Efficient calculation of the Greeks for exponential Lévy processes: An application of Measure Valued Differentiation

Georg Ch. Pflug[1] and Philipp Thoma [2]

Abstract

Monte Carlo simulation methods have become more and more important in the financial sector in the past years. In this paper, we introduce a new simulation method for the estimation of the derivatives of prices of financial contracts with respect to (w.r.t.) certain distributional parameters, called the "Greeks". In particular, we assume that the underlying financial process is a Lévy-type process in discrete time.

Our method is based on the Measure Valued Differentiation (MVD) approach, which allows to represent derivatives as differences of two processes, called the phantoms. We discuss the applicability of MVD for different types of option payoffs in combination with different types of models of the underlying and provide a framework for the applicability of MVD for path-dependent payoff functions, as Lookback Options or Asian Options.

Keywords: Lévy-Processes, estimation of the Greeks, Measure Valued Differentiation, exotic options, Lookback option, Asian options.

[1]University of Vienna, Department of Statistics and Operations Research, Oskar Morgenstern Platz 1, 1090 Vienna, Austria and IIASA, Laxenburg; georg.pflug@univie.ac.at

[2]University of Vienna. This author was supported by the Jubilaeumsfondsprojekt AB13377 of the Austrian National Bank; Philipp.Thoma@oenb.at

1   Introduction

In Finance, simulation methods have become standard for pricing various, but especially exotic financial contracts under various model assumptions. This paper focuses on the estimation of sensitivities of option prices w.r.t. changes of distributional parameters, the so called "Greeks", when the underlying stochastic model follows some Markov process of exponential Lévy-type. Price sensitivities are used in measuring and managing risk and find also application in hedging-strategies.

In the classical Black-Scholes framework, there exist closed formulas for the Greeks of plain Vanilla European Call or Put Options. But if we turn to more complicated payoff functions (eg Exotic Options) or if we consider other driving Lévy processes than Brownian motion, numerical approaches relying on Monte Carlo methods are needed to estimate both, the values and their sensitivities.

Suppose that $S_\theta(t)$ describes the random price process of the underlying and H is the payoff function of an option on this underlying with maturity T. θ is some parameter governing the distribution of the process $S_\theta(\cdot)$. The current (time 0) fair option price is

$$\mathbb{E}_Q[e^{-rT}H(S_\theta(T))],$$

where Q is a probability law which makes the discounted process

$$\tilde{S}_\theta(t) = e^{-rt}S_\theta(t) \tag{1.1}$$

a martingale. Our goal is to calculate

$$\frac{\partial}{\partial\theta}\mathbb{E}_Q[e^{-rT}H(S_\theta(T))]. \tag{1.2}$$

Since closed form expressions do not exist (except for very special cases), estimations of (1.2) with high accuracy are needed.

There exist several Monte Carlo techniques for estimating the Greeks. Each of those methods has its individual advantages and disadvantages, depending on the model of the stock price and the structure of the payoff function.

The most simple is the *finite difference (FD) approximation*: The price is estimated under $\theta + h$ and under θ and the difference of the estimates is divided by h. Even if the variance of the FD estimate is reduced by taking highly correlated estimates, there is still the bias issue: FD estimates are biased. The more sophisticated methods are subdivided into methods which

assume parameterized integrands

$$\theta \mapsto \int H(\theta, \omega) d\nu(\omega)$$

and methods with parametrized integrators

$$\theta \mapsto \int H(\omega) d\nu_\theta(\omega),$$

see Pflug (1996). The *pathwise method (Infinitesimal Pertubation Analysis IPA)* is a method of the first sort, see Glasserman (2004). Other direct methods belong to the second group, as the *Likelihood Ratio method* or *Score Function method*, see Glasserman (2004), the *Malliavin calculus*, (Sanz-Sole (2005), Montero and Kohatsu-Higa (2003)), *Algorithmic Differentiation*, Capriotti (2011) and the *Measure Valued Differentiation*, (Heidergott et al (2010), Heidergott and Vasquez-Abad (2006), Heidergott et al (2008), Heidergott and Leahu (2010), Vasquez-Abad and Heidergott (2008)).

The main contribution of this article is to demonstrate how the MVD method may be applied for a variety of different stock-price models and different, notably nondifferentiable payoff functions. While pathwise derivation methods work well for smooth payoff functions, they fail for nonsmooth ones. On the other hand, likelihood ratio methods require the knowledge of the density. Often, the explicit form of the density is unknown. Glasserman and Liu (2010) present an approximative method, which approximated the density from its characteristic function or cumulant function. This *saddle point approximation* results in a certain bias, which has to be controlled. Our measure-valued differentiation method is generally unbiased and does not require the knowledge of the density or the characteristic function of the Markov transition operator.

In this paper, we treat exclusively models with discrete time, a generalization for continuous processes will be treated in a subsequent paper. It is organized as follows. In Section 2 we discuss some Markov models for the underlying, in particular Lévy-type models. Section 3 introduces the measure valued differentiation method and its extension for Markov processes. Section 4 comprised some numerical examples about the performance of the estimation method for path-independent payoffs. The last section demonstrates how the method can be applied to the sensitivity estimation for path-dependent payoffs, such as Lookback and Asian Options.

2   Exponentials of Lévy processes

In this paper, we assume that the price process of the underlying is Markovian. While our methodology applies to any Markov processes, which can be unbiasedly simulated in discrete steps (such as continuous time processes with discrete random jumps), we restrict the presentation and the examples to exponentials of Lévy processes, since these processes are widely used models for stocks prices in Finance.

A Lévy process has independent and stationary increments. It is fully characterized by the characteristic function φ of its (infinitely divisible) increment distribution

$$\mathbb{E}\left[\exp\left(iu\left(X(t+1) - X(t)\right)\right)\right] \;=\; \varphi(u).$$

By the Lévy-Khintchine formula, φ has to be of the form

$$\varphi(u) = \exp\left[iau - \frac{\sigma^2}{2}u^2 + \int_{\mathbb{R}-\{0\}} (e^{iux} - 1 - iux\mathbf{1}_{\{|x|<1\}})\, d\nu(x)\right]$$

see eg Schoutens (2003). By the property of infinite divisibility, the characteristic function of the increment in time interval Δ is just the Δ-th power of φ

$$\mathbb{E}\left[\exp\left(iu\left(X(t+\Delta) - X(t)\right)\right)\right] \;=\; (\varphi(u))^{\Delta}.$$

We assume that the price process of the underlying is an *exponential of a Lévy process*, ie is modeled as

$$S(t) = S(0)\cdot\exp(X(t))$$

with $X = \{X(t), t \geq 0\}$ being a Lévy process with $X(0) = 0$ and $S(0)$ is today's price. Notice that both processes $X(t)$ and $S(t)$ are Markovian.

Change of measure.   The fundamental theorem of asset pricing states that the absence of arbitrage opportunity is equivalent to the existence of a (not necessarily unique) equivalent measure $\mathbb{Q}$ under which $\tilde{S}(t) = e^{-rt}S(t)$ is a martingale.

Exponentials of Lévy processes typically allow several equivalent martingale measures and one has to select one for pricing purposes. In literature, one finds the following selection principles: (1) the Mean Correcting Martingale Measure, (2) the Minimal Martingale Measure, (3) the Variance Optimal Martingale Measure, (4) the Esscher Martingale Measure, (5) the Minimal Entropy Martingale Measure, (6) the Utility Based Martingale Measure (see

the overview by Miyahara (2005). In this paper, we concentrate on the mean correcting martingale measure, which can be obained in a very simple way by just changing the parameter values of the original Lévy process, see Schoutens (2003). Notice that we consider only discrete-time processes, for which the mean-correcting martingale measure is automatically equivalent.

