\begin{aligned} \Psi\left(s, \bar\theta_j, ..., \bar\theta_n\right) &= \psi\left(s, \bar\theta_j, \Psi\left(t_{j+1} - t_j, \bar\theta_{j+1}..., \bar\theta_n\right)\right) \\ \Psi\left(s\right) &= 0 \end{aligned}$$

and

$$\tilde\lambda_t = \lambda - \zeta\bar\theta_{n_t}\rho - \zeta^2\,\Psi\left(\tau_t - t, \bar\theta_{n_t}, ..., \bar\theta_n\right) \quad \textit{and} \quad \tilde\mu_t = \frac{\lambda\mu}{\tilde\lambda_t}\,.$$

Proof. We define the function Φ iteratively by

$$\begin{aligned} \Phi\left(s, \bar\theta_j, ..., \bar\theta_n\right) &= \phi\left(s, \bar\theta_j, \Psi\left(t_{j+1} - t_j, \bar\theta_{j+1}, ..., \bar\theta_n\right)\right) \\ &\quad + \Phi\left(t_{j+1} - t_j, \bar\theta_{j+1}, ..., \bar\theta_n\right) \\ \Phi\left(s\right) &= 0 \end{aligned}$$

Denote

$$D(t, X_t, V_t) = \left.\frac{d\mathbb{P}_{\bar\theta}}{d\mathbb{P}}\right|_{\mathcal{F}_t}\,.$$

Then

$$\begin{aligned} D(t, X_t, V_t) &= \frac{e^{\sum_{j=1}^{n_t - 1} \bar\theta_j \left(X_{t_j} - X_{t_{j-1}}\right)}}{\mathbb{E}\left[e^{\sum_{j=1}^{n} \bar\theta_j \left(X_{t_j} - X_{t_{j-1}}\right)}\right]}\,\mathbb{E}\left[e^{\sum_{j=n_t}^{n} \bar\theta_j \left(X_{t_j} - X_{t_{j-1}}\right)}\,\middle|\,\mathcal{F}_t\right] \\ &= \frac{e^{\sum_{j=1}^{n_t - 1} \bar\theta_j \left(X_{t_j} - X_{t_{j-1}}\right) + \Phi\left(\tau_t - t, \bar\theta_{n_t}, ..., \bar\theta_n\right)}}{e^{\Phi\left(t_1, \bar\theta_1, ..., \bar\theta_n\right) + \Psi\left(t_1, \bar\theta_1, ..., \bar\theta_n\right) V_0 + \bar\theta_1 X_0}}\,e^{\Psi\left(\tau_t - t, \bar\theta_{n_t}, ..., \bar\theta_n\right) V_t + \bar\theta_{n_t} X_t}\,. \end{aligned}$$

The dynamics of $D(t, X_t, V_t)$ can then be expressed using Itō's Lemma as

$$\begin{aligned} dD(t, X_t, V_t) &= D(t, X_t, V_t)\left(\bar\theta_{n_t} dX_t + \Psi\left(\tau_t - t, \bar\theta_{n_t}, ..., \bar\theta_n\right) dV_t\right) + ...\,dt \\ &= D(t, X_t, V_t)\sqrt{V_t}\left(\bar\theta_{n_t} dW_t^1 + \zeta\,\Psi\left(\tau_t - t, \bar\theta_{n_t}, ..., \bar\theta_n\right) dW_t^2\right)\,. \end{aligned}$$

By Girsanov's theorem,

$$d\begin{pmatrix}\tilde{W}_t^1\\ \tilde{W}_t^2\end{pmatrix} = d\begin{pmatrix}W_t^1\\ W_t^2\end{pmatrix} - \sqrt{V_t}\begin{pmatrix}\bar{\theta}_{n_t} + \zeta\rho\,\Psi\left(\tau_t - t, \bar{\theta}_{n_t}, ..., \bar{\theta}_n\right)\\ \bar{\theta}_{n_t}\rho + \zeta\,\Psi\left(\tau_t - t, \bar{\theta}_{n_t}, ..., \bar{\theta}_n\right)\end{pmatrix} dt$$

is a 2-dimensional Brownian motion under the measure $\mathbb{P}_\theta$. Replacing W in eq. (6.3) by $\tilde{W}$ gives the result. □

Remark 6.4. Prop. 6.3 shows that the time-dependent Esscher transform changes a classical Heston process into a Heston process with time-inhomogeneous drift.

hestonsatisfies

Remark 6.5 (Asymptotic optimality for the Heston model)**.** The limiting Laplace exponent of the Heston model is finite and continuous on the bounded interval, $J = [u_-, u_+]$, where

$$u_\pm = \frac{\left(\frac{1}{2} - \frac{\lambda}{\zeta}\rho\right) \pm \sqrt{\left(\frac{1}{2} - \frac{\lambda}{\zeta}\rho\right)^2 + \frac{\lambda^2}{\zeta^2}(1-\rho^2)}}{(1-\rho^2)},$$

which contains a neighborhood of zero. The function w (stable equilibrium of the second Riccati equation) is given by

$$w(u) = \frac{(\lambda - u\rho\zeta) - \sqrt{(\lambda - u\rho\zeta)^2 - \zeta^2(u^2 - u)}}{\zeta^2}, \tag{6.8}$$

wheston.eq

so that

$$w(u_-) = \frac{1}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u_-\right) \qquad \text{and} \qquad w(u_+) = \frac{1}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u_+\right).$$

and we see that Assumption 3 is verified and thus asymptotic optimality for Asian and European put options is guaranteed in the Heston model only when $\rho = 0$. Nevertheless, since the actual variance reduction problem is itself unsolvable, our goal is to find a good candidate measure that we can test numerically. The fact that we do not have the full theory to justify it is therefore not problematic.

