Importance sampling applied to Greeks for jump-diffusion models with stochastic volatility

Sergio De Diego[1] Eva Ferreira[1] Eulalia Nualart[2]

September 2016

Abstract

The aim of this paper is to develop a variance reduction technique, based on importance sampling in conjunction with the stochastic Robbins-Monro algorithm, for option prices of jump-diffusion models with stochastic volatility. This is done by combining the work developed by Arouna [2, 3] for pricing diffusion models, and extended by Kawai [19, 20] for Lévy processes without Brownian component. We apply this technique to improve the numerical computation of derivative price sensitivities for general Lévy processes, allowing both Brownian and jump parts. Numerical examples are performed for the Black-Scholes and Heston models with jumps, and the Barndorff-Nielsen and Shephard model to illustrate the efficiency of this numerical technique. The numerical results support that the proposed methodology improves the efficiency of the usual Monte Carlo procedures.

1 Introduction

Monte Carlo simulation techniques are widely used for pricing and hedging complex financial assets. Due to the increasing complexity of the models and the number of state variables, since Monte Carlo estimations have rate of convergence of order $N^{-1/2}$, in most cases the implementation of variance reduction techniques is needed to increase this ratio of convergence. In this sense, there are different techniques, such as control variates (see Glasserman [18] and the references therein), localization technique (see Fournier and *al.* [14]) or importance sampling technique (see Su and Fu [24] and Arouna [2, 3]). However, as noted in [2], importance sampling has not been widely used in Finance. The main idea of the importance sampling technique is to determine a change of drift to reduce the variance of the sample, and therefore obtain more efficient estimates.

[1]Departamento de Economía Aplicada III, Facultad de Ciencias Económicas y Empresariales, Universidad del País Vasco, Avda. Lehendakari, Aguirre, 83, 48015 Bilbao, Spain
e-mail: sergio.dediego@ehu.eus, eva.ferreira@ehu.eus
Eva Ferreira's research is supported by MINECO-FEDER ECO2014-51914-P, IT-IT793-13 and Sergio De Diego was supported by a Basque Country Government assistantship.

[2]Department of Economics and Business, Universitat Pompeu Fabra and Barcelona Graduate School of Economics, Ramón Trias Fargas 25-27, 08005 Barcelona, Spain
e-mail: eulalia.nualart@upf.edu
Eulalia Nualart's research is supported in part by the European Union programme FP7-PEOPLE-2012-CIG under grant agreement 333938.

[12]The authors would like to thank the anonymous referees for their comments and suggestions that helped to improve the paper.

Although Arouna [2] was not the first to use this technique for financial models (see also [24]), he proposes the use of the importance sampling technique in conjunction with the stochastic Robbins-Monro algorithm for the numerical computation of option prices of diffusion models. The idea is to use this stochastic algorithm to find the optimal drift in the Girsanov transform, that is, the change of probability measure under which the variance of the estimate is minimal.

Subsequently, Kawai [19, 20] extends Arouna's work by applying this technique to Lévy processes without Brownian component. In this case, the change of probability measure is performed using the Esscher transform. In this approach the parameter used in the measure transform may be restricted into a specific compact set. This differs from the Brownian case, where the drift can take any real value, depending on the structure of the Lévy measure.

In this paper we generalize these works considering general diffusion models with independent Brownian and jump parts, and stochastic volatility. The purpose is to reduce the variance of the estimates by changing the probability measure with the use of the Esscher transform. In a recent the paper by Alaya and al. [1], they also apply the importance sampling technique to a general class of Lévy process, using instead the statistical Romberg method to simulate the process. In our case, we also extend the application of the importance sampling technique for the numerical computation of derivative price sensitivities, known as Greeks.

Specifically, let $(S_t, t \geq 0)$ denote the price of an underlying asset under the risk-neutral probability P, adapted to the filtration of an independent Brownian motion and jump process. Consider a path-dependent and square integrable payoff $h(S_t, t \leq T)$ with maturity $T > 0$. The price of a contingent claim is expressed as

$$F = e^{-rT} \mathrm{E_P}[h(S_t, t \leq T)],$$

where $r > 0$ is the risk-free interest rate. The Greeks are the derivatives of the price F with respect to the different parameters of the model dynamics of S_t. The most common Greeks are the Delta and the Gamma, which are the first and second derivatives of F with respect to S_0, respectively.

If z denotes a generic parameter of the model, the simplest method in order to compute numerically the derivative of F with respect to z is to apply the finite difference scheme given by

$$\frac{\partial F}{\partial z} \approx \frac{F(z+\epsilon) - F(z-\epsilon)}{2\epsilon},$$

for sufficiently small ϵ. Although this method is easy to implement, the finite difference estimator has a large bias and variance, especially for discontinuous payoffs.

Different alternative methods to compute these derivatives have been proposed in the literature (see Glasserman [18] for a review). More recently, [16] proposes the so called COS method, based on the Fourier-cosine series. In this paper, we concentrate on the Malliavin calculus approach. This method, introduced by Fournier and *al.* [14, 15], consists of using the integration by parts formula of the Malliavin calculus to obtain closed formulas for the Greeks as the expectation of the original payoff multiplied by some explicit weights. That is, the sensitivity with respect to z obtained via the Malliavin calculus has the form

$$\frac{\partial F}{\partial z} = e^{-rT} \mathrm{E_P}\left[h(S_t, t \leq T)\pi\right], \tag{1}$$

where π is a random variable that depends on the Brownian motion and jump component of the asset price along the interval $[0, T]$. These closed formulas allow a direct Monte Carlo simulation for discontinuous payoffs.

While the pioneer papers only dealt with the case where S_t satisfies a diffusion processes on the Wiener space (see [22] for a review), much has been done in order to extend this technique in the presence of price jumps. See for example [4, 6, 7, 12, 13, 21].

However, an important computational problem arises from this technique. As explained in [14, 15], the weights π that appear in (1) increase the variance with respect to the variance of the random variable $h(S_t, t \leq T)$, which makes the numerical computation very unstable. In order to overcome this problem, Fournier and *al.* [14, 15] propose a localization technique that consists on applying the Malliavin calculus, or integration by parts, to a piece of the functional whose expectation has to be computed. This allows several probabilistic expressions to be written that share the same expectation and the aim is thus to seek a reduced variance. Two problems arise in the use of this approach: a) one needs to perform different integration by parts depending on the model considered, and b) the choice of the localization parameter remains unclear. In fact, most papers applying this localization procedure (for example [4, 6, 14, 15]) consider a random choice of the localization parameter. For a digital call option, the localization parameter choice is illustrated in [15], which shows how this is not an easy task. A numerical procedure for the choice of this parameter is proposed in [11]. Other papers (for example [12, 21]) only show that the numerical simulation of the Greeks obtained via the Malliavin calculus (Monte Carlo with no variance reduction) outperforms the finite difference scheme. In fact, there is no optimal selection procedure of the location parameter for this task, so this methodology is not suitable for variance reduction to compute the Greeks.

In this paper, we apply the variance reduction technique exposed in Section 2 to improve the numerical computation of the Greeks formulas obtained via the Malliavin calculus, with no need for any localization procedure. We present numerical examples such as the Black-Scholes model with jumps, the Heston model with jumps, and the Barndorff-Nielsen and Shephard model. The formulas for the Greeks via the Malliavin calculus for these models are computed in [12] and [6]. We compare the numerical computation of these formulas with and without the use of the variance reduction technique, and we observe that the variance is significantly reduced.

The paper is organized as follows. Section 2 introduces the variance reduction technique based on the Esscher transform. In particular, we prove the convergence in our setting of the algorithm that combines the Robbins-Monro algorithm with the projection method introduced by Chen and *al.* In Section 3 we apply this technique to several Greeks obtained via the Malliavin calculus, and present some numerical examples. Finally, Section 4 concludes.

2 Variance reduction technique

2.1 Importance sampling technique

Let $B = (B_t, t \geq 0)$ be a d-dimensional Brownian motion and let $Z = (Z_t, t \geq 0)$ be an m-dimensional pure jump Lévy process with no drift and independent of B, with Lévy measure ν. That is, Z is an m-dimensional process with stationary and independent increments, continuous in probability, and such that $Z_0 = 0$, a.s.

By the independence of B and Z, we can assume that they are defined on the product

filtered probability space

$$(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathrm{P}) = (\Omega_1 \times \Omega_2, \mathcal{F}_1 \otimes \mathcal{F}_2, (\mathcal{F}_t^1 \otimes \mathcal{F}_t^2)_{t\geq 0}, \mathrm{P}_1 \times \mathrm{P}_2), \tag{2}$$

where $(\Omega_1, \mathcal{F}_1, (\mathcal{F}_t^1)_{t\geq 0}, \mathrm{P}_1)$ and $(\Omega_2, \mathcal{F}_2, (\mathcal{F}_t^2)_{t\geq 0}, \mathrm{P}_2)$ are the filtered probability spaces of B and Z, respectively. We denote by E_{P}, $\mathrm{E}_{\mathrm{P}_1}$ and $\mathrm{E}_{\mathrm{P}_2}$ the expectations under the probability measures P, P_1 and P_2, respectively.