2.1 Examples of Lévy processes

We now turn to some examples of Lévy-processes modeling the logprice of stocks.

The Geometric Brownian Motion (GBM). The GBM is used in the Black Scholes model. The pertaining logprice process is a Brownian motion with constant drift μ and constant diffusion σ^2

$$X(t) = \left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W(t)$$

where $W(t)$ is a standard Brownian motion. Thus the increment distribution is

$$X(t+1) - X(t) \sim N\left(\left(\mu - \frac{1}{2}\sigma^2\right), \sigma^2\right)$$

with characteristic function $\varphi(u) = \exp[iu(\mu - \sigma^2/2) - \sigma^2 u^2/2]$. The stock prices then follow the process

$$S(t) = S(0)\exp\left[\left(\mu - \frac{1}{2}\sigma^2\right)t + \sigma W(t)\right]$$

and under the (unique) equivalent martingale measure the discounted process $\tilde{S}(t)$ is distributed as

$$\tilde{S}(t) \sim e^{-rt}S(0)\exp\left[(r - \frac{1}{2}\sigma^2)t + \sigma W(t)\right]. \tag{2.1}$$

The Gamma model. The Gamma(a, b)-process is the Lévy process with increment distribution

$$X(t+1) - X(t) \sim \text{Gamma}(a, b),$$

ie has a distribution with density $\frac{1}{b^a \Gamma(a)}x^{a-1}\exp(-x/b)\mathbf{1}_{\{x \geq 0\}}$ and characteristic function $\varphi(u) = (1 - iub)^{-a}$.

The Variance Gamma model. If $X_1(\cdot)$ is a Gamma(a, b_1)-process and $X_2(\cdot)$ is an independent Gamma(a, b_2) process, then the difference

$$X(t) = X_1(t) - X_2(t)$$

5

is a Variance Gamma process (Schoutens (2003)), a process whose increment distribution has characteristic function

$$\varphi(u) = (1-iub_1)^{-a}(1+iub_2)^{-a} = (1-iu(b_1-b_2)+u^2b_1b_2)^{-a} = (1-iub+u^2c)^{-a}.$$

Here we have set $b = b_1 - b_2$ and $c = b_1 \cdot b_2$. We denote this process by $\mathrm{VG}(a,b,c)$. Under the mean-correcting martingale measure mentioned earlier the discounted process satisfies

$$\tilde{S}(t) \sim e^{-rt}e^{t(r+a\log(1-b-c))}S(0)\exp(X(t)).$$

The Compound Poisson model. The Compound Poisson process is defined as

$$X(t) = \sum_k^{N(t)} Z_k \ \ t \geq 0.$$

where $N(t)$ is a Poisson process with intensity λ and (Z_k) are i.i.d. jump sizes, which are independent from $N(t)$, (Schoutens (2003)). The increment distribution is

$$X(t+1) - X(t) \sim \sum_k^{N(1)} Z_k$$

with characteristic function

$$\varphi(u) = \exp\left(\lambda\left(\varphi_Z(u) - 1\right)\right).$$

where φ_Z is the characteristic function of the jump distribution.

We consider in particular jump distributions which are mixtures of Gamma distributions on the positive resp. negative half line, ie

$$Z_k = \begin{cases} V_k^+ & \text{with probability } p \\ V_k^- & \text{with probability } 1-p \end{cases}$$

with

$$\begin{aligned} V_k^+ &\sim \mathrm{Gamma}(a_1, b_1) \\ V_k^- &\sim \mathrm{Gamma}(a_2, b_2). \end{aligned}$$

independent of each other. This jump distribution has characteristic function $\varphi_Z(u) = p(1 - iub_1)^{-a_1} + (1 - p)(1 + iub_2)^{-a_2}$.

For this process we write $\mathrm{CP}(\lambda, a_1, b_1, a_2, b_2, p)$. If $X(\cdot) \sim \mathrm{CP}(\lambda, a_1, b_1, a_2, b_2, p)$, then the martingale corresponding to its exponential satisfies

$$\tilde{S}(t) \sim e^{-rt}e^{mt}S(0)\exp(X(t))$$

with $m = r - \lambda\left(p(1 - b_1)^{-a_1} + (1 - p)(1 + b_2)^{-a_2} - 1\right).$

3 The Measure Valued Differentiation method

3.1 Introduction

The Measure Valued Differentiation method (MVD) is based on a weak notion of differentiability of probability measures. Let θ be the parameter of interest. We want to calculate or estimate

$$\frac{d}{d\theta}\mathbb{E}[H(S_\theta)] = \int H(\omega)\frac{d}{d\theta}\nu_\theta(d\omega).$$

where ν_θ is the distribution of S_θ under θ.

Definition. Let $\mathcal{H}$ be a set of mappings $H : \mathbb{R} \mapsto \mathbb{R}$ which are absolutely integrable with respect to ν_θ for any $\theta \in \Theta \subseteq \mathbb{R}, \mathcal{H} \subset L^1(\nu_\theta, \Theta)$, the family of all functions which are integrable w.r.t. all $\nu_\theta, \theta \in \Theta$. A function $\theta \mapsto \nu_\theta$ mapping an open subset of $\mathbb{R}$ into the family of all probability measures is called $\mathcal{H}$-differentiable (weakly differentiable w.r.t. $\mathcal{H}$), if a finite signed measure ν'_θ exists, such that

$$\forall H \in \mathcal{H}: \quad \lim_{h\to 0}\frac{1}{h}\left(\int H(s)\nu_{\theta+h}(ds) - \int H(s)\nu_\theta(ds)\right) = \int H(s)\nu'_\theta(ds).$$

Let c_θ be a constant and ν_θ^+ and ν_θ^- two probability measures such that

$$\int H(s)\nu'_\theta(ds) = c_\theta\left(\int H(s)\nu_\theta^+(ds) - \int H(s)\nu_\theta^-(ds)\right)$$

for all $H \in \mathcal{H}$, then the triplet $(c_\theta, \nu_\theta^+, \nu_\theta^-)$ is called a *weak derivative triplet* of ν_θ.

The probability measures ν_θ^+ and ν_θ^- can be obtained by decomposing $d\nu_\theta/d\theta$ into a difference between two densities. Such a decomposition can always be found by the Jordan-Hahn decomposition of signed measures, but any other decomposition may also do the job.

Let us summarize the method: Suppose that S_θ is distributed according to ν_θ and that $H \in \mathcal{H}$. Let us further assume that $\theta \mapsto \nu_\theta$ is weakly differentiable with triplet $(c_\theta, \nu_\theta^+, \nu_\theta^-)$. Then the fundamental equation

$$\frac{\partial}{\partial\theta}\mathbb{E}[H(S_\theta)] = c_\theta\left[\mathbb{E}\left[H(S_\theta^+)\right] - \mathbb{E}\left[H(S_\theta^-)\right]\right] \tag{3.1}$$

holds, provided that the random variables realize the weak derivative, ie they satisfy $S_\theta^+ \sim \nu_\theta^+$ and $S_\theta^- \sim \nu_\theta^-$. In our terminology, we call the random variables S_θ^+ resp. S_θ^- the *positive* resp. *negative realization* of the MVD.

There are numerous papers, where the theory of MVD has been established, (Pflug (1996), Heidergott and Vasquez-Abad (2006), Heidergott and Hordijk (2004), Vasquez-Abad and Heidergott (2008)). The interchange of expectation and differentiation is widely justified, because probability densities tend to be smooth functions of their parameters. The condition for unbiasedness - the integrability condition- is a simple growth condition on the payoff function H, see (Heidergott et al (2010)).