6.1.1. *Numerical results for European put options.* In this case, the asymptotically optimal variance reduction measure of Proposition 5.5 is supported by the single point $\{T\}$ and given (with $\epsilon = 1$) by

$$\frac{d\mathbb{P}_{\bar{\theta}^*}}{d\mathbb{P}} = \frac{e^{\bar{\theta}^* X_T}}{\mathbb{E}\left[e^{\bar{\theta}^* X_T}\right]}.$$

For $\bar{\theta} \in \mathbb{R}$,

$$\begin{aligned}&\widehat{H}(\bar{\theta}) + Th(\bar{\theta})\\ &= \log\left(\frac{K}{1-\bar{\theta}}\right) - \bar{\theta}\log\left(\frac{-\bar{\theta}\,K/S_0}{1-\theta}\right) + T\,\mu\,\frac{\lambda}{\zeta}\left(\frac{\lambda}{\zeta} - \rho\,\bar{\theta} - \frac{\gamma(\bar{\theta})}{\zeta}\right).\end{aligned} \tag{6.9}$$

eq:HHatHestonEuro

In order to obtain $\bar{\theta}^*$, we therefore differentiate (6.9) with respect to $\bar{\theta}$ and solve the resulting equation numerically.

We simulate $N = 10000$ trajectories of the Heston model with parameters $\lambda = 1.15$, $\mu = 0.04$, $\zeta = 0.2$, $\rho = -0.4$ and initial values $V_0 = 0.04$ and $S_0 = 1$, under both $\mathbb{P}$, eq. (6.3), and $\mathbb{P}_{\bar{\theta}^*}$, eq. (6.7), with $n = 1$ and $t_1 = T$, using a standard Euler scheme with 200 discretization steps. For the $\mathbb{P}$-realisations $X^{(i)}$, we calculate the

European put price as $\frac{1}{N}\sum_{j=1}^{N}\left(K - S_0\, e^{X_T^{(i)}}\right)_+$ and for the $\mathbb{P}_{\bar\theta^*}$-realisations $X^{(i,\bar\theta^*)}$, as

$$\frac{e^{\phi\left(T,\bar\theta^*,0\right)+\psi\left(T,\bar\theta^*,0\right)V_0}}{N}\sum_{j=1}^{N} e^{-\theta\, X_T^{(i,\bar\theta^*)}}\left(K - S_0 e^{X_T^{(i,\bar\theta^*)}}\right)_+ . \tag{6.10}$$

eq:EstimatorTheta

Each time, we compute the $\mathbb{P}_{\bar\theta^*}$-standard deviation, the variance ratio and the adjusted variance ratio, i.e. the variance ratio divided by the ratio of simulation time. The latter measures the actual efficiency of the method, given the fact that simulating under the measure change takes in general slightly more time.

In Table 1, we fix the strike to the value $K = 1$ and let the maturity T vary from 0.25 to 3, whereas in Tables 2 and 3, we fix maturity to $T = 1$ and to $T = 3$, while we let the strike K vary between 0.25 and 1.75. We calculate each time the price, the standard error, the variance ratio adjusted and not adjusted by the ratio of simulation times.

T	Price	Std. error	Var. ratio	Adj. ratio	Time, s
0.25	0.0395	$3.72\cdot10^{-4}$	2.46	2.14	20.2
0.5	0.0550	$4.54\cdot10^{-4}$	3.12	2.83	19.9
1	0.0780	$5.59\cdot10^{-4}$	3.92	3.66	19.5
2	0.111	$7.20\cdot10^{-4}$	4.21	3.89	19.7
3	0.134	$8.48\cdot10^{-4}$	4.19	3.79	19.8

TABLE 1. The variance ratio as function of the maturity for at-the-money European put options.

tab:PutVarianceRe

K	Price	Std. error	Var. ratio	Adj. ratio	Time, s
0.5	0.00014	$7.65\cdot10^{-6}$	26.6	24.5	18.4
0.75	0.00794	$1.34\cdot10^{-4}$	6.53	5.91	18.7
1	0.0773	$5.60\cdot10^{-4}$	3.96	3.65	18.5
1.25	0.261	$8.62\cdot10^{-4}$	4.20	3.78	18.9
1.5	0.502	$7.92\cdot10^{-4}$	5.84	5.36	18.6
1.75	0.749	$6.84\cdot10^{-4}$	8.45	7.29	19.7

TABLE 2. The variance ratio as function of the strike for the European put option with maturity $T = 1$.