By the Lévy-Khintchine representation theorem, the characteristic function of Z is uniquely given by

$$\mathrm{E}_{\mathrm{P}_2}\left[e^{i\langle u, Z_t\rangle}\right] = \exp\left(t\int_{\mathbf{R}_0^m}\left(e^{i\langle u,z\rangle} - 1 - i\langle u,z\rangle \mathbf{1}_{(0,1]}(|z|)\right)\nu(dz)\right), \qquad u \in \mathbf{R}^m,$$

where $\mathbf{R}_0^m = \mathbf{R}^m \setminus \{0\}$, and ν is a positive σ-finite measure on $\mathbf{R}_0^m$ such that $\int_{\mathbf{R}_0^m}(|z|^2 \wedge 1)\nu(dz) < \infty$. Here $\langle\cdot,\cdot\rangle$ and $|\cdot|$ denote respectively the Euclidean scalar product and norm.

We are interested in numerically computing by Monte Carlo simulation quantities of the form

$$V = \mathrm{E}_{\mathrm{P}}[f(S_t, 0 \leq t \leq T)] = \mathrm{E}_{\mathrm{P}}[g(B_t, Z_t, 0 \leq t \leq T)], \tag{3}$$

where $T > 0$ is fixed, f and g are real-valued functions and $(S_t, t \in [0,T])$ is a stochastic process adapted to the filtration $\mathcal{F}_t$.

In order to improve the efficiency of the Monte Carlo simulation of V, our aim is to apply a variance reduction technique that consists of computing the optimal change of measure that attains the minimal variance of V. That is, we change the law of the processes $(B_t, t \in [0,T])$ and $(Z_t, t \in [0,T])$ by considering two parameters $\mu \in \mathbf{R}^d$ and $\theta \in \Lambda_1$, where

$$\Lambda_1 = \left\{u \in \mathbf{R}^m : \mathrm{E}_{\mathrm{P}_2}[e^{\langle u, Z_1\rangle}] < \infty\right\} = \left\{u \in \mathbf{R}^m : \int_{|z|>1} e^{\langle u,z\rangle}\nu(dz) < \infty\right\}.$$

We assume that $\mathrm{Leb}(\Lambda_1) > 0$ and we observe that Λ_1 contains the origin and is convex. We consider the probability measures $\mathrm{Q} = \mathrm{Q}_1 \times \mathrm{Q}_2$ on $(\Omega, \mathcal{F})$, where Q_1 and Q_2 are given by the Esscher transforms

$$\left.\frac{d\mathrm{Q}_1}{d\mathrm{P}_1}\right|_{\mathcal{F}_t^1} = \frac{e^{\langle \mu, B_t\rangle}}{\mathrm{E}_{\mathrm{P}_1}[e^{\langle \mu, B_t\rangle}]} = e^{\langle \mu, B_t\rangle - \frac{1}{2}|\mu|^2 t}, \qquad \left.\frac{d\mathrm{Q}_2}{d\mathrm{P}_2}\right|_{\mathcal{F}_t^2} = \frac{e^{\langle \theta, Z_t\rangle}}{\mathrm{E}_{\mathrm{P}_2}[e^{\langle \theta, Z_t\rangle}]} = e^{\langle \theta, Z_t\rangle - t\varphi(\theta)}, \tag{4}$$

where

$$\varphi(\theta) = \log \mathrm{E}_{\mathrm{P}_2}[e^{\langle \theta, Z_1\rangle}] = \frac{1}{t}\log \mathrm{E}_{\mathrm{P}_2}[e^{\langle \theta, Z_t\rangle}].$$

Note that $\varphi(\theta)$ is continuous and $\nabla\varphi(\theta)$ exists for $\theta \in \Lambda_1$.

We use E_{Q}, $\mathrm{E}_{\mathrm{Q}_1}$ and $\mathrm{E}_{\mathrm{Q}_2}$ to denote the expectations under Q, Q_1 and Q_2, respectively. Observe that, under Q, the process $B_t - \mu t$ is a d-dimensional Brownian motion, and the law of the process Z_t under Q can be determined from its characteristic function

$$\begin{aligned}
&\mathrm{E}_{\mathrm{Q}_2}\left[e^{i\langle u, Z_t\rangle}\right] = \mathrm{E}_{\mathrm{P}_2}\left[e^{i\langle u, Z_t\rangle} e^{\langle \theta, Z_t\rangle - t\varphi(\theta)}\right] \\
&= \exp\left(it\langle u, \int_{|z|\leq 1} z(e^{\langle z,\theta\rangle} - 1)\nu(dz)\rangle + t\int_{\mathbf{R}_0^m}\left(e^{i\langle u,z\rangle} - 1 - i\langle u,z\rangle \mathbf{1}_{(0,1]}(|z|)\right)e^{\langle z,\theta\rangle}\nu(dz)\right),
\end{aligned}$$

for $u \in \mathbf{R}^m$. That is, under Q, Z_t is still an m-dimensional Lévy process with Lévy density $e^{\langle \theta, z\rangle}\nu(dz)$ and drift $\int_{|z|\leq 1} z(e^{\langle z,\theta\rangle} - 1)\nu(dz)$.

We can now rewrite the quantity (3) under the new probability measure Q, that is,

$$V = \mathrm{E_Q}[g(B_t, Z_t, 0 \leq t \leq T)e^{-\langle \mu, B_T\rangle + \frac{1}{2}|\mu|^2 T - \langle \theta, Z_T\rangle + T\varphi(\theta)}]. \tag{5}$$

At this point, the goal is to find the optimal parameters $(\mu^*, \theta^*) \in \mathbf{R}^d \times \Lambda_1$ that minimize the variance of V, which is given by

$$\mathrm{Var}(V) = H(\mu, \theta) - V^2, \tag{6}$$

where

$$\begin{aligned} H(\mu,\theta) &= \mathrm{E_Q}[g^2(B_t, Z_t, 0 \leq t \leq T)e^{-2\langle \mu, B_T\rangle + |\mu|^2 T - 2\langle \theta, Z_T\rangle + 2T\varphi(\theta)}] \\ &= \mathrm{E_P}[g^2(B_t, Z_t, 0 \leq t \leq T)e^{-\langle \mu, B_T\rangle + \frac{1}{2}|\mu|^2 T - \langle \theta, Z_T\rangle + T\varphi(\theta)}]. \end{aligned} \tag{7}$$

Observe that since V in (6) does not depend on (μ, θ), it suffices to minimize the second moment $H(\mu, \theta)$ in order to minimize the variance of V. The next result therefore gives sufficient conditions that ensure that the function $H(\mu, \theta)$ has a unique minimum.

Proposition 1. *Assume that the following conditions hold:*

(1) $\mathrm{P}(g(B_t, Z_t, 0 \leq t \leq T) > 0) \neq 0$.

(2) *There exists* $p > 1$ *such that* $\mathrm{E_P}[|g(B_t, Z_t, 0 \leq t \leq T)|^{2p}] < \infty$.

Consider the set

$$\Lambda_2 = \Lambda_1 \cap \left\{ u \in \mathbf{R}^m : \int_{|z|>1} |z|^{2q} e^{-q\langle u, z\rangle}\nu(dz) < \infty,\ \frac{1}{q} + \frac{1}{p} = 1 \right\},$$

and assume that $\mathrm{Leb}(\Lambda_2) > 0$. *Then, the set* Λ_2 *is convex, and the function* $H(\mu, \theta)$ *is twice differentiable and strictly convex in* $\mathbf{R}^d \times \Lambda_2$.

Proof. The convexity of Λ_2 is proved as in [19] applying Hölder's inequality. We next show the second statement. Consider the functions

$$F(\mu, \theta) = (T\mu - B_T)G, \qquad J(\mu, \theta) = (T\nabla\varphi(\theta) - Z_T)G, \tag{8}$$

where

$$G = g^2(B_t, Z_t, 0 \leq t \leq T)e^{-\langle \mu, B_T\rangle + \frac{1}{2}|\mu|^2 T - \langle \theta, Z_T\rangle + T\varphi(\theta)}.$$

If $|\mu| \leq K$, using Hölder's inequality twice (first with respect to P_1 and second with respect to P_2, using the independence between Z and B) and proceeding as in the proof of Proposition 1 in [2], we obtain that

$$\mathrm{E_P}\left[|F(\mu,\theta)|\right] \leq C_K \mathrm{E_P}[|g(B_t, Z_t, 0 \leq t \leq T)|^{2p}]^{1/p} e^{T\varphi(\theta)} \mathrm{E_{P_2}}\left[e^{-q\langle \theta, Z_T\rangle}\right]^{1/q},$$

which is finite if $\theta \in \Lambda_2$. Similarly,

$$\mathrm{E_P}\left[|J(\mu,\theta)|\right] \leq C_K \mathrm{E_P}[|g(B_t, Z_t, 0 \leq t \leq T)|^{2p}]^{1/p} e^{T\varphi(\theta)} \mathrm{E_{P_2}}\left[|T\nabla\varphi(\theta) - Z_T|^q e^{-q\langle \theta, Z_T\rangle}\right]^{1/q},$$

which is finite if $\theta \in \Lambda_2$. Therefore, from hypothesis (2) and the dominated convergence theorem, we obtain that $H(\mu, \theta)$ is differentiable and

$$\nabla H(\mu, \theta) = \left(\mathrm{E_P}\left[F(\mu, \theta)\right], \mathrm{E_P}\left[J(\mu, \theta)\right]\right). \tag{9}$$