3.2 Examples of Weak Derivative triplets

Alternatively to the differentiation of measures, one may also differentiate the characteristic functions: Suppose that φ_θ is the characteristic function of ν_θ and suppose further that its derivative w.r.t. θ can be written as

$$\frac{\partial}{\partial \theta}\varphi_\theta(u) = c_\theta(\varphi_\theta^+(u) - \varphi_\theta^-(u))$$

where φ_θ^+ and φ_θ^- are characteristic functions as well. Then we have found the $\mathcal{H}$-weak derivative triplet, where $\mathcal{H}$ is the space of all bounded functions, which are continuous ν_θ-almost everywhere. For extension, suppose that $\mathcal{H}_p$ is the space of all ν_θ-a.e. continuous functions, which do not grow faster than $|x|^p$. If we assume that $(\theta, u) \mapsto \varphi_\theta(u)$ is jointly differentiable in both variables, but at least p-times w.r.t. u, then weak $\mathcal{H}_p$-differentiability holds.

Before listing some examples, we review important relationships between some probability distributions.

- If $X \sim \mathrm{Gamma}(1/2, 2\sigma^2)$ (ie a multiple of a χ^2 distribution with one degree of freedom), then $\sqrt{X}$ is distributed as the absolute value of a $N(0, \sigma^2)$ distribution.

- If $X \sim \mathrm{Gamma}(1, 2\sigma^2)$, (ie a multiple of a χ^2 distribution with two degrees of freedom), then $\sqrt{X}$ is distributed according to a $\mathrm{Raleigh}(\sigma^2)$ distribution, which is a Weibull distribution with exponent 2. The $\mathrm{Raleigh}(\sigma^2)$ distribution has density $\frac{x}{\sigma^2}\exp(-\frac{x^2}{2\sigma^2})\mathbf{1}_{\{x \geq 0\}}$ and characteristic function

$$\varphi(u) = 1 + iu\sigma\sqrt{2\pi}\exp\left(-\frac{u^2\sigma^2}{2}\right)\Phi(iu\sigma),$$

  where Φ is the analytic continuation of the normal Gaussian distribution function.

Distribution ν_θ	c_θ	Positive Part ν_θ^+	Negative Part ν_θ^-
Poisson(θ)	1	Poisson(θ)+1	Poisson(θ)
Normal(θ, σ^2)	$1/\sigma\sqrt{2\pi}$	θ+Rayleigh(σ)	θ−Rayleigh(σ)
Normal(μ, θ^2)	$1/\theta$	ds-Maxwell(μ, θ^2)	Normal(μ, θ^2)
Gamma(a, θ)	a/θ	Gamma($a+1, \theta$)	Gamma(a, θ)
Gamma(a, b)+ θ	$1/b$	Gamma(a, b) + θ	Gamma($a-1, b$) + θ
Exponential(θ)	$1/\theta$	Exponential(θ)	Erlang(θ,2)

Table 1: Examples of Weak derivatives. θ is the varying parameter.

- If $X \sim$ Gamma($3/2, 2\sigma^2$), (ie a multiple of a χ^2 distribution with three degrees of freedom), then $\sqrt{X}$ is distributed according to a Maxwell(σ^2) distribution. The Maxwell(σ^2) distribution has density

$$\sqrt{\frac{2}{\pi}} \frac{x^2}{\sigma^3} \exp\left(-\frac{x^2}{2\sigma^2}\right) \mathbf{1}_{\{x \geq 0\}}$$

and characteristic function

$$\varphi(u) = \sqrt{\frac{2}{\pi}} iu\sigma + 2\exp\left(-\frac{u^2\sigma^2}{2}\right) \Phi(iu\sigma)(1 - u^2\sigma^2).$$

The symmetrized and shifted version of it is called double sided Maxwell distribution (ds-Maxwell(μ, σ^2)). It has density

$$\frac{1}{\sigma^3\sqrt{2\pi}} (x-\mu)^2 \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

and characteristic function

$$\varphi(u) = (1 - u^2\sigma^2)\exp\left(iu\mu - \frac{u^2\sigma^2}{2}\right).$$

It is not difficult to identify various weak derivative triplets using the above relationships. The most important examples, to which we refer in this paper are collected in Table 1. We can see that one of the realizations is often distributed as the nominal part, which reduces the computational effort.

Remark 1. For a given continuous transform $\mathbb{T}$, the image measure (push-forward measure) $\nu^{\mathbb{T}}$ of measure ν is defined as

$$\nu^{\mathbb{T}}(A) = \nu(\mathbb{T}^{-1}(A)).$$

If $\theta \mapsto \nu_\theta$ is differentiable in the measure valued sense with derivative triplet $(c_\theta, \nu_\theta^+, \nu_\theta^-)$, then also $\theta \mapsto \nu_\theta^{\mathbb{T}}$ is differentiable and has triplet $(c_\theta, (\nu_\theta^+)^{\mathbb{T}}, (\nu_\theta^-)^{\mathbb{T}})$. To put this into the level of realizations, if the triplet $(c_\theta, X_\theta^+, X_\theta^-)$ realizes $(c_\theta, \nu_\theta^+, \nu_\theta^-)$, then $(c_\theta, \mathbb{T}(X_\theta^+), \mathbb{T}(X_\theta^-))$ realizes $(c_\theta, (\nu_\theta^+)^{\mathbb{T}}, (\nu_\theta^-)^{\mathbb{T}})$.

This fact is especially important when one considers the transformation of a Lévy process to an exponential Lévy process using the transform $\mathbb{T}(w) = S(0) \exp(w)$. If the two random variables X^+ resp. X^- realize the two parts of the MVD of ν_θ, then $S(0) \exp(X^+)$ resp. $S(0) \exp(X^-)$ realize the two parts of the MVD of $\nu_\theta^{\mathbb{T}}$, ie of the exponential model.

3.3 The Measure Valued Derivative for Markov processes

As was already said, we consider only discrete time homogeneous Markov processes with infinitely divisible increment distribution in this paper. Suppose that Δ is the time increment. With a slight abuse of notation, we write now

$$S_\theta(i) \qquad \text{for} \qquad S_\theta(i\Delta)$$

bearing in mind that the maturity time T corresponds to the step n with $T = n \cdot \Delta$.

To the process $S_\theta(i), i = 0, \ldots, n$ we associate the transition operator

$$\mathbb{P}_\theta(w, A) = P\{S_\theta(i+1) \in A | S_\theta(i) = w\},$$

the starting distribution [3] γ and the payoff function H.

Introduce the following notations:

$\gamma \mathbb{P}_\theta$ for the measure $(\gamma \mathbb{P}_\theta)(A) = \int \mathbb{P}_\theta(w, A) \, d\gamma(w)$,
$\mathbb{P}_\theta H$ for the function $(\mathbb{P}_\theta H)(u) = \int H(w) \mathbb{P}_\theta(u, dw)$,
$\mathbb{P}_\theta^2(w, A)$ for the two-step transition $\mathbb{P}_\theta^2(w, A) = \int \mathbb{P}_\theta(v, A) \mathbb{P}_\theta(w, dv)$,
$\mathbb{P}_\theta^n(w, A)$ for the n-step transition.