tab:PutVarianceRe

K	Price	Std. error	Var. ratio	Adj. ratio	Time, s
0.25	$7.1\cdot10^{-5}$	$1.84\cdot10^{-5}$	92.0	70.9	23.1
0.5	0.00418	$6.05\cdot10^{-5}$	16.1	16.0	20.0
0.75	0.0369	$3.43\cdot10^{-4}$	6.67	6.00	20.4
1	0.133	$8.51\cdot10^{-4}$	4.24	4.15	20.2
1.25	0.300	$1.34\cdot10^{-3}$	3.61	3.13	21.3
1.5	0.517	$1.60\cdot10^{-3}$	3.47	3.30	19.9
1.75	0.755	$1.64\cdot10^{-3}$	3.89	3.53	19.9

TABLE 3. The variance ratio as function of the strike for the European put option with maturity $T = 3$.

tab:PutVarianceRe

In all the cases, we can see that the variance ratio is quite significant for deep out of the money options and less significant, yet still very interesting, when the option is at or in the money. This corresponds to the natural behaviour of variance reduction techniques that involve measure changes, as the measure change increases the exercise probability. Note that the simulation time is only slightly larger when simulating with the measure change, while the time required for the optimization procedure is negligible compared with the simulation time. In Figure 6.1, we fix the maturity to $T = 1.5$ and plot the empirical variance of the estimator (6.10) as a function of $\bar{\theta}^*$. Our method provides $\bar{\theta}^* = -0.457$ as the candidate asymptotically optimal measure change. We can therefore see that our candidate $\bar{\theta}^*$ is very close to the optimal one.

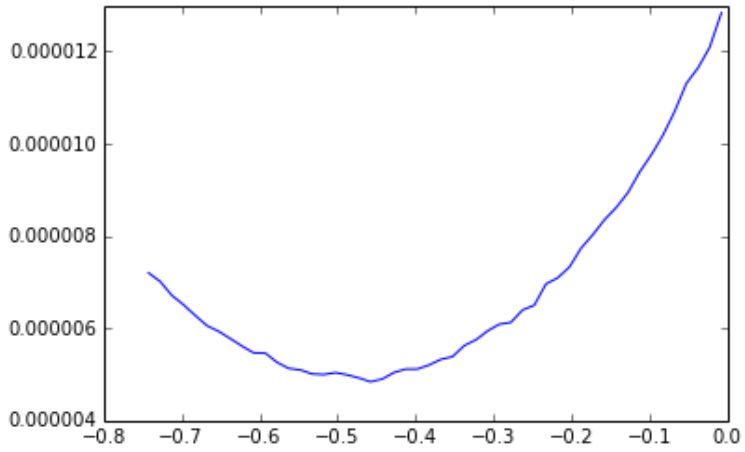

FIGURE 6.1. The variance of the Monte-Carlo estimator as a function of θ.

fig:VarianceHesto

6.1.2. *Numerical results for Asian put options.* We now consider the case of a (discretized) Asian put option with log payoff (6.1) and discretization dates $t_j = \frac{j}{n}T$. By Prop. 5.5, the asymptotically optimal variance reduction measure (with $\epsilon = 1$) is given by

$$\frac{d\mathbb{P}_{\bar{\theta}^*}}{d\mathbb{P}} = \frac{e^{\sum_{j=1}^n \bar{\theta}_j^* (X_{t_j} - X_{t_{j-1}})}}{\mathbb{E}\left[e^{\sum_{j=1}^n \bar{\theta}_j^* (X_{t_j} - X_{t_{j-1}})}\right]},$$

where $\bar{\theta}^*$ is computed by minimizing

$$\log\left(\frac{K}{1-\bar{\theta}_1}\right) - \sum_{m=1}^n (\bar{\theta}_m - \bar{\theta}_{m+1}) \log\left(\frac{-(\bar{\theta}_m - \bar{\theta}_{m+1})\, nK/S_0}{1-\bar{\theta}_1}\right) + \frac{T}{n}\sum_{j=1}^n h\left(\bar{\theta}_j\right) ,$$

see Prop. 6.1 and eq. (6.6). By differentiating with respect to $\bar{\theta}_j$, we obtain, for $j = 2, ..., n$,

$$\begin{aligned} 0 &= \partial_{\bar{\theta}_j}\left\{\widehat{H}(\bar{\theta}) + \frac{T}{n}\sum_{m=1}^n h\left(\bar{\theta}_m\right)\right\} \\ &= \frac{T\, h'\left(\bar{\theta}_j\right)}{n} - \log\left[-(\bar{\theta}_j - \bar{\theta}_{j+1})\right] + \log\left[-(\bar{\theta}_{j-1} - \bar{\theta}_j)\right] , \end{aligned} \tag{6.11}$$

eq:MinimizationEq

while, for $j = 1$, we have

$$\begin{aligned} 0 &= \partial_{\bar\theta_1}\left\{\widehat{H}(\bar\theta) + \frac{T}{n}\sum_{m=1}^{n} h\left(\bar\theta_m\right)\right\} \\ &= \log\left(1-\bar\theta_1\right) - \log(n\,K/S_0) + \frac{T}{n}\,h'\left(\bar\theta_1\right) - \log\left[-(\bar\theta_1-\bar\theta_2)\right]\,. \end{aligned} \tag{6.12}$$