By an analogous argument, H is twice differentiable and its Hessian matrix is given by

$$\begin{pmatrix} \mathrm{E_P}\left[(TI_m + (T\mu - B_T)(T\mu - B_T)^{\mathrm{T}})G\right] & \mathrm{E_P}\left[(T\mu - B_T)(T\nabla\varphi(\theta) - Z_T)^{\mathrm{T}}G\right] \\ \mathrm{E_P}\left[(T\nabla\varphi(\theta) - Z_T)(T\mu - B_T)^{\mathrm{T}}G\right] & \mathrm{E_P}\left[(T\mathrm{Hess}(\varphi(\theta)) + WG\right] \end{pmatrix},$$

where $W = (T\nabla\varphi(\theta) - Z_T)(T\nabla\varphi(\theta) - Z_T)^{\mathrm{T}})$, I_m denotes the $m \times m$ identity matrix and A^{T} denotes the transpose of a matrix A.

Then, using hypothesis (1), we get that this Hessian matrix is positive definite and thus H is strictly convex on $\mathbf{R}^d \times \Lambda_2$. □

Remark 1. *If there exists* $(\mu^*, \theta^*) \in \mathbf{R}^d \times \Lambda_2$ *such that*

$$\nabla H(\mu^*, \theta^*) = 0, \tag{10}$$

then, under the hypotheses of Proposition 1, since H *is strictly convex, we conclude that* (μ^*, θ^*) *is the unique minimum of* H *in* $\mathbf{R}^d \times \Lambda_2$.

2.2 Robbins-Monro algorithm and projection method

The aim of this section is to apply the Robbins-Monro algorithm adapted by Chen and *al.* in order to compute the minimum (μ^*, θ^*) of the function $H(\mu, \theta)$ numerically. Since in some cases it is difficult to check condition (10), we need to distinguish between unconstrained and constrained algorithms in order to ensure the existence and uniqueness of this minimum. The algorithms for each case are discussed, with the corresponding propositions to state the conditions for the convergence.

2.2.1 Unconstrained algorithms

The Robbins-Monro (RM) algorithm is an stochastic algorithm that allows to find numerically the minimum of a function that is defined by the expectation of a random variable. Due to the exponential growth of our gradient function (9), if this method is applied directly, we will obtain a explosive numerical behavior. Alternatively, Chen and Zhu [9] and Chen, Guao and Gao [10] propose a new algorithm consisting of truncating the RM algorithm at randomly varying bounds, in order to avoid the explosion of the algorithm. As pointed out in [19], in order to apply these techniques, we need to assume the uniqueness of the minimum of the objective function, in the whole space. These are called unconstrained algorithms. That is, we assume (i) $\Lambda_2 = \mathbf{R}^m$ and (ii) there exists a unique minimum $(\mu^*, \theta^*) \in \mathbf{R}^d \times \mathbf{R}^m$ of $H(\mu, \theta)$, defined in (7). For assumption (ii) to hold, by Proposition 1, it suffices to assume hypotheses (1) and (2), and that

$$\lim_{|\mu|+|\theta|\uparrow+\infty} H(\mu, \theta) = +\infty. \tag{11}$$

The main steps of the Robbins-Monro algorithm adapted by Chen described in [10] are as follows:

1. Consider two independent sequences $(B_t^n, t \in [0, T])_{n\geq 0}$ and $(Z_t^n, t \in [0, T])_{n\geq 0}$, which are iid copies of the processes $(B_t, t \in [0, T])$ and $(Z_t, t \in [0, T])$.

2. Choose a sufficiently large constant $M > 0$, and two increasing sequence of positive numbers $(u_n)_{n\geq 0}$ and $(v_n)_{n\geq 0}$ tending to infinity such that $u_0 > M$ and $v_0 > M$.

3. Choose an initial condition $(\mu_0, \theta_0) \in \mathbf{R}^d \times \mathbf{R}^m$.

4. Define the sequence $(\mu_n, \theta_n)_{n\geq 1}$ by

$$(\mu_{n+1}, \theta_{n+1}) = \begin{cases} (\mu_n, \theta_n) - \frac{\beta}{n+1}(F_{n+1}, J_{n+1}), & \text{if } |\mu_n - \frac{\beta}{n+1}F_{n+1}| \leq u_{\sigma(n)} \text{ and} \\ & \quad |\theta_n - \frac{\beta}{n+1}J_{n+1}| \leq v_{\sigma(n)}, \\ (\mu_0, \theta_0), & \text{otherwise,} \end{cases} \tag{12}$$

where $\sigma(n)$ is the number of projections after n iterations. That is,

$$\sigma(n) = \sum_{k=0}^{n-1} \mathbf{1}_{\left\{|\mu_n - \frac{\beta}{n+1}F_{n+1}| \geq u_{\sigma(n)}, |\theta_n - \frac{\beta}{n+1}J_{n+1}| \geq v_{\sigma(n)}\right\}}, \quad \sigma(0) = 0, \tag{13}$$

$\beta > 0$, and

$$(F_{n+1}, J_{n+1}) = (F(\mu_n, \theta_n, (B_t^{n+1}, Z_t^{n+1} \in [0,T])), J(\mu_n, \theta_n, (B_t^{n+1}, Z_t^{n+1} \in [0,T]))),$$

where it should be recalled that F and J are defined in (8).

Under all the conditions mentioned above, it is shown in [10] that the sequence (θ_n, μ_n) defined by this algorithm converges to the unique minimum (μ^*, θ^*) of the function H in $\mathbf{R}^d \times \mathbf{R}^m$. However, several remarks are in order to choose the different parameters to apply this algorithm in practice.

Remark 2. *1. The term $\frac{\beta}{n+1}$ can be replaced by a general sequence of positive number $(\gamma_n)_{n\geq 0}$ satisfying that*

$$\sum_{n\geq 0} \gamma_n = +\infty, \quad \text{and} \quad \sum_{n\geq 0} \gamma_n^2 < +\infty.$$

For example in [2], the author uses the series $\frac{\beta}{n+\alpha}$, with $\alpha > 0$.

2. If there can be an intuitive idea of a sufficiently small neighborhood of (μ^*, θ^*), then it is natural to choose the initial condition (μ_0, θ_0) in this neighborhood. Otherwise, a natural choice is to start the algorithm at $(0, 0)$, as it is the case in all our numerical simulations.

3. Observe that in [10] the authors show the convergence of the algorithm for any sequence u_n and v_n increasing towards infinity and $M > 0$ a sufficiently large constant. However, in practice, these sequences need to be chosen according to the growth of the variables F_{n+1} and J_{n+1}. In the next subsection we give an example of the choice of these sequences for the case of a Poisson process.

4. Observe that in [10] the authors choose two arbitrary points in order to re-initialize the algorithm when the two conditions in (12) are not satisfied. Then, the choice of the constant $M > 0$ depends on the choice of these points. However, in practice we have observed that it is enough to re-initialize the algorithm at the initial value, and to choose a large enough $M > 0$ as it will be discussed in the numerical examples.

2.2.2 Example: Poisson process

Assume that the Lévy process Z is an m-dimensional Poisson process denoted $N = (N_t, t \geq 0)$ with intensity $\lambda > 0$. In this case,

$$\varphi(\theta) = \langle \lambda, e^\theta - 1 \rangle, \quad \nabla\varphi(\theta) = \lambda e^\theta, \quad \Lambda_1 = \Lambda_2 = \mathbf{R}^m.$$

Observe that we have used the abuse of notation $\langle \lambda, e^\theta - 1\rangle = \lambda \sum_{i=1}^m (e^{\theta_i} - 1)$. Of course, the case $\lambda \in \mathbf{R}^m$ can also be considered.

This is an example of unconstrained algorithm. Therefore, if g satisfies conditions (1) and (2) of Proposition 1, together with (11), then, H has a unique minimum (μ^*, θ^*) in $\mathbf{R}^d \times \mathbf{R}^m$.

The next result show the convergence of the Robbins-Monro algorithm adapted by Chen and *al.* in this particular case.

Proposition 2. *Assume condition* (1) *of Proposition 1, condition* (11)*, and replace condition* (2) *in Proposition 1 by the stronger condition*

(2') *There exists* $p > 1$ *such that* $\mathrm{E_P}[|g(B_t, N_t, 0 \leq t \leq T)|^{4p}] < \infty$.