Using this notation, we write for the expected payoff at maturity time $T = n \cdot \Delta$

$$\mathbb{E}\left[H(S_\theta(n))\right] = \gamma \, \mathbb{P}_\theta^n \, H.$$

Definition. The Markov transition $\mathbb{P}_\theta(\cdot, \cdot)$ is called weakly $\mathcal{H}$-differentiable, if there is a signed transition $\mathbb{P}_\theta'$ such that for all functions $h \in \mathcal{H}$ and every point mass δ_w (ie the probability distribution concentrated on the point w)

$$\frac{1}{s}|\delta_w \mathbb{P}_{\theta+s} h - \delta_w \mathbb{P}_\theta h - s \cdot \delta_w \mathbb{P}_\theta' h| \to 0$$

---

[3]While γ is in most cases just the point mass $\delta_{S(0)}$ at $S(0)$, we allow here some slight generalization.

as $s \to 0$.

The finite signed transition $\mathbb{P}_\theta'$ may be decomposed as

$$\mathbb{P}_\theta'(w, A) = c_\theta(w)[\mathbb{P}_\theta^+(w, A) - \mathbb{P}_\theta^-(w, A)], \tag{3.2}$$

where $\mathbb{P}_\theta^+$ and $\mathbb{P}_\theta^-$ are regular Markov transitions. [4]

Remark 2. For exponentials of processes with independent increments, the situation simplifies considerably. Since $S_\theta(t + \Delta) = S_\theta(t) \cdot \exp(X_\theta(t + \Delta) - X_\theta(t))$, by weakly differentiating the increment distribution $X_\theta(t + \Delta) - X_\theta(t)$, giving a positive realization, a negative realization and a constant c_θ, these can be transformed by the exponential transform (see Remark 2) and this triplet does not depend on the previous state $S_\theta(t)$. Thus due to the independent increment property of Lévy processes, the measure valued differentiation of increments can be done very efficiently.

Recall now the Leibnitz rule for the derivation of a power of operators

$$(\mathbb{P}_\theta^n)' = \sum_{i=1}^n \mathbb{P}_\theta^{i-1} \mathbb{P}_\theta' \mathbb{P}_\theta^{n-i}. \tag{3.3}$$

Using the decomposition (3.2) this can be written in terms of MVD

$$\frac{\partial}{\partial \theta} \mathbb{E}\left[H(S_\theta(n))\right] = \gamma \left[\sum_{i=1}^n \mathbb{P}_\theta^{i-1} c_\theta \left(\mathbb{P}_\theta^+ - \mathbb{P}_\theta^-\right) \mathbb{P}_\theta^{n-i}\right] H \tag{3.4}$$

for bounded, continuous H, see Heidergott et al (2010).

Formula (3.4) can be used in different ways to construct unbiased estimates of $\frac{\partial}{\partial \theta} \gamma \mathbb{P}_\theta^n H$ as we introduce them below: the exact estimate MVDe, the randomized estimate MVDr and the compromise estimate MVD(k). The randomized estimators make use of the identity

$$\sum_{i=1}^n a_i = \mathbb{E}[n a_\tau]$$

for a random τ, which is uniformly distributed on the integers $1, \ldots, n$. For a compact presentation, we explain the compromise estimate MVD(k) first, since the other estimates are special cases. The estimate MVD(k) generates k phantom pairs (see below), where k is a divisor of n.

The compromise estimate MVD(k). Suppose that $n = k \cdot q$.

---

[4]Notice that $c(\cdot)$ can be chosen in a measurable way, if (3.2) is the decomposition in the positive and negative part. If another decomposition is chosen, one has to require measurability.

1. Sample k random uniform times $\tau_1, ..., \tau_k$ in $(1, ..., q)$, ..., $(q \cdot (i-1) + 1, ..., q \cdot i)$, ..., $((k-1) \cdot q + 1, ..., k \cdot q = n)$

2. Sample $S_\theta(0)$ from the starting distribution γ.

3. Sample n steps with transition $\mathbb{P}_\theta$, giving $S_\theta(0), \dots, S_\theta(n)$

4. For all $i = \tau_1, ..., \tau_k$ sample one transition step from $S_\theta(i-1)$ with transition $\mathbb{P}_\theta^+$ and one with transition $\mathbb{P}_\theta^-$, giving $S_{\theta,i-1}^+(i)$ resp.$S_{\theta,i-1}^-(i)$ using a coupling technique (see below). Store $c_\theta(i)$.

5. Continue these processes $S_{\theta,i-1}^+(l)$ resp.$S_{\theta,i-1}^-(l)$, $l = i+1, ..., n$ using transition $\mathbb{P}_\theta$. It is advisable to use for both phantom processes the same multiplicative increments as they were generated in step 2. The processes $(S_{\theta,i-1}^+(\cdot), S_{\theta,i-1}^-(\cdot))$ are called *positive phantoms*, resp. *negative phantoms*, together they are the *phantom pairs*.

6. The unbiased estimate is

$$m \cdot \sum_{i=1}^{k} c_\theta(i) \left[H(S_{\theta,i-1}^+(n)) - H(S_{\theta,i-1}^-(n)) \right]. \qquad (3.5)$$

7. The final estimate is the arithmetic mean of N independent replications of estimate (3.5).

The MVD(k) algorithm includes the extreme cases: The *exact algorithm* MVDe=MVD(n), where $k = n$, ie all n possible starting points for the phantom process are considered; and the *fully randomized algorithm* MVDr=MVD(1), where only one random starting point is considered.

MVDr tends to have a higher variance, because of introducing an additional random variable τ. On the other hand, the MVDe algorithm has smaller variance but because of summing up n elements, the computational effort is much higher. The MVD(k) algorithm is a compromise between the randomized estimator and the exact estimator.

The extreme algorithms MVDe and MVDr illustrated in Figure 3.3.

3.4   Variance reduction via coupling

The variance of a MVD estimator of the form $c \cdot (H(S^+) - H(S^-))$ is given by

$$c^2 \cdot \left(Var \left[H(S^+) \right] + Var \left[H(S^-) \right] - 2Cov \left[H(S^+), H(S^-) \right] \right).$$

An appropriate choice of the positive and negative phantoms may cause positive correlation and therefore lead to a reduced variance of the estimator.
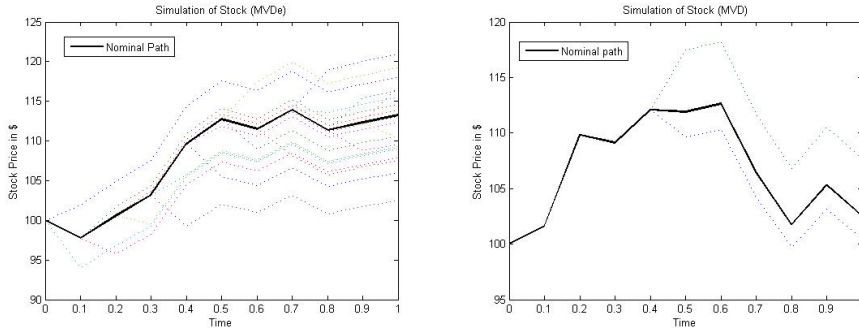
Figure 3.1: Left: MVDe with $t = 1$, $n = 10$ (all phantom pairs); Right: MVDr, with $\tau = 4$ (one randomly selected phantom pair).

Example. Sensitivity estimation w.r.t. μ in the Gaussian case:

Consider a normal distributed random variable $X \sim N(\mu, \sigma^2)$. When calculating the sensitivity w.r.t. μ, the variable splits up into $X^+ \sim \mu + \text{Rayleigh}(\sigma)$ and $X^- \sim \mu - \text{Rayleigh}(\sigma)$ according to the weak derivative triplet, see Table 1. For coupling we have just to use the same $\text{Rayleigh}(\sigma)$ for the positive and negative part, see Heidergott et al (2008).