eq:MinimizationEq

Taking the exponential in eqs. (6.11) and (6.12), we obtain

$$\begin{aligned} \bar\theta_2 - \bar\theta_1 &= (1 \quad - \bar\theta_1 \quad)\; e^{\frac{T}{n}h'(\bar\theta_1)}\cdot\frac{S_0}{n\,K} \\ \bar\theta_3 - \bar\theta_2 &= (\bar\theta_2 \quad - \bar\theta_1 \quad)\; e^{\frac{T}{n}h'(\bar\theta_2)} \\ \vdots \quad &= \qquad\quad \vdots \\ \bar\theta_n - \bar\theta_{n-1} &= (\bar\theta_{n-1} - \bar\theta_{n-2})\; e^{\frac{T}{n}h'(\bar\theta_{n-1})} \\ -\bar\theta_n &= (\bar\theta_n \quad - \bar\theta_{n-1})\; e^{\frac{T}{n}h'(\bar\theta_n)}\,. \end{aligned}$$

Finally, let $\mathcal{T}$ be the real-valued function defined by

$$\bar\theta_n \mapsto \mathcal{T}(\bar\theta_n) = (1-\bar\theta_1)\mathrm{e}^{\frac{T}{n}h'(\bar\theta_1)}\cdot\frac{S_0}{n\,K} - \bar\theta_2 - \bar\theta_1\,,$$

where $\bar\theta_{n-1} = \bar\theta_n + \bar\theta_n\, e^{-\frac{T}{n}h'(\bar\theta_n)}$ and iteratively,

$$\bar\theta_{j-2} = \bar\theta_{j-1} - (\bar\theta_j - \bar\theta_{j-1})\, e^{-\frac{T}{n}h'(\bar\theta_{j-1})}\,, \qquad j = n, ..., 3\,.$$

Solving numerically the equation $\mathcal{T}(\bar\theta_n) = 0$, we find the optimal $\bar\theta^*$.

As before, we simulate $N = 10000$ trajectories of the Heston model with parameters $\lambda = 1.15$, $\mu = 0.04$, $\zeta = 0.2$, $\rho = -0.4$ and initial values $V_0 = 0.04$ and $S_0 = 1$, under both $\mathbb{P}$, eq. (6.3), and $\mathbb{P}_{\bar\theta^*}$, eq. (6.7), with $n = 200$ and $t_j = \frac{j}{n}T$, using a standard Euler scheme with 200 discretization steps. For the $\mathbb{P}$-realisations $X^{(i)}$, we calculate the Asian put price as

$$\frac{1}{N}\sum_{j=1}^{N}\left(K - \frac{S_0}{n}\sum_{j=1}^{n} e^{X^{(i)}_{t_j}}\right)_+ \tag{6.13}$$

and for the $\mathbb{P}_{\bar\theta^*}$-realisations $X^{(i,\bar\theta^*)}$, as

$$\frac{e^{\Phi\left(t_1,\bar\theta^*{}_1,...,\bar\theta^*{}_n\right)+\Psi\left(t_1,\bar\theta^*{}_1,...,\bar\theta^*{}_n\right)V_0}}{N}\sum_{j=1}^{N} e^{-\sum_{j=1}^{n}\bar\theta^*_j\,(X^{(i,\theta^*)}_{t_j} - X^{(i,\theta^*)}_{t_{j-1}})}\left(K - \frac{S_0}{n}\sum_{j=1}^{n} e^{X^{(i,\bar\theta^*)}_{t_j}}\right)_+\,. \tag{6.14}$$

Each time, we compute the $\mathbb{P}_{\bar\theta^*}$-standard deviation and the adjusted and non-adjusted variance ratios. In Table 4, we fix the maturity to $T = 1.5$ and let the strike K vary between 0.6 and 1.3.

K	Price	Std. error	Var. ratio	Adj. ratio	Time, s
0.6	$3.466 \cdot 10^{-5}$	$4.13 \cdot 10^{-6}$	16.9	14.6	19.9
0.7	0.000562	$2.60 \cdot 10^{-5}$	5.77	4.77	21.1
0.8	0.00414	$9.64 \cdot 10^{-5}$	4.36	3.77	20.1
0.9	0.0185	0.00024	3.48	3.09	20.6
1	0.0558	0.00043	3.49	3.07	20.1
1.1	0.120	0.00057	3.69	3.20	20.1
1.2	0.206	0.00062	4.27	3.80	19.7
1.3	0.301	0.00059	5.30	4.41	21.0

TABLE 4. The variance ratio as function of the strike for the Asian put option. $\lambda = 1.15$, $\mu = 0.04$, $\zeta = 0.2$, $\rho = -0.4$, $S_0 = 1$, $V_0 = 0.04$, $T = 1.5$, $N = 10000$, 200 discretization steps. tab:AsianPutVaria

The conclusion is the same as for the European put option. Indeed, the variance ratio explodes when the option moves away from the money. Due to the time-dependence of the measure change, the adjusted variance ratio is consistently around 13% below its non-adjusted version. The adjusted variance ratio remains however very interesting, with values above 3 around the money.