We can then choose two increasing sequence of positive numbers $(u_n)_{n\geq 0}$ *and* $(v_n)_{n\geq 0}$ *such that the sequence* (μ_n, θ_n) *defined in* (12) *converges a.s. to the unique minimum* (μ^*, θ^*) *of* H *in* $\mathbf{R}^d \times \mathbf{R}^m$.

Proof. Let $\mathcal{F}_T^n$ denote the σ-algebra generated by the random vectors

$$(\theta_k, \mu_k, (B_t^k, Z_t^k, t \in [0,T]), k \leq n).$$

Since (μ_n, θ_n) are $\mathcal{F}_T^n$-measurable and $(B_t^{n+1}, Z_t^{n+1}, t \in [0,T])$ are independent of $\mathcal{F}_T^n$, we have that

$$\nabla H(\mu_n, \theta_n) = (\mathrm{E_P}[F_{n+1}|\mathcal{F}_T^n], \mathrm{E_P}[J_{n+1}|\mathcal{F}_T^n]).$$

Consider the $\mathcal{F}_T^n$-martingale $(M_n)_{n\geq 0}$, where $M_0 = (0,0)$ and for $n \geq 1$,

$$M_n = \sum_{i=0}^{n-1} \frac{\beta}{i+1}(F_{i+1} - \mathrm{E_P}[F_{i+1}|\mathcal{F}_T^i], J_{i+1} - \mathrm{E_P}[J_{i+1}|\mathcal{F}_T^i]).$$

Given the results in [2, 3], it suffices to choose the sequences u_n and v_n such that the martingale M_n converges a.s. as n tends to ∞. The brackets process clearly satisfies that

$$\langle M_n \rangle \leq \sum_{i=0}^{n-1} \left(\frac{\beta}{n+1}\right)^2 \mathrm{E_P}[|F_{i+1}|^2 + |J_{i+1}|^2|\mathcal{F}_T^i].$$

Moreover, since (μ_n, θ_n) is $\mathcal{F}_T^n$-measurable and $(B_t^{n+1}, N_t^{n+1}, t \in [0,T])$ are independent of $\mathcal{F}_T^n$, we have that

$$\mathrm{E_P}[|F_{i+1}|^2 + |J_{i+1}|^2|\mathcal{F}_T^i] = s^2(\mu_i, \theta_i),$$

where

$$s^2(\mu,\theta) = \mathrm{E_P}[|F(\mu,\theta,(B_t,N_t,t\in[0,T]))|^2] + \mathrm{E_P}[|J(\mu,\theta,(B_t,N_t,t\in[0,T]))|^2].$$

To ease the exposition, we assume in the rest of the proof that $T = 1$. Following as in [2], for any $p > 1$, we have that

$$\mathrm{E_P}[|F(\mu,\theta)|^2] = C_1(p,q)e^{9|\mu|^2}\mathrm{E_P}\left[|e^{-2\langle\theta,N_1\rangle + 2\langle\lambda,e^\theta-1\rangle}|^{\frac{pq}{q-1}}\right]^{\frac{q-1}{pq}}.$$

We now treat the expectation on the rhs. We have

$$\mathrm{E_P}\left[|e^{-2\langle\theta,N_1\rangle+2\langle\lambda,e^\theta-1\rangle}|^{p'}\right]^{\frac{1}{p'}} = e^{2\langle\lambda,e^\theta-1\rangle+\frac{1}{p'}\langle\lambda,e^{-2p'\theta}-1\rangle} \leq e^{2\lambda m(e^{|\theta|}-1)+\frac{1}{p'}\lambda m(e^{2p'|\theta|}-1)}.$$

On the other hand, for any $p>1$,

$$\mathrm{E_P}[|J(\mu,\theta)|^2] \leq C_2(p,q)\mathrm{E_P}\left[|e^{-2\langle\mu,B_1\rangle+1|\mu|^2}|^{\frac{pq}{q-1}}\right]^{\frac{q-1}{pq}} (\mathrm{E_P}[f^{\frac{p}{p-1}}(B_1,N_1)])^{\frac{p-1}{p}},$$

where

$$f(B_1,N_1) = |\lambda e^\theta - N_1|^2 e^{-2\langle\theta,N_1\rangle+2\langle\lambda,e^\theta-1\rangle}.$$

It is easy to see that

$$\mathrm{E_P}\left[|e^{-2\langle\mu,B_1\rangle+|\mu|^2}|^{\frac{pq}{q-1}}\right]^{\frac{q-1}{pq}} = e^{(1+\frac{2pq}{q-1})|\mu|^2}$$

and

$$(\mathrm{E_P}[f^{\frac{p}{p-1}}(B_1,N_1)])^{\frac{p-1}{p}} \leq c(p)e^{2\lambda m(e^{|\theta|}-1)}\left((\lambda m e^{|\theta|})^{\frac{2p}{p-1}}e^{\lambda m(e^{\frac{2p}{p-1}|\theta|}-1)} + e^{\frac{4p^2}{(p-1)^2}|\theta|^2}\right)^{\frac{p-1}{p}}.$$

Therefore, taking $p=q=2$, we obtain that

$$s^2(\mu,\theta) \leq ce^{2\lambda m(e^{|\theta|}-1)+9|\mu|^2}\left(e^{\frac{1}{4}\lambda m(e^{8|\theta|}-1)} + (\lambda m e^{|\theta|})^2 e^{\frac{1}{2}\lambda m(e^{4|\theta|}-1)} + e^{8|\theta|^2}\right) \leq ce^{(4\lambda m+8)e^{8|\theta|}+9|\mu|^2}.$$

Since $|\mu_n| \leq u_n$ and $|\theta_n| \leq v_n$, it suffices to choose

$$u_n = \sqrt{c_1 \ln n} \quad \text{and} \quad v_n = \frac{1}{8}\ln(c_2 \ln n),$$

so that

$$\lim_{n\to\infty}\langle M_n\rangle \leq c \lim_{n\to\infty}\sum_{i=0}^{n-1}\frac{n^{9c_1+4\lambda m c_2}}{n^2} < \infty.$$

□

2.2.3 Constrained algorithms

As pointed out in [19], if we apply the Robbins-Monro stochastic algorithm constrained on the convex set $\mathbf{R}^d \times \Lambda_2$, where $\Lambda_2 \neq \mathbf{R}^m$, we cannot ensure that the limit of the sequence constructed with the algorithm will converge to a point (μ^*,θ^*) such that $\nabla H(\mu^*,\theta^*)=0$. In order to ensure this fact, Proposition 1 shows that it is sufficient that $(\mu,\theta) \in \mathbf{R}^d \times \Lambda_2$ exists such that $\nabla H(\mu,\theta)=0$. Since in general this is not clear form a theoretical point of view, we need to apply in those cases the Robbins-Monro algorithm adapted by Chen and *al.* constrained on the set Λ_2, and then numerically to analyze the behavior of the gradient in the applications.

In this case, the steps of the Robbins-Monro algorithm that differ from the unconstrained algorithm defined in Section 2.2.2. are steps 2. and 4. which need to be replaced by

2. Choosing a sufficiently large constant $M > 0$, and an increasing sequence of positive numbers $(u_n)_{n\geq 0}$ tending to infinity such that $u_0 > M$.

4. Defining the sequence $(\mu_n, \theta_n)_{n\geq 1}$ by

$$(\mu_{n+1}, \theta_{n+1}) = \begin{cases} (\mu_n, \theta_n) - \frac{\beta}{n+1}(F_{n+1}, J_{n+1}), & \text{if } |\mu_n - \frac{\beta}{n+1}F_{n+1}| \leq u_{\sigma(n)} \text{ and} \\ & \theta_n - \frac{\beta}{n+1}J_{n+1} \in \Lambda_2, \\ (\mu_0, \theta_0), & \text{otherwise,} \end{cases}$$

where $\sigma(n)$ is the number of projections after n iterations. That is,

$$\sigma(n) = \sum_{k=0}^{n-1} \mathbf{1}_{\left\{|\mu_n - \frac{\beta}{n+1}F_{n+1}| \geq U_{\sigma(n)},\, \theta_n - \frac{\beta}{n+1}J_{n+1} \notin \Lambda_2\right\}}, \quad \sigma(0) = 0.$$

We then have the following analogue of Proposition 2 in this case.

Proposition 3. *Assume that the following conditions hold:*

(1) $\mathrm{P}(g(B_t, Z_t, 0 \leq t \leq T) > 0) \neq 0$.

(2) *There exists* $p > 1$ *such that* $\mathrm{E_P}[|g(B_t, Z_t, 0 \leq t \leq T)|^{4p}] < \infty$.

We can then choose the increasing sequence of positive numbers $(u_n)_{n\geq 0}$ *such that the sequence* (μ_n, θ_n) *converges a.s. to a point* (μ^*, θ^*) *in* $\mathbf{R}^d \times \Lambda_2$. *Moreover,* $H(\mu^*, \theta^*) \leq H(0, 0)$.