Example. Sensitivity w.r.t. σ in the Gaussian case:

Consider $X \sim N(\mu, \sigma^2)$ and we are interested in the sensitivity w.r.t. the volatility σ. According to Table 1 we have that $X^+ \sim \mu + \text{ds-Maxwell}(0, \sigma^2)$ and $X^- \sim \mu + N(0, \sigma^2)$. According to Heidergott et al (2008), we use following coupling method: First simulate $V \sim \text{ds-Maxwell}(0, \sigma^2)$ and set $X^+ = \mu + V$. Then sample an independent uniform random variable U on $(0, 1)$ and set $X^- = \mu + UV$. Then X^- is normally distributed with parameters μ and σ and is positively correlated with X^+.

Example. Sensitivity w.r.t. λ in the Poisson case:

As it was already shown, a pair realizing the weak derivative in the Poisson case can be chosen as $X^- \sim \text{Poisson}(\lambda)$ and $X^+ = X^- + 1$.

Example. Sensitivity w.r.t. b in the Gamma case:

If X^- is generated according to $\text{Gamma}(a, b)$, then $X^+ \sim \text{Gamma}(a+1, b)$ may be generated as $X^+ = X^- + V$, where $V \sim \text{Exponential}(1/b)$.

4   Numerical results

In this section we present numerical results of sensitivity estimates. In particular, we consider two types of payoff functions:

- The plain vanilla call payoff

$$H(s) = \max(s - K, 0),$$

- The digital payoff

$$H(s) = \mathbf{1}_{\{s > K\}}.$$

Since the second payoff is not differentiable, the pathwise differentiation method fails in this case. For comparison, we have also considered the plain vanilla payoff, showing the the results are reliable also in the differentiable case.

Remark 3. Notice that in case that the derivative is taken w.r.t. the interest rate r, ie if $\theta = r$, also the derivative w.r.t. the discount factor e^{-rT} has to be taken into account. In this case

$$\frac{\partial}{\partial r} \mathbb{E}[e^{-rT} H(S_r(n))] = -Te^{-rT} \mathbb{E}[H(S_r(n))] + e^{-rT} \frac{\partial}{\partial r} \gamma \, \mathbb{P}_r^n \, H. \qquad (4.1)$$

If however the sensitivity w.r.t. some other parameter is searched for, we have just to calculate

$$\frac{\partial}{\partial \theta} \mathbb{E}[e^{-rT} H(S_\theta(n))] = e^{-rT} \frac{\partial}{\partial \theta} \gamma \, \mathbb{P}_\theta^n \, H. \qquad (4.2)$$

In all the numerical examples we show, the time unit is 1 year and the maturity time was set to $T = 1$. The elementary time step was set to one trading day thus we chose $n = 252$ and $\Delta = 1/n$.

Remark 4. When considering the mean-correcting martingale measure, one has to take also a derivative w.r.t. the exponential of the location parameter into account. For instance, using the previous result about derivatives w.r.t location parameter of the Gamma distribution, we obtain that the weak derivative of $\mathbb{E}[H(S(n)e^{m(\theta)})]$ w.r.t. θ, where $S(n) = \exp[X(n)]$, with $X(n) \sim$ Gamma(a, b) is

$$\frac{m(\theta)'}{b}[H(S(n)e^{m(\theta)}) - H(\exp(V^-)e^{m(\theta)})]$$

where $V^- \sim$ Gamma$(a - 1, b)$.

4.1   The Geometric Brownian Motion

4.1.1   Rho (ρ): Sensitivity w. r. t. r

$$\rho = \frac{\partial}{\partial r} \mathbb{E}\left[e^{-rT} H(S_r(T))\right].$$

Renaming the process $S_\theta(i\Delta)$ simply $S_\theta(i)$ we arrive at $\log S(i+1) - \log S(i) = V$ with $V \sim N(n^{-1}(r - \sigma^2/2), n^{-1}\sigma^2)$. According to the weak derivative triplet of a Normal distribution, see Table 1, the positive resp. negative realizations of the weak derivative are

$$[\log S(i+1) - \log S(i)]^+ = (r - \sigma^2/2)/n + V^+$$
$$[\log S(i+1) - \log S(i)]^- = (r - \sigma^2/2)/n - V^-$$

with $V^+ = V^- \sim \text{Rayleigh}(\sigma^2/n)$ and $c_r = \frac{1}{\sigma\sqrt{2\pi n}}$.

The Rho (ρ) for a Plain Vanilla Option. It is well known that the exact value of Rho of an option on a non-dividend-paying stock can be calculated as

$$\rho = KTe^{-rT}\Phi(d_2)$$

with

$$d_2 = \frac{\log(S(0)/K) + (r - \sigma^2/2)T}{\sigma\sqrt{T}},$$

where $\Phi(x)$ is the standard normal cumulative probability function, see Hull (2005).

Table 2 shows the performance of IPA (=Pathwise method), the FD (=Finite Difference method), MVDr (=randomized phantom estimator), MVDe (=Measure Valued Differentiation with total differentiation) and MVD(6) (Measure Valued Differentiation with differentiation at 6 random points) when calculating the ρ for a plain Vanilla Option, whereas the Measure Valued Derivatives were computed according to (4.1).

A single simulation of each estimator is based on $N = 200$ replications of the stockprice path. [5] To compare the computational effort of the methods we use the "work-normalized variance" (WNV) which is given by the product of the variance and the expected work per run, see Glynn and Whitt (1992).

In this case, the IPA method performs better, but MVD(6) is not much worse and is applicable for a much larger class of payoff functions.

The Rho (ρ) for a Digital Option. With the same inputs as we had before for the Vanilla Option, we obtain the results shown in Table 3 for the ρ of the Digital Option. The value of BS-Rho was taken out of www.mathfinance.de/optioncalculator.php.

4.1.2 Delta (Δ): Sensitivity w. r. t. $S(0)$

$$\Delta = \frac{\partial}{\partial S(0)} \mathbb{E}\left[e^{-rT}H(S_{S(0)}(T))\right].$$

---

[5]Throughout the paper each estimator was simulated 300 times to get the arithmetic mean of the estimator, the arithmetic mean of the computational time and the variance.

	$\widehat{\rho}$	Variance	computation time	WNV
exact value	56.379			
IPA	56.058	11.090	0.794	8.809
MVDr	55.901	20.736	0.824	17.088
MVDe	56.673	11.051	3.533	39.036
MVDc(6)	56.393	15.195	0.923	14.027
FD	63.135	1006.390	1.574	1584.51

Table 2: The Rho for a Vanilla Option in the BS-model, $N = 200$, $S(0) = 100$, $r = 0.01$, $\sigma = 0.05$, $K = 100$.

	$\widehat{\rho}$	Variance	computation time	WNV
exact value	7.15			
MVDr	7.070	4.508	0.783	3.528
MVDe	7.135	2.015	4.459	8.986
MVDc(6)	7.162	2.752	0.869	2.391
FD	6.37	23.773	1.5298	36.368

Table 3: The Rho for a Digital Option in the BS-model, $N = 200$

Notice that the parameter $S(0)$ only influences the first step

$$\log S(1) \sim N\left(\log S(0) + \left(r - \frac{1}{2}\sigma^2\right)\frac{1}{n}, \sigma^2 \frac{1}{n}\right).$$

According to the weak derivative triplet of a normal distribution

$$
\begin{aligned}
[\log S(1)]^+ &= \log S(0) + (r - \sigma^2/2)/n + V^+ \\
[\log S(1)]^- &= \log S(0) + (r - \sigma^2/2)/n - V^-
\end{aligned}
$$

with $V^+ = V^- \sim \text{Rayleigh}(\sigma^2/n)$ and $c_{S(0)} = \frac{\sqrt{n}}{S(0)\sigma\sqrt{2\pi}}$.