6.2. **European put options in the Heston model with negative exponential jumps.** We now consider the Heston model with negative exponential jumps

$$\begin{aligned} dX_t &= \left(\delta - \frac{V_t}{2}\right) dt + \sqrt{V_t}\, dW_t^1 + dJ_t\,, \quad X_0 = 0 \\ dV_t &= \lambda(\mu - V_t)\, dt + \zeta\sqrt{V_t}\, dW_t^2\,, \qquad V_0 = V_0 \\ d\langle W^1, W^2\rangle_t &= \rho\, dt\,, \end{aligned} \tag{6.15}$$

eq:HestonWithJump

where W^1, W^2 are standard $\mathbb{P}$-Brownian motions and $(J_t)_{t\geq 0}$ is an independent compound Poisson process with constant jump rate r and jump distribution Neg-Exp(α), i.e. the Lévy measure of $(J_t)_{t\geq 0}$ is $\nu(dx) = r\,\alpha e^{\alpha x}\mathbb{1}_{\{x<0\}}dx$. The martingale condition on $S = S_0\, e^X$ imposes $\delta = \frac{r}{\alpha+1}$. The Laplace transform of (X_t, V_t) is

$$\mathbb{E}\left(e^{uX_t + wV_t}\right) = e^{\phi(t,u,w) + \psi(t,u,w)V_0 + uX_0}\,,$$

where ϕ, ψ satisfy the Riccati equations

$$\begin{aligned} \partial_t\phi(t,u,w) &= F(u,\psi(t,u,w)) \qquad \phi(0,u,w) = 0 \\ \partial_t\psi(t,u,w) &= R(u,\psi(t,u,w)) \qquad \psi(0,u,w) = w \end{aligned} \tag{6.16}$$

eq:JumpRiccatiEqu

for $F(u,w) = \lambda\mu\, w + \tilde{\kappa}(u)$, where $\tilde{\kappa}(u) = \frac{ru(u-1)}{(\alpha+1)(\alpha+u)}$, and

$$R(u,w) = \frac{\zeta^2}{2}\, w^2 + \zeta\rho\, uw - \lambda w + \frac{1}{2}(u^2 - u)\,.$$

Again, a standard calculation shows that the solution of the generalized Riccati equations (6.16) is

$$\begin{aligned} \psi(t,u,w) &= \frac{1}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u\right) - \frac{\gamma}{\zeta^2}\,\frac{\tanh\left(\frac{\gamma}{2}\,t\right) + \eta}{1 + \eta\,\tanh\left(\frac{\gamma}{2}\,t\right)} \\ \phi(t,u,w) &= \mu\,\frac{\lambda}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u\right) t - 2\mu\frac{\lambda}{\zeta^2}\,\log\Big(\cosh\left(\frac{\gamma}{2}t\right) + \eta\,\sinh\left(\frac{\gamma}{2}t\right)\Big) + t\tilde{\kappa}(u)\,, \end{aligned} \tag{6.17}$$

eq:GRiccatiSoluti

where $\gamma = \gamma(u) = \zeta\sqrt{\left(\frac{\lambda}{\zeta} - \rho u\right)^2 + \frac{1}{4} - \left(u - \frac{1}{2}\right)^2}$ and $\eta = \eta(u,w) = \frac{\lambda - \zeta\rho u - \zeta^2 w}{\gamma(u)}$. Furthermore, for the Heston model with negative jumps, the function h is given by

$$h(u) = \mu\,\frac{\lambda}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u\right) - \mu\,\frac{\lambda}{\zeta^2}\,\gamma(u) + \tilde{\kappa}(u)\,. \tag{6.18}$$

eq:HestonWithJump

hestonsatisfies

Remark 6.6 (Asymptotic optimality for the Heston model with jumps)**.** The limiting Laplace exponent of the Heston model is finite and continuous on the bounded interval, $J = [u_-, u_+]$, where

$$u_\pm = \frac{\left(\frac{1}{2} - \frac{\lambda}{\zeta}\,\rho\right) \pm \sqrt{\left(\frac{1}{2} - \frac{\lambda}{\zeta}\,\rho\right)^2 + \frac{\lambda^2}{\zeta^2}\,(1-\rho^2)}}{(1-\rho^2)} \vee (-\alpha)\,,$$

which contains a neighborhood of zero. The function w has the same form (6.8) as in the Heston model without jumps, so that

$$w(u_-) = \frac{1}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u_-\right)\mathbf{1}_{u_- > -\alpha} + w(-\alpha)\mathbf{1}_{u_- = -\alpha} \qquad \text{and} \qquad w(u_+) = \frac{1}{\zeta}\left(\frac{\lambda}{\zeta} - \rho u_+\right).$$

Thus, Assumption 3 is verified and asymptotic optimality holds for Asian and European put options when $\rho = 0$ and jumps are sufficiently small, namely

$$\alpha > \sqrt{\frac{1}{4} + \frac{\lambda^2}{\zeta^2}} - \frac{1}{2}.$$

Let us now study the effect of the Esscher transform on the dynamics of the Heston model with jumps.

mpsDynamicsPTheta

Proposition 6.7. *Let $\mathbb{P}_{\bar{\theta}}$ be the measure given by*

$$\frac{d\mathbb{P}_{\bar{\theta}}}{d\mathbb{P}} = \frac{e^{\bar{\theta}\,X_T}}{\mathbb{E}\left[e^{\bar{\theta}\,X_T}\right]}\,.$$