Proof. We assume that $T = 1$ to ease the exposition. As in Proposition 2, for any $p > 1$, we have that

$$\mathrm{E_P}[|F(\mu, \theta)|^2] = C_1(p, q)e^{9|\mu|^2}\mathrm{E_P}\left[|e^{-2\langle\theta, Z_1\rangle + 2\varphi(\theta)}|^{\frac{pq}{q-1}}\right]^{\frac{q-1}{pq}},$$

which is finite since $\theta \in \Lambda_2$. On the other hand, for any $p > 1$,

$$\mathrm{E_P}[|J(\mu, \theta)|^2] \leq C_2(p, q)\mathrm{E_P}\left[|e^{-2\langle\mu, B_1\rangle + |\mu|^2}|^{\frac{pq}{q-1}}\right]^{\frac{q-1}{pq}} \left(\mathrm{E_P}[f^{\frac{p}{p-1}}(B_1, Z_1)]\right)^{\frac{p-1}{p}},$$

where

$$f(B_1, Z_1) = |\nabla\varphi(\theta) - Z_1|^2 e^{-2\langle\theta, Z_1\rangle + 2\varphi(\theta)}.$$

It is easy to see that

$$\mathrm{E_P}\left[|e^{-2\langle\mu, B_1\rangle + |\mu|^2}|^{\frac{pq}{q-1}}\right]^{\frac{q-1}{pq}} = e^{(1 + \frac{2pq}{q-1})|\mu|^2},$$

and

$$\left(\mathrm{E_P}[f^{\frac{p}{p-1}}(B_1, Z_1)]\right)^{\frac{p-1}{p}} \leq C.$$

Therefore, taking $p = q = 2$, we obtain that

$$s^2(\mu, \theta) \leq ce^{9|\mu|^2}.$$

In this case, since $|\mu_n| \leq u_n$ it suffices to choose $u_n = \sqrt{\frac{1}{10}\ln n}$, so that

$$\lim_{n\to\infty} \langle M_n\rangle \leq c \lim_{n\to\infty} \sum_{i=0}^{n-1} \frac{n^{9/10}}{n^2} < \infty.$$

The last statement follows due to the strict convexity of H in $\mathbf{R}^d \times \Lambda_2$. □

2.2.4 Example: Compound Poisson process

Assume that the Z is one-dimensional and has Lévy density $w(z)dz$, where $w(z) = \gamma\eta e^{-\eta z}\mathbf{1}_{z>0}$. That is, Z is a compound Poisson process of the form $Z_t = \sum_{i=1}^{N_t} J_i$, where N_t is a Poisson process with parameter $\gamma > 0$, and J_i are iid exponential random variables with parameter $\eta > 0$.

In this case, $\Lambda_1 = \{u \in \mathbf{R} : u < \eta\}$, $\varphi(\theta) = \frac{\gamma\theta}{\eta-\theta}$, $\varphi'(\theta) = \frac{\eta\gamma}{(\eta-\theta)^2}$, and

$$\Lambda_2 = \left\{ u \in \mathbf{R} : -\frac{\eta}{q} < u < \eta \right\}.$$

Under Q_2 defined in (4), Z_t is a compound Poisson process with drift $\Gamma = \int_0^1 z(e^{\theta z} - 1)w(z)dz$, and Lévy density $\tilde{w}(z) = e^{z\theta}w(z)$. For all $\theta \in \Lambda_1$, $\int \tilde{w} = \frac{\eta\gamma}{\eta-\theta}$. That is, under Q_2, Z_t can be written as

$$Z_t = \Gamma t + \sum_{k=1}^{N_t} P_k,$$

where N_t is a Poisson process with parameter $\frac{\eta\gamma}{\eta-\theta}$ and P_k are iid exponential random variables with parameter $\eta - \theta$.

3 Application to numerical computation of Greeks

3.1 General algorithm

We denote by S_t the price of an underlying asset at time $t > 0$ under the risk-neutral probability measure P. We assume that S_t is adapted to the filtration defined in (2). We consider a square integrable payoff $h(S_t, t \le T)$ with maturity $T > 0$ and contingent claim

$$F = e^{-rT}\mathrm{E}_\mathrm{P}[h(S_t, t \le T)],$$

where $r > 0$ denotes the risk-free interest rate. We denote by z a generic parameter of the model dynamics of S_t. We assume that the techniques of the Malliavin calculus can be applied to compute the sensitivity with respect to z, which has the form

$$V_z = \frac{\partial F}{\partial z} = \mathrm{E}_\mathrm{P}[f(S_t, 0 \le t \le T)] = \mathrm{E}_\mathrm{P}[g(B_t, Z_t, 0 \le t \le T)],$$

as in (3). Then, in order to compute numerically this quantity, following Section 2, we proceed with the algorithm as follows:

1. Run the Robbins-Monro algorithm defined in Section 2.2.1 (or 2.2.3 if it is a constrained algorithm) to obtain a numerical estimation $(\widehat{\mu^*}, \widehat{\theta^*})$ of (μ^*, θ^*).

2. Insert this estimation in formula (5) for V_z under Q:

$$V_z = \mathrm{E}_\mathrm{Q}[g(B_t, Z_t, 0 \le t \le T)e^{-\langle\widehat{\mu^*}, B_T\rangle + \frac{1}{2}|\widehat{\mu^*}|^2 T - \langle\widehat{\theta^*}, Z_T\rangle + T\varphi(\widehat{\theta^*})}],$$

   and from this formula compute a numerical estimation of V_z applying the usual Monte Carlo method.

The implementation of this algorithm adds an extra computational cost to the numerical computation of the Greeks. As explained in [2], the numerical cost of this algorithm is equivalent to $Cm(n+N)$, where C is a constant, m is the number of step discretizations, and n and N are respectively the number of the Robbins-Monro algorithm and Monte Carlo runs. As we observe in the numerical examples, the standard deviation ratio of the estimation (with and without the variance reduction) goes up to 7, which corresponds to a variance reduction of 50. Thus as in [2, 20], this gain justifies the additional cost, since at most 50% simulated paths are used to estimate the optimal drift in addition to those used for the standard Monte Carlo simulation. Note also that multilevel Monte Carlo methods could be applied in order to reduce this computational cost (see for example [17]).

3.2 Numerical examples

3.2.1 Black-Scholes model with jumps

Assume that the price S_t of the underlying asset under the risk-neutral probability measure P follows the following jump diffusion model

$$dS_t = rS_t dt + \sigma S_t dB_t + (\alpha - 1)S_t(dN_t - \lambda dt), \qquad S_0 = x, \tag{14}$$

where $r > 0$ is the risk-free constant interest rate, $\alpha > 0$, B_t is a standard Brownian motion, N_t is a standard Poisson process with intensity $\lambda > 0$ independent of the Brownian motion, x is fixed, and σ is a positive constant.

Consider a square integrable payoff $h(S_T)$ with maturity $T > 0$. Using the Malliavin calculus technique it is shown in [12] that for a contingent claim

$$F = e^{-rT}\mathrm{E_P}[h(S_T)],$$

the Delta and Gamma of this model are given by

$$\Delta = \frac{\partial F}{\partial x} = e^{-rT}\mathrm{E_P}\left[h(S_T)\pi_\Delta\right], \qquad \Gamma = \frac{\partial^2 F}{\partial x^2} = e^{-rT}\mathrm{E_P}\left[h(S_T)\pi_\Gamma\right],$$

where

$$\pi_\Delta = \frac{B_T}{x\sigma T}, \quad \pi_\Gamma = \frac{1}{x^2\sigma T}\left(\frac{B_T^2}{\sigma T} - B_T - \frac{1}{\sigma}\right).$$

We now apply the algorithm exposed in Section 3.1 to numerically compute the values of Δ and Γ. By Itô's formula, the unique solution to equation (14) is explicit and given by

$$S_t = x\exp\left((r - (\alpha-1)\lambda - \sigma^2/2)t + \sigma B_t + N_t\log(\alpha)\right).$$

Moreover, we are in the case of an unconstrained algorithm (Section 2.2.1) and the underlying jump process is a Poisson process (Example 2.2.2). Therefore, the function $H(\mu,\theta)$ we need to minimize is given by

$$H(\mu,\theta) = e^{-rT}\mathrm{E_P}[h^2(S_T)\pi_{\Delta,\Gamma}^2 e^{-\langle\mu,B_T\rangle + \frac{1}{2}|\mu|^2 T - \langle\theta,N_T\rangle + T\langle\lambda,e^\theta - 1\rangle}],$$

where $(\mu,\theta) \in \mathbf{R}\times\mathbf{R}$.

Tables 1 and 2 present the values obtained for the Delta and the Gamma of a European call with payoff function $h(x) = (x-K)_+$. DeltaMC and GammaMC are computed with a direct run of $N = 90,000$ Monte Carlo simulations, while DeltaRM and GammaRM

include the variance reduction technique with $n = 20,000$. The StdRatio is the ratio of the standard deviations of both estimators. DeltaRef and GammaRef are the reference values. See for example [8] and the references therein for analytical formulas of the Greeks for the Black-Scholes and Heston models with jumps.