Since only the first transition is parameter-sensitive, all three presented algorithms coincide in this case.

The Delta (Δ) for a Plain Vanilla Option. The exact value of Delta of an option on a non-dividend-paying stock can be calculated by

$$\Delta = \Phi(d_1)$$

with

$$d_1 = \frac{\log(S(0)/K) + (r + \sigma/2)T}{\sigma\sqrt{T}}$$

	$\widehat{\Delta}$	Variance	computation time	WNV
exact value	0.589			
IPA	0.588	0.0013	0.7819	0.0010
MVDr	0.594	0.0022	0.7612	0.0017
FD	0.950	1122.625	1.498	1682.507

Table 4: The Delta for a Vanilla Option in the BS-model, $N = 200$.

	$\widehat{\Delta}$	Variance	computation time	WNV
Delta	0.078			
MVDr	0.078	0.0005	0.760	0.0004
FD	-0.238	22.844	1.519	34.707

Table 5: The Delta for a Digital Option in the BS-model, $N = 200$.

where $\Phi(x)$ is the cumulative probability function for a standard normal variable, see Hull (2005).

With the same inputs as we had before while calculating the Rho(ρ), we obtain the results shown in Table 4 for the Delta (Δ) of the Vanilla Option.

The Delta (Δ) for Digital Options. With the same inputs as we had before for the Vanilla Option, we obtain the results shown in Table 5 for the Δ of the Digital Option. The value of BS-Delta was taken out of www.mathfinance.de/optioncalculator.php.

4.1.3 Vega (ν): Sensitivity with respect to σ

The realizations V, V^+, V^- are the same as for Rho. With probability p,

$$p = \frac{\frac{|\mu'_\sigma|}{\sigma_\sigma \sqrt{2\pi}}}{\frac{|\mu'_\sigma|}{\sigma_\sigma \sqrt{2\pi}} + \frac{\sigma'_\sigma}{\sigma_\sigma}} = \frac{|\mu'_\sigma|}{|\mu'_\sigma| + \sigma'_\sigma \sqrt{2\pi}} = \frac{\frac{\sigma}{n}}{\frac{\sigma}{n} + \sqrt{\frac{2\pi}{n}}}.$$

the positive and negative parts are

$$[\log S(i+1) - \log S(i)]^+ = (r - \sigma^2/2)/n - V^+$$
$$[\log S(i+1) - \log S(i)]^- = (r - \sigma^2/2)/n + V^-$$

	$\widehat{\nu}$	Variance	computation time	WNV
exact value	38.897			
IPA	38.923	19.393	0.789	15.293
MVDr	39.031	6889.428	0.771	5311.487
MVDe	38.715	83.938	5.635	473.001
MVDc(6)	38.943	1328.352	0.862	1144.891
FD	38.428	1186.400	1.5324	1818.002

Table 6: The Vega for a Vanilla Option in the BS-model, $N = 200$.

and with probability $1 - p$

$$\begin{aligned}
[\log S(t_{i+1}) - \log S(t_i)]^+ &= (r - \sigma^2/2)/n + W \\
[\log S(t_{i+1}) - \log S(t_i)]^- &= (r - \sigma^2/2)/n + UW.
\end{aligned}$$

with $W \sim$ ds-Maxwell$(0, \sigma^2/n)$ and $U \sim$ Uniform$[0, 1]$. For c_σ we have

$$c_\sigma = |\mu'_\sigma| \frac{1}{\sigma_\sigma \sqrt{2\pi}} + \frac{\sigma'_\sigma}{\sigma_\sigma} = \frac{|\mu'_\sigma| + \sigma'_\sigma \sqrt{2\pi}}{\sigma_\sigma \sqrt{2\pi}} = \frac{\frac{\sigma}{n} + \sqrt{\frac{2\pi}{n}}}{\sigma \sqrt{\frac{2\pi}{n}}}.$$

The Vega (ν) for a Plain Vanilla Option.

The exact value for Vega for an European Call or Put Option on a non-dividend-paying stock is defined by

$$\nu = S(0)\sqrt{T}\Phi(d_1)$$

with

$$d_1 = \frac{\log(S(0)/K) + (r + \sigma^2/2)T}{\sigma\sqrt{T}},$$

with Φ being the standard normal distribution, see Hull (2005), p. 336. With the input arguments of the previous examples, we get results shown in Table 6.

4.2   The Variance Gamma model

4.2.1   Sensitivity w.r.t. b_1

The log returns follow $\log S(i+1) - \log S(i) = V_1 - V_2$ with $V_1 \sim$ Gamma $(a/n, b_1)$, $V_2 \sim$ Gamma $(a/n, b_2)$ and

$$\begin{aligned}
[\log S(i + 1) - \log S(i)]^+ &= V_1^+ - V_2 \\
[\log S(i + 1) - \log S(i)]^- &= V_1 - V_2
\end{aligned}$$

	Estimator	Variance	computation time	WNV
MVDr	-27.64	108.59	9.80	1064.71
MVDe	-27.12	3.05	64.47	30.76
FD	11.69	14429.34	18.44	266077.03

Table 7: Sensitivity w.r.t. b_1 of a Vanilla Option in the VG-model, $N = 200$.

with $V_1^+ \sim \text{Gamma}(a/n + 1, b_1)$ and $c_{b_1} = \frac{a}{b_1 n}$. Under the mean-correcting martingale measure we have to calculate also the derivative w.r.t. the exponential of the correcting location parameter. According to Remark 4, we calculate

$$
\begin{aligned}
[\log S(T) - \log S(0)]^+ &= W_1 - W_2 \\
[\log S(T) - \log S(0)]^- &= W_1^- - W_2
\end{aligned}
$$

with

$$
\begin{aligned}
W_1 &\sim \text{Gamma}(a, b_1) \\
W_2 &\sim \text{Gamma}(a, b_2) \\
W_1^- &\sim \text{Gamma}(a - 1, b_1)
\end{aligned}
$$

and $c = m'_{b_1}/b = \frac{-a}{b_1(1 - b_1)}$.

Sensitivity w.r.t. b_1 for a Plain Vanilla Call Option.

In a simulation with input parameters: $r = 0.01$, $a = 1$, $b_1 = 0.01$, $b_2 = 0.01$, $S(0) = 100$, $K = 100$, $n = 252$, $h = 0.01$ we get results shown in Table 7.

4.3   The Compound Poisson model

4.3.1   Sensitivity w.r.t. b_1

In the case of a Compound Poisson model the log returns follow

$$
\log S(i + 1) - \log S(i) = \sum_{k=1}^{N(\frac{1}{n})} (p \cdot V_1 - (1 - p) \cdot V_2)
$$

	Estimator	Variance	computation time	WNV
MVDr	31.104	116.727	0.259	30.28
MVDe	31.351	5.642	0.924	5.22
FD	31.176	1148.64	0.244	279.95

Table 8: Sensitivity w.r.t. b_1 of a Vanilla Option in the CP-model, $N = 200$.

with $V_1 \sim \text{Gamma}(a_1, b_1)$, $V_2 \sim \text{Gamma}(a_2, b_2)$. Further

$$[\log S(i+1) - \log S(i)]^+ = \sum_{k=1}^{N(\frac{1}{n})} \left(p \cdot V_1^+ - (1-p) \cdot V_2 \right)$$

$$[\log S(i+1) - \log S(i)]^- = \sum_{k=1}^{N(\frac{1}{n})} \left(p \cdot V_1 - (1-p) \cdot V_2 \right)$$

with $V_1^+ \sim \text{Gamma}(a_1 + 1, b_1)$ and $c_{b_1} = a_1/b_1$.
Sensitivity w.r.t. b_1 for a Plain Vanilla Call Option.