Under $\mathbb{P}_{\bar{\theta}}$, the dynamics of the $\mathbb{P}$-Heston process with jumps (X_t, V_t) becomes

$$\begin{aligned}
dX_t &= \delta dt + \left(\bar{\theta} + \zeta\rho\,\psi\left(T-t,\bar{\theta},0\right) - \frac{1}{2}\right)V_t\,dt + \sqrt{V_t}\,d\tilde{W}^1_t + dJ_t\,, \quad X_0 = 0\\
dV_t &= \tilde{\lambda}_t\left(\tilde{\mu}_t - V_t\right)dt + \zeta\sqrt{V_t}\,d\tilde{W}^2_t\,, \qquad V_0 = V_0\\
d\left\langle \tilde{W}^1, \tilde{W}^2\right\rangle_t &= \rho\,dt\,,
\end{aligned} \tag{6.19}$$

eq:HestonWithJump

where $\tilde{W}$ is 2-dimensional correlated $\mathbb{P}_{\bar{\theta}}$-Brownian motion, ϕ and ψ are given in (6.17),

$$\tilde{\lambda}_t = \lambda - \zeta\bar{\theta}\rho - \zeta^2\,\psi\left(T-t,\bar{\theta},0\right) \quad \text{and} \quad \tilde{\mu}_t = \frac{\lambda\mu}{\tilde{\lambda}_t}$$

and $(J_t)_{t\geq 0}$ is a compound Poisson process with jump rate $\frac{r\alpha}{\alpha+\bar{\theta}}$ and jump distribution Neg-Exp$(\alpha+\bar{\theta})$ under $\mathbb{P}_{\bar{\theta}}$.

Proof. Denote

$$D(t,X_t,V_t) = \left.\frac{d\mathbb{P}_{\bar{\theta}}}{d\mathbb{P}}\right|_{\mathcal{F}_t} = \frac{e^{\phi\left(T-t,\bar{\theta},0\right)}}{e^{\phi\left(T,\bar{\theta},0\right)+\psi\left(T,\bar{\theta},0\right)V_0}}\,e^{\psi\left(T-t,\bar{\theta},0\right)V_t + \bar{\theta}\,X_t}\,.$$

The dynamics of $D(t, X_t, V_t)$ can then be expressed using Itō's Lemma as

$$\begin{aligned} dD(t, X_t, V_t) &= D(t, X_t, V_t)\left(\bar\theta dX_t + \psi\left(T-t, \bar\theta, 0\right) dV_t\right) + \ldots\, dt \\ &= D(t, X_t, V_t)\Big[\sqrt{V_t}\left(\bar\theta dW_t^1 + \zeta\,\psi\left(T-t, \bar\theta, 0\right) dW_t^2\right) + \bar\theta\left(\delta dt + dJ_t\right)\Big] \end{aligned}$$

and Girsanov's theorem then shows that

$$d\begin{pmatrix} \tilde W_t^1 \\ \tilde W_t^2 \end{pmatrix} = d\begin{pmatrix} W_t^1 \\ W_t^2 \end{pmatrix} - \sqrt{V_t}\begin{pmatrix} \bar\theta + \zeta\rho\,\psi\left(T-t, \bar\theta, 0\right) \\ \bar\theta\rho + \zeta\,\psi\left(T-t, \bar\theta, 0\right) \end{pmatrix} dt$$

is a 2-dimensional Brownian motion under the measure $P_{\bar\theta}$. Replacing W in eq. (6.3) by $\tilde W$ gives eq. (6.19). In order to finish the proof, it remains to show that the jump process $(J_t)_{t\geq 0}$ has the desired distribution under $\mathbb{P}_{\bar\theta}$. Let us calculate the $\mathbb{P}_{\bar\theta}$-Laplace transform of J_t:

$$\begin{aligned} \mathbb{E}^{\mathbb{P}_{\bar\theta}}\left[e^{uJ_t}\right] &= \frac{\mathbb{E}\left[e^{uJ_t}\,\mathbb{E}\left[e^{\bar\theta X_T}\,\middle|\,\mathcal{F}_t\right]\right]}{\mathbb{E}\left[e^{\bar\theta X_T}\right]} \\ &= \frac{e^{\phi\left(T-t,\bar\theta,0\right)}}{\mathbb{E}\left[e^{\bar\theta X_T}\right]}\mathbb{E}\left[e^{uJ_t + \psi\left(T-t,\bar\theta,0\right)V_t + \bar\theta X_t}\right]. \end{aligned}$$