INSERT TABLE 1 AND TABLE 2 ABOUT HERE

The Robbins-Monro algorithm is implemented with initial value $(\mu_0, \theta_0) = (0, 0)$. In Proposition 2 the sequences u_n and v_n for which this stochastic algorithm converges are computed. That is,

$$u_n = \sqrt{c_1 \ln n} + M \quad \text{and} \quad v_n = \frac{1}{8} \ln(c_2 \ln n) + M,$$

for $M > 0$ a sufficiently large constant. In practice, for this example it is sufficient to choose $M = 50$ as in [2]. The choice of this value depends on the growth of the function inside the expectation of the function $H(\mu, \theta)$. As likewise remarked in [2], the value of the ratio K/S_0 plays an important role in the simulations, also in the presence of jumps. We have observed much less effect in the choice of the other parameters. In particular, since we are implementing the Robbins-Monro algorithm with $n = 20,000$, we have observed no significant effect on the choice of the parameter β as soon as it ranges in the interval $[1, 1000]$ (see also [2]).

We observe in Tables 1 and 2 that the values of DeltaRM and GammaRM are closer to the reference value than the others. Moreover, the value of the StdRatio goes up to 7, which shows the efficiency of the method in this example.

In Figures 1 and 2 we compute the value of Delta and Gamma for several Monte Carlo paths. We clearly observe that the variance reduction technique with the Robbins-Monro algorithm (Malliavin Reduction) gives a more accurated convergence to the reference value (exact) than the crude Monte Carlo method (Malliavin).

INSERT FIGURE 1 AND FIGURE 2 ABOUT HERE

3.2.2 Heston model with jumps

We now assume that the underlying asset dynamics under a risk-neutral probability measure P is given by the stochastic volatility model

$$\begin{cases} dS_t = rS_t dt + \sqrt{V_t} S_t dB_t + (\alpha - 1) S_t (dN_t - \lambda dt), & S_0 = x, \\ dV_t = \kappa(\theta - V_t) dt + \sigma \sqrt{V_t} dW_t, & V_0 = y, \end{cases}$$

where B_t and W_t are two correlated standard Brownian motions with $\mathrm{E}[W_t B_t] = \rho t$, and κ, θ and σ are positive constants.

Consider now a square integrable payoff $h(S_T)$ with maturity $T > 0$. Given the results in [12], for a contingent claim

$$F = e^{-rT} \mathrm{E_P}[h(S_T)],$$

the Delta and Gamma of this model are given by

$$\Delta = e^{-rT} \mathrm{E_P}\left[h(S_T)\pi_\Delta\right], \qquad \Gamma = e^{-rT} \mathrm{E_P}\left[h(S_T)\pi_\Gamma\right],$$

where

$$\pi_\Delta = \frac{1}{xT}\left(\int_0^T \frac{dB_t}{\sqrt{V_t}} - \frac{\rho}{\sqrt{1-\rho^2}} \int_0^T \frac{dW_t}{\sqrt{V_t}} \right),$$

and

$$\pi_\Gamma = \pi_\Delta^2 - \frac{1}{x}\pi - \frac{1}{x^2T^2}\frac{1}{1-\rho^2}\int_0^T \frac{dt}{V_t}.$$

In order to apply the setting of Section 2 to numerically compute the values of these Greeks, observe that we are again in the case of an unconstrained algorithm and the underlying jump process is a Poisson process. However, in general there is no close formula for the solution to this equation, thus need to proceed with a discretization scheme. For example, the Euler scheme given by

$$\begin{aligned} S_{t_{i+1}} &= S_{t_i} + rS_{t_i}\delta + \sigma(Y_{t_i})S_{t_i}X_i\sqrt{\delta} + (\alpha-1)S_{t_i}(\Delta N_i - \lambda\delta), \\ Y_{t_{i+1}} &= Y_{t_i} + b(Y_{t_i})\delta + a(Y_{t_i})(\rho X_i + \sqrt{1-\rho^2}X_{i+m})\sqrt{\delta}, \end{aligned}$$

where $0 = t_0 < t_1 < \cdots < t_m = T$, $\delta = t_{i+1} - t_i = \frac{T}{m}$, $(X_1, \ldots, X_m, X_{m+1}, \ldots, X_{2m})$ is a sequence of $2m$ iid Normal(0,1) random variables, and $(\Delta N_i = N_{t_{i+1}} - N_{t_i}, i = 1, ..., m)$ is a sequence of m iid Poisson$(\delta\lambda)$ random variables.

Therefore, the function we need to minimize is

$$H(\mu,\theta) = e^{-rT}\mathrm{E_P}[h^2(S_T)\pi_{\Delta,\Gamma}^2 e^{-\langle\mu,B_T\rangle + \frac{1}{2}|\mu|^2T - \langle\theta,N_T\rangle + T\langle\lambda,e^\theta-1\rangle}],$$

where $(\mu,\theta) \in \mathbf{R}^{2m} \times \mathbf{R}^m$.

As for the Black-Scholes model, Tables 3 and 4 present different values obtained for the Delta and Gamma of a European call. We use the same number of runs of the Monte Carlo and Robbins Monro algorithms, and $m = 100$ for the discretization. For the Robbins-Monro algorithm we again set $(\mu_0, \theta_0) = (0, 0)$ as initial value and choose $M = 100$ for the truncation which is larger than the last example due to the stochastic volatility (as in [2]).

INSERT TABLES 3 AND 4 ABOUT HERE

Again, we obtain more accurated values of the Greeks and the StdRatio takes values roughly from 2 to 4. Finally, in Figures 3 and 4 we also observe better convergences to the reference value.

INSERT FIGURES 3 AND 4 ABOUT HERE

3.2.3 Barndorff-Nielsen and Shephard model

The Barndorff-Nielsen and Shephard (BNS) model was introduced in [5] and can be used, for example, to fit high-frequency stock price data. We consider this model with no leverage effect. In this case, it is shown in Nicolato and Venaros [23] that there exists a risk neutral measure P on $(\Omega, \mathcal{F})$, under which the asset price S_t writes as

$$dS_t = S_t(rdt + \sigma_t dB_t), \quad S_0 = x,$$

where $r > 0$ is the constant market interest rate, B_t is a standard Brownian motion, and the stochastic volatility σ_t^2 is given by the Lévy-Ornstein-Uhlenbeck process

$$d\sigma_t^2 = -\lambda\sigma_t^2 dt + dZ_{\lambda t}, \quad \sigma_0^2 > 0,$$

where $\lambda > 0$ is the mean-reversion rate, and Z_t is a pure jump Lévy process with no drift and positive increments, independent of the Brownian motion, with Lévy measure ν having density w with respect to the Lebesgue measure, that is, $\nu(dz) = w(z)dz$.

Consider now a square integrable payoff $h(S_T)$, where h is a locally integrable functions whose set of discontinuities has Lebesgue measure zero. Then, Benth and *al.* [6] obtain formulas for the Greeks for a contingent claim

$$F = e^{-rT}\mathrm{E_P}[h(S_T)].$$

For example, they show that the Delta of this model is given by

$$\Delta = e^{-rT}\mathrm{E_P}\left[h(S_T)\pi\right], \quad \pi = \frac{1}{xT}\int_0^T \frac{1}{\sigma_t} dB_t.$$

We assume that the invariant distribution of σ_t^2 is a Gamma distribution with parameters $\gamma > 0$ and $\eta > 0$. In this case, it is shown in [23] that Z_t is the compound Poisson process of Example 2.2.4.

We apply the variance reduction technique of Section 2. As pointed out in Example 2.2.4, this is a constrained algorithm, so we need to analyze the numerical behavior of the gradient when applying the Robbins-Monro algorithm.

As for the Heston model, we consider the Euler discretization scheme given by

$$\begin{aligned} S_{t_{i+1}} &= S_{t_i} + rS_{t_i}\delta + \sigma_{t_i} S_{t_i} X_i \sqrt{\delta}, \\ \sigma^2_{t_{i+1}} &= \sigma^2_{t_i} - \lambda\sigma^2_{t_i}\delta + \sum_{k=1}^{\Delta N_i} J_k^i, \end{aligned}$$

where $0 = t_0 < t_1 < \cdots < t_m = T$, $\delta = t_{i+1} - t_i = \frac{T}{m}$, $(X_1, \ldots, X_m)$ is a sequence of m iid Normal(0,1) random variables, $(\Delta N_i, i = 1, ..., m)$ is a sequence of m iid Poisson$(\delta\gamma\lambda)$ random variables, and $(J_k^i, i = 1, ..., m)$ is a sequence of m iid exponential random variables with parameter η.