A simulation with input parameters: $r = 0.01$, $\lambda = 10$, $a_1 = 0.01$, $b_1 = 0.01$, $a_2 = 0.01$, $b_2 = 0.01$, $p = 0.5$, $S(0) = 100$, $K = 100$, $n = 252$, $h = 0.01$ yields results according to Table 8.

5 MVD for path-dependent payoff functions

In this section, we demonstrate how the MVD method can be modified to allow the sensitivity estimation of Lookback and Asian Options. These options are *path dependent*, ie the payoff depends on the whole history of the underlying and not just on its value at maturity.

5.1 Lookback Options

A Lookback Option is based on the maximum value of the underlying. Let $S_\theta(\cdot)$ be the Markovian price process of the underlying. Define the maximum process as $M_\theta(i) = \max_{k \leq i} S_\theta(k)$. The payoff of a Lookback Option is $H(M_\theta(n))$ and its fair price is
$\mathbb{E}[e^{-rT} H(M_\theta(n))]$. Notice that the maximum process M_θ is not Markovian itself, but it is the first component of the Markovian pair $(M_\theta(\cdot), S_\theta(\cdot))$. While the evolution of the second component is as before and given by the transition

of $X_\theta = \log(S(t))$, the transition of the first components is given by

$$M_\theta(1) = S_\theta(1)$$

$$M_\theta(i+1) = \begin{cases} S_\theta(i+1) & \text{if } S_\theta(i+1) > M_\theta(i) \\ M_\theta(i) & \text{otherwise.} \end{cases}$$

Let $\mathbb{P}_\theta^{(M)}$ be the transition operator of this two-dimensional process. We aim an unbiased estimate

$$\frac{\partial}{\partial\theta}\mathbb{E}[e^{-rT}H(M_\theta(n))] = \frac{\partial}{\partial\theta}e^{-rT}\gamma\left(\mathbb{P}_\theta^{(M)}\right)^n H.$$

Here - of course - H applies only to the first component of the Markov process.

The MVDr algorithm for the Lookback Option.

1. Sample random uniform time τ in $\{1,...,n\}$. Sample $S_\theta(1),...,S_\theta(\tau)$ and calculate $M_\theta(1),\ldots,M_\theta(\tau)$.

2. Then at time τ do
   Use the positive resp. negative part of the derivative of $\mathbb{P}_\theta$ to make one transition from $S_\theta(\tau-1)$ to $S_\theta^+(\tau)$ and $S_\theta^-(\tau)$ respectively. Store c_θ.
   Let $M_\theta^+(\tau) = \max(M(\tau-1), S_\theta^+(\tau))$ and $M_\theta^-(\tau) = \max(M(\tau-1), S_\theta^-(\tau))$.

3. Continue the processes $S_\theta^+(\cdot)$ and $S_\theta^-(\cdot)$ using the transition $\mathbb{P}_\theta$ and calculate $(M_\theta^+(\ell), S_\theta^+(\ell))$ resp. $(M_\theta^-(\ell), S_\theta^-(\ell)$ for $\ell = \tau+1,\ldots,n$ with

$$M_\theta^+(\ell) = \begin{cases} S_\theta^+(\ell) & \text{if } S_\theta^+(\ell) > M_\theta^+(\ell-1) \\ M_\theta^+(\ell-1) & \text{otherwise} \end{cases}$$

resp.

$$M_\theta^-(\ell) = \begin{cases} S_\theta^-(\ell) & \text{if } S_\theta^-(\ell) > M_\theta^-(\ell-1) \\ M_\theta^-(\ell-1) & \text{otherwise} \end{cases}$$

4. The unbiased estimate is

$$n \cdot c_\theta \cdot [H(M_\theta^+(n)) - H(M_\theta^-(n))]. \tag{5.1}$$

5. The final estimate is the arithmetic mean of N independent replications of estimate (5.1).

To see the correctness of the algorithm notice that $\mathbb{T}(x) = \max(x,c)$ is continuous therefore by Remark 2 it may be composed with the weak derivative pair in the given manner.

As before, the MVDr algorithm constructs just one phantom pair. In complete analogy to the standard option case, the MVDe resp. the MVD(k) estimates can be constructed based on n resp. k pairs of phantoms.

	Estimator	Variance	computation time	WNV
Rho	52.90			
MVDr	52.849	19.557	0.761	14.873
MVDe	52.933	5.236	5.620	29.430
MVDc(6)	52.947	6.481	0.853	5.528

Table 9: The Rho for a Lookback Option in the BS-model , $N = 200$.

	Estimator	Variance	computation time	WNV
MVD	1.01	0.0014	0.76	0.0011
FD	1.40	1256.46	1.54	1933.6

Table 10: The Delta for a Lookback Option in the BS-model , $N = 200$.

Example (BS-model: The Rho (ρ) for a Lookback Option). In a simulation, with $S(0) = 100, r = 0.01, \sigma = 0.05, K = 100, T = 1$ and a sample size of $N = 200$ we get results shown in Table 9, whereas the Measure Valued Derivatives were computed according to 4.1, just replacing $S_r(n)$ by $M_r(n)$. The value for Rho was taken from www.wolframalpha.com.

Example (BS-model: The Delta (Δ) for a Lookback Option). In a simulation, with $S(0) = 100$, $r = 0.01$, $\sigma = 0.05$, $K = 100$ and a sample size of $N = 200$ we get results shown in Table 10, where S^+ respectively S^- were computed as described on page 21, and were implemented in the algorithm for Lookback Options to obtain M^+ respectively M^-.

Example (VG-model: Sensitivity w.r.t. b_1 for a Lookback Option). Table 11 shows the results with following input parameters: $T = 1$, $S(0) = 100$, $K = 100$, $r = 0.01$, $a = 1$, $b_1 = 0.01$, $b_2 = 0.01$, $n = 252$, $h = 0.01$ (for FD).

5.2 Asian Options

An Asian Option is based on the average value of the underlying. Let $S_\theta(\cdot)$ be the Markovian price process of the underlying. Define the average process as $A_\theta(k) = \frac{1}{k} \sum_{i=1}^{k} S_\theta(i)$. The payoff of an Asian Option is $H(A_\theta(n))$ and its fair price is $\mathbb{E}[e^{-rT} H(A_\theta(n))]$. Notice that the average process A_θ is not Markovian itself, but it is the first component of the Markovian pair $(A_\theta(\cdot), S_\theta(\cdot))$.

	Estimator	Variance	computation time	WNV
MVDr	85.27	51.99	9.75	506.74
MVDe	84.58	2.47	20.99	51.93
FD	89.32	217.70	18.53	4034.30

Table 11: Variance Gamma process (Lookback), Sensitivity w.r.t. b_1, $N = 200$

While the evolution of the second component is as before and given by the transition of $X_\theta = \log(S(t))$, the transition of the first components is given by

$$A_\theta(i) = \frac{i}{i+1} A_\theta(i-1) + \frac{1}{i+1} S_\theta(i). \tag{5.2}$$

Let $\mathbb{P}_\theta^{(A)}$ be the transition operator of this two-dimensional process. We aim an unbiased estimate

$$\frac{\partial}{\partial \theta} \mathbb{E}[e^{-rT} H(A_\theta(n))] = \frac{\partial}{\partial \theta} e^{-rT} \gamma \left(\mathbb{P}_\theta^{(A)}\right)^n H.$$

Here - of course - H applies only to the first component of the Markov process.