By independence of the jumps,

$$\mathbb{E}\left[e^{uJ_t + \psi\left(T-t,\bar\theta,0\right)V_t + \bar\theta X_t}\right] = e^{\bar\theta\delta\, t}\,\mathbb{E}\left[e^{(u+\bar\theta)J_t}\right]\,\mathbb{E}\left[e^{\psi\left(T-t,\bar\theta,0\right)V_t + \bar\theta(X_t - \delta\, t - J_t)}\right],$$

where $\mathbb{E}\left[e^{(u+\bar\theta)J_t}\right] = e^{-rt\frac{u+\bar\theta}{u+\bar\theta+\alpha}}$. Furthermore, $(X_t - \delta\, t - J_t,\, V_t)_{t\geq 0}$ is a standard Heston process without jumps. Therefore comparing (6.5) and (6.17), we find that

$$\mathbb{E}\left[e^{\psi\left(T-t,\bar\theta,0\right)V_t + \bar\theta(X_t - \delta\, t - J_t)}\right] = e^{\phi\left(t,\bar\theta,\psi\left(T-t,\bar\theta,0\right)\right) - t\frac{r\bar\theta(\bar\theta-1)}{(\alpha+1)(\alpha+\theta)} + \psi\left(t,\bar\theta,\psi\left(T-t,\bar\theta,0\right)\right)V_0}.$$

Using the fact that $\psi(t, \bar\theta, \psi\left(T-t, \bar\theta, 0\right)) = \psi\left(T, \bar\theta, 0\right)$ and

$$\phi\left(T-t, \bar\theta, 0\right) + \phi(t, \bar\theta, \psi\left(T-t, \bar\theta, 0\right)) = \phi\left(T, \bar\theta, 0\right)$$

(see eq. (2.1) in (Keller-Ressel, 2011)), we finally obtain

$$\begin{aligned} \mathbb{E}^{\mathbb{P}_{\bar\theta}}\left[e^{uJ_t}\right] &= e^{\bar\theta\delta\, t - rt\frac{u+\bar\theta}{u+\bar\theta+\alpha} - t\frac{r\bar\theta(\bar\theta-1)}{(\alpha+1)(\alpha+\bar\theta)}} \\ &= e^{\bar\theta\frac{r}{\alpha+1}\,t - rt\frac{u+\bar\theta}{u+\bar\theta+\alpha} - t\frac{r\bar\theta(\bar\theta-1)}{(\alpha+1)(\alpha+\bar\theta)}} = e^{-\frac{r\alpha}{\alpha+\bar\theta}\,t\,\frac{u}{u+(\alpha+\bar\theta)}}, \end{aligned}$$

which is indeed the Laplace transform of a compound Poisson process with jump rate $\frac{r\alpha}{\alpha+\bar\theta}$ and Neg-Exp$(\alpha+\bar\theta)$-distributed jumps. □

6.2.1. *Numerical results for the European put option.* Similarly to the case of the Heston model without jumps, we find the optimal $\bar\theta^*$ by minimizing numerically

$$\begin{aligned} &\widehat{H}(\bar\theta) + Th(\bar\theta)\, dt \\ &= \log\left(\frac{K}{1-\bar\theta}\right) - \bar\theta\log\left(\frac{-\bar\theta\, K/S_0}{1-\bar\theta}\right) + T\,\mu\,\frac{\lambda}{\zeta}\left(\frac{\lambda}{\zeta} - \rho\bar\theta - \frac{\gamma(\bar\theta)}{\zeta}\right) + T\,\tilde\kappa(\bar\theta). \end{aligned} \tag{6.20}$$

eq:HHatHestonWith

We simulate $N = 10000$ trajectories of the Heston model with jumps with parameters $\lambda = 1.1$, $\mu = 0.7$, $\zeta = 0.3$, $\rho = -0.5$, $r = 2$, $\alpha = 3$ and initial values $V_0 = 1.3$ and $S_0 = 1$, under both $\mathbb{P}$, eq. (6.15), and $\mathbb{P}_{\bar\theta^*}$, eq. (6.19), using a standard Euler scheme with 200 discretization steps. For the $\mathbb{P}$-realisations $X^{(i)}$, we calculate the standard Monte-Carlo estimator of the European put price and for

the $\mathbb{P}_{\bar{\theta}^*}$-realisations $X^{(i,\bar{\theta}^*)}$, we use (6.10) where ϕ and ψ are given in (6.17) and compute the same statistics as in the previous examples. In Table 5, we fix the strike to the value $K = 1$ and let the maturity T vary from 0.25 to 3, whereas in Tables 6 and 7, we fix the maturity to $T = 1$ and to $T = 3$, while we let the strike K vary between 0.25 and 1.75.

T	Price	Std. error	Var. ratio	Adj. ratio	Time, s
0.25	0.0945	$9.96 \cdot 10^{-4}$	3.28	3.00	23.6
0.5	0.147	$1.28 \cdot 10^{-3}$	3.20	2.99	24.5
1	0.215	$1.61 \cdot 10^{-3}$	2.95	2.77	24.7
2	0.309	$2.04 \cdot 10^{-3}$	2.61	2.43	24.7
3	0.374	$2.30 \cdot 10^{-3}$	2.40	2.20	25.0

TABLE 5. The variance ratio as function of the maturity for the European put option in the Heston model with jumps. tab:PutVarianceRe

K	Price	Std. error	Var. ratio	Adj. ratio	Time, s
0.25	0.00606	$7.83 \cdot 10^{-5}$	11.6	10.4	25.8
0.5	0.0377	$4.03 \cdot 10^{-4}$	5.42	5.28	24.7
0.75	0.105	$9.44 \cdot 10^{-4}$	3.76	3.19	27.3
1	0.215	$1.61 \cdot 10^{-3}$	2.93	2.89	26.1
1.25	0.369	$2.26 \cdot 10^{-3}$	2.65	2.46	25.4
1.5	0.550	$2.80 \cdot 10^{-3}$	2.43	2.24	24.9
1.75	0.766	$3.05 \cdot 10^{-3}$	2.57	2.44	24.6