Therefore, the function we need to minimize is

$$H(\mu, \theta) = e^{-rT}\mathrm{E_P}[h^2(S_T)\pi^2 e^{-\langle \mu, B_T\rangle + \frac{1}{2}|\mu|^2 T - \langle \theta, Z_T\rangle + T\sum_{i=1}^m \frac{\gamma\theta_i}{\eta - \theta_i}}],$$

where $(\mu, \theta) \in \mathbf{R}^m \times \Lambda_2^m$, and

$$\Lambda_2 = \left\{ u \in \mathbf{R} : -\frac{\eta}{q} < u < \eta \right\}.$$

The Robins-Monro algorithm is initialized at $(\mu_0, \theta_0) = (0, 0)$ and we choose $M = 100$ for the truncation. Table 5 contains the different values obtained for the Delta of a European call, where we again observe a clear gain in the accuracy of the estimation and the reduction of the variance.

INSERT TABLE 5 ABOUT HERE

Finally, Figures 5 presents the convergence of one of the cases in the table, and we have numerically checked that the gradient evaluated at the Robbins-Monro sequence (μ_n, θ_n) indeed converges to zero.

INSERT FIGURE 5 ABOUT HERE

4 Conclusions

In this paper we have developed a procedure to reduce the variance when numerically computing the Greeks obtained via the Malliavin calculus for jump-diffusion models with stochastic volatility. These include Heston with jumps and the Barndorff-Nielsen and Shephard models. The procedure is validated theoretically and outperforms the usual Monte Carlo methodology. The simulation study shows that the standard deviation gets substantially diminished and that we obtain more accurate values for the Greeks. Although the implementation of this algorithm adds an extra computational cost to the numerical computation of the Greeks, the numerical examples in Section 3 show that the gain in variance reduction compensates the relatively small additional cost.

References

[1] Alaya, M.B., Hajji, K. and Kebaier A. (2016) Importance sampling and statistical Romberg method for Lévy processes, *Stochastic Processes and their Applications*, **126**, 1901-1931. DOI: 10.1016/j.spa.2015.12.008.

[2] Arouna, B. (2004) Robbins-Monro algorithms and variance reduction in finance, *Journal of Computational Finance*, **7**, 35–61. DOI: 10.21314/JCF.2003.111.

[3] Arouna, B. (2004) Adaptative Monte Carlo method, a variance reduction technique, *Monte Carlo Methods and Applications*, **10**, 1–24. DOI: 10.1515/156939604323091180.

[4] Bavouzet, M.P., and Messaoud, M. (2006) Computation of Greeks using Malliavin's calculus in jump type market models, *Electronic Journal of Probability*, **11**, 276–300. DOI: 10.1214/EJP.v11-314.

[5] Barndorff, O.E., and Shephard, N. (2001) Non-Gaussian Ornstein-Uhlenbeck-based models and some of their use in financial economics, *Journal of the Royal Statistical Society: Series B*, **63**, 167–241. DOI: 10.1111/1467-9868.00282.

[6] Benth, F.E., Groth, M. and Wallin, O. (2010) Derivative-free Greeks for the Barndorff-Nielsen and Shephard stochastic volatility model, *Stochastics*, **82**, 291–313. DOI: 10.1080/17442501003629554.

[7] Benth, F.E., Di Nunno, G., and Khedher, A. (2011) Robustness of option prices and their deltas in markets modelled by jump-diffusions. *Communications in Stochastic Analysis*, **5**, 285–307.

[8] Broadie, M., and Kaya, 0. (2004) Exact simulation of option Greeks under stochastic volatility and jump diffusion models. *Proceedings of the 2004 Winter Simulation Conference*, **2**, 1607–1615. DOI: 10.1109/WSC.2004.1371506.

[9] Chen, H.F., and Zhu, Y.M. (1986) Stochastic approximation procedure with randomly varying truncations. *Scientia Sinica Series A*, **29**, 914–926.

[10] Chen, H.F., Guo, L. and Gao, A.J. (1988) Convergence and robustness of the Robbins-Monro algorithm truncated at randomly varying bounds. *Stochastic Processes and their Applications*, **27**, 217–231. DOI: 10.1016/0304-4149(87)90039-1.

[11] Cufaro Petroni, N. and Sabino, P. (2013) Multidimensional quasi-Monte Carlo Malliavin Greeks. *Decisions in Economics and Finance*, **36**, 199–224. DOI:10.1007/s10203-011-0125-z

[12] Davis, M.H.A. and Johansson, M.P. (2006) Malliavin Monte Carlo Greeks for jump diffusions, *Stochastic Processes and Their Applications*, **116**, 101–129. DOI:10.1016/j.spa.2005.08.002.

[13] El-Khatib, Y., and Privault, N. (2004) Computations of Greeks in a market with jumps via the Malliavin calculus, *Finance and Stochastics*, **8**, 161–179. DOI:10.1007/s00780-003-0111-6.

[14] Fournier, E., Lasry, J. M., Lebuchoux, J., Lions, P.L. and Touzi, N. (1999) Application of Malliavin calculus to Monte Carlo methods in finance, *Finance and Stochastics*, **3**, 391–412. DOI:10.1007/s007800050068.

[15] Fournier, E., Lasry, J. M., Lebuchoux, J. and Lions, P.L. (2001) Application of Malliavin calculus to Monte Carlo methods in finance II, *Finance and Stochastics*, **5**, 201–236. DOI:10.1007/PL00013529

[16] Fang, F., and Oosterlee, C. W. (2009) A novel pricing method for European options based on Fourier-cosine series expansions. *SIAM Journal on Scientific Computing*, **31**, 826–848. DOI:10.1137/080718061.

[17] Giles, M.B. (2008) Multilevel Monte Carlo path simulation, *Operations Research*, **56**, 607–617. DOI 10.1287/opre.1070.0496.

[18] Glasserman, P. (2004) *Monte Carlo Methods in Financial Engineering*. Springer-Verlag, New-York. DOI: 10.1007/978-0-387-21617-1.

[19] Kawai, R. (2008) Optimal importance sampling parameter search for Lévy processes via stochastic approximation, *SIAM Journal on Numerical Analysis* , **47**, 293–307. DOI. 10.1137/070680564.

[20] Kawai, R. (2008) Adaptive Monte Carlo variance reduction for Lévy processes with two-time-scale stochastic approximation, *Methodology and Computing in Applied Probability*, **10**, 199–223. DOI:10.1007/s11009-007-9043-5.

[21] Kawai, R., and Takeuchi, A. (2013) Computation of Greeks for asset price dynamics driven by stable and tempered stable processes, *Quantitative Finance*, **13**, 1303–1316. DOI: 10.1080/14697688.2011.589403.

[22] Kohatsu-Higa, A. and Montero, M. (2004) Malliavin calculus in finance, *In: Handbook of computational and numerical methods in finance*, Birkhäuser Boston, 111–174.

[23] Nicolato, E. and Vernardos, E. (2003) Option pricing in stochastic volatility models of the Ornstein-Uhlenbeck type, *Mathematical Finance*, **13**, 445–466. DOI: 10.1111/1467-9965.t01-1-00023.

[24] Su, Y. and Fu, M.C. (1998) Optimal importance sampling in securities pricing, *Journal of Computational Finance*, **5**, 27–50.

5 Tables and Figures

σ	K/S_0	**DeltaMC**	**DeltaRef**	**DeltaRM**	**StdRatio**
0.1	0.8	0.9376	0.9510	0.9504	1.0406
	1	0.8168	0.8115	0.8104	3.5242
	1.2	0.2093	0.2089	0.2091	6.1567
0.3	0.8	0.8538	0.8536	0.8520	4.0177
	1	0.6522	0.6540	0.6534	5.8366
	1.2	0.4258	0.4295	0.4287	6.1948

Table 1: European call under Black-Scholes model with jumps. Parameters: S_0=50, $r = 0.05$, T=1, $\alpha = 0.5$, $\lambda = 0.1$, $\beta = 1000$.

σ	K/S_0	**GammaMC**	**GammaRef**	**GammaRM**	**StdRatio**
0.1	0.8	-0.0009	0.0001	0.0003	1.3707
	1	0.0115	0.0114	0.0113	2.2844
	1.2	0.0383	0.0366	0.0365	7.2398
0.3	0.8	0.0045	0.0052	0.0050	2.6818
	1	0.0105	0.0104	0.0104	3.8576
	1.2	0.0132	0.0127	0.0127	6.7382

Table 2: European call under Black-Scholes model with jumps. Parameters: S_0=100, $r = 0.1$, T=1, $\alpha = 0.5$, $\lambda = 0.1$, $\beta = 1000$.

λ	K/S_0	**DeltaMC**	**DeltaRef**	**DeltaRM**	**StdRatio**
0.1	0.8	0.5653	0.5531	0.5508	3.2562
	1	0.4545	0.4460	0.4475	3.9337
	1.2	0.3512	0.3476	0.3491	3.7939
1	0.8	0.5598	0.5551	0.5598	2.1040
	1	0.4911	0.4852	0.4887	2.2720
	1.2	0.4188	0.4300	0.4268	2.5998

Table 3: European call for Heston with jumps. Parameters: $S_0 = 50, r = 0.05, \kappa = 2, \theta = 0.4, \sigma = 0.2, \rho = 0.5, \alpha = 0.5, T = 1, \beta = 1, V_0 = 0.04$.