The MVDr algorithm for the Asian Option.

1. Sample random uniform time τ in $\{1, ..., n\}$. Sample $S_\theta(1), ..., S_\theta(\tau-1)$ and calculate $A_\theta(1), \ldots, A_\theta(\tau-1)$.

2. Then at time τ do: Use the positive resp. negative part of the derivative of $\mathbb{P}_\theta$ to make one transition from $S_\theta(\tau-1)$ to $S_\theta^+(\tau)$ and $S_\theta^-(\tau)$ respectively. Store c_θ. Let

$$A_\theta^+(\tau) = \frac{1}{\tau+1} \left(\underbrace{\tau \cdot A(\tau-1)}_{\sum_{i=0}^{\tau-1} S_\theta(i)} + S_\theta^+(\tau) \right)$$

resp.

$$A_\theta^-(\tau) = \frac{1}{\tau+1} \left(\underbrace{\tau \cdot A_\theta(\tau-1)}_{\sum_{i=0}^{\tau-1} S_\theta(i)} + S_\theta^-(\tau) \right).$$

3. Continue the processes $S_\theta^+(\cdot)$ and $S_\theta^-(\cdot)$ using the transition $\mathbb{P}_\theta$ and calculate $(A_\theta^+(\ell), S_\theta^+(\ell))$ resp. $(A_\theta^-(\ell), S_\theta^+(\ell)$ for $\ell = \tau+1, \ldots, n$ with

$$A_\theta^+(\ell) = \frac{\ell}{\ell+1} \left(A_\theta^+(\ell-1) + \frac{1}{\ell} S_\theta^+(\ell) \right)$$

23

resp.

$$A_\theta^-(\ell) = \frac{\ell}{\ell+1}\left(A_\theta^-(\ell-1) + \frac{1}{\ell}S_\theta^-(\ell)\right).$$

4. The unbiased estimate is

$$n \cdot c_\theta\left[H(A_\theta^+(n)) - H(A_\theta^-(n))\right]. \qquad (5.3)$$

5. The final estimate is the arithmetic mean of N independent replications of estimate (5.3).

Also in this case, one may use the variants MVDe and MVD(k).

Example (BS-model: The Rho (ρ) for an Asian Option). The payoff function of a plain vanilla Asian Call Option is defined as

$$H(S(T)) = e^{-rT}\max\left(0, \frac{1}{T}\sum_{t=1}^{T}S(t) - K\right).$$

With $S(0) = 100$, $r = 0.01$, $\sigma = 0.05$, $K = 100$, $T = 1$ we get results shown in Table 12, where the value for the BS-Rho was taken from www.wolframalpha.com.

	Estimator	Variance	computation time	WNV
BS-Rho	27.111			
MVDr	27.30	8.12	0.7617	6.18
MVDe	27.14	2.80	6.3196	17.68
MVDc(6)	26.99	3.01	0.8968	2.70

Table 12: BS-Rho (Asian) via Average process, $N = 200$.

6 Relation to Portfolio Risk Optimization and Conclusion

The "Greeks" can be seen as measures of the risk against changes of parameters of the underlying processes. In a portfolio hedging situation, one may try to compose a portfolio, which is as little sensitive w.r.t. parameter changes as possible. Suppose that there are m contracts in a portfolio, where the value $V^{(i)}$ of the i-th contract depends on parameters $(\theta_1, \ldots, \theta_\ell)$. Denote the parameter sensitivity w.r.t. θ_j of the i-th contract by $V^{(i)}_{\theta_j}$. With the

portfolio weights x_i, the objective is to make $\sum_{i=1}^{m} x_i V_{\theta_j}^{(i)}$ as close to 0, e.g. by minimizing

$$\left(\sum_{i=1}^{m} x_i V_{\theta_j}^{(i)} \right)^2.$$

One sees that this requires to estimate $V_{\theta_j}^{(i)}$, which is the topic of this paper.

We have presented the fundamentals of Measure Valued Differentiation (MVD) and provided new MVD-type algorithms to derive the "Greeks" for Lévy-type processes in discrete time. We have tested MVD within the classical Black-Scholes framework, obtaining consistent results for the Greeks.

The strength of MVD solidifies in the case of discontinuous payoff functions (ie Digital Option) and for Lévy type, where pathwise differentiation fails. In cases where MVD and the Score Function method can be applied, MVD has typically smaller variance, see Heidergott at al (2008). Also, for MVD the knowlegde of the explicit form of the transition density is not required in contract to the SF method. With the compromise estimator MVD(k) one can balance computational effort against precision.

References

Capriotti, L. (2011). Fast Greeks by algorithmic differentiation. *The Journal of Computational Finance* Vol. 14, No. 3, pp. 3-35

Fu, M. Variance-Gamma and Monte Carlo. In: Advances in Mathematical Finance (M. Fu, R.A. Jarrow, J.-Y. Yen and R.J. Elliott eds.) pp. 21-35, Birkhauser, Boston, 2007 Glasserman, P. (2004). Monte Carlo Methods in Financial Engineering. Springer Verlag

Glasserman, P. and Liu, Z. (2010). Estimating Greeks in simulating Lévy-driven models. *The Journal of Computational Finance*, Vol. 14, No. 2, pp. 3-55

Glynn, P. and Whitt, W. (1992). The asymptotic efficiency of simulation estimators. *Operations Research*, Vol. 40, No. 3, pp. 505-520

Heidergott, B. and Hordijk, A. (2004). Single-run gradient estimation via measure-valued differentiation. *Automatic Control, IEEE Transactions*, Vol.49, No.10, pp. 1843- 1847

Heidergott, B. and Leahu, H. (2010) Weak differentiability of product measures. *Mathematics of Operations Research*, Vol. 35, pp. 27-51

Heidergott, B. and Vazquez-Abad, F.J. (2006). Measure-valued differentiation for random horizon problems. *Markov Processes Related Fields*, Vol. 12, pp. 509-536

Heidergott, B., Vazquez-Abad, F.J., Pflug, G.Ch. and Farenhorst-Yuan, T. (2010). Gradient estimation for discrete-event systems by measure-valued differentiation. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, Vol. 20 (1), Association for Computing Machinery

Heidergott, B., Vazquez-Abad, F.J. and Volk-Makarewicz, W. (2008). Sensitivity estimation for gaussian systems. *European Journal of Operations Research*, Vol. 187, pp. 193-207

Hull, J.C. (2005). *Fundamentals of Futures and Options Markets*. Pearson Education, fifth edition

Miyahara, Y. (2005). *Martingale measures for the geometric Lévy models.* Graduate School of Economics Nagoya City University. Montero, M. and Kohatsu-Higa, A. (2003). *Malliavin Calculus applied to Finance.* Physica, Vol. 320, pp. 548-570

Pflug, G.Ch. (1996). *Optimization of Stochastic Models.* Kluwer Academic Publishers

Sanz-Sole, M. (2005). *Malliavin Calculus.* EPFL Press

Schoutens, W (2003). *Lévy Processes in Finance.* Wiley Series in Probability and Statistics

Vazquez-Abad, F.J. and Heidergott, B. (2008). Measure-valued differentiation for Markov Chains. *Journal of Optimization Theory and Applications*, Vol. 136, pp. 187-209