TABLE 6. The variance ratio as function of the strike for the European put option with maturity $T = 1$ in the Heston model with jumps. tab:PutVarianceRe

K	Price	Std. error	Var. ratio	Adj. ratio	Time, s
0.25	0.0280	$2.69 \cdot 10^{-4}$	5.19	4.99	24.8
0.5	0.108	$8.60 \cdot 10^{-4}$	3.32	3.05	25.1
0.75	0.226	$1.58 \cdot 10^{-3}$	2.68	2.56	26.3
1	0.374	$2.31 \cdot 10^{-3}$	2.39	2.20	27.0
1.25	0.545	$3.01 \cdot 10^{-3}$	2.20	2.19	25.2
1.5	0.730	$3.66 \cdot 10^{-3}$	2.09	1.94	24.6
1.75	0.932	$4.27 \cdot 10^{-3}$	1.97	1.83	24.8

TABLE 7. The variance ratio as function of the strike for the European put option with maturity $T = 3$ in the Heston model with jumps. tab:PutVarianceRe

When adding negative jumps to the Heston model, one can see that the variance ratio diminishes. When the options are out of the money however it is still sufficiently high to make it interesting to use in applications. In Figure 6.2, we fix the maturity to $T = 1.5$ and plot again the empirical variance of the estimator as a function of $\bar{\theta}^*$ for the Heston model with jumps. The method provides $\bar{\theta}^* = -0.312$

as candidate asymptotically optimal measure change which is, as in the continuous case, very close to the optimal one.

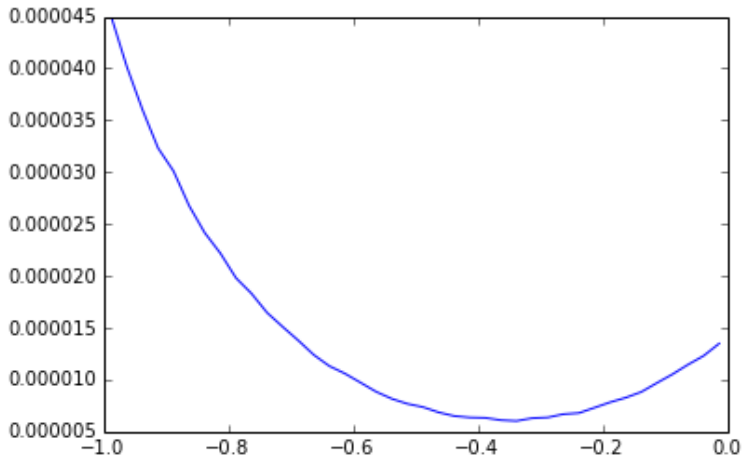

FIGURE 6.2. The variance of the Monte-Carlo estimator as a function of $\bar{\theta}^*$ for the Heston model with jumps.

fig:VarianceHesto

REFERENCES

Barndorff-Nielsen, O. E. (1997). Processes of normal inverse Gaussian type. *Finance and Stochastics*, 2(1):41–68.

Barndorff-Nielsen, O. E. and Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(2):167–241.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche mark options. *The Review of Financial Studies*, 9(1):69–107.

Dembo, A. and Zeitouni, O. (1998). *Large Deviations Techniques and Applications*. Springer, Application of Mathematics, second edition.

Duffie, D., Filipovic, D., and Schachermayer, W. (2003). Affine processes and applications in finance. *The Annals of Applied Probability*, 13(3):984–1053.

Dupuis, P. and Wang, H. (2004). Importance sampling, large deviations, and differential games. *Stochastics: An International Journal of Probability and Stochastic Processes*, 76(6):481–508.

Dupuis, P. and Wang, H. (2007). Subsolutions of an Isaacs equation and efficient schemes for importance sampling. *Mathematics of Operations Research*, 32(3):723–757.

Ekeland, I. and Temam, R. (1999) *Convex Analysis and Variational Problems*. Society for Industrial and Applied Mathematics.

Genin, A. and Tankov, P. (2020). Optimal importance sampling for Lévy processes. *Stochastic Processes and their Applications*, 130(1): 20–46.

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

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

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

Jacquier, A., Keller-Ressel, M., and Mijatović, A. (2013). Large deviations and stochastic volatility with jumps: asymptotic implied volatility for affine models. *Stochastics: An International Journal of Probability and Stochastic Processes*, 85(2):321–345.

Keller-Ressel, M. (2011). Moment explosions and long-term behaviour of affine stochastic volatility models. *Mathematical Finance*, Volume 21(1).

Léonard, C. (2000). Large deviations for Poisson random measures and processes with independent increments. *Stochastic Processes and their Applications*, 85:93–121.

Robertson, S. (2010). Sample path large deviations and optimal importance sampling for stochastic volatility models. *Stochastic Processes and their Applications*, 120(1):66–83.

Rockafellar, R. T. (1971). Integrals which are convex functionals, ii. *Pacific Journal of Mathematics*, 39(2).