λ	K/S_0	**GammaMC**	**GammaRef**	**GammaRM**	**StdRatio**
0.1	0.8	0.0171	-0.0173	-0.0171	2.1517
	1	-0.0088	-0.0091	-0.0088	2.4984
	1.2	-0.0043	-0.0035	-0.0038	2.8132
1	0.8	-0.0241	-0.0246	-0.0250	1.6967
	1	-0.0173	-0.0173	-0.0173	1.8948
	1.2	-0.0119	-0.0117	-0.0119	1.9958

Table 4: European call for Heston with jumps. Parameters: $S_0 = 50, r = 0.05, \kappa = 0.1, \theta = 0.4, \sigma = 0.3, \rho = 0.5, \alpha = 0.5, T = 1, \beta = 1, V_0 = 0.05$.

K/S_0	**DeltaMC**	**DeltaRef**	**DeltaRM**	**StdRatio**
0.8	0.8134	0.8006	0.8090	3.7985
1	0.5909	0.5965	0.5886	3.3446
1.2	0.3981	0.3838	0.3853	3.5117

Table 5: European call for Barndorff-Nielsen and Shephard model. Parameters: $S_0 = 100, r = 0.01, \eta = 1, \sigma_0^2 = 0.1, \gamma = 3, T = 1, \beta = 1, \lambda = 0.1$.

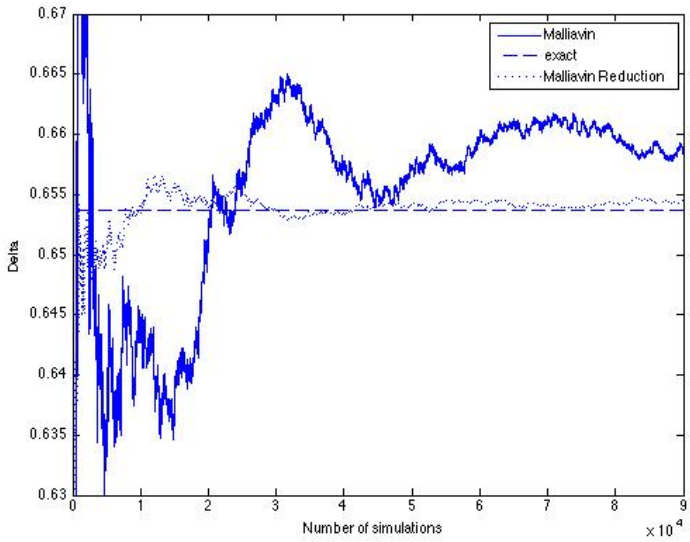

Figure 1: European call under Black-Scholes model with jumps. Parameters: $S_0 = 50$, $K = 50$, $r = 0.05$, $T = 1$, $\sigma = 0.3$, $\alpha = 0.5$, $\lambda = 0.1$, $\beta = 1$.

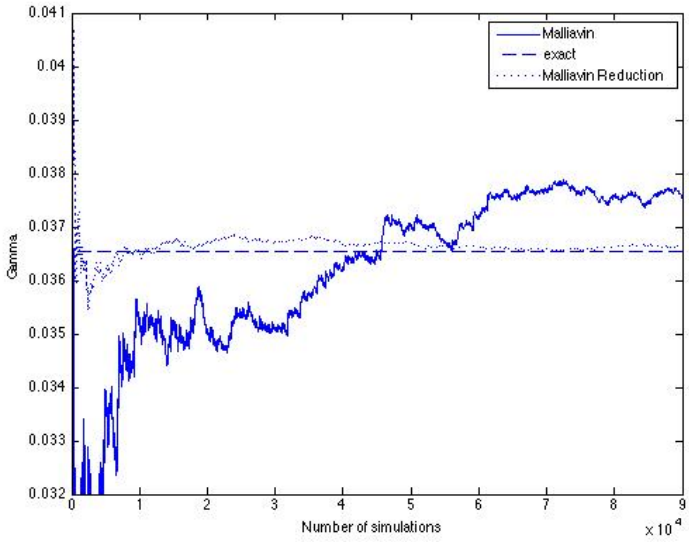

Figure 2: European call under Black-Scholes model with jumps. Parameters: $S_0 = 100$, $K = 120$, $r = 0.1$, $T = 1$, $\sigma = 0.1$, $\alpha = 0.5$, $\lambda = 0.1$, $\beta = 100$.

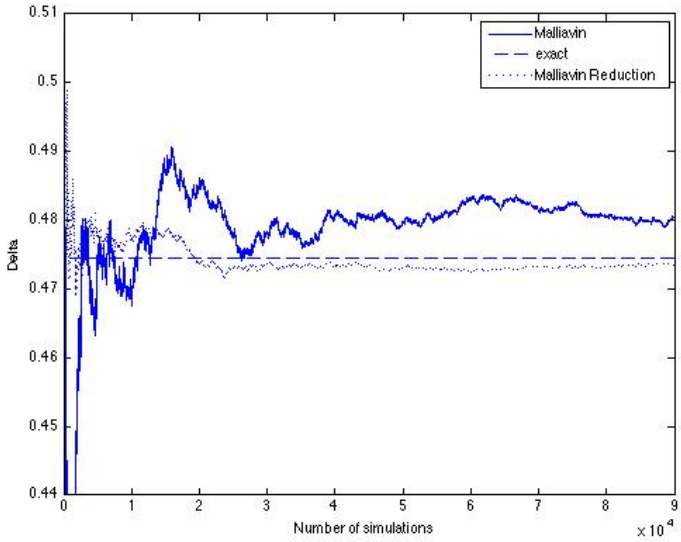

Figure 3: European call under Heston model with jumps. Parameters: $S_0 = 50$, $K = 50$, $r = 0.05$, $\kappa = 2$, $\rho = 0.5$, $\alpha = 0.5$, $\theta = 0.4$, $V_0 = 0.04$, $\beta = 10$, $\lambda = 0.1$, $T = 1$, $\sigma = 0.1$.

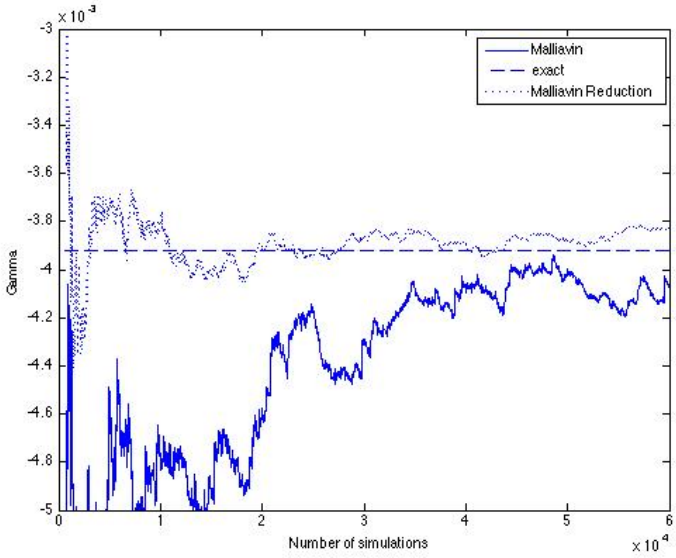

Figure 4: European call for Heston model with jumps. Parameters $S_0 = 50$, $K = 60$, $r = 0.05$, $\kappa = 2$, $\rho = 0.5$, $\alpha = 0.5$, $\theta = 0.4$, $V_0 = 0.04$, $\beta = 1$, $\lambda = 0.1$, $T = 1$, $\sigma = 0.2$.

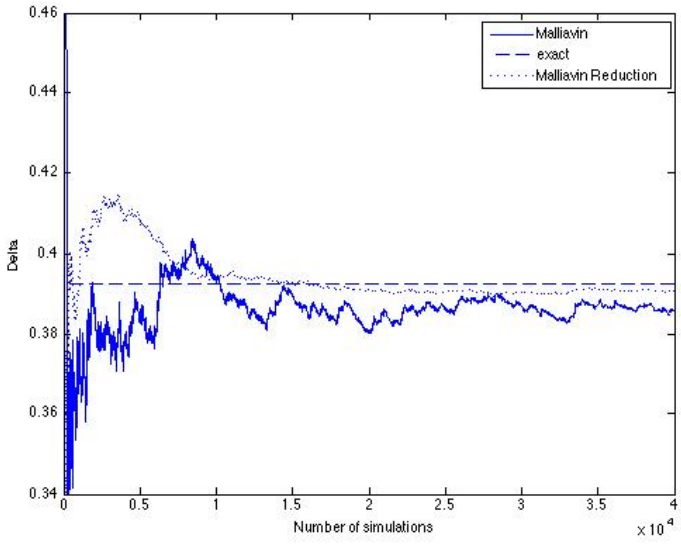

Figure 5: European call for Barndorff-Nielsen and Shephard model. Parameters $S_0 = 100$, $K = 120$, $\sigma_0^2 = 0.1$, $r = 0.01$, $T = 1$, $\lambda = 0.1$, $\gamma = 3$, $\eta = 1$, $\beta = 1